Shayne C. Gad

Integrated Safety and Risk Assessment for Medical Devices and Combination Products

Springer

Integrated Safety and Risk Assessment for Medical Devices and Combination Products

Shayne C. Gad

Integrated Safety and Risk Assessment for Medical Devices and Combination Products

Springer

Shayne C. Gad
Principal
Gad Consulting Services
Raleigh, NC, USA

ISBN 978-3-030-35240-0 ISBN 978-3-030-35241-7 (eBook)
https://doi.org/10.1007/978-3-030-35241-7

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Contents

Chapter 1
Introduction: History and Where We Are Headed

Shayne C. Gad

Abstract While medical devices have been derived and used since at least ancient Egypt, means of verifying their biologic safety to patients (biocompatibility) and regulations requiring and governing such pre-use evaluation (testing) are much more recent. Less than a century has seen the modern approach, with testing dictated by type, and duration, of patient contact are much more recent. Such requirements first arise in the 1960s due to concerns with materials migrated from a device into the patient body. The science and complexity of testing involved are continuously evolving (accelerated by concerns as to the safety of silicones in the late 1980's) and have also served to drive the growth of the medical device market (now nearly a third the size of the pharmaceutical market) and the innovations and complexity of devices and device/drug combinations.

Keywords Adverse effects on patients · Biocompatibility · Biodegradation of material or device · Breast implants · Center for Devices and Radiological Health · Constituent materials in the device · Cooper Committee · Cumulative duration of contact · Dalkon Shield · Dr. John Autian · Food, Drug, and Cosmetic Act · IDE · Leachables · Medical Device Amendments · Medical device industry · Patient contact mode and duration · Patient exposure parameters · Safe Medical Devices Act · Tripartite

The medical device industry in the Unites States and worldwide is immense in its economic impact, scope (between 92,000 and 145,000 different devices are produced in the United States by ~12,000 different manufacturers employing some 370,000 people; it is believed that ~2100 of these manufacturers are development stage companies without products yet on the market), and importance to the health of the world's citizens (The Wilkerson Group 2013; MDDI 2013; Nugent 1994). The assessment of the safety to patients using the multitude of items produced by this industry is dependent on schemes and methods which are largely particular to

S. C. Gad, PhD, DABT (✉)
Gad Consulting Services, Raleigh, NC, USA
e-mail: scgad@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_1

these kinds of products, not as quantitative or modern as those employed for foods, drugs, and pesticides, and continue to be in a state of flux. Regulation of the preclinical safety evaluation of such devices is, in fact, fairly recent. It is only with the Medical Device Amendments (to the Food, Drug, and Cosmetic Act) of 1992 that devices have come to be explicitly regulated at all and with the Safe Medical Devices Act of 1993, the Medical Device Amendments of 1992, and subsequent laws that regulation of devices for biocompatibility became rigorous. The FDA's publication of their "Use of ISO-10993" document in June of 2016 marks the most recent regulatory guidance (FDA 2016).

The causes behind this timing are reviewed in the case histories presented in the last chapter of this book.

For purposes of this book, the safety we are concerned with is that related to the biological and chemical interactions of devices with patients' bodies and not that due to mechanical or structural malfunction (such as structural failure of heart valves and pacemakers). Such safety, also referred to as biocompatibility, only became of general concern to the public with publicity around plasticizers in devices and increased mortality with cardiovascular stents. Earlier cases of perceived significant risk on the part of devices (the Dalkon Shield intrauterine device, silicones in breast implants, latex present in gloves, and a wide range of other devices) have largely faded from public and professional memory by the beginning of the twenty-first century, to be replaced by phthalates, BPA (bisphenol amine), and heavy metals.

1.1 Biocompatibility

A medical device that is adequately designed for its intended use should be safe for that use. The device should not release any harmful substances into the patient which can lead to adverse effects over the period of patient contact. Some manufacturers believe that biocompatibility is sufficiently indicated if their devices are made of "medical grade material," ASTM standard metals, or materials approved by FDA as direct or indirect food additives. The term "medical grade" does not have an accepted legal or regulatory definition and can be misleading and assigned without biocompatibility testing. Likewise, the existence of a Material Master File (MMF) does not provide any assurance as to what biocompatibility data (or of what quality) is available in the file. More to the point, as the extent of required data and testing is expanded by regulatory antibiotics, what constitutes adequate testing is a moving target as time passes.

There is no universally accepted definition for biomaterial or biocompatibility. Yet the manufacturer who ultimately markets a device will be required by FDA to demonstrate biocompatibility of the product as part of the assurance of its safety and effectiveness. The device manufacturer (and not those providing the constituent materials or parts) is responsible for understanding biocompatibility tests and selecting currently accepted methods which best demonstrate:

- The lack of adverse biological response from the constituent materials in the device
- The absence of adverse effects on patients

Diversity of the materials used, types of medical devices, nature and duration of patient exposure, and potential harms present an enormous challenge to design and conduct well-defined biocompatibility testing programs. Experience gained in one application area is not necessarily transferable to another application. The same applies to different or sometimes slightly different (variable) materials. Biodegradation and interaction of materials complicate safety considerations, as does the increased scope of combination device/drug products (CFR 1992).

Biocompatibility describes the state of a biomaterial within a physiological environment without the material adversely affecting the tissue or (if there is systemic exposure, the body) the body adversely affecting the material. Biocompatibility is the end product of chemical and physical interactions between the material and the tissue/body and the biological response to these reactions. Unlike with drugs or biologics, adverse effects can be due not only to chemical effects but to physical effects associated with surface characteristics of a device (Gad and Gad-McDonald 2015).

Biocompatibility tests are used to predict and therefore avoid significant adverse reactions and establish the absence of any harmful effects of the component material. Such tests help to determine the potential risk which the material may pose to the patient. The proper use of biocompatibility tests can lead to the rejection of potentially harmful materials from use in devices while permitting safe materials to be used for manufacturing the device.

Any biocompatibility statement is useful only when it is considered in the proper context. A statement such as "polycarbonate is biocompatible" lacks precision and can lead to misunderstanding. Any statement of biocompatibility should include information on the type of device, intended conditions of use, degree and duration of patient contact, and the potential of the device to cause harm. Manufacturers should avoid using the term "biocompatible" without clearly identifying the environment in which it is used and any limitations on such use. Conditions of manufacturing, packaging, and cleaning can also be critical.

The need for biocompatibility testing and the extent of such testing that should be performed depend on numerous factors which are presented and considered in Chap. 2. These factors include the type of device, intended use, liability, degree and duration of patient contact, nature of the components, nature of potentially expressed patient population (does it include pediatric patients), and potential of the device to cause harm (Gad and Gad-McDonald 2015). There are no universal tests to satisfy all situations, and there is no single test which can predict biological performance of the material or device and reliably predict the safety of the device. The types and intended uses of medical devices determine the types and number of tests required to establish biocompatibility. Biological tests should be performed under the condition which simulates the actual use of the product (including sterilization mode and packaging) or material as closely as possible and should demonstrate the biocompatibility of a material or device for a specific intended use or range of uses. These tests will be more extensive for a new material than for those materials that have an established history of long and safe uses.

All materials used in the manufacture of a medical device should be considered for evaluation of their suitability for intended use if they have direct or indirect patient contact (DiSilvo 2009). Consideration should always be given to the possibility of the

release of toxic substances from the base materials, as well as any contaminants which might remain after the manufacturing process or sterilization. The extent of these investigations will vary depending on previously known information (prior art) and initial screening tests.

1.1.1 Fundamentals of Biocompatibility Tests

Biocompatibility is generally demonstrated by tests utilizing fundamental toxicological principles which provide information on the potential toxicity of materials in the clinical application. Many classical toxicological tests, however, were developed for a pure chemical agent and are not relevant to biocompatibility testing of devices constructed from multiple materials. In addition, medical devices are an unusual test subject in toxicity testing. As will be discussed, a biomaterial is a complex entity, and the material toxicity is mediated by both physical and chemical properties. Toxicity from biomaterial often comes from leachable components or contaminants introduced during manufacture, and the chemical composition of a material is often not known. Toxicological information on the material and its chemical composition is seldom available, and the possible interactions among the components in any given biological test system are seldom known.

Biocompatibility cannot be defined by any single test. It is highly unlikely that any single parameter will be able to ensure biocompatibility. Therefore, it is necessary to test as many biocompatibility parameters as appropriate to develop a matrix of information for assessment. It is also important to test as many samples as possible. Therefore, suitable positive and negative controls should produce a standard response index for repeated tests (Boutrand 2012). Additionally, it is important to make use of exaggerated conditions, such as using higher levels of exposure, exaggerated temperature of extraction, and longer contact durations or multiple other factors more severe than the actual use conditions. Identifying and subsequently ensuring an acceptable exposure level that is multiple factors below the lowest toxic level is the general, and expected, practice.

Historically, basic biocompatibility tests are short-term tests to establish acute or short-term toxicity. Data from these short-term tests should not be stretched to cover the areas where no test results are available, and indeed longer-term and more rigorous tests are now being required. A complication for biocompatibility testing compared with pharmaceuticals is that all testing must be performed before there is any clinical evaluation or use.

Biocompatibility testing should be designed to assess the potential adverse effects under actual use conditions or specific conditions close to the actual use conditions. The physical and biological data obtained from biocompatibility tests should be correlated to the device and its use. Accuracy, reproducibility, and interpretability of tests depend on the method and equipment used and the investigator's skill and experience.

There are several toxicological principles which the investigator must consider before planning biocompatibility testing programs. Biocompatibility depends on the tissue that contacts the device. For example, the requirements for a blood-contacting device would be different from those applicable to a urethral catheter. Also, the degree and nature of required biocompatibility assurance depend on the nature, extent, and duration of contact with the human body. Some materials, such as those used in orthopedic implants, are meant to last for a long period in the patient. In this case, a biocompatibility testing program needs to show that the implant as introduced into the body does not adversely affect the body during the long period of use (Greco 1994). The possibility of biodegradation of material or device cannot be ignored, and evaluation of such is now required by ISO-10993 guidances. Biodegradation by the body can change an implant's safety and effectiveness (USP 2006). The leachables from plastic used during a hemodialysis procedure may be very low, but the patient who is dialyzed three times a week may be exposed to a total of several grams during their lifetime. The foreign body response mounted by the body has acute, midterm, and long-term components which are generally predictable. Therefore, cumulative effects (chronicity) should be assessed.

Two materials having the same chemical composition but different physical characteristics may not induce the same biological response. The nature of the tissue to device interface (is the device surface smooth textured or rough) is very important. Also, past biological experiences with seemingly identical materials also have possible limited toxicity. Toxicity can arise from leachable components of the material previously used without adverse effect due to differences in formulation and manufacturing procedures.

Empirical correlation between biocompatibility testing results and actual toxicity findings in humans and the extrapolation of the quantitative result from short-term in vitro tests to quantitate toxicity at the time of use are controversial. These need careful and scientifically sound interpretation and adjustment. The control of variation in biological susceptibility and resistance to obtain a biological response range for toxic effect and host factors which determine the variability of susceptibility in toxicological response adjustment to susceptibility in the human population also need careful attention.

The challenge of a biocompatibility assessment is to create and use knowledge to reduce the degree of unknowns and to help make the best possible decisions. The hazard presented by a substance, with its inherent toxic potential, can only be manifested when fully evaluated in a patient. Therefore risk, which is actual or potential harm, is a function of toxic hazard and exposure. The safety of any leachables contained in the device or on its surface can be evaluated by determining the total amount of potentially harmful substance, estimating the amount reaching the patient tissues, assessing the risk of exposure, and performing the risk versus benefit analysis. When the potential harm from the use of biomaterial is identified from the biocompatibility tests, this potential must be compared against the availability of a suitable alternate material.

1.2 Scope of Devices and the Medical Device Market

According to section 201(h) of the Food, Drug, and Cosmetic Act, a medical device is defined as an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component, part, or accessory that is:

- Recognized in the official *National Formulary*, or the *United States Pharmacopoeia* (USP 2013), or any supplement to them.

  Intended for use in the diagnosis of disease, in man or other animals, or

- Intended to affect the structure or any function of the body or man or other animals, and which does not achieve any of its primary intended purposes through chemical action within or on the body of man or other animals, and which is not dependent upon being metabolized for the achievement of any of its principal intended purposes (CDRH 1992).

Under this definition, historically devices could be considered as belonging to one of nine categories (North American Industrial Classification System): surgical and medical instruments, ophthalmic, dental, lab apparatus, irradiation, specialty devices, medical/surgical supplies, in vitro diagnostics, and electromedical.

The top twenty medical devices by revenues in 1999 were:

1. Incontinence supplies
2. Home blood glucose-monitoring products
3. Wound closure products
4. Implantable defibrillators
5. Soft contact lenses
6. Orthopedic fixation devices
7. Pacemakers
8. Examination gloves
9. Interventional cardiovascular coronary stents
10. Arthroscopic accessory instruments
11. Prosthetic knee joint implants
12. Lens care products
13. Prosthetic hip joint implants
14. Multiparameter patient-monitoring equipment
15. Mechanical wound closure
16. Wound suture products
17. Absorbable polymers
18. Hearing aids
19. Wheelchair and scooter/mobility aids industry
20. Peritoneal dialysis sets

1.3 History

As has previously been reviewed by Hutt (1989), the regulation of medical devices has followed a different history than that of drugs. Medical devices go back to at least the Egyptians and Etruscans. Problems with fraudulent devices in the United States date back to the late 1700s, though no legislative remedy was attempted until the 1900s. In fact, the legislative history of the 1906 Food and Drug Act contains no references to devices. Devices continued to be regulated under the postal fraud statutes. Such regulation was evidently ineffectual, as fraudulent devices flourished during this period. Starting in 1926, the Food and Drug Administration (FDA) monitored such devices and assisted the US Postal Service in its regulatory actions. Medical devices were covered in the 1938 Act, but only in regard to adulteration and misbranding. Over the intervening years, various committees which examined medical device regulation consistently came to similar conclusions: that the FDA has inadequate authority and resources to regulate the medical device industry. As part of the agreement that resulted in passage of the 1962 amendments, however, all references to medical devices were deleted. The need and demand for increased regulation continued to grow. In 1967, President Lyndon Johnson supported the proposed Medical Device Safety Act, which nevertheless was not well received by Congress. In fact, no legislation pertaining to medical device safety was passed until 1976.

In 1969, at the request of then President Richard Nixon, the Department of Health, Education, and Welfare (HEW) established a Study Group in Medical Devices, also known as the Cooper Committee, because it was chaired by the Director of the National Heart and Lung Institute, Dr. Theodore Cooper. Its report in 1970 concluded that a different regulatory approach was needed to deal with medical devices. This report initiated the chain of events that culminated in the Medical Device Amendment of 1976. In the interim, the Bureau of Medical Devices and Diagnostic Products was created in 1979. Remarkably, the 1976 Amendment retained the essential provisions of the Cooper Committee Report regarding inventory and classification of all medical devices by class: Class I (general controls), Class H (performance standards), or Class III (premarket approval). These classifications are discussed in greater detail later in this chapter. These remain the essential regulations applicable to medical devices. Both the Drug Price Competition and Patent Restoration Act of 1984 and the Orphan Drug Act of 1983 contained language that made the provisions of the laws applicable to medical devices but did not have provisions unique to medical devices. The recent perceptions, revelations, and controversy surrounding silicone breast implants will probably cause additional changes in the regulation of devices.

As a consequence, 1978 brought guidelines for investigational device exemptions (IDEs, the equivalent of INDAs for drugs). These requirements, as shall be seen later, effectively excluded a wide range of medical devices from regulation by establishing an exemption for those new or modified devices which are equivalent to existing devices. The year 1990 saw the passage of the Safe Medical Devices Act, which made premarketing requirements and postmarketing surveillance more rigor-

ous. The actual current guidelines for testing started with the USP guidance on biocompatibility of plastics. A defined regulatory approach sprang from the tripartite agreement, which is a joint intergovernmental agreement between the United Kingdom, Canada, and the United States (with France having joined later). After lengthy consideration, the FDA announced acceptance of International Standards Organization (ISO) 10,993 guidelines for testing (ASTM 1990; FAO 1991; MAPI 1992; O'Grady 1990; Spizizen 1992) under the rubric of harmonization. This is the second major trend operative in device regulation: the internationalization of the market place with accompanying efforts to harmonize regulations. Under ICH (International Conference on Harmonization) great strides have been made for drugs in this area.

Independent of FDA initiatives, the USP and ASTM have promulgated test methods and standards for various aspects of establishing the safety of drugs (such as the 2013 standards for measurement of heavy metals in extractable materials from devices), which were, in effect, regulations affecting the safety of drugs and devices. Most of the actual current guidelines for the conduct of nonclinical safety evaluations of medical devices have evolved from such quasi-agency actions (such as the USP's 1965 promulgation of biological tests for plastics and ongoing American National Standards Institute (ANSI) standard promulgation).

Public concerns about three specific device safety issues served to increase regulatory scrutiny. The first of these, the Dalkon Shield, was an intrauterine contraceptive device produced by the A. H. Robbins Corporation (Sivin 1993). Its use was associated with unacceptable rates of pregnancy, pelvic inflammatory disease, and death in women who used it. The device was withdrawn from the market in 1974 and in 1988 Robbins reached a $3.3 billion settlement in response to a class action suit (Nocera 1995).

The second case is that of silicone-filled breast implants, which have been purported to cause a range of autoimmune and neurologic effects on some women who have them. Though the validity of these claims remains unproven or disproven, litigation over them drove the primary manufacturer (Dow Corning) into bankruptcy and led to the removal of these products from the market (though, in 2006, they have returned to the market). Since the late 1980s concern has grown about allergic responses to latex in devices. Several deaths have been blamed on anaphylactic responses to such effects (Lang 1996). The third was associated with toxic shock syndrome (TSS) caused by super absorbant tampons.

1.4 Nonspecific Regulatory Considerations

A broad scope review of regulatory toxicology is presented in Gad (2001). Some necessary to understand regulations beyond those covered in Chap. 2 requires review here, however.

1.4.1 Good Laboratory Practices

The original promulgation of GLPs was by the US FDA in 1978 in response to a variety of cases which led the agency to conclude that some of the data that it had obtained in support of product approvals were not trustworthy. Subsequently, other regulatory agencies and authorities in the United States and across the world have either promulgated their own version of similar regulations or required adherence to the set generated by the US FDA or another body. The EEC requirement for compliance with GLPs for safety tests has recently been reinforced in a modification of Directive 75/318/EEC (Regulatory Affairs Focus 1996; ISO 1990; European Community 1991). The FDA last revised the GLP regulations in 1989 (FDA 1986).

The GLPs require that all pivotal preclinical safety studies, that is, those that are used and regulatorily required to make decisions as to the safety of the product (in our case, a device), be conducted under a well-defined protocol utilizing procedures set forth in written standard operating procedures by trained (as established by documentation) personnel under the direction of a study director. All work must be reviewed by an independent Quality Assurance Unit (QAU). The regulations require rigorous attention to record keeping, but do not dictate how actual studies are designed or conducted in a technical sense (Gad and Taulbee 1996).

1.4.2 Animal Welfare Act (AWA)

Gone are the days when the biomedical research scientist could conduct whatever procedures or studies that were desired using experimental animals. The Animal Welfare Act (APHIS 1989) (and its analogues in other countries) rightfully requires careful consideration of animal usage to ensure that research and testing uses as few animals as possible in as humane a manner as possible. As a start, all protocols must be reviewed and approved by an Institutional Animal Care and Use Committee (IACUC) prior to animals being ordered or a study being initiated. Such review takes time, but should not serve to hinder good science. When designing a study or developing a new procedure or technique, the following points should be kept in mind:

1. Will the number of animals used be sufficient to provide the required data, yet not constitute excessive use? It ultimately does not reduce animal use to utilize too few animals to begin with and then have to repeat the study.
2. Are the procedures employed the least invasive and traumatic available? This practice is not only required by regulations but is also sound scientific practice, since any induced stress will produce a range of responses in test animals that can mask or confound the chemically induced effects.

1.4.3 Regulations Versus Law

A note of caution must be inserted here. The law (the statute promulgated by Congress) and the regulations (the documents written by the regulatory authorities to enforce the laws) are separate documents. The sections in the law do not necessarily have numerical correspondence with regulation. For example, the regulations on the PMA process is described in 21 CFR 312 (FDA 2013), but the law describing the requirement for a PMA process is in Section 515 of the FDLI. Because the regulations rather than the laws themselves have a greater impact on toxicological practice, greater emphasis is placed on regulation in this chapter. For a complete review of FDA law, the reader is referred to the monographs by Food and Drug Law Institute in FDLI (2013).

Laws authorize the activities and responsibilities of the various federal agencies. All proposed laws before the US Congress are referred to committees for review and approval. The committees responsible for FDA oversight are summarized on Table 1.1. This table also highlights the fact that authorizations and appropriations (the funding necessary to execute authorizations) are handled by different committees. Figure 1.1 presents the organization of the Center for Devices and Radiological Health (CDRH). As can be seen by the organizational structure presented in the figure, the categorization of devices for division review purposes is functionally based.

1.4.4 Organizations Regulating Drug and Device Safety in the United States

The agency formally charged with overseeing the safety of drugs and devices in the United States is the FDA. It is headed by a commissioner who reports to the Secretary of the Department of Health and Human Services (DHHS) and has a tremendous range of responsibilities covering almost a third of the economy of the United States. Medical devices are overseen by the CDRH, headed by a director.

Table 1.1 Congressional committees responsible for FDA oversight

Authorization	
Senate	All public health service agencies are under the jurisdiction of the Labor and Human Resources Committee
House	Most public health agencies are under the jurisdiction of the Health and the Environmental Subcommittee of the House Energy and Commerce Committee
Appropriation	
Senate	Unlike most other public health agencies, the FDA is under the jurisdiction of Agriculture, Rural Development, and Related Agencies Subcommittee of the Senate Appropriations Committee
House	Under the jurisdiction of the Agriculture, Rural Development, and Related Agencies Subcommittee of the House Appropriations Committee

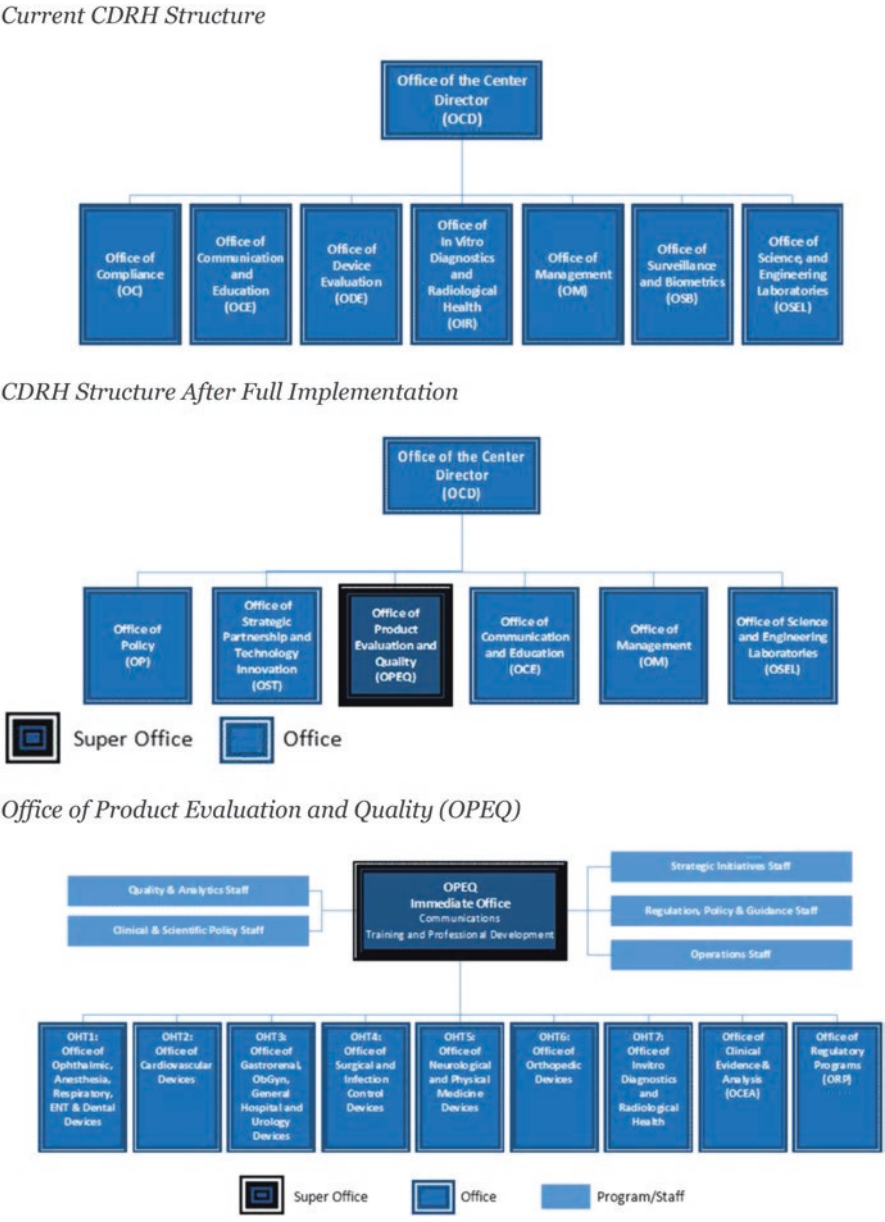


Fig. 1.1 Center for Devices and Radiologic Health (CDRH) Organizational Structure

Drugs are overseen primarily by the Center for Drug Evaluation and Research (CDER) (though some therapeutic or healthcare entities are considered as biologically derived and therefore regulated by the Center for Biologics Evaluation and Research, or CBER). There are also "combination products" (part drug, part device)

which may be regulated by either or both CDER/CBER and CDRH, depending on the principal mode of action (PMOA) of the product.

Most of the regulatory guidance for a toxicologist involved in assessing the biocompatibility of devices is with the appropriate part of the CDRH, though for combination products, the two centers charged with drugs or biologicals may also come into play. Within the CDRH there is a range of groups (called divisions) which focus on specific areas of use for devices (such as general and restorative devices; cardiovascular, respiratory, and neurological devices; ophthalmic devices; reproductive, abdominal, ear, nose, and throat, and radiological devices; and clinical laboratory devices). Within each of these, there are engineers, chemists, pharmacologists/toxicologists, statisticians, and clinicians.

There is also at least one nongovernmental body which must review and approve various aspects of devices, setting forth significant "guidance" for the evaluation of safety of devices. This is the USP, and its responsibilities and guidelines are presented later in Chap. 2.

Modern regulation of the biological safety of medical devices and the materials that they are composed of begins in the late 1950s with concern over the potential risks arising from chemical moieties in plastics migrating into drugs. Prior to this time, most drugs and infusion solutions had been stored and dispensed or delivered from glass containers. This was advanced by the works of Dr. John Autian, who founded the Drug-Plastic Research Laboratory at the College of Pharmacy at the University of Texas. His initial publication on the toxicology of phthalate esters (Calley et al. 1966) lead to the testing and plastics designations section (for medical "closures") in the *United States Pharmacopeia*. The resulting testing requirements are shown in Table 1.2.

These testing guidelines, being all that was available, were used to evaluate the biological safety of medical devices and nonmetal biomaterials.

The next step was the development of the tripartite, originally developed jointly by Canada, the United Kingdom, and the United States (a group subsequently joined

Table 1.2 Innovative areas of medical device (The Wilkerson Group 2013)

Rank	Product	Revenue growth rate (%) (years)	Specialty
1	Fibrin sealants	174.6 (95–02)	Wound care
2	Solid artificial organs	141.2 (95–02)	Transplant/ implant
3	Left ventricular assist devices	96.0 (95–02)	Cardiovascular
4	Skin substitute products	63.1 (97–04)	Wound care
5	Refractive surgical devices	54.4 (98–05)	Ophthalmic
6	Gynecologic falloposcopes	49.5 (95–00)	Endoscopic/MIS
7	PTMR products	47.8 (00–04)	Cardiovascular
8	Bone growth substitutes and growth factors	47.0 (97–04)	Orthopedics
9	Growth factor dressings	46.0 (97–04)	Wound care
10	Vascular stent-grafts	46.0 (97–04)	Cardiovascular

Source: Frost & Sullivan

by France) in 1986 (FDA 1986). These guidelines first presented a classification of devices by type and duration of patient exposure.

With a few years of exposure, this guidance evolved into the ISO 10993 system.

1.5 Potential Patient Exposure Parameters (Routes, Regimens, Quantities, and Durations) as a Principal Determinant of Risk

Unlike drugs, food additives, pesticides, biologics, industrial chemicals, or consumer products, the biologic safety (biocompatibility) of medical devices is not determined relative to known administered doses of substances nor for the most part (see the chapter on leachables and extractables [L&Es] and determination of qualified safety levels – which for devices are called tolerable exposures or TEs – for the exception to this) are the precise chemical entities to which patients (or cellular or animal models) are exposed/identified (Gad and Schuh 2018; Gad and Gad-McDonald 2015).

Rather, we use defined biological test systems to evaluate effects in terms of responses to define contact between the device and potential patients. That is, we use bioassays.

The potential interactions between a medical device and patients are determined by three factors (which are incorporated into the ISO 10992-1 testing matrix).

1.5.1 What Is the Type or Route of Patient Exposure?

Which patient tissues have contact with a device is overwhelming the determinant what happens at this direct tissue/device surface interface that presents potential adversity. While there are exceptions (genotoxicity, pyrogenicity and for the most part sensitization), physical and chemical interactions by which the host and device modify each other occur at this interface or very near it.

Devices may have more than one type of tissue contact, which complicates evaluation of potential interactions.

1.5.2 How Much of the Device Contacts Patient Tissues?

The measurement here is not (generally) of the mass of the device, but rather of the surface area.

When the test in question consists of direct device to tissue contact (such as with implantation), the device itself determines the quantity of surface to tissue contact (such as in implantation studies). That said, in many cases, what is tested is an

Table 1.3 Volume/surface area

Form of material	Thickness	Amount of *sample* for each 20 mL of extracting medium[a]	Subdivided into
Film or sheet	<0.5 mm	Equivalent of 120 cm^2 total surface area (both sides combined)	Strips of about 5 × 0.3 cm
	0.5–1 mm	Equivalent of 60 cm^2 total surface area (both sides combined)	
Tubing	<0.5 mm (wall)	Length (in cm) = 60 cm^2/(sum of ID and OD circumferences)	Sections of about 5 × 0.3 cm
	0.5–1 mm (wall)	Length (in cm) = 60 cm^2/(sum of ID and OD circumferences)	
Slabs, tubing, and molded items	>1 mm	Equivalent of 60 cm^2 total surface area (all exposed surfaces combined)	Pieces up to about 5 × 0.3 cm
Elastomers	>1 mm	Equivalent of 25 cm^2 total surface area (all exposed surfaces combined)	Do not subdivide[b]

[a]When surface area cannot be determined due to the configuration of the specimen, use 0.1 g of elastomer or 0.2 g of plastic or other polymers for every 1 mL of extracting fluid
[b]Molded elastomeric closures are tested intact

extract solution derived from the actual device. In these situations, testing practices and guidelines call for determining the potential surface area having contact with patient tissues and then using a guideline (ISO 10993)-prescribed volume of one or more vehicles (solvents, really) to be used in performing extractions so as to provide a liquid which can be used in subsequent actual tests (Table 1.3).

In some cases, the shape of a device component having patient contact is so irregular that it is not possible to accurately calculate a surface area, so rather the weight of the device determines the volume of extraction solution (Table 1.4).

In most cases, two separate extraction fluids are used – a polar (such as water, saline, or ethanol in water) and a nonpolar (such as hexane). See Table 1.1 for a list of extraction fluids. These are intended to simulate the principal physicochemical components of the body – water (~67% of body volume on average) and lipids. The original USP list of solvents was more extensive, as it was intended to reflect the range of solvents which were used in the formulation of medicants in contrast with the plastic and elastomer containers ("closures") for drugs. This broader range of solvents is still reflected in the (mouse) acute systemic toxicity test.

An exception is in the case of mammalian in vitro genotoxicity tests, where extraction is directly into culture medium with serum. Here, the underlying thought is that the medium stands in place of blood, which would serve to transfer any potential genotoxic moiety from the surface of the device to a potential susceptible target tissue.

Table 1.4 Extraction fluids

Common name: Cottonseed oil (CSO)
Chemical name: NA
Molecular weight: NA
Formula: Mixture of natural products; glycerides of palmitic, olive, and linoleic acids
Density: 0.915–0.921 g/ml
Volatility: Low
Solubility/miscibility: Soluble in ether, benzene, chloroform, and DMSO. Slightly soluble in ethanol
Biological considerations: Orally, serves as energy source (and therefore can alter food consumption and/or body weight). Prolonged oral administration has been associated with enhanced carcinogenesis
Chemical compatibility/stability considerations: Thickens upon prolonged exposure to air. Available in USP grade.
Uses (routes): In extractions and as a vehicle for oral, dermal, vaginal, rectal and subcutaneous administration
Common name: DMSO/dimethyl sulfoxide
Chemical name: Sulfinylbis[methane]; CAS #67–68-5
Molecular weight: 78.13
Formula: C_2H_6OS
Density: 1.100 g/ml at 20 °C
Volatility: Medium
Solubility/miscibility: Soluble in water, ethanol, acetone, ether, oils
Biological considerations: Oral LD50 (rats) = 17.9 ml/kg. Repeated dermal exposure can defat skin. Repeated oral exposure can produce corneal opacities. Not cytotoxic to cells in primary culture at less than 0.05% (V/V). Intraperitoneal LD50 (mice) = 11.6 ml/kg
Chemical compatibility/stability considerations: Very hydroscopic liquid. Combustible
Uses (routes): All, as a carrier at up to 5% to enhance absorption
Common name: Ethanol; EtOH
Chemical name: Ethyl alcohol; CAS #64–17-5
Molecular weight: 46.07
Formula: C_2H_5OH
Density: 0.789 g/ml
Volatility: High, but declines when part of mixture with water
Solubility/miscibility: Miscible with water, acetone, and most other vehicles
Biological considerations: Orally, will produce transient neurobehavioral intoxication. Oral LD50 (rats) = 13.0 ml/kg. Intravenous LD50 (mice) = 5.1 ml/kg
Chemical compatibility/stability considerations: Flammable colorless liquid available USP grade
Uses (routes): Extraction solvent vehicle for dermal and oral, though can be used in lower concentrations for most other routes. Volume of oral instillation should be limited to 5 ml/kg
Common name: Polyethylene glycol (PEG)
Chemical name: NA
Molecular weight: 400 (approximate average, range 380–420)
Formula: $H(OCH_2CH_2)_nOH$

(continued)

Table 1.4 (continued)

Density: 1.128 g/ml
Volatility: Very low
Solubility/miscibility: Highly soluble in water. Soluble in alcohol and many organic solvents
Biological considerations: Employed as water-soluble emulsifying/dispersing agents. Oral LD50 (mice) = 23.7 ml/kg. Oral LD50 (rats) = 30 ml/kg
Chemical compatibility/stability considerations: Do not hydrolyze or deteriorate on storage and will not support mold growth. Clear, viscous liquid
Uses (routes): As extraction solvent for oral administration as a vehicle full strength or mixed with water. Total dosage of PEG-400 should not exceed 5–10 ml
Common name: Saline
Chemical name: Physiological saline; isotonic salt solution
Molecular weight: 18.02
Formula: 0l9% NaCl in water (weight to volume)
Density: As water
Volatility: Low
Solubility/miscibility: As water
Biological considerations: No limitations – preferable to water in parenteral applications
Chemical compatibility/stability considerations: None
Uses (routes): Extraction solvent all except periocular

Source: Gad and Chengelis (1992); Lewis (2012)

1.5.3 What Is the Cumulative Duration of Contact of a Device with a Patient?

The cumulative duration of contact is critical in determining both the potential risk to patients and the extent of testing required. Very short-term exposures generally require just the basic three tests (cytotoxicity, irritation in the appropriate tissue, and sensitization). With longer duration of exposure, the range and scope of potential interactions between host and devices increase, calling for a more extensive range of tests.

The “continental divide” is 30 days, after which exposure is considered “permanent.” The basis for this is by this time, the body’s adaptive immune system has had time to fully respond to the surface of the device and any moieties which may be released from the device into the body.

Notice that duration is defined as cumulative if the identical type of device is sequentially replaced with new units on a regular basis (such as occurs with catheters or wound dressings) then it is as if a single device was left in place for the entire time the device type had patient contact.

Note also that by definition, implanted devices have “permanent” durations of contact. It is important to differentiate that components/tools (such as guidewires or tracers) which are used to put an implant in place do not have permanent contact (rather their contact is less than 24 hours); the implanted devices themselves are permanent.

The first special case is that of resorbable devices. These are almost always permanent, as it takes more than 30 days for the device to be (effectively) dissolved into the body and have much greater potential to generate/release chemical components that are distributed throughout the body.

The second special case is that of respiratory devices – devices meant to support patient breathing and in some cases to administer/infuse drug materials by the pulmonary route. The direct patient contact with these devices is limited to external skin where the devices generally touch the face, and the epithelial tissue on the inside of the nose and/or mouth; however, hair flow through the devices into breathing channels has the opportunity to pick up and carry on materials from the interior surface of the device as it passes through, progressing perhaps all the way into a patients deep dungs. A further complication is that the devices have significant use in the very young (neonates, pediatrics, and juveniles) and very old and in individuals who are already significantly compromised in their breathing.

References

APHIS. (1989). Animal and Plant Health Inspection Service, United States Department of Agriculture. *Federal Register, 54*(168), 36112–36163.

ASTM. (1990). Standardization in Europe: A success story. *ASTM Standardization News*, *38*, February 1990, pp. 1–4.

Boutrand, J.-P. (Ed.). (2012). *Biocompatibility and performance of medical devices*. Cambridge: Woodhead Publishing.

Calley, D., Autian, J., & Guess, W. L. (1966). Toxicology of a series of phthalate esters. *Journal of Pharmaceutical Sciences, 55*(2), 158–162.

CDRH. (1992). *Regulatory requirements for medical devices: A workshop manual*. Rockville: Center for Device and Radiological Health, HHS Publication FDA 92–4165.

CFR. (1992). FDA's policy statement concerning cooperative manufacturing arrangements for licensed biologics. *Federal Register, 57*, 55544.

Di Silvio, L. (Ed.). (2009). *Cellular response to biomaterials*. Cambridge: Woodhead Publishing.

European Committee for Standardization. (1991). *CEN Annual Report 1991*. Brussels.

FAO. (1991). *Report of the FAO/WHO Conference on Food Standards, Chemicals in Food and Food Trade (in cooperation with GATT)*, Vol. 1, Rome, March 18–27.

FDA. (2016). *ISO 10993-1. Biological evaluation of medical devices. Part1: Evaluation and testing within a risk management process*. Department of Health and Human Services. Silver Springs, MD.

FDA. (1986). *Tripartite biocompatibility guidance for medical devices*. Rockville: Toxicology Subgroup of the Tripartite Sub-Committee on Medical Devices, Food and Drug Administration, Center for Devices and Radiological Health (FDA CDRH).

FDA. (2013). *Good laboratory practice regulations*: *21 CFR 58.1–58.219*.

FDLI. (2013). *Compilation of food and drug Laws, volumes I, II, III and supplement*. Washington, D.C.: Food and Drug Law Institute.

Gad, S. C. (2001). *Regulatory toxicology* (2nd ed.). Philadelphia: Taylor & Francis.

Gad, S. C., & Chengelis, C. P. (Eds.) (1992) *Animals Models in Toxicology*. Marcel Dekker, New York. p. 884.

Gad, S. C., & Gad-McDonald, S. E. (2015). *Biomaterials, medical devices, and combination products: Biocompatibility testing and safety assessment*. Boca Raton: CRC Press.

Gad, S. C., & Schuh, J. C. L. (2018). Points to consider for toxicologic pathologists evaluating biomaterials and medical devices. *Toxicologic Pathology, 46*(4), 366–371.

Gad, S. C., & Taulbee, S. (1996). *Handbook of data recording, maintenance and management for the biomedical sciences*. Boca Raton: CRC Press.

Greco, R. S. (1994). *Implantation biology: The host response and biomedical devices*. Boca Raton: CRC Press.

Hutt, P. B. (1989). A history of government regulation and misbranding of medical devices. *Food Drug Cosmetic Law Journal, 44*(2), 99–117.

ISO. (1990). *ISO 9000 international standards for quality management vision 2000—A strategy for international standards' implementation in the quality arena during the 1990s (2nd ed., compendium)*. Brussels: EEC.

Lang, L. A. (1996). A review of latex hypersensitivity. *Toxic Substance Mechanisms, 15*, 1–11.

Lewis, R. J. (2012) *SAX'S Dangerous Properties of Industrial Materials*, 12th Ed., Wiley-Interscience, Hoboken, NJ.

MAPI. (1992). The European community's new approach to regulation of product standards and quality assurance (ISO 9000): What it means for U.S. manufacturers. *MAPI Economic Report ER-218*.

MDDI. (2013). Industry snapshot. *Medical Device & Diagnostics Industry*, July 2013, 52–62.

Nocera, J. (1995). Fatal litigation. *Fortune*, October 23, 1995, 60–82.

Nugent, T. N. (1994). *Health care products & services basic analysis*. New York: Standard & Poor's Industry Surveys.

O'Grady, J. (1990). Interview with Charles M. Ludolph. *ASTM Standardization News*, *26*, February, 1990, pp. 3–6.

Regulatory Affairs Focus. (1996). European update. *Regulatory Affairs Focus, 1*(4), 8.

Sivin, I. (1993). Another look at the Dalkon Shield: Meta-analysis underscores the problems. *Contraception, 48*, 1–12.

Spizizen, G. (1992). The ISO 9000 standards: Creating a level playing field for international quality. *National Productivity Review, 11*(3), 331–346.

The Wilkerson Group. (2013). *Forces reshaping the performance and contribution of the U.S. medical device industry*. Washington, D.C.: Health Industry Manufacturers Association.

USP. (2006). *The United States pharmacopoeia <88>, biological reactivity tests, in vivo, USP 29*. Rockville: United States Pharmacopoeial Convention.

USP. (2013). *The United States pharmacopoeia, 36 revision*. Rockville: United States Pharmacopoeial Convention.

Chapter 2
Regulatory Aspects and Strategy in Medical Device and Biomaterials Safety Evaluation

Shayne C. Gad

Abstract While there is some blurring along the edges, the definition of a medical device is fundamentally a product with a therapeutic effect or application that does not depend on interaction with the body chemically or metabolically. In vitro diagnostics are which test samples taken from the body and radiologic contrivance (such as X-rays, ultrasound devices, MRIs and such), which use energy prom outside the body for either diagnostic or therapeutic purposes, are regulated as devices. Regulations as to what testing must be alone (prior to exposure of patients to a device) have been very much harmonized for major markets (the United States, Europe, Japan, China, and many others) by the adoption of ISO (International Standards Organizations) guidelines though there are variations in how such testing is evaluated and interpreted. The key elements are the novelty of the device and its constituent materials (expressed by classifying a device as Class I, needs minimal premarket testing, to Class II, extensive nonclinical and then clinical testing is required) and by the type and duration of patient contact.

Keywords 510(k) · "De novo" 510(k) route · Biologic tests · CE mark(ings) · Class I, II, III · Classification of devices · EC Medical Device Directive · Exposures of patients to the devices · External devices · Externally communicating devices · GLP regulations · IDE · In vitro · In vivo · Internal devices · ISO · Japan · Ligature material · MHW requirements · Nature and cumulative duration of contact · Nitinol implant · Noncontact devices · Pre-amendment devices · Premarket approval (PMA) · Risk assessment · Safety margins · Sterilized · Technical dossier · Uncertainty factors (UCFs) · USP (*United States Pharmacopeia*) · Wound dressing

S. C. Gad, PhD, DABT (✉)
Gad Consulting Services, Raleigh, NC, USA
e-mail: scgad@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_2

As discussed in Chap. 1, in the United States (according to 201(h) of the Food, Drug, and Cosmetic Act), a medical device is defined as an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar or related article, including a component, part, or accessory that is:

(a) Recognized in the official National Formulary, the United States Pharmacopoeia (USP), or any supplement to them, and
(b) Intended for use in the diagnosis of disease or other condition, or in the cure, mitigation, treatment, or prevention of disease, in man or other animals, or intended to affect the structure or any function of the body or man or other animals, and which does not achieve any of its primary intended purposes through chemical action within or on the body of man or other animals, and which is not dependent upon being metabolized for the achievement of any of its principal intended purposes (CDRH 1992). This same operational definition generally applies across the major global markets.

2.1 Regulatory Basis

2.1.1 Regulations: General Considerations for the United States

The US regulations for medical devices derive from seven principal laws:

Federal Food, Drug, and Cosmetic Act of 1938
Medical Device Amendments of 1976
Safe Medical Devices Act of 1990
Medical Device Amendments of 1992
FDA Modernization Act of 1997 (Section 204)
Blue Book Memos—ODE Guidance Memoranda of 1997
Use of International Standard ISO-10993 (2013)

The US federal regulations that govern the testing, manufacture, and sale of medical devices are covered in Chap. 1, Title 21 of the Code of Federal Regulations (21 CFR). These comprise nine 6 × 8-inch volumes which stack 8 inches high. This title also covers foods, veterinary products, medical devices, and cosmetics. As these topics will be discussed elsewhere in this book, here we will briefly review those parts of 21 CFR that are applicable to medical devices (Gad 2001; Heller 1999).

Of most interest to a toxicologist working in this arena would be Chap. 1, Subchapter A (Parts 1–78), which covers general provisions, organization, etc. The good laboratory practices (GLPs) are codified in 21 CFR 58. The regulations applicable to medical devices are covered in Subchapter H, Parts 800–895 of 21 CFR. As discussed earlier, the term medical device covers a wide variety of products: contact lenses, hearing aids, intrauterine contraceptive devices, syringes, catheters, drip bags, orthopedic prostheses, etc. The current structure of the law was established by the Medical Device Amendment of 1976. Products on the market on the day the amendment was passed were assigned to one of three classes (I, II, or III), based on

the recommendation of advisory panels. Medical device classification procedure is described in Part 860. Class I products (the least risk burdened) were those for which safety and effectiveness could be reasonably assured by general controls. Such devices are available over the counter to the general public. Class II products were those for which a combination of general controls and performance standards was required to reasonably assure safety and effectiveness. Class II devices are generally available only with a doctor's prescription, but may be used at home. Class III products were those for which general controls and performance standards were inadequate; these were required to go through a premarket approval process. All devices commercially distributed after May 28, 1976 ("pre-amendment Class III devices"), which are not determined to be substantially equivalent to an existing marketed device are automatically categorized as Class III and require the submission of a PMA. Please note that these are classifications for regulatory purposes only and are distinct from the classification (HIMA/PHRMA) of product types (e.g., internal versus external) discussed elsewhere in this chapter. Kahan (1995) provides a detailed overview of what comprises general controls, performance standards, and such.

As with the subchapter on drugs, much of the subchapter on medical devices in the regulations concerns categorizations and specifics for a wide variety of devices. For a toxicologist involved in new product development, the parts of highest interest are 812 and 814. As with drugs, devices must be shown to be safe and effective when used as intended, and data must be provided to demonstrate such claims. In order to conduct the appropriate clinical research to obtain these data, a sponsor applies to the agency for an IDE, as described in 21 CFR 812. As stated in this section, "an approved investigational device exemption (IDE) permits a device that would otherwise be required to comply with a performance standard or to have premarket approval to be shipped lawfully for the purpose of conducting investigations of that device." Given the broad range of products that fall under the category of medical devices, the toxicological concerns are equally broad; testing requirements to support an IDE are vaguely mentioned in the law, even by FDA standards. In this regard, the law simply requires that the IDE application must include a report of prior investigations which "shall include reports of all prior clinical, animal and laboratory testing." There is no absolute written requirement for animal testing, only a requirement that such testing must be reported. It should be noted here that unlike with the clinical evaluation of drugs, in most cases all requested biocompatibility testing must be performed before any clinical testing is initialized.

There are, of course, standards and conventions to be followed in designing a safety package to support an IDE, and these are discussed in a subsequent section of this chapter. The expansion and increased sophistication of ISO guidances have tended to shift the balance toward an increasing set of required pre-IDE biocompatibility tests.

In order to obtain a license to market a device, a sponsor either submits a 510(k) premarket notification or applies for a premarket approval (PMA), as described in 21 CFR 814. Like an NDA, a PMA application is a very extensive and detailed document that must include, among other things, a summary of clinical laboratory studies submitted in the application 921 CFR 814.20(b)(3)(v)(A), as well as a section containing results of the nonclinical laboratory studies with the device, including

microbiological, toxicological, immunological, biocompatibility, stress, wear, shelf life, and other laboratory or animal tests as appropriate. As with drugs, these tests must be conducted in compliance with the GLP regulations. Under the language of the law, a sponsor submits a PMA, which the FDA then "files." The acceptance for filing of an application means that "FDA has made a threshold determination that the application is sufficiently complete to permit substantive review." Reasons for refusal to file are listed in 814.44(e) and include items such as an application that is not complete and has insufficient justification for the omission(s) present. The agency has 45 days from receipt of an application to notify the sponsor as to whether or not the application has been filed. The FDA has 180 days after filing of a complete PMA (21 CFR 814,40) to send the applicant an approval order, an "approved" letter or a "not approved" letter, or an order denying approval. An "approval order" is self-explanatory and is issued if the agency finds no reason (as listed in 814.45) for denying approval. An "approved" letter 814.44(e) means the application substantially meets requirements, but some specific additional information is needed. A "not approved" letter, 814.45(f), means that the application contains false statements of fact and does not comply with labeling guidelines or that nonclinical laboratory studies were not conducted according to GLPs, etc. Essentially, an order denying approval means that the sponsor must do substantially more work and must submit a new application for PMA for the device in question. 510(k) premarket approval submissions are less extensive than PMAs, but must still include appropriate preclinical safety data. 510(k)s are supposed to be approved in 90 days.

An alternative new is the "de novo" 510(k) route, filed for devices for which there is a lack of a suitable predicate, but for which a determination of "no significant risk" has been made.

Actual review and approval times historically have been much longer than the statutory limits (The Grey Sheet 1996). For 2016, 59.7% of 510(k)s were initially submitted to and cleared by FDA (versus other regulatory bodies). The average total review time for Class III products in the United States cleared by 510(k) was 177 days. Less than 20% of 510(k)s took less than 3 months to clear. Current average PMA review times are projected to be 16 months (Eisenhart 2017). See Chap. 1 for a discussion of general regulatory considerations (such as good laboratory practices) which are applicable to all safety evaluation studies.

2.2 Regulations Versus Law

A note of caution must be inserted here. The law (the document passed by Congress) and the regulations (the documents written by the regulatory authorities to enforce the laws) are separate documents. The sections in the law do not necessarily have numerical correspondence. For example, the regulations on the PMA process are described in 21 CFR 312, but the law describing the requirement for a PMA process is in Section 515 of the FDCA. Because the regulations rather than the laws themselves have a greater impact on the practice of nonclinical safety evaluation,

Table 2.1 FDA device categories and suggested biological testing (FDA 1995)

FDA	INSERT WEBSITE
CDER	INSERT WEBSITE
CBER	INSERT WEBSITE
CDRH	INSERT WEBSITE

greater emphasis is placed on regulation in this chapter. For a complete review of FDA law, the reader is referred to the monographs by Food and Drug Law Institute (FDLI) in 1995, 1996, and 1998 (Table 2.1).

2.3 Organizations Regulating Device Safety in the United States

The agency formally charged with overseeing the safety of devices and diagnostics in the United States is the Center for Devices and Radiological Health (CDRH) of the FDA. It is headed by a commissioner who reports to the Secretary of the Department of Health and Human Services (DHHS) and has a tremendous range of responsibilities. Medical devices are specifically overseen by the CDRH, headed by a director. Drugs are overseen primarily by the Center for Drug Evaluation and Research (CDER) (though some therapeutic or healthcare entities are considered as biologically derived and therefore regulated by the Center for Biologics Evaluation and Research, or CBER). There are also "combination products" (part drug, part device) which may be regulated by either or both CDER/CBER and CDRH, depending on what the principal mode of action (PMOA) is determined to be by the FDA (CFR 1992).

2.4 Classification of Devices

In the United States, in accordance with the 1976 Medical Device Amendment, devices are categorized as below:

- Class I—General controls (equivalent to OTC)
- Class II—Performance standards and special controls (distribution is licensed healthcare professional controlled)
- Class III—Premarket approval (clinical use only)
- Pre-amendment devices

In Europe, there is a lengthy set of rules in the EC Medical Device Directive (Council Directive 1993) to place devices in Classes I, IIa, IIb or III. Class I is the minimum grade and Class II the maximum (The Final Draft Guidelines on Medical Device Classification, n.d.). This classification determines the extent of supporting data that is required to obtain marketing approval.

In the USA, the FDA Center for Devices and Radiological Health recognizes three classes of medical device, and this system is based on whether the product was on the market prior to the passage of the 1976 Medical Device Amendments. If a new device is substantially equivalent to a pre-amendment device, then it will be classified the same as that device. This means that for Class I and II products, no premarket approval is necessary. Class III products need pre-marketing approval, and all new devices which are not substantially equivalent to existing products fall automatically into Class III.

Japan (MHLW 2012) and Korea have a somewhat different three class system. Class I includes products that have no body contact and would not cause any damage to the human body if they failed, for example, X-ray film. These products need pre-marketing approval in terms of medical device regulations, although they may need to be tested under industrial guidelines like those of the OECD. Class II products have external contact with the body, Class III have internal contact, and both need additional testing. Figure 2.1 presents the MHLW scheme for device classification.

Most of the regulatory interaction of a toxicologist involved in assessing the biocompatibility of devices is with the appropriate part of the CDRH, though for combination products, the two centers charged with drugs or biologicals may also come into play. Within the CDRH there is a range of groups (called divisions) which focus on specific areas of use for devices (such as general and restorative devices; cardiovascular, respiratory, and neurological devices; ophthalmic devices; reproductive, abdominal, ear, nose, and throat and radiological devices; and clinical laboratory devices). Within each of these there are engineers, chemists, pharmacologists/toxicologists, statisticians, and clinicians.

There is also at least one nongovernmental body which must review and approve various aspects of devices, setting forth significant "guidance" for the evaluation of safety of devices. This is the USP, and its responsibilities and guidelines are presented later in this chapter.

The other two major regulatory organizations to be considered are the International Standards Organization (ISO), with ISO 10993 standards (ISO, various dates) (ISO 1992, 2008, 2017), and the Japanese Ministry of Health and Welfare (MHW) with its guidelines (MHLW 2012).

2.5 Toxicity Testing: Medical Devices

In a statutory sense, historically any item promoted for a medical purpose which does not rely on chemical action to achieve its intended effect is a medical device (as discussed earlier). In vitro diagnostic tests are also regulated as medical devices. The regulation of devices under these definitions has had a different history than that of drugs—it has not been as strict and it has evolved at a slower rate. However, the requirements for the safety evaluation and biocompatibility evaluation of devices have rapidly been becoming more sophisticated and closer to that for new drugs.

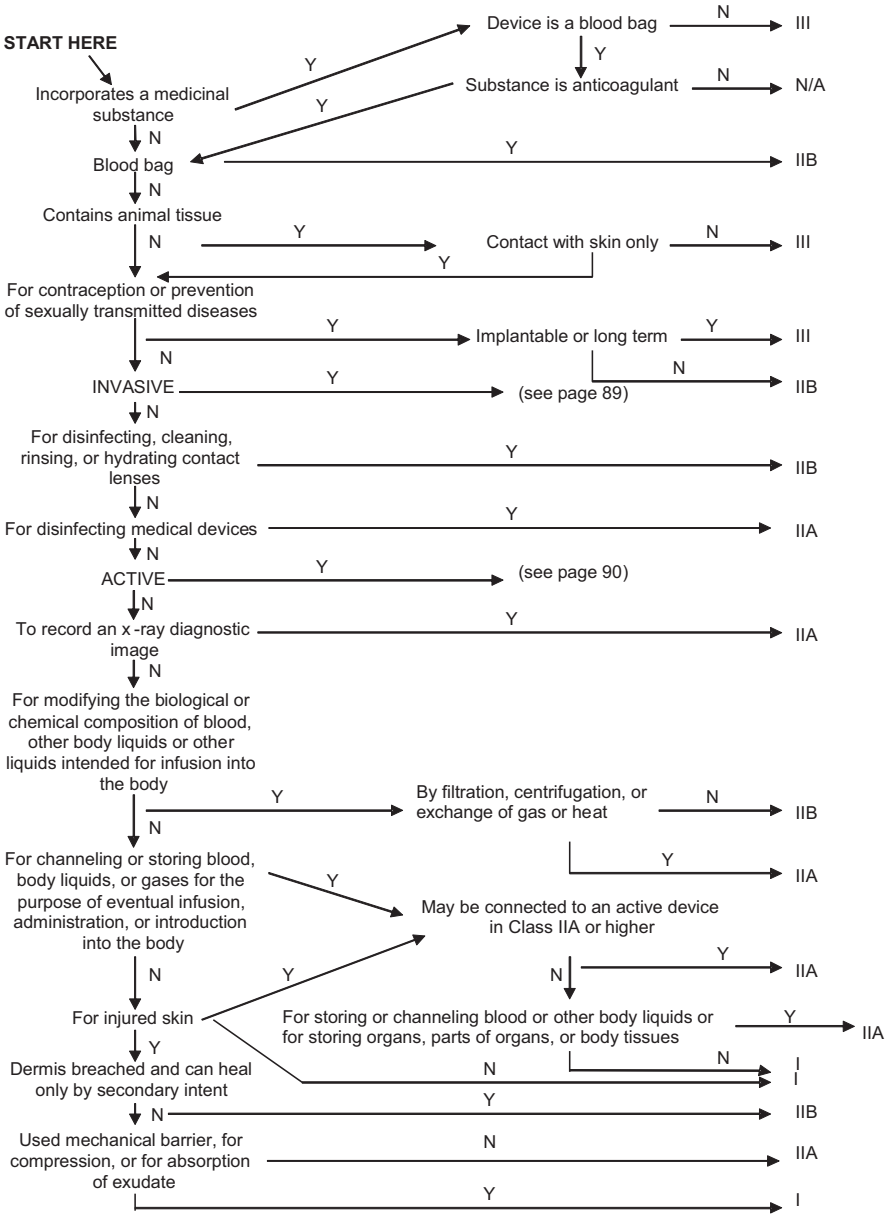


Fig. 2.1 Medical device classification flowchart

(continued)

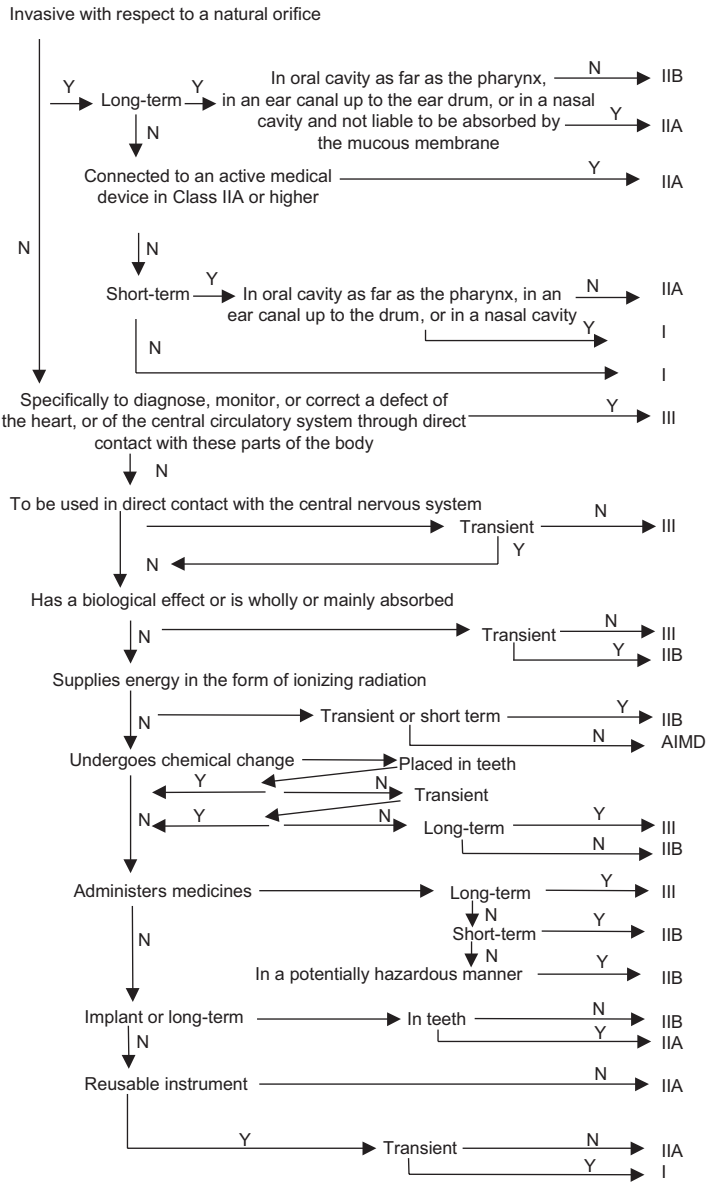


Fig. 2.1 (continued)

The safety concerns are, however, also somewhat different. Toxicologic safety concerns for devices (as opposed to concerns of mechanical safety, such as disintegration of heart valves) are called biocompatibility concerns.

Medical devices are classified as being in three different classes and are regulated accordingly. Class III devices are subject to the greatest degree of regulation

and include devices which are implanted in the body, support life, prevent health impairment, or present an unreasonable risk of illness or injury. These are subject to premarketing approval. Class II and Class I devices are subject to lesser control, required only to comply with general controls and performance standards.

There are several governing schemes for dictating what testing must be done on new Class III devices in the general case, with each developed and proposed by a different regulatory organization at different times over the last few years. ISO has attempted to harmonize these requirements so that different (or duplicate) testing would not need to be performed to gain device approval in different national markets. As discussed in the last chapter of this book, there are also specialized testing requirements for some device types such as contact lenses (FDA 1997) and tampons (CDRH 1995c). The ISO effort has generally been successful and parallels that of ICH for drugs (though ISO is, it should be noted, an NGO and not a governmental regulatory body). Where differences exist, they are highlighted in this volume as specific requirements and designs are presented.

As with drugs, all safety testing for devices must be conducted in conformity with GLPs (FDA 1987; Fries 1999; and Gad and Taulbee 1996). Table 2.2 presents the existing FDA CDRH requirements for device characterization and testing (CDRH 1995a, b). The exact nature of the test protocols is based on recommendations by USP, ISO, and others. It should be noted that Class I devices, if new, are also subject to the ISO guidelines. It should also be noted that the FDA generally (but not strictly) now adheres to the ISO guidance on test requirements (see Tables 2.13 and 2.14) (FDA 2016).

Additional concerns with devices are considerations of their processing after production. For example, concerns have risen about the potential for allergies to develop to latex components and for male reproductive effects for DEHP leaching from medical devices have led to the requirement that all such devices in either of these categories be appropriately labeled.

Devices which have systemic exposure need to be sterilized. Radiation and heat can be used for some devices, but others cannot be sterilized in these. Ethylene oxide or other chemical sterilants must be used, raising concerns that residual sterilants may present problems. At the same time, devices with exposure to the fluid path must be demonstrated to be neither pyrogenic nor hemolytic in their final manufactured form.

1. The selection of material(s) to be used in device manufacture and its toxicological evaluation should, initially, take into account full characterization of the material, for example, formulation, known and suspected impurities, and processing.
2. The material(s) of manufacture, the final product, and possible leachable chemicals or degradation products should be considered for their relevance to the overall toxicological evaluation of the device.
3. Tests to be utilized in the toxicological evaluation should take into account the bioavailability of the bioactive material, i.e., nature, degree, frequency, duration, and conditions of exposure of the device to the body. This principle may lead to the categorization of devices which would facilitate the selection of appropriate tests.

Table 2.2 FDA device categories and suggested biological testing (FDA 2016)

Device Categories			Initial Evaluation								Supplemental Evaluation	
	Body contact	Contact duration	Cytotoxicity	Sensitization	Irritation or intracutaneous reacitivity	Systemic toxicity (acute)/ pyrogenicity	Subchronic toxicity	Genotoxicity	Implantation	Hemocompatibility	Chronic toxicity	Carcinogenicity
Surface Devices	Skin	A	●	●	●							
		B	●	●	●							
		C	●	●	●							
	Mucosal membrane	A	●	●	●							
		B	●	●	●	0	0		0			
		C	●	●	●	0	●	●	0		0	
	Breached/ compromised surface	A	●	●	●	0						
		B	●	●	●	0	0		0			
		C	●	●	●	0	●	●	0		0	
External Communicating Devices	Blood path indirect	A	●	●	●	●				●		
		B	●	●	●	●	0			●		
		C	●	●	0	●	●	●	0	●	●	●
	Tissue/bone dentin communicating	A	●	●	●	0						
		B	●	●	0	0	0	●	●			
		C	●	●	0	0	0	●	●		0	●
	Circulating blood	A	●	●	●	●		0		●		
		B	●	●	●	●	0	●	0	●		
		C	●	●	●	●	●	●	0	●	●	●

Implant Devices	Bone/tissue	A	●	●	●	0						
		B	●	●	0	0	0	●	●			
		C	●	●	0	0	0	●	●		●	●
	Blood	A	●	●	●	●			●	●		
		B	●	●	●	●	0	●	●	●		
		C	●	●	●	●	●	●	●	●	●	●

A = Limited exposure (≤24 h); B = Prolonged exposure (24 h – 30 days); C = Permanent contact (>30 days); ● = FDA and ISO evaluation tests; 0 = Additional tests for FDA

1. For these devices with possible leachables or degradation products, e.g., absorbable surfaces, hemostatic agents, etc., testing for pharmacokinetics may be required
2. Reproductive and developmental toxicity tests may be required for certain materials used for specialized indications
3. Considerations should be given to long-term biological tests where indicated in the table taking into account the nature and mobility of the ingredients in the materials used to fabricate the device

4. Any in vitro or in vivo experiments or tests must be conducted according to recognized good laboratory practices followed by evaluation by competent informed persons.
5. Full experimental data, complete to the extent that an independent conclusion could be made, should be available to the reviewing authority, if required.
6. Any change in chemical composition, manufacturing process, physical configuration, or intended use of the device must be evaluated with respect to possible changes in toxicological effects and the need for additional toxicity testing.
7. The toxicological evaluation performed in accordance with this guidance should be considered in conjunction with other information from other nonclinical tests, clinical studies, and post-market experiences for an overall safety assessment.

2.6 Device Categories: Definitions and Examples

The fundamental basis for evaluating device biocompatibility is based on nature and cumulative duration of exposures of patients to the devices.

A. Noncontact Devices

Devices that do not contact the patient's body directly or indirectly; examples include in vitro diagnostic devices.

B. External Devices

1. *Intact surfaces* Devices that contact intact external body surfaces only; examples include electrodes, external prostheses, and monitors of various types.
2. *Breached or compromised surfaces* Devices that contact breached or otherwise compromised external body surfaces; examples include ulcer, burn and granulation tissue dressings or healing devices, and occlusive patches.

C. Externally Communicating Devices

1. *Intact natural channels* Devices communicating with intact natural channels; examples include contact lenses, urinary catheters, intravaginal and intraintestinal devices (sigmoidoscopes, colonoscopes, stomach tubes, gastroscopes), endotracheal tubes, and bronchoscopes.
2. *Blood path, indirect* Devices that contact the blood path at one point and serve as a conduit for fluid entry into the vascular system; examples include solution administration sets, extension sets, transfer sets, and blood administration sets.
3. *Blood path, direct* Devices that contact recirculating blood; examples include intravenous catheters, temporary pacemaker electrodes, oxygenators, extracorporeal oxygenator tubing and accessories, and dialyzers, dialysis tubing, and accessories.

D. Internal Devices

1. *Bone* Devices principally contacting bone; examples include orthopedic pins, plates, replacement joints, bone prostheses, and cements.
2. *Tissue and tissue fluid* Devices principally contacting tissue and tissue fluid or mucus membranes where contact is prolonged; examples include pacemakers, drug supply devices, neuromuscular sensors and stimulators, replacement tendons, breast implants, cerebrospinal fluid drains, artificial larynx, vas deferens valves, ligation clips, tubal occlusion devices for female sterilization, and intrauterine devices.
3. *Blood* Devices principally contacting blood; examples include permanent pacemaker electrodes, artificial arteriovenous fistulae, heart valves, vascular grafts, blood monitors, internal drug delivery catheters, and ventricular assist pumps.

2.7 Biological Tests

Also required to properly utilize the tables is a knowledge of the objectives of the specified biological tests. These can be considered as follows (Gad and Chengelis 1998; Goering and Galloway 1989):

Sensitization Assay Estimates the potential for sensitization of a test material and/or the extracts of a material using it in an animal and/or human. ISO (ISO, 1992 and 1996) and MHW procedures are contrasted in Table 2.3.

Irritation Tests Estimates the irritation potential of test materials and their extracts, using appropriate site or implant tissue such as skin and mucous membrane in an animal model and/or human. ISO and MHW procedures are contrasted in Table 2.4 and for eye irritation in Table 2.5.

Table 2.3 Differences between sensitization test procedures required by ISO 10993-10 and the MHW guidelines

ISO 10993-10	MHW 1995
Sample preparation: Extraction in polar and/or nonpolar solvents	Two extraction solvents, methanol and acetone, recommended
Extraction ratio: Extraction ratio is dependent on thickness of device or representative portion	Specific extraction ratios: 10:1 (volume solvent/weight sample)
Extract used for testing. If extraction is not possible, the adjuvant and patch test can be utilized	Residue obtained from extraction is redissolved and used for testing. (If residue does not dissolve in DMSO, or a sufficient amount of residue is not obtained, the adjuvant and patch test is recommended). Sufficient amount of residue: 0.1–0.5% (weight residue/weight test material)

Table 2.4 Differences in intracutaneous reactivity test procedures required by ISO 10993-10 and the MHW guidelines

ISO 10993-10	MHW
Number of test animals: Three rabbits for 1–2 extracts	Two rabbits for each extract
Number of test/control injections per extract: Five test and five control injections	10 test and five control injections
Evaluation of responses: Quantitative comparison of responses of test and control responses	Qualitative comparison of test and control responses

Table 2.5 Differences in eye irritation testing procedures outlined in ISO 10993-10 and the MHW guidelines

ISO 10993-10	MHW 1995
Time of exposure: 1 second	30 seconds
Grading scale: Classification system for grading ocular lesions	Draize or McDonald-Shadduck scale

Table 2.6 Differences between cytotoxicity test procedures specified by ISO 10993-5 and the MHW guidelines (MHW 1995)

ISO 10993-10	MHW 1995
Number of cells per dish: 0.5–1 million cells	40–200 cells per dish
Extraction ratio: 60 cm^2 per 20 ml if thickness 80.5 mm 120 cm^2 per 20 ml if thickness 70.5 mm 4 g per 20 ml	5 cm^2/ml or 1 g/10 ml
Exposure period: Typically 24–72 h (2 h for filter diffusion test)	6–7 days
Toxicity determination: Visual grading and/or quantitative assessments	Quantification of surviving colonies
Positive controls: Materials providing a reproducible cytotoxic response (e.g., organotin-impregnated polyvinyl chloride)	Segmented polyurethane films containing 0.1% zinc diethyldithiocarbamate and 0.25% zinc dibutyldithiocarbamate

Cytotoxicity With the use of cell culture techniques, this test determines the lysis of cells (cell death), the inhibition of cell growth, and other toxic effects on cells caused by test materials and/or extracts from the materials. ISO and MHW procedures are contrasted in Table 2.6.

Table 2.7 Comparison of grading scales used to score responses of test animals to ASTM and ISO/USP procedures

	ASTM	ISO/USP
Response	Description	
Normal, no symptoms	Mouse exhibits no adverse physical symptoms after injection	
Slight	Mouse exhibits slight but noticeable symptoms of hypokinesis, dyspnea, or abdominal irritation after injection	
Moderate	Mouse exhibits definite evidence of abdominal irritation, dyspnea, hypokinesis, ptosis, or diarrhea after injection (Weight usually drops to between 15 and 17 g)	
Marked	Mouse exhibits prostration, cyanosis, tremors, or severe symptoms of abdominal irritation, diarrhea, ptosis, or dyspnea after injection (Extreme weight loss; weight usually less than 15 g)	
Dead, expired	Mouse dies after injection	
	Interpretation	*Interpretation*
	The test is considered negative if none of the animals injected with the test article extracts shows a significantly greater biological reaction than the animals treated with the control article If two or more mice show either marked signs of toxicity or die, the test article does not meet the requirements of the test If any animals treated with a test article shows slight signs of toxicity, and not more than one animal shows marked signs of toxicity or dies, a repeat test using freshly prepared extract should be conducted using groups of 10 mice each. A substantial decrease in body weight for all animals in the group, even without other symptoms of toxicity, requires a retest using groups of 10 mice each. In the repeat test, the requirements are met if none of the animals injected with the test article shows a substantially greater reaction than that observed in the animals treated with the control article	The test is considered negative if none of the animals injected with the test article shows a significantly greater biological reaction than the animals treated with the control article If two or more mice die, or show signs of toxicity such as convulsions or prostration, or if three or more mice lose more than 2 g of body weight, the test article does not meet the requirements of the test If any animal treated with a test article shows only slight signs of biological reaction, and not more than one animal shows gross signs of biological reaction or dies, a repeat test should be conducted using groups of 10 mice. On the repeat test, all 10 animals must not show a significantly greater biological reaction than the animals treated with the control article

Acute Systemic *Toxicity* Estimates the harmful effects of either single or multiple exposures to test materials and/or extracts, in an animal model, during a period of less than 24 hours. ISO and MHW procedures are contrasted in Table 2.7.

Table 2.8 Differences in hemolysis test procedures recommended by ISO 10993-4 and the MHW guidelines

ISO 10993-4	MHW 1995
Hemolysis can be assessed by any of several validated methods to assay hemoglobin in plasma	Hemolytic index is assessed by measuring hemoglobin at 1, 2, and 4 h by spectrophotometric methods The hemolysis over this period is expressed as a percentage of the positive control

Table 2.9 Comparison of pyrogen test procedures required by ISO 10993-11 and the MHW guidelines

ISO 10993-11	MHW 1995
Number of animals: Three rabbits required; comparison of febrile response in test animals to baseline temperature for evaluation of pyrogenicity potential	Three rabbits (test) required; comparison to baseline temperature is evaluated as index of pyrogenicity potential
Test duration: Test measurement intervals: every 30 min for 3 h	Test measurement intervals: every hour for 3 h
Evaluation: Cutoff for positive febrile response: 0.5 °C	Cutoff for positive febrile response: 0.6 °C

Hematocompatibility Evaluates any effects of blood contacting materials on hemolysis, thrombosis, plasma proteins, enzymes, and the formed elements using an animal model. Traditionally, hemolysis, which determines the degree of red blood cell lysis and the separation of hemoglobin caused by test materials and/or extracts from the materials in vitro, has been "the" representative test employed. A broader range of primary tests (adding evaluations of thrombosis, coagulation, platelets, and immunology aspects) is currently recommended. ISO and MHW procedures for hemolysis are contrasted in Tables 2.8 and 2.9.

Implantation Tests Evaluates the local toxic effects on living tissue, at both the gross level and microscopic level, to a sample material that is surgically implanted into appropriate animal implant site or tissue, e.g., muscle and bone, for 7–90 days. ISO and MHW procedures are contrasted in Table 2.10.

Genotoxicity The application of mammalian or non-mammalian cell culture techniques for the determination of gene mutations, changes in chromosome structure and number, and other DNA or gene toxicities caused by test materials and/or extracts from materials. Selected tests representing gene mutation tests (Ames or mouse lymphoma), chromosomal aberration tests (CHO), and DNA effects tests (mouse micronucleus and sister chromatid exchange) should generally be employed. ISO and MHW procedures are contrasted in Table 2.11.

Subchronic Toxicity The determination of harmful effects from multiple exposures to test materials and/or extracts during a period of 1 day to less than 10% of the total life of the test animal (e.g., up to 90 days in rats).

Table 2.10 Differences in ISO 10993-3 and the MHW guidelines for assessing the effects of device or material implantation

ISO 10993-3	MHW 1995
Time point(s) of assessment: Sufficient to achieve steady state(e.g., 2, 4, 6, and 12 weeks)	7 days and 4 weeks
Number of animals: At least three per time period of assessment	At least four per time period
Number of samples of evaluation: At least eight per time period for test and control	No minimum number specified
Evaluation criteria: Comparative evaluation of responses to test and control materials	If more than two of the four test sites in each animal exhibit a significant response compared to control sites, the test is considered positive

Table 2.11 Differences in genotoxicity testing procedures required by ISO 10993-3 and the MHW guidelines

ISO 10993-10	MHW 1995
Extraction vehicles: A physiological medium is used and, where appropriate, a solvent (e.g., dimethylsulfoxide)	Recommends methanol and acetone as extracting vehicles
Extraction: Extract test material and test the extract or dissolve material in solvent and conduct test. The conditions of extraction should maximize the amount of extractable substances, as well as subject the test device or material to the extreme conditions it may be exposed to, without causing significant degradation. Extraction ratio is dependent on thickness of test material	Extract at room temperature at a ratio of 10:1 (solvent/material) and obtain residue (at least 0.10.5% [weight of residue/weight of test material]), redissolve in appropriate solvent and test residue If sufficient residue is unobtainable, extract test material (in ethanol, acetone, or DMSO at 10 g of test material per 20 ml for the Ames mutagenicity assay and in cell culture medium at 120 cm^3 or 4 g/20 ml for the chromosomal aberration assay), at 37 °C for 48 hours and test extract. The Ames mutagenicity assay is conducted with a volume of 200 μl per plate.

Chronic Toxicity The determination of harmful effects from multiple exposures to test materials and/or extracts during a period of 10% to the total life of the test animal (e.g., over 90 days in rats).

Carcinogenesis Bioassay The determination of the tumorigenic potential of test materials and/or extracts from either single or multiple exposures, over a period of the total life (e.g., 2 years for rat, 18 months for mouse, or 7 years for dog).

Pharmacokinetics To determine the metabolic processes of absorption, distribution, biotransformation, and elimination of toxic leachables and degradation products of test materials and/or extracts.

Reproductive and Developmental Toxicity The evaluation of the potential effects of test materials and/or extracts on fertility, reproductive function, and prenatal and early postnatal development.

The tests for leachables such as contaminants, additives, monomers, and degradation products must be conducted by choosing appropriate solvent systems that will yield a maximal extraction of leachable materials to conduct biocompatibility testing. Chapter 3 addresses the issues behind sampling, sample preparation, and solvents.

The effects of sterilization on device materials and potential leachables, as well as toxic by-products, as a consequence of sterilization should be considered. Therefore, testing should be performed on the final sterilized product or representative samples of the final sterilized product. Table 2.11 presents the basis for test selection under the tripartite agreement.

2.8 United States Pharmacopoeial Testing

The earliest guidance on what testing was to be done on medical devices was that provided in the USP and other pharmacopoeias. Each of the major national pharmacopoeias offers somewhat different guidance. The test selection system for the USP (presented in Table 2.11), which classified plastics as Classes I through VI, is now obsolete and replaced in usage by the other guidelines presented here. But the actual descriptions of test types, as provided in the USP (and presented in the appropriate chapters later in this book), are still very much operative (USP 2007).

There are British, European, and Japanese pharmacopoeias, of which the latter requires the most attention due to some special requirements still being operative if product approval is desired.

2.9 ISO Testing Requirements

The European Economic Community adopted a set of testing guidelines for medical devices under the aegis of ISO (ISO 1992; The Gray Sheet 1992). The ISO 10993 guidelines for testing provide a unified basis for international medical device biocompatibility evaluation, both in terms of test selection (as presented in Tables 2.12 and 2.13) and test design and interpretation (Table 2.14).

The international standards specified methods of biological testing of medical and dental materials and devices and their evaluation in regard to their biocompatibility. Because of the many materials and devices used in these areas, the standard offers a guide for biological testing.

Table 2.12 Classification of plastics (USP XXIII)

Plastic classes[a]						Tests to be conducted			
I	II	III	IV	V	VI	Test material	Animal	Dose	Procedures[b]
x	x	x	x	x	X	Extract of sample in sodium chloride inspection	Mouse	50 ml/kg	A (iv)
x	x	x	x	x	x		Rabbit	0.2 ml/animal at each of 10 sites	B
x	x	x	x	x		Extract of sample in 1 in 20	Mouse	50 ml/kg	A (iv)
x	x	x	x	x		Solution of alcohol in sodium chloride injection	Rabbit	0.2 ml/animal at each of 10 sites	
	x		x	x		Extract of sample in polyethylene glycol 400	Mouse	10 g/kg	A (ip)
		x	x				Rabbit	0.2 ml/animal at each of 10 sites	
		x	x	x	x	Extract of sample in vegetable oil	Mouse	50 ml/kg	A (ip)
			x	x	x		Rabbit	0.2 ml/animal at each of 10 sites	B
		x			x	Implant strips of sample	Rabbit	4 strips/animal	C

[a]Tests required for each class are indicated by "x" in appropriate rows
[b]Legend: A (ip), systemic injection test (intraperitoneal); A(iv), systemic injection test (intravenous); B, intracutaneous (intracutaneous); C, implantation test (intramuscular implantation)
The table lists the biological tests that might be applied in evaluating the safety of medical devices and/or polymers. This does not imply that all the tests listed under each category will be necessary or relevant in all cases. Tests for devices made of metals, ceramics, biological materials, etc. are not included here but are under consideration
Categorization of medical devices is based on body contact and contact duration

2.10 MHW Requirements

The Japanese ISO test selection guidelines vary from those of FDA and ISO and are summarized in Table 2.16 (MHLW 2012).

Actual test performance standards also vary, as shown in Tables 2.4, 2.5, 2.6, 2.7, 2.8, 2.9, 2.10 and 2.11.

Committees dealing with materials and devices must decide on tests and test series relevant to the respective materials and devices, It is the responsibility of the product committees to select adequate test methods for products. The standard contains animal tests, but tries to reduce those tests to the justifiable minimum. Relevant international and national regulations must be observed when animals are used.

ISO 10993 is based on existing national and international specifications, regulations, and standards wherever possible. It is open to regular review whenever new research work is presented to improve the state of scientific knowledge. Tables 2.4 and 2.5 provide the test matrices under ISO 10993. Subsequently, specific guidance on individual test designs, conduct, and interpretation has been provided as subparts 2–11 of ISO-10993 (Table 2.14) (AAMI 2006).

Table 2.13 ISO initial evaluation tests

Device categories		Biological tests								
Body contact duration A—limited exposure B—prolonged or repeated exposure C—permanent contact		Cytotoxicity	Sensitization	Irritation or Intracutaneous	Acute systemic toxicity	Subchronic toxicity	Mutagenicity	Pyrogenicity	Implantation	Hemocompatability
Surface devices										
Skin	A	x	x	x						
	B	x	x	x						
	C	x	x	x						
Mucous membranes	A	x	x	x						
	B	x	x	x						
	C	x	x	x		x	x			
Breached surface	A	x	x	x						
	B	x	x	x						
	C	x	x	x		x	x			
Externally communicating										
Blood path indirect	A	x	x	x				x		x
	B	x	x	x				x		x
	C	x	x		x	x	x	x		x
Tissue/bone communicating	A	x	x	x						

	B	x	x				x		x	
	C	x	x				x		x	
Internal devices										
Circulating blood	A	x	x	x	x			x		x
	B	x	x	x	x		x	x		x
	C	x	x	x	x	x	x	x		x
Implant devices										
Bone/tissue	A	x	x	x						
	B	x	x				x		x	
	C	x	x				x		x	
Blood	A	x	x	x	x			x	x	x
	B	x	x	x	x		x	x	x	x
	C	x	x	x	x	x	x	x	x	x

Table 2.14 ISO special evaluation tests

Device categories		Biological tests			
Body contact duration A—limited exposure B—prolonged or repeated exposure C—permanent contact (time limits to added)		Chronic toxicity	Carcinogenicity	Reproductive/ developmental	Degradation
Surface devices					
Skin	A				
	B				
	C				
Mucous membranes	A				
	B				
	C				
Breached surface	A				
	B				
	C				
Externally communicating					
Blood path indirect	A				
	B				
	C	x	x		
Tissue/bone communicating	A				
	B				
	C		x		
Internal devices					
Circulating blood	A				
	B				
	C	x	x		
Bone/tissue	A				
	B				
	C	x	x		
Blood	A				
	B				
	C	x	x		

2.11 CE Marking of Devices

After June 14, 1998, all medical products distributed in Europe have had to bear the CE mark. ISO 9000 certification supplements and supports an assessment of conformity to the Medical Devices Directive (MDD), which must be performed by a certification body appointed by the EU member states. (Haindl 1997) To qualify for the CE mark, manufacturers of Class IIa, IIb, and III devices must be certified by a notified body (which is recognized by the national health authorities) to Annex II, V, or VI of the MDD (also known as 93/42/EEC) and comply with the essential require-

Table 2.15 ANSI/AAMI/ISO standards

	ISO Designations	Year Issued
Evaluation and testing	10993-1	2003
Animal welfare requirements	10993-2	2006
Tests for genotoxicity, carcinogenicity, and reproductive toxicity	10993-3	2003
Selection of tests for interactions with blood	10993-4	2002 and A1/2006
Tests for in vitro cytotoxicity	10993-5	1999
Tests for local effects after implantation Tests for irritation and delayed-type hypersensitivity	10993-6 BE78	1995/(R)2001 2002
Ethylene oxide sterilization residuals	10993-7	1995/(R)2001
Canceled	10993-8	-----
Framework for identification and quantification of potential degradation products	10993-9	1999/(R)2005
Tests for systemic toxicity	10993-11	2006
Sample preparation and reference materials	10993-12	2002
Identification and quantification of degradation products from polymeric devices	10993-13	1999/(R)2004
Identification and quantification of degradation products from ceramics	10993-14	2001
Identification and quantification of degradation products from metals and alloys	10993-15	2000
Toxicokinetic study design for degradation products and leachables from medical devices	10993-16	1997/(R)2003
Establishment of allowable limits for leachable substances	10993-17	2002
Physiochemical, morphological, and topographical characterization of materials Chemical characterization of materials	10993-19 BE83	2006 2006
Principles and methods for immunotoxicology testing of medical devices	10993-20	2006
Clinical investigation of medical devices for human subjects—Part 1: general requirements	14155-1	2003
Clinical investigation of medical devices for human subjects—Part 2: clinical investigation plans	14155-2	2003

ments of the directive. Manufacturers of active implantables and IVDs have separate directives to contend with. When auditing for compliance, the notified body will check a number of items in addition to a manufacturer's QA system, including technical files, sterility assurance measures, subcontracting procedures, recall and vigilance systems, and declarations of conformity. Depending on the classification and certification route, some devices will also require an EC-type examination or a design review by the notified body.

Manufacturers of Class I products, who require minimal interaction with a notified body, appear to be the clear winners in this scheme, but even they must deal with a number of vague or confusing requirements (see Table 2.16). Simply classifying their products according to the dictates of 93/42/EEC, Annex IX, can be a

Table 2.16 Japanese MHW test selection guidelines

Device categories		Initial evaluation						Supplemental evaluation					
Surface devices	Body contact	Contact duration	Cytotoxicity	Sensitization	Irritation or intracutaneous reactivity	Systemic toxicity (acute)	Subchronic toxicity	Genotoxicity	Pyrogen	Implantation	Hemocompatibility	Chronic toxicity	Carcinogenicity
		A	X	X	X								
	Skin	B	X	X	X								
		C	X	X	X								
		A	X	X	X								
	Mucosal membrane	B	X	X	X								
		C	X	X	X		X	X					
	Breached/	A	X	X	X								
	compromised	B	X	X	X								
	surface	C	X	X	X		X	X					
External communicating devices		A	X	X	X	X			X		X		
	Blood path indirect	B	X	X	X	X			X		X		
		C	X	X		X	X	X	X		X	X	X
	Tissue/bone dentin communicating	A	X	X	X								
		B	X	X				X		X			
		C	X	X				X		X			X
		A	X	X	X	X			X		X		
	Circulating blood	B	X	X	X	X		X	X		X		
		C	X	X	X	X	X	X	X		X	X	X
Implant devices		A	X	X	X								
	Bone/tissue	B	X	X				X		X			
		C	X	X				X		X		X	X
		A	X	X	X	X			X	X	X		
	Blood	B	X	X	X	X		X	X	X	X		
		C	X	X	X	X	X	X	X	X	X	X	X

A = temporary contact (<24 h) B = short- and medium-term contact (24 h and 29 days) C = long-term contact (>30 days)

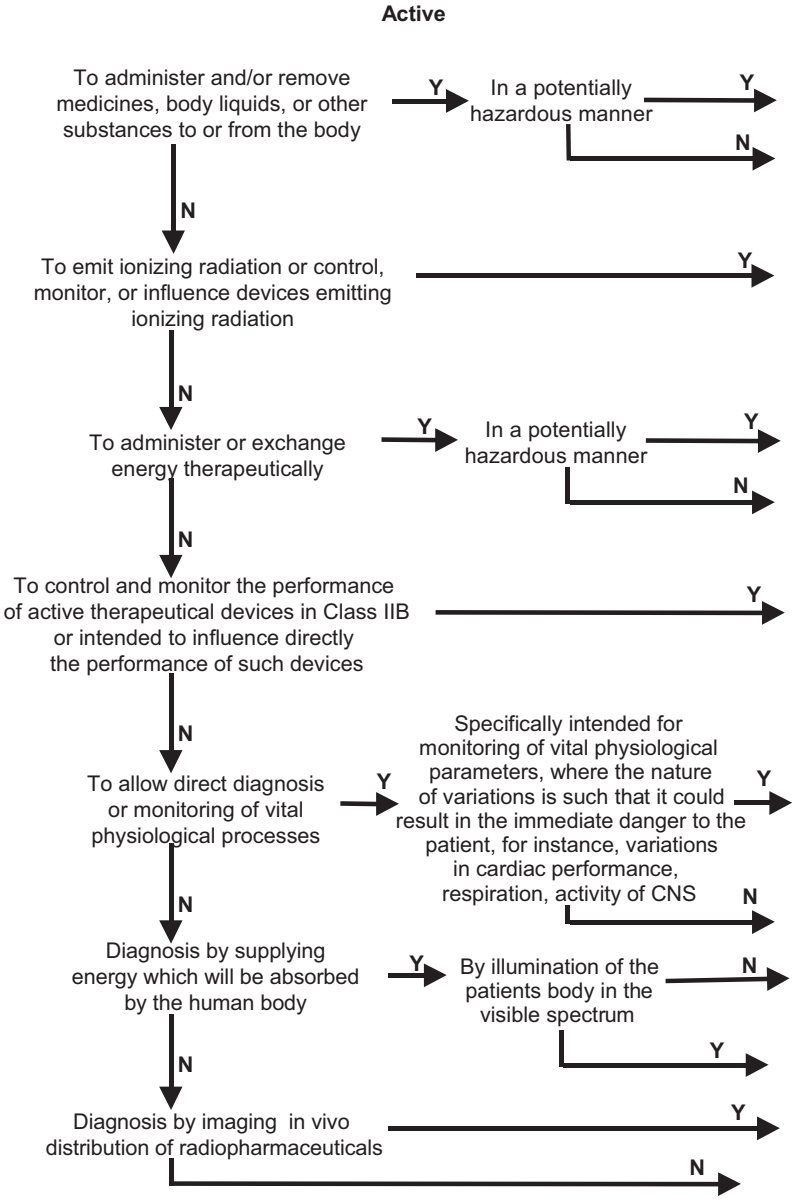


Fig. 2.2 Medical device classification flowchart

tricky affair, and faulty classification can lead to bigger problems. The simplified flowcharts in Fig. 2.2 should help manufacturers determine whether their products qualify as Class I devices. For more difficult products, manufacturers may need to refer to a consultant or obtain a suitable software program.

Classification is based on the intended and declared use of a product, not solely on its salient features. The Class I designation usually—but not always—excludes sterile products and measuring devices that measure physiological parameters or require a high degree of accuracy. So, for example, a reusable scalpel is Class I, but a sterile scalpel is Class IIa; a scalpel blade for the reusable device is Class I, but if it is supplied sterile, it is Class IIa; a scalpel blade for the reusable device is Class I but if it is supplied sterile, it is Class IIa. A stethoscope, a simple graduated syringe (not for injection pumps), and a measuring spoon for administering an expectorant are not considered measuring devices, although a hand-driven blood-pressure gage and a digital thermometer are.

All of the classification rules are included in the directive, but they're not easy to understand. An EC working group has drawn up a separate paper known as MEDDEV 10/93 to explain the rules and provide some practical guidelines. For example, the directive stipulates that reusable surgical instruments belong in the Class I designation as long as they are not intended for more than an hour of continuous use. According to this definition, items such as scissors and tweezers, even if they are used in a 6-hour operation, are still considered Class I devices because they are not used continuously during that time.

Even if a Class I product is supplied sterile, the manufacturer must issue a self-declaration of conformity. In this case, the manufacturer needs only to certify the QC system governing those aspects of manufacture concerned with securing and maintaining sterile conditions. If the device is packaged and sterilized by a company that works with a certified process, then the manufacturer must only validate the process for the particular device and submit the results to a notified body. The manufacturer still needs certification by a notified body in regard to the performance aspects relating to sterility and measurement function; the notified body will also want to inspect the manufacturer's facility Nonetheless, the procedure is far less complicated than a full production audit.

All manufacturers applying for CE marking privileges—including manufacturers of Class I devices—must prepare the proper technical documentation; appoint a "responsible person" within the EEC; design product labels and labeling according to 93/42/EEC, Annex I, paragraph 13; and sign a declaration of conformity. The technical dossier should not pose a major problem for manufacturers familiar with device master files. A list of required dossier contents is given in Table 2.17. For biological material testing, Europe uses the ISO 10993 (EN 30993) protocols, but test results according to the tripartite agreement (or USP XXIII) are accepted. Every electrical device must also be proven to comply with the EMC requirements defined in the MDD; suppliers of preassembled electrical components may have the appropriate test results already available. Reformatting an existing device master file is not necessary, only creating an index that cross-references the essential requirements of the directives with the device file contents. The master file is a controlled document, as defined in ISO 9000, and manufacturers would do well to regard it as highly confidential.

The technical dossier is closely linked to the responsible person, a representative in the EEC governed by European law and authorized by the manufacturer to oversee routine regulatory affairs. Specifically, the responsible person must ensure com-

Table 2.17 Which products are Class I?

The classification of a product refers to its intended use. The following is a simplified listing of Class I products:
• Noninvasive (and nonactive) devices that do not modify the biological or chemical composition of blood or liquids intended for infusion; store blood, body liquids, or tissues for administration; or connect to an active medical device
• Dressings intended only as a mechanical barrier or for absorption of exudates
• Invasive products for use in natural body orifices and stomas for no longer than 1 h or in the oral or nasal cavity or ear canal for up to 30 days
• Surgical invasive products if they are reusable instruments and not intended for continuous use of more than 1 h
• Active devices that administer neither energy nor substances to the body nor are made for diagnosis
Class I products cannot:
• Incorporate medicinal products (drugs) or animal tissue
• Be intended for contraception or the prevention of sexually transmitted diseases

pliance with the European vigilance system, which covers both post-market surveillance and adverse incident reporting. For example, if a patient were injured by a device, or if a patient would have been injured had the caregiver not intervened, the responsible person would have to investigate the incident together with the device's manufacturer and file a report with the competent authorities. Moreover, the European authorities must be able to obtain the master file in case of trouble; therefore, the manufacturer must either store the file or its abbreviated form with the responsible person or draw up a contractual agreement that gives the agent the right to access the master file without delay if required by the authorities. The agent must be available all year, as the time frame for notification could be as short as 10 days. Ideally, the responsible person should be familiar with the national regulation in all member states.

The simplest way to maintain a European address will be to appoint a distributor as their responsible person, although this course is not without potential problems. The selected distributor does not need certification as long as the manufacturer's name and CE mark are on the product labeling. The name of the responsible person must also appear on the label, package insert, or outer packaging, even if the product is sold by a completely different distributor in another country. There is no official rule or proposal regarding how many responsible persons a manufacturer should have, but each one must appear on the labeling; therefore, appointing more than one is of limited use. The responsible person should be selected with great care; device master files (Table 2.17) must be made available to the responsible person in the event of patient injury or near injury, and many distributors are potential competitors. Class I devices, by nature, will rarely lead to patient injury, but manufacturers should still consider labeling issues when choosing a representative. It's easy to change distributors, but changing the responsible person means changing all the product labeling. As an alternative, manufacturers can contract with a professional agency to serve as a representative completely independent from any distribution network.

The issue of labeling is itself a source of contention. Not all countries have decided yet whether they will insist on having their own language on device labels. Many countries have rather imprecise rules, dictating that their national language must appear only if necessary. Manufacturers can reduce potential trouble by using the pictograms and symbols defined in the harmonized European standard EN 980. For instructions of use, manufacturers are advised to use all 12 languages, used in the European Economic Area. The requirements for labeling are presented in Annex I, paragraph 13, of the MDD; some devices may be subject to additional requirements outlined in product standards.

Class I products fall under the jurisdiction of local authorities, but who serves as those authorities may differ from country to country. In Germany, for example, there are no clear-cut regulations that define the competence of the local authorities, except in the case of danger to the patient. European product liability laws more or less give the consumer the right to sue anybody in the trade chain. Normally, claims would be filed against the manufacturer, but it is possible that there will be claims against a responsible person. This is a rather new legal situation, and the rules will be determined by court decisions. It is hoped that Class I products will not instigate many court actions, but clearly, even manufacturers of Class I devices will have a host of new concerns under the CE marking scheme.

2.12 Risk Assessment

The reality is that not all materials used on devices are entirely safe. Generally, if one looks long enough at small enough quantities, some type of risk can be associated with every material. Risk can be defined as the possibility of harm or loss. Health risk, of course, is the possibility of an adverse effect on one's health. Risk is sometimes quantified by multiplying the severity of an event times the probability the event will occur, so that:

$$\text{Rist} = \text{serverity} \times \text{probability}$$

While this equation appears useful in theory, in practice it is difficult to apply to the biological safety of medical devices. The process known as health-based risk assessment attempts to provide an alternative strategy for placing health risks in perspective (Stark 1998; AAMI 1998).

2.13 Standards and Guidances

A paradigm for the risk assessment process has been detailed in a publication prepared by the US National Academy of Sciences (Hayes 2014). Although devised primarily for cancer risk assessment, many of the provisions also apply to the

assessment of other health effects. The major components of the paradigm are (1) hazard identification, (2) dosage-response assessment, (3) exposure assessment, and (4) risk characterization (Ecobichon 1992).

The general approach to risk assessment was adapted to medical devices via the draft CEN standard *Risk Analysis*, published in 1993,[1] and more recently via the ISO standard, *ISO 14538—Method for the Establishment of Allowable Limits for Residues in Medical Devices Using Health-Based Risk Assessment*, published in 1996.[2] At the present time, the FDA is also working to develop a health-based risk assessment protocol adapted to medical devices. Informally called the Medical Device Paradigm, the document is not yet generally available (Brown and Stratmeyer 1997).[3]

Some manufacturers may object that regulators are once again attempting to impose a "drug model" on medical devices. However, we shall see in the following pages that judicious application of these risk assessment principles can provide a justification for using materials that carry with them some element of risk, and that may, under traditional biocompatibility testing regimes, be difficult to evaluate or be deemed unsuitable for medical device applications.

2.14 Method

Hazard Identification The first step in the risk assessment process is to identify the possible hazards that may be presented by a material. This is accomplished by determining whether a compound, an extract of the material, or the material itself produces adverse effects and by identifying the nature of those effects. Adverse effects are identified either through a review of the literature or through actual biological safety testing.

Dose-Response Assessment The second step is to determine the dose-response of the material—that is, what is the highest weight or concentration of the material that will not cause an effect? This upper limit is called the *allowable limit*. There are numerous sources in the literature of data from which to determine allowable limits; some will be more applicable than others, and some may require correction factors.

Exposure Assessment The third step is to determine the exposure assessment by quantifying the *available dose* of the chemical residues that will be received by the patient. This is readily done by estimating the number of devices to which a patient

[1] *CEN BTS 3/WG 1—Risk Analysis* is available through the British Standards Institute.

[2] Available from the Association for the Advancement of Medical Instrumentation, 3330 Washington Blvd., Ste. 400, Arlington, VA 22201.

[3] Draft copies of the Medical Device Paradigm may be obtained by contacting Dr. Melvin Stratmeyer, FDA Center for Devices and Radiological Health, HFZ-112, Division of Life Sciences, Office of Science and Technology, FDA, Rockville, MD 20857.

is likely to be exposed in a sequential period of use (for instance, during a hospital stay) or over a lifetime. For example, a patient might be exposed to 100 skin staples following a surgical procedure or to two heart valves in a lifetime; thus, the amount of residue available on 100 skin staples or two heart valves would be determined.

Risk Characterization Characterizing the risk constitutes the final step of the process. The allowable limit is compared with the estimated exposure: if the allowable limit is greater than the estimated exposure by a comfortable safety margin, the likelihood of an adverse event occurring in an exposed population is small, and the material may be used.

2.15 Case Studies

We can best get a sense of how these standards work by looking at some actual medical case studies that illustrate the risk assessment process (Stark 1997).

Nitinol Implant Nitinol is an unusual alloy of nickel and titanium that features the useful property of "shape memory." A nitinol part can be given a particular shape at a high temperature and then cooled to a low temperature and compressed into some other shape; the compressed part will subsequently deploy to its original shape at a predetermined transition temperature. This feature is particularly beneficial for vascular implant applications in which the shape of the device in its compressed state eases the insertion process. The nitinol deploys as it is warmed by the surrounding tissue, expanding to take on the desired shape of a stent, filter, or other devices. The transition temperature depends on the alloy's relative concentrations of nickel and titanium: a typical nickel concentration of 55–60% is used in medical devices, since this gives a transition temperature at approximately the temperature of the body (37 °C).

Hazard Identification One concern with using nitinol in implant applications is the potential release of nickel into the body. Although nickel is a dietary requirement, it is also highly toxic—known to cause dermatitis, cancer subsequent to inhalation, and acute pneumonitis from inhalation of nickel carbonyl and to exert a toxic effect on cellular reproduction. It is a known sensitizer, with approximately 5% of the domestic population allergic to this common metal, probably through exposure from costume jewelry and clothing snaps. *The biocompatibility question at hand is whether or not* in vivo *corrosion of nitinol releases unsafe levels of nickel.*

Dose-Response Assessment A search of the world medical literature revealed that the recommended safe level of exposure to nickel in intravenous fluids is a maximum of 35 μg/day (Stark 1997). This value can be taken as an allowable limit of nickel exposure for a 70-kg (154-lb) adult.

The intravenous fluid data are based on subjects that are comparable to the patients who will be receiving nitinol implants. The data are for humans (not animals), for ill patients (not healthy workers or volunteers), and for similar routes of exposure (intravenous fluid and tissue contact). For these reasons, no safety correction factor need be applied to the allowable limit of exposure.

Exposure Assessment The available dose of nickel from nitinol implants can be estimated from data found in the literature. In one study, dental arch wires of nitinol were extracted in artificial saliva and the concentration of nickel measured in the supernatant. Corrosion reached a peak at day 7 and then declined steadily thereafter. The average rate of corrosion under these conditions was 12.8 μg/day/cm^2 over the first 28 days.

Risk Characterization A comparison of the available dose with the allowable limit for intravenous fluid levels shows that there is approximately a threefold safety margin, assuming that the implanted device is a full 1 cm^2 in surface area. (Devices with less surface area will contribute even less to the nickel concentration and have an even larger safety margin.) Considering the high quality of the data, a threefold safety margin is sufficient to justify using nitinol in vascular implants.

Wound Dressings Today's wound dressings are highly engineered products, designed to maintain the moisture content and osmatic balance of the wound bed so as to promote optimum conditions for wound healing. Complex constructions of hydrocolloids and superabsorbers, these dressings are sometimes used in direct tissue contact over full-thickness wounds that penetrate the skin layers.

Hazard Identification There have been reports in the literature of patients succumbing to cardiac arrest from potassium overload, with the wound dressing as one of the important contributors of excess potassium in the bloodstream. The effects of potassium on cardiac function are well characterized. Normal serum levels for potassium are 3.8–4 milliequivalents per liter. As the potassium concentration rises to 5–7 mEq/L, a patient can undergo cardiac arrest and die. *The biocompatibility issue to be explored is whether or not a wound dressing formulation might release dangerous levels of potassium if used on full-thickness wounds.*

Dose-Response Assessment An increase of approximately 1 mEq/L of potassium is unlikely to provoke mild adverse events in most patients. Assuming that the average person's blood volume is 5 L, a one-time dose of 5 mEq of potassium may begin to cause adverse reactions. This value can be considered to be the allowable limit of potassium for most patients.

Exposure Assessment Let us suppose that each dressing contains 2.5 g of potassium bicarbonate. Since the molecular weight of potassium bicarbonate is 100 g/mole, each dressing contains 0.025 mole of sodium bicarbonate or 0.025 mEq of potassium ion. If a patient were to use four dressings in a day, the available dose of potassium would be 0.1 mEq/day.

Risk Characterization Comparing the available dose of potassium (0.1 mEq) to the allowable limit (5 mEq) shows that there is a 50-fld safety margin. Considering that patients may be small in size, may have kidney impairment, or may receive potassium from additional sources such as intravenous fluids, this safety margin is too small, and so the dressing should be reformulated.

Ligature Material A manufacturer purchases commercial black fishing line to use as a ligature in a circumcision kit. Because the ligature is not "medical grade," a cytotoxicity test is routinely conducted as an incoming inspection test. It was assumed that a negative cytotoxicity test would be associated with an acceptable incidence of skin irritation.

Hazard Identification A newly received lot of the fishing line failed the cytotoxicity test. The extraction ration of this material—of indeterminate surface area—was 0.2 g/ml, with a 0.1-ml aliquot of sample extract being applied to a culture dish. Thus, 0.2 g/ml × 0.1 ml = 0.02 g represents a toxic dose of fishing line.

Dose-Response Assessment A titration curve was obtained on the sample extract. If the sample was diluted 1:2, the test was still positive; however, if the sample was diluted 1:4, the test was negative. Thus, 0.02 g/4 = 0.005 g of fishing line, the maximum dose that is not cytotoxic. This value was called the allowable limit of fishing line.

Exposure Assessment Each circumcision kit contained about 12 in. of line, but only about 4 in. of material was ever in contact with the patient. Since an 8-yd line was determined to weigh 5 g, the available dose of fishing line was calculated to be 5 g/288 in. × 4 in. = 0.07 g.

Risk Characterization A comparison of the available dose (0.07 g) with the allowable limit (0.005 g) convinced the manufacturer to reject the lot of fishing line.

2.16 Sources of Data

Data for calculating the allowable limit of exposure to a material can come from many sources, most of them promulgated by industrial and environmental hygienists and related agencies (Hayes 2014).

Threshold limit values (TLVs) are time-weighted average concentrations of airborne substances. They are designed as guides to protect the health and well-being of workers repeatedly exposed to a substance during their entire working lifetime (7–8 h/day, 40 h/wk). TLVs are published annually by the American Conference of Governmental Industrial Hygienists (ACGIH 1986). Biological Exposure Indices (BEIs) are also published annually by ACGIH. These are the maximum acceptable

concentrations of a substance at which a worker's health and well-being will not be compromised.

Other published guides include Workplace Environmental Exposure Levels (WEELs), from the American Industrial Hygiene Association (1980); Recommended Exposure Limits (RELs), from the US National Institute for Occupational Safety and Health; and Permissible Exposure Limits (PELs), from the US Occupational Safety and Health Administration. In the United States, PELs have the force of law.

Another important limit measurement, Short-Term Exposure Limits (STELs) are defined as the maximum concentration of a substance to which workers can be exposed for a period of up to 15 min continuously, provided that no more than four excursions per day are permitted, and with at least 60 min between exposure periods. The STEL allows for short-term exposures during which workers will not suffer from irritation, chronic or irreversible tissue damage, or narcosis of sufficient degree to increase the likelihood of injury, impair self-rescue, or materially reduce work efficiency. Some substances are given a "ceiling"—an airborne concentration that should not be exceeded even momentarily. Examples of substances having ceilings are certain irritants whose short-term effects are so undesirable that they override consideration of long-term hazards.

2.17 Uncertainty Factors (UCFs)

An uncertainty factor is a correction that is made to the value used to calculate an allowable limit. It is based on the uncertainty that exists in the applicability of the data to actual exposure conditions. Typically, uncertainty factors range in value from 1 to 10. For example, a correction factor of 10 might be applied for data obtained in animals rather than humans or to allow for a different route of exposure. In other words, for every property of available data that is different from the actual application, a correction factor of between 1 and 10 is applied. If our first example had been of a small amount of data obtained in animals by a different route of exposure, an uncertainty factor of 1000 might be applied.

2.18 Safety Margins

A safety margin is the difference or ration between the allowable limit (after correction by the uncertainty factor) and the available dose. How large does a safety margin need to be? Generally, a safety margin of 100× or more is desirable, but this can depend on the security of the risk under consideration, the type of product, the business risk to the company, and the potential benefits of product use.

2.19 End Note

Two things need to be kept in mind when considering guideline testing requirements.

First, FDA has many "tools" available should the situation arise. The most obvious are the device-specific case devices guidelines such as those for breast implants, cardiovascular guidewires, ocular lenses, and tampons (to name but a few). It is beyond the scope of this chapter to cite and explain each of these, but all four of the above are examples I have had to consider and incorporate in responses during the last year. Also included are past guidances that are still active and sometimes still relevant and referenced to by the FDA, for example, the CDRH (2006) methods of evaluating immunotoxicity.

Second, regulation and requirements for biocompatibility testing are continuing far to evolve at an even increased rate. Change is the only eternal and testing data from even 3 years ago may no longer be acceptable (Tables 2.18, 2.19 and 2.20).

Table 2.18 Contents of a device master file

1. EC declaration of conformity and classification according to Annex IX of the MDD
2. Name and address of the manufacturer's European responsible person
3. Product description, including:
• All variants
• Intended clinical use
• Indications/contraindications
• Operating instructions/instructions for use
• Warnings/precautions
• Photographs highlighting the product
• Photographs highlighting the usage
• Brochures, advertising, catalog sheets, marketing claims (if available)
• Product specifications including:
• Parts list, list of components
• Specifications of materials used, including data sheets
• List of standards applied
• Details of substance(s) used (in the event of drug-device combination)
• QA specifications (QC specs, in-process controls, etc.), etc.
• Labeling, accompanying documents, package inserts (DIN EN 289, prEN 980)
• Instruction for use (prEN 1041)
• Service manual
• Product verification, including:
• Testing data and reports, functionality studies, wet lab or benchtop testing
• Materials certificates/reports on biological tests
• EMC testing and certificates
• Validation of the packaging/aging studies
• Compatibility studies (connection to other devices)
• Risk analysis (DIN EN 1441)
• Clinical experience
4. List of requirements (Annex I) indicating cross-reference with documentation

Table 2.19 An overview of the classification of medical devices (Rules 1–12)

Class	Noninvasive devices				Invasive devices				Additional rules for active devices			
						Surgically invasive devices						
	Others	Channeling or storing substances for introduction into the body	Biological or chemical modification of liquids for infusion	Contact with injured skin	Body orifices	Transient use	Short-term use	Long-term use, implantable devices	Therapeutic devices for administration or exchange of energy	Diagnostic devices to supply energy, vital physiological processes; radiopharmaceutical imaging	Devices for administration or removal of substances to or from the body	Others
I	Regular	Regular		Mechanical barrier; compression; absorption of exudates	Transient use; ENT short-term	Reusable surgical instruments				Illumination of human body in visible spectrum		Regular
IIa		Body substances; connections to AMD ≥ IIa	Filtration; centrifugation; gas or heat exchange	Regular	Short-term use ENT, long-term connection to AMD ≥ IIa	Regular	Regular	Placed in teeth	Regular	Regular	Regular	
IIb			Regular	Wounds with breached dermis, healing by secondary intent	Long-term use	Ionizing radiation; biological effect; absorbed; potential hazard of medicine delivery system	Ionizing radiation; chemical change (except in teeth); medicine administration	Regular	Potentially hazardous (nature, density, site of energy); Class IIb ATD monitor control development	Immediate danger to heart, respiration, CNS; ionizing radiation including control monitoring	Potentially hazardous (substances, part of body, mode of applications)	
III						Heart; CCS	Heart; CCS; CNS; biological effect; absorbed	Heart; CCS; CNS; biological effect; absorbed; chemical change; medicine administration				
Rule	1	2	3	4	5	6	7	8	9	10	11	12

Key: *AMD* active medical device, *ATD* active therapeutic device, *CCS* central circulatory system, *CNS* central nervous system, *ENT* ear, nose, and throat

Table 2.20 Reduced scheme for combined characteristics of active and nonactive devices

Class	Noninvasive devices				Invasive devices				Additional rules for active devices			
					Surgically invasive devices							
	Others	Channeling or storing substances for introduction into the body	Biological or chemical modification of liquids for infusion	Contact with injured skin	Body orifices	Transient use	Short-term use	Long-term use, implantable devices	Therapeutic devices for administration or exchange of energy	Diagnostic devices to supply energy, vital physiological processes; radiopharmaceutical imaging	Devices for administration or removal of substances to or from the body	Others
I										Illumination of human body in visible spectrum		Regular
IIa		Body substances	Filtration; centrifugation; gas or heat exchange	Regular	Short-term use ENT, long-term	Regular	Regular		Regular	Regular	Regular	
IIb			Regular	Wounds with breached dermis, healing by secondary intent	Long-term use	Biological Effects			Potentially hazardous (nature, density, site of energy); Class IIb ATD monitor control development	Immediate danger to heart, respiration, CNS; ionizing radiation including control monitoring	Potentially hazardous (substances, part of body, mode of applications)	
III						Heart; CCS	Heart; CCS; CNS; biological effects					
Rule	1	2	3	4	5	6	7	8	9	10	11	12

Key: *ATD* active therapeutic device, *CCS* central circulatory system, *CNS* central nervous system, *ENT* ear, nose, and throat

References

AAMI. (1998). *AAMI standards: Reduced devices – Risk management – Part 1: Applications*. AAMI/ISO 14971-1. Arlington: AAMI.

AAMI. (2006). *AAMI standards: Vol. 4, biological evaluation of medical devices*. Arlington: AAMI.

ACGIH. (1986). *Documentation of the threshold limit values for substances in workroom air* (5th ed.). Cincinnati: American Conference of Governmental Industrial Hygienists.

AIHA. (1980). *Hygienic guide series* (Vol. Vols. I and II). Akron: American Industrial Hygiene Association.

Brown, R. P., & Stratmeyer, M. (1997). Proposed approach for the biological evaluation of medical device materials. In *Proceedings of the Medical Design and Manufacturing East 97 Conference and Exposition* (pp. 205-9–205-18). Santa Monica: Canon Communications.

CDRH. (1992). *Regulatory Requirements for Medical Devices: A Workshop Manual*, Center for Device and Radiological Health, HHS Publication FDA 92-4165, August 1992.

CDRH. (1995a). *Premarket notification (510(k)) guidance document for contact lens car products*. Washington, DC: Center for Device and Radiological Health, Food and Drug Administration.

CDRH. (1995b). *Testing guidelines for class III soft (hydrophilic) contact Lens solutions*. Washington, DC: Center for Device and Radiological Health, Food and Drug Administration.

CDRH. (1995c). *Draft guidance for the content of premarket notifications for menstrual tampons*. Washington, DC: Center for Device and Radiological Health, Food and Drug Administration.

CDRH. (2006). *Immunotoxicity testing guidance. U.S. Department of Health and Human Services: FDA*. Rockville: Silver Spring.

CFR. (1992). FDA's policy statement concerning cooperative manufacturing arrangements for licensed biologics, *Federal Register*, 57:55544, November 25, 1992.

Council Directive 93/42/EEC of 14 June 1993 Concerning Medical Devices, *Official J Eur Communities*, 36 (July 12): 1, 1993.

Ecobichon, D. J. (1992). *The basis of toxicology testing*. Boca Raton: CRC Press.

Eisenhart S. (2017). EMERGO Study: FDA 510(k) Submissions from US Companies on the Decline. *EMERGO United States Medical Device QA/RA Blog*, Austin, TX. Available online at: https://www.emergobyul.com/blog/2017/03/emergo-study-fda-510k-submissions-us-companies-decline. Last Accessed 6 July 2018.

FDA. (1995). *EPA Bluebook Memorandum #G95: Use of International Standard ISO-10993*. Biological evaluation of Medical Devices Part 1: Evaluation and Testing = Food and Drug Administration, Washington, DC.

FDA. (2016) *Use of International Standard ISO 10993-1, "Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process"*, CDRH, June 16, 2016.

FDLI. (1995). *Compilation of food and drug laws*, Volumes I and II, Volume III (1996), Supplement (1998). Washington, DC: Food and Drug Law Institute.

"Final Draft Guidelines on Medical Device Classification," MEDDEV 10/93, Brussels, European Commission, October 1993.

Food and Drug Administration (1987). Good Laboratory Practice Regulations: Final Rule. *Federal Register*, Part VI, Vol. 52, No. 172. September 4, 1987.

Food and Drug Administration. (1997). Premarket Notification [510k] Guidance Document for Class II Daily Wear Contact Lenses. Accessible at: http://www.fda.gov/downloads/MedicalDevices/.../ucm080960.pdf. Washington, DC.

Fries, R. C. (1999). *Medical device quality assurance and regulatory compliance*. New York: Marcel Dekker.

Gad, S. C. (2001). *Regulatory toxicology* (2nd ed.). London: Taylor and Francis.

Gad, S. C., & Chengelis, C. P. (1998). *Acute toxicology*. La Jolla: Academic Press.

Gad, S. C., & Taulbee, S. (1996). *Handbook of data recording, maintenance and management for the biomedical sciences*. Boca Raton: CRC Press.

Goering, P. L., & Galloway, W. D. (1989). Toxicology of medical device material. *Fundamental and Applied Toxicology, 13*, 193–195.

Haindl, H. (1997). CE marking via self-declaration. *Medical Device and Diagnostic Industry, 19*, 86–90.

Hayes, A. W. (2014). *Principles and methods of toxicology* (6th ed., pp. 26–58). New York: Macmillan.

Heller, M. A. (1999). *Guide to medical device regulation* (Vol. Vols. 1 & 2). Washington, DC: Thompson Publishing Group.

ISO. (1992). *Biological evaluation of medical and dental materials and devices*. Brussels: International Organization for Standardization (ISO).

ISO. (1996). *ISO 14538: Risk and Hazard assessment of medical devices*. Brussels: International Organization for Standardization (ISO).

ISO. (2008). *ISO 7405: Evaluation of biocompatibility of medical devices used in dentistry*. Brussels: International Organization for Standardization (ISO).

ISO. (2017). *ISO 18562:1-4 biocompatibility evaluation of breathing gas pathways in healthcare application*. Brussels: International Organization for Standardization (ISO).

ISO (various dates). *Biological evaluation of medical devices. ISO 10993, parts 1–20*. Geneva: International Organization for Standardization, various dates.

Kahan, J. S. (1995). *Medical devices–obtaining FDA market clearance*. Watham: Parexel.

MHLW. (2012). *Basic principles of biological safety evaluation required for application for approval to market medical devices*. Yakuji: Nippo, LTD..

MHW. (1995). *Japanese Guidelines for Basic Biological Tests for Medical Devices and Materials*, notification no. 99, (Tokyo: Pharmaceutical Affairs Bureau, Ministry of Health and Welfare, July 27, 1995).

Stark, N. J. (1997). Case studies: Using the world literature to reduce biocompatibility testing. In *Proceedings of the Medical Design and Manufacturing East 97 Conference and Exposition* (pp. 205-1–205-7). Santa Monica: Canon Communications.

Stark, N. J. (1998). Conducting health-based risk assessment of medical materials. *Medical Plastics and Biomaterials, 5*, 18–25.

The Gray Sheet. (1992). EC 'Medical Devices' Directive Slated for Adoption in Mid-1993, EC Commission Official Says: CEN Estimate Development of 92 Standards for Directive, M-D-D-1 Reports, *The Gray Sheet*, October 12, 1992.

The Gray Sheet. (1996). European Union Class III Device Approvals Average 240 Days or Less, HIMA Survey Says: Study Release Intended to Bolster Support for FDA Reform Legislation, *The Gray Sheet*, February 26, 1996, pp. 7–8.

USP. (2007). *Biological tests-plastics, The United States Pharmacopoeia* (XXX ed., pp. 1235–1238). Rockville: United States Pharmacopoeial Convention.

Chapter 3
Biocompatibility Testing: The Biologic Tests

Shayne C. Gad

Abstract Most biocompatibility testing requirements are drawn from a common set of designs—from those guidelines called out by the ISO-10993-1 table. All devices require three basic end points to be addressed—cytotoxicity, irritation and sensitization. Most studies have as a predicate the preparation of extracts using polar and nonpolar vehicles (sample preparations). This chapter starts examining the details for sample preparation and then presents similar examinations of each major standard test type. In each case, the performance and results evaluating are followed by examination of common issues and routes to their resolutions. Finally, uncommon biologic test types are reviewed.

Keywords Autoimmunity · Bacterial mutation tests · Bruce Ames · Carcinogenicity · Chronic toxicity · Clinical pathology · Clotting time · Coagulation · Complement activation · Cytogenetic assay · Cytotoxicity · Dermal irritation · Endotoxin · Exhaustive extraction · Genotoxicity · Guinea pig maximization test (GMPT) · Hemocompatibility · Hemolysis · Histopathology · Host-resistant assays · Human peripheral blood lymphocytes · Humoral (innate) immunity · Immunopathologic · Immunotoxicity · Immunotoxicology · Implantation · Implantation as a method for other end points · Intracutaneous reactivity · Japanese (MHW) extraction methods · Local lymph node assay (LLNA) · Material-mediated pyrogenicity · Microbiome · Ocular irritation · Pyrogenicity · Reference materials · Sample preparation · Sensitization · Subacute toxicity · Subchronic toxicity · Systemic toxicity · T-cell (adaptive) lymphoproliferation · Thrombogenicity · USP rabbit pyrogen test

Here we consider how to actually perform the traditional biologic test components of a biocompatibility assessment, starting with selection of test articles (samples) and the preparation or the derivation of extracts to use in testing. That is, how to select actual individual items (usually in triplicate) for testing (sampling) and how

S. C. Gad, PhD, DABT (✉)
Gad Consulting Services, Raleigh, NC, USA
e-mail: scgad@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_3

to prepare the items selected for the testing process (sample preparation). ISO 10993-12 addresses both these issues (as well as the selection and use of reference materials) in summary fashion.

This chapter focuses on the utilitarian aspects of test selection and performance, as well as pointing out and addressing commonly encountered issues in test performance and interpretation. A previous volume (Gad and Gad McDonald 2016) by the authors provides an extensive coverage of theory and background.

3.1 Sample Preparation (ISO 10993-12)

How samples are prepared for testing once selected is as critical as what samples are tested. With devices, much of sample preparation centers around the derivation of extracts for use in tests where inclusions of an intact solid device are either inappropriate or not physically positive.

Sample test materials or test components should be in final "market ready" form—cleaned and sterilized just as a final product supplied to a patient would be. The manufacturer should choose a biologically relevant solvent system that is listed in the ISO guidance, does not deform the physical structure of the device, and will yield a quantity of extract sufficient to perform tests for biocompatibility. The observed biological response results from the combination of the concentration of the substances that reach the target cells and the intrinsic activity of the substances upon these cells. No single simple extraction procedure would simulate the effect of exposure to the physiological environment. For example, serum contains electrolytes, a variety of fats, and nitrogenous compounds. However, the extraction solution should be chemically and biologically uncomplicated so that the extraction solution itself does not interfere with subsequent tests. The extraction solution should include an appropriate combination of polar and nonpolar solvents, and the extraction procedure should occur in a static condition as well as under agitation. The ratio of product to extraction solution and the time for the extraction process should reflect the quantities and dwell times occurring in the anticipated use of the device.

Unless extraction is for direct testing in a mammalian cell culture (in which case culture media should be employed), at least two solvents, polar and nonpolar, are used to obtain soluble extracts for biocompatibility tests preferable at an elevated temperature. Polar solvent can be water, saline solution, and/or water or saline solution with alcohol. Nonpolar solvent can be cottonseed oil (CSO), sesame oil, and/or polyethylene glycol (DEG) solution. The most commonly employed solvents for extraction are characterized in Table 3.1. Culture medium may also be used for cell-based test systems such as in cytotoxicity testing.

Extracting conditions should attempt to exaggerate the clinical use conditions so as to define the potential toxicological hazard without causing significant changes such as fusion or melting of the material pieces or altering the chemical structure. If a device component melts or loses shape during extraction, the resulting extractant solution cannot be used in a valid test.

Table 3.1 Commonly utilized extractants and vehicles

Common name: **Cottonseed oil (CSO)** Chemical name: NA Molecular weight: NA Formula: Mixture of natural products; glycerides of palmitic, olive, and linoleic acids Density: 0.915–0.921 g/ml Volatility: Low Solubility/miscibility: Soluble in ether, benzene, chloroform, and DMSO. Slightly soluble in ethanol Biological considerations: Orally, serves as energy source (and therefore can alter food consumption and/or body weight). Prolonged oral administration has been associated with enhanced carcinogenesis Chemical compatibility/stability considerations: Thickens upon prolonged exposure to air. Available in USP grade Uses (routes): In extractions and as a vehicle for oral, dermal, vaginal, rectal, and subcutaneous administration
Common name: **DMSO/dimethyl sulfoxide** Chemical name: Sulfinylbis [methane]; CAS #67-68-5 Molecular weight: 78.13 Formula: C_2H_6OS Density: 1.100 g/ml at 20 ° C Volatility: Medium Solubility/miscibility: Soluble in water, ethanol, acetone, ether, oils Biological considerations: Oral LD50 (rats) = 17.9 ml/kg. Repeated dermal exposure can defat skin. Repeated oral exposure can produce corneal opacities. Not cytotoxic to cells in primary culture at less than 0.05% (V/V). Intraperitoneal LD50 (mice) = 11.6 ml/kg. Chemical compatibility/stability considerations: Very hydroscopic liquid. Combustible Uses (routes): All, as a carrier at up to 5% to enhance absorption
Common name: **Ethanol; EtOH** Chemical name: Ethyl alcohol; CAS #64-17-5 Molecular weight: 46.07 Formula: C_2H_5OH Density: 0.789 g/ml Volatility: High, but declines when part of mixture with water Solubility/miscibility: Miscible with water, acetone, and most other vehicles Biological considerations: Orally, will produce transient neurobehavioral intoxication. Oral LD50(rats) = 13.0 ml/kg. Intravenous LD50 (mice) = 5.1 ml/kg Chemical compatibility/stability considerations: Flammable colorless liquid available USP grade Uses (routes): Extraction solvent vehicle for dermal and oral, though can be used in lower concentrations for most other routes. Volume of oral instillation should be limited to 5 ml/kg
Common name: **Polyethylene glycol (PEG)** Chemical name: NA Molecular weight: 400 (approximate average, range 380–420) Formula: $H(OCH_2CH_2)_nOH$ Density: 1.128 g/ml Volatility: Very low Solubility/miscibility: Highly soluble in water. Soluble in alcohol and many organic solvents Biological considerations: Employed as water-soluble emulsifying/dispersing agents. Oral LD50 (mice) = 23.7 ml/kg. Oral LD50 (rats) = 30 ml/kg Chemical compatibility/stability considerations: Do not hydrolyze or deteriorate on storage and will not support mold growth. Clear, viscous liquid Uses (routes): As extraction solvent for oral administration as a vehicle full strength or mixed with water. Total dosage of PEG-400 should not exceed 5–10 ml

(continued)

Table 3.1 (continued)

Common name**: Saline**
Chemical name: Physiological saline; isotonic salt solution
Molecular weight: 18.02
Formula: 0 l9% NaCl in water (weight to volume)
Density: As water
Volatility: Low
Solubility/miscibility: As water
Biological considerations: No limitations—Preferable to water in parenteral applications
Chemical compatibility/stability considerations: None
Uses (routes): Extraction solvent all except perocular

Source: Gad and Chengelis (1998); Lewis (1993)

The results derived from tests where the conditions of extraction were exaggerated need to be viewed in light of these exaggerations. Judgment needs to be used in interpreting the results as to their appropriateness to the actual use conditions and device potential toxicity.

The concentration of any endogenous or extraneous substances in the extract, and hence the amount exposed to the test cells, depends on the interfacial area, condition of the sample surface, the extraction volume, pH, chemical solubility, osmolarity, agitation, temperature, and other factors. These conditions should each be carefully considered. It should also be remembered that for solid polymer and elastomer components, unfinished areas (such as are exposed if an elastomeric closure or stopper piece is cut) are likely to have more leachable materials than are present in a more fully cured or finished surface.

Use of a lipophilic (non-polar, such as CSO) and a hydrophilic (polar, generally saline) solvent system simulates likely physiologic extraction conditions in use. The addition of ethanol and PEG provides a fair representation of potential extraction conditions when the device is in extended contact with a drug or therapeutic solution (Autian 1977).

General points or guidance for extraction include (AAMI 2014):

- The extraction shall be performed in sterile, chemically inert containers by using aseptic techniques.
- The extraction time and temperature are dependent on the physicochemical characteristics of the material and extraction vehicle. Expected (by the regulators) conditions are (AAMI 2014; ISO 2012):

  (a) 37 °C for 72 hours (the current preferred extraction method only for situations where there is a component that would be damaged by higher temperatures, such as protein in culture media).
  (b) 72 hours at 50 °C (unless it has a protein or the temperature will physically deform your device or a component thereof, this is the usual case). As will be seen in (d) below, for extractables and in situations where a device is implanted for a long time, “exhaustive extraction” (not just 72 hours) at 50 °C is expected.
  (c) 1 hour at 121 °C (for metals and ceramics).

(d) Exhaustive extraction at 40 °C for leachables and extractables (L&E) assessments. This involved extracting in successive aliquots of solvent until there is no more than a 10% increase in quantities/concentrations of extracted materials of that for the initial period.

Extraction conditions should simulate as closely as possible the conditions under which the device will normally be used. Therefore, item (a) gives the preferred conditions for extraction, and generally the other conditions are when a device is intended for steam or thermal sterilization or resterilization (such as with surgical instruments).

The recommended conditions should be applied according to the device characteristics and specific conditions of use.

Extraction procedures using culture medium with serum can only be used under the conditions specified in (a) above (i.e., not less than 24 hours at 37 °C). It should be noted that such media has both lipophilic and hydrophilic component.

When agitation is considered to be appropriate, the method should be specified and reported. Elastomeric materials should never be cut prior to extraction, and polymers should not be cut into small pieces before extraction except in extraordinary circumstances. The ratio between the surface area of the material and the volume of extraction vehicle shall be no more than 6 cm^2/ml. The surface area shall be calculated on the basis of the overall sample dimensions, not taking into account surface irregularity and porosity. However, the actual surface characteristics should be considered in the interpretation of the test results. If the surface area is indeterminate, then 0.1 g/ml to 0.2 g/ml shall be used.

Liquid extracts shall, if possible, be used immediately after preparation.

If an extract is stored, then the stability of the extract under the conditions of storage should be verified with appropriate methods.

If the extract is filtered, centrifuged, or processed by other methods prior to being applied to the cells, this must be included in the final report.

For use in direct contract tests (such as implantation studies), materials which have various shapes, sizes or physical states (i.e., liquid or solid) may be tested without modification in the cytotoxicity assays.

The preferred sample of a solid specimen should have at least one flat surface. Adjustments shall be made for other shapes and physical states.

Japanese (MHW) extraction methods are test method dependent.

The sterility of the test specimen shall conform to the requirements in the USP (2013).

Test materials from sterilized devices which are normally supplied non-sterile but are sterilized before use shall be sterilized by the method recommended by the manufacturer and handled aseptically throughout the extraction and test procedure.

Test materials from devices not required to be sterile in use shall be used as supplied and handled aseptically throughout the extraction and test procedure.

Liquids shall be tested by either (a) direct deposition or (b) deposition onto a biologically inert absorbent matrix (filter discs have been found to be suitable).

If appropriate, materials classed as super-absorbent shall be wetted with culture medium prior to testing.

The USP (2013) provides specific guidance for use in preparing extraction solutions for use in biological reactivity tests. These are as follows.

3.1.1 Apparatus

For the tests includes the following.

3.1.1.1 Autoclave

Use an autoclave capable of maintaining a temperature of 121 ± 2.0°, equipped with a thermometer, a pressure gauge, a vent cock, a rack adequate to accommodate the test containers above the water level, and a water-cooling system that will allow for cooling of the test containers to about, but not below, 20° immediately following the heating cycle.

3.1.1.2 Oven

Use an oven, preferably a forced-circulation model, that will maintain operating temperatures of 50° or 70° within ±2°.

3.1.1.3 Extraction Containers

Use only containers, such as ampuls or screw-cap culture test tubes, of Type I glass. If used, culture test tubes are closed with screw caps having suitable elastomeric liners. The exposed surface of the elastomeric liner is completely protected with an inert solid disk 0.05 mm to 0.075 mm in thickness. A suitable disk may be fabricated from a polytetrafluoroethylene (polytef) resin.

3.1.2 Preparation of Apparatus

Cleanse all glassware thoroughly with chromic acid cleansing mixture, or if necessary with hot nitric acid, followed by prolonged rinsing with water. Clean cutting utensils by an appropriate method (e.g., successive cleaning with acetone and methylene chloride) prior to use in subdividing a specimen. Clean all other equipment by thorough scrubbing with a suitable detergent and prolonged rinsing with water.

Render containers and equipment used for extraction, and in transfer and administration of test material, sterile and dry by a suitable process. [If ethylene oxide is used as the sterilizing agent, allow adequate time for completing degassing.]

3.2 Procedure

3.2.1 Preparation of Sample

Both the *Systemic Injection Test* and the *Intracutaneous Test* may be performed using the same extract, if desired, or separate extracts may be made for each test. Select and subdivide into portions a Sample of the size indicated in Table 3.2. Note that ISO 10993 (Part 12) guidance is equivalent but stated differently, as present in Table 3.3. Remove particulate matter, such as lint and free particles, by treating each

Table 3.2 Surface area of specimen to be used[a]

Form of material	Thickness	Amount of *Sample* for each 20 mL of extracting medium	Subdivided into
Film or sheet	<0.5 mm 0.5 to 1 mm	Equivalent of 120 cm^2 total surface area (both sides combined) Equivalent of 60 cm^2 total surface area (both sides combined)	Strips of about 5 × 0.3 cm
Tubing	<0.5 mm (wall) 0.5 to 1 mm (wall)	Length (in cm) = 60 cm^2/(sum of ID and OD circumferences) Length (in cm) = 60 cm^2/(sum of ID and OD circumferences)	Sections of about 5 × 0.3 cm
Slabs, tubing, and molded items	>1 mm	Equivalent of 60 cm^2 total surface area (all exposed surfaces combined)	Pieces up to about 5 × 0.3 cm
Elastomers	>1 mm	Equivalent of 25 cm^2 total surface area (all exposed surfaces combined)	Do not subdivide[b]

[a]When surface area cannot be determined due to the configuration of the specimen, use 0.1 g of elastomer or 0.2 g of plastic or other polymers for every 1 mL of extracting fluid
[b]Molded elastomeric closures are tested intact

Table 3.3 Extraction ratios[a]

Form/material	Thickness	Surface area/volume
Non-absorbent	Not applicable	6 cm^2/ml
Absorbents and hydrocolloids[b]	Not applicable	$\frac{0.1\text{g}}{(1\text{ml} + \text{absorption capacity})}$
Film, sheet, or tubing wall[c]	<0.5 mm 0.5 to 1 mm	6 cm^2/ml 3cm^2/ml
Slabs, tubing, and molded items	>1 mm	3cm^2/ml
Elastomers[d]	>1 mm	1.25 cm^2/ml
Indeterminate surface area	Not applicable	0.2 gm sample/ml or 0.1 gm elastomer/ml

[a]ISO 10993—Part 12: Sample Preparation and Reference Materials
[b]Based on a technique developed by NJ Stark
[c]May be subdivided into strips or sections
[d]Do not subdivide: cut edges have different extraction properties than outer surfaces

subdivided Sample or Negative Control as follows: Place the Sample into a clean, glass-stoppered, 100-mL graduated cylinder of Type I glass, and add about 70 mL of water for injection. Agitate for about 30 seconds, and drain off the water, repeat this step, and dry those pieces prepared for the extraction with vegetable oil in an oven at a temperature not exceeding 50°. [*Note*—Do not clean the Sample with a dry or wet cloth or by rinsing or washing with an organic solvent, surfactant, etc.]

3.2.2 Preparation of Extracts

Place a properly prepared Sample to be tested in an extraction container, and add 20 mL of the appropriate extracting medium. Repeat these directions for each extracting medium required for testing. Also prepare one 20-mL blank of each medium for parallel injections and comparisons. Extract by heating in an autoclave at 121° for 60 minutes, in an oven at 70° for 24 hours, or at 50° for 72 hours. Allow adequate time for the liquid within the container to reach the extraction temperature.

Sample preparation actions and extraction conditions should not in any instance cause physical changes such as fusion or melting of the Sample pieces, which result in a decrease in the available surface area. A slight adherence of the pieces can be tolerated. Always add the cleaned pieces individually to the extracting medium. If culture tubes are used for autoclave extractions with vegetable oil, seal screw caps adequately with pressure-sensitive tape.

Cool to about room temperature but not below 20°, shake vigorously for several minutes, and decant each extract immediately, using aseptic precautions, into a dry, sterile vessel. Store the extracts at a temperature between 20° and 30° and do not use for tests after 24 hours. Of importance are the contact of the extracting medium with the available surface area of the plastic and the time and temperature during extraction; the proper cooling, agitation, and decanting process; and the aseptic handling and storage of the extracts following extraction.

One should be aware that close reading requirements under MHW Notification 99 is required. Otherwise, exaggerated extraction conditions may be inappropriately and erroneously employed (such as in the case of biologically derived materials such as collagen).

3.2.3 Reference Materials

In nearly every biocompatibility test, reference materials are used to serve as experimental controls. Negative controls, in the form of blanks, are used in most biological evaluations where test article extracts are prepared. The use of these blanks provides the basis for a comparison of the effects of the test material extract with a validated negative test result. Japanese (MHW 2000) guidelines consistently refer to these as Standard Reference Materials (SRMs).

A number of materials have been used extensively in biological testing as negative or positive controls. High-density polyethylene, obtained from the *United States Pharmacopeia*, is a standard negative control. The nonreactive plastic can be implanted into living tissue and the results compared with those for a test material that has been similarly implanted. Likewise, a polyvinyl chloride formulation containing organotin additives serves well as a positive control.

It is preferable to evaluate medical devices in their final product form. The reasoning is simple—the biological testing must incorporate everything involved in making the device. Obviously, the constituent materials must be safe for patient contact; equally important to device biocompatibility are the processes and materials used during manufacturing. For most devices, the use of fluid extracts of the test materials prepared in a fashion to mimic or exaggerate the expected clinical conditions is the most appropriate technique for determining the potential effects of chemical leachables.

3.3 Cytotoxicity Testing (ISO 10993-5)

The cell culture including cytotoxicity methods is a long-established screening method for device or biomaterial biocompatibility screening and may be a fair predictor of biocompatibility when used together with other appropriate tests (Wilsnack 1976; Gad 2000). Several highly specialized cell culture methods are available to evaluate the biocompatibility of the raw materials used in the manufacturing of the device or auditing the production of the manufacturing process. Such cytotoxicity offers several advantages:

- It is simple, rather inexpensive, and easy to perform.
- It allows testing of a biomaterial on human tissue.
- It is sensitive to toxic material.
- It is easy to manipulate and allow more than one end point investigation.
- It can be used to construct a concentration-response curve.
- It can give quick and quantitative results and allows direct access or direct observation or measurement.

Despite these advantages, cytotoxicity use is currently limited to screening the biomaterials and finished medical devices. It is one of the three tests required for all medical devices, irregardless of route or duration of patient exposure. But it is essential that cytotoxicity results should be used in conjunction with other tests to assess.

The objective of cell cytotoxicity testing is to screen the biocompatibility of the polymer and elastomer portions of medical devices or medical device components using mammalian cell cultures. Cytotoxicity is a useful method for screening material and frequently serves as a quality control mechanism for audit or batch testing programs (11). It is one of the oldest assays designed specifically to screen plastics for toxicity (Rosenbluth et al. 1965). Given the extreme sensitivity of this test, materials

found to be cytotoxic must be assessed along with the results of in vivo studies and others to evaluate the relevant risk to human health. Unlike the other studies utilized in biocompatibility testing, cytotoxicity is not a clear pass/fail test in the eyes of regulators. Failure in cytotoxicity is generally grounds for performing confirmatory tests such as an implantation or intracutaneous reactivity and acute systemic toxicity (Barile 1994). There are clear differences in the sensitivity of different tissues (eye > muscle > skin) to cytotoxicity effects (ISO Japan 2009).

3.3.1 Background

Historically, cytotoxicity was originally adapted in the early 1950s for the use of screening plastics used in pharmaceutical containers. The concern was for potential toxic entities leaking from the containers into the drugs at a time before our current analytical chemistry methods were available to serve the purpose. As such the plastics were called "closures."

The great majority of toxic compounds are chemically stable and produce their characteristic effects by interference with biochemical or physiological homeostatic mechanisms (Gad 2000, Di Silvio 2010). This means that an understanding of the pharmacodynamics of extractable toxics is essential. In the case of drugs, it has been estimated that some 80 percent of the adverse reactions are the result of exaggerated pharmacological responses rather than "off target hits." Many adverse events are the consequence of the disturbance of normal physiology and do not result in cell death. This is one reason that cytotoxicity assays on their own cannot provide a full assurance of safety. At the same time, the intact organism has available extensive defensive mechanisms not available to the cell in cultures.

Cytotoxicity, the causing of cell death, is often the consequence of exposure to a harmful chemical, but the number of cells which must be killed before the function of a tissue or organism is noticeably impaired is highly variable. Some cell types, notably the epithelia including the liver, have the ability to regenerate in response to insult, while others, most notably neurons, have little such capability. Some organs, such as the liver, lung, and kidney, have a substantial reserve capacity in excess of normal requirements, and normal body function can be maintained in the presence of marked organ impairment.

Cytotoxicity assays measure loss of some cellular or intercellular structure and/or functions, including cell death. They are generally simple to perform and reproducible and have a clearly defined end point. However, specificity in the prediction of end points in standard tests in vivo (such as lethality, irritation, and implantation) is extremely variable for a variety of reasons, including the fact that the assay systems are continually exposed, whereas in vivo there are biological and biochemical protective measures in operation which limit the duration of exposure. Some assays may not be universally capable of detecting all chemical classes of irritants because of the transitory nature and threshold of in vivo expression from the end point used.

A variety of cell lines have been used including corneal epithelial cells, lung fibroblasts, Chinese hamster ovary cells, canine renal cells, HeLa cells, and microorganisms. The Japanese recommend the V79 cell line.

Differentiated cells are used to evaluate the effects materials may have on specific tissues. Differentiated cells are generally non-fibroblastic cells which are different from transformed and fibroblastic cells such as L-cells used in ISO cytotoxicity test methods. Differentiated cells have organ-specific or tissue-specific functions and have specific biological end points or measurable characteristics. Liver cells, which are differentiated cells, have all or some liver functions.

To perform the test in differentiated cells is important for at least two reasons. First, the tissue-type-specific features of differentiated cells may modulate the effects of chemicals on the fundamental properties of cells. Second, it is important to determine the effects of chemicals on specific cell functions or responses. Culture systems for growing epithelial, liver, or embryonic cells have been developed only recently. The number of available differentiated cells for biocompatibility testing is currently small, but there is significant development in this area.

3.3.2 Neutral Red Uptake (NRU) Assay

In this procedure, cells, usually mouse fibroblasts or Chinese hamster ovary cells, are exposed to the test material and then to neutral red. Retention of neutral red indicates cell viability. Bagley et al. (2008) found that, in general, the concentration of test material required to reduce neutral red uptake decreased as the in vivo irritant potential of the test material increased. Neutral red is a vital dye, which is preferentially absorbed and endocytosed by viable cells and internalized inside lysosomes. In this respect, it is considered as an indicator of lysosome and cell integrity. The basis of the test is that a cytotoxic material, regardless of site or mechanism of action, can interfere with this process and result in a reduction of the number of viable cells. NRU has been shown to correlate well with the number of cells in the culture and has also been found to be more accurate than MTT assay (Ciapetti et al. 1996). Attention needs to be paid to the technical aspects of incubation as emphasized by Blein et al. (1991) who found that correlation with materials with an extreme pH was underestimated because of the buffering effect of the culture medium and that volatile materials were also underestimated probably because of loss of material.

3.3.3 MTT

The MTT assay introduced in 1983 is a colorimetric assay that is used to determine the cell proliferation, viability, and cytotoxicity. It is advantageous, as it can be rapidly performed on a microtiter plate assay and read on a spectrophotometer ELISA

plate reader at the absorbance of 570 nm. Traditionally, the reduction of the tetrazolium salts to their equivalent formazan precipitates was used for the histochemical demonstration of the activities of oxidative and non-oxidative enzymes in mitochondrial using both light and electron microscopy. Currently, in the MTT assay, the tetrazolium MTT (3-[4,5-dimethylthiazol-2-yl]-2,5-diphenyltetrazolium bromide) is reduced in a mitochondrial-dependent reaction to an insoluble purple formazan by cleavage of the tetrazolium ring by succinate dehydrogenase within the cell since it cannot pass through the cell membrane. By addition of spectrophotometric grade dimethylsulfoxide (DMSO), isopropanol, or other suitable solvents, the formazan is solubilized and liberated and is readily quantified colorimetrically. Cytotoxic concentration is generally determined as the concentration that kills 50% of the cells, generally known as the IC_{50}.

Several inter-test comparisons have been undertaken. Sina et al. (1992) and Ciapetti et al. (1996) compared leucine incorporation, MTT dye reduction, and neutral red uptake in corneal epithelial cells and Chinese hamster lung fibroblasts. None of the end point target cell combinations accurately predicted in vivo eye irritation in this series, but the MTT dye reduction method gave the best overall correlation.

In vitro assays such as cytotoxicity serve to assess acute toxic effects, and in far larger organisms are most effective in predicting local tissue irritation effects. Such experiments using cultured cells utilize a variety of methods based on either a fresh isolate from fragments of tissue or cell suspension (primary cell culture) which grow to confluence and then age and die or single cell clones (continuous cell culture) which have an indefinite capacity to grow and replicate. The continuous cell lines have the advantage of being consistent, reliable, and reproducible. They act as a standard with a documented history; they have fewer biological variables and may be tuned to particular toxicity concerns by using a variety of tissues and species with a range of doses and exposure periods. As a result, these methods can be very efficient in screening and are often more sensitive than acute toxicity tests in animals. Early cell culture methods were markedly subjective screening to merely estimate the numbers of living or dead cells, but now morphological analysis by electron microscopy reveals a spectrum of microcellular changes; and cell function tests measure biochemical parameters indicating the nature of cell stress.

Although many modifications have been made, cell culture tests are of four main types: gel diffusion, direct contact, extract dilution, and cell function tests.

Gel diffusion uses agar or agarose to cover a cell monolayer. A sample of the material or extract is placed on top of the gel providing a concentration gradient of diffusibles. Agarose allows a faster diffusion of uncharged molecules and is as sensitive as the rabbit intramuscular implantation test.

Direct contact of the test material onto a culture layer is more sensitive than the rabbit intramuscular implantation test, but care must be taken to avoid physical damage to the cells by pressure or movement of the sample.

Extracts may be serially diluted in the nutrient media and provide a quantitative comparison with reference extracts. Inevitably the correlation with animal tests will depend on the nature of the eluants.

Cell function tests are a very precise way of registering cellular response to any insult. In particular, inhibition of cell growth can be measured with considerable sensitivity.

With increasing complexity of test methodology, the results may be less reproducible than previously, and increasing sensitivity may not assist the accurate prediction of risk to humans as the impact of a material on the body systems may be much less intense than to cells in a multi-well test system.

Several tests are available to test cytotoxicity by direct contact. These include:

- ASTM F813 Practice for Direct Contact Cell Culture Evaluation of Materials for Medical Devices.
- ASTM F895 Test Method for Agar Diffusion Cell Culture Screening for Cytotoxicity.
- ASTM F1027 Standard Practice for Assessment of Tissue and Cell Compatibility of Orificial Prosthetic Materials and Devices.
- NIH Publication No. 85-2185 Guidelines for Blood-Material Interactions.
- HIMA Report: Guidelines for the Preclinical Safety Evaluation of Materials Used in Medical Devices.
- Others including many device-specific toxicity guidance documents on toxicity testing.

In addition, the agar overlay tissue culture method and fluid medium tissue culture method can be used for direct contact cytotoxicity testing. In the fluid medium method, the test material or device is placed directly on the growing monolayer cell surface. In the agar overlay method, the solid test sample is place on or in the agar layer containing the vital stain such as neutral red over the growing monolayer of cells. The response is evaluated grossly and microscopically and graded according to the zone index, the size of the cytopathic area, the lysis index, and percent of cell lysis.

Proper cytotoxicity testing should include at least one test with extract and one direct contact test if feasible.

In addition, differentiated cells are used to evaluate the effects materials may have on specific tissues. Differentiated cells are generally non-fibroblastic cells which are different from transformed and fibroblastic cell lines such as L929 cells used in ISO cytotoxicity test methods. Differentiated cells have organ-specific or tissue-specific functions and have specific biological end points or measurable characteristics. Liver cells, which are differentiated cells, have all or some liver functions.

The three specific tests prescribed by ISO (and USP) are presented below.

3.3.4 Agar Diffusion Test

This test is designed for elastomeric closures in a variety of shapes. The agar layer acts as a cushion to protect the cells from mechanical damage while allowing the diffusion of leachable chemicals from the polymeric specimens. Extracts of materials that are to be tested are applied to a piece of filter paper:

Sample Preparation—Use extracts, prepared as directed, or use portions of the test specimens having flat surfaces not less than 100 mm^2 in surface area.

Procedure—Prepare the monolayers in 60-mm diameter plates using 7 mL of *cell culture preparation*. Aspirate the culture medium from the monolayers, and replace it with serum-supplemented culture medium containing not more than 2% of agar. Place the flat surfaces of *Sample Preparation*, USP negative control plastic RS (to provide a *Negative Control*), and either USP positive bioreaction extract RS or USP positive bioreaction solid RS (to provide a *Positive Control*) in duplicate cultures in contact with the solidified agar surface. Incubate all cultures for not less than 24 hours at 37 ± 1°, preferably in a humidified incubator containing 5 ± 1% of carbon dioxide. Note that in contrast to the now preferred extraction at 50 °C for 72 hours, culture medium with serum is limited to 37 °C. Each culture must be examined under a microscope, using cytochemical stains, if possible.

Interpretation of Results—The biological reactivity (cellular degeneration and malformation) is described and rated on a scale of 0 to 4 (Table 3.4). Measure the responses obtained from the *Negative Control* and the *Positive Control*. The test system is suitable of the observed response corresponds to the labeled biological reactivity grad of the relevant reference standard. Measure the response obtained from the *Sample Preparation*. The *Sample* meets the requirements of the test if none of the cell culture exposed in the *Sample* show greater than a mild reactivity (Grade 2). Repeat the test if the suitability of the system is not confirmed.

3.3.5 Direct Contact Test

This test is designed for materials in a variety of shapes. The procedure allows for simultaneous extraction and testing of leachable chemicals from the specimen with a serum-supplemented medium. The procedure is not appropriate for very low- or high-density materials that could cause mechanical damage to the cells.

3.3.5.1 Sample Preparation

Use portions of the test specimen having flat surfaces not less than 100 mm^2 in surface area.

Table 3.4 Reactivity grades for direct contact and agar diffusion test

Grade	Reactivity	Description of reactivity zone
0	None	No detectable zone around or under specimen
1	Slight	Zone limited to area under specimen
2	Mild	Zone extends less than 0.5 cm beyond specimen
3	Moderate	Zone extends 0.5–1.0 cm beyond the specimen
4	Severe	Zone extends greater than 1.0 cm beyond specimen but does not involve entire dish

3.3.5.2 Procedure

Prepare the monolayers in 35-mm-diameter plates using 2 mL of cell suspension. Aspirate the culture medium from the cultures, and replace it with 0.8 mL of fresh culture medium. Place a single *Sample Preparation*, USP negative control plastic RS (to provide a *Negative Control*), and USP positive bioreaction solid RS (to provide a *Positive Control*) in each of duplicate cultures. Incubate all cultures for not less than 24 hours at 37 ± 1° in a humidified incubator preferable containing 5 ± 1% of carbon dioxide. Examine each culture around each *Sample*, *Negative Control*, and *Positive Control*, under a microscope, using cytochemical stains if desired.

3.3.5.3 Interpretation of Results

Proceed as directed for *Interpretation of Results* under *Agar Diffusion Test* using Table 3.5. The *Sample* meets the requirements of the test if none of the cultures treated with the *Sample* shows greater than a mild reactivity (Grade 2). Repeat the test if the suitability of the system is not confirmed.

3.3.6 Elution Test

This test is designed for the evaluation of extracts of polymeric materials. The procedure allows for extraction of the specimens at physiological or nonphysiological temperatures for varying time intervals. It is appropriate for high-density materials and for dose-response evaluations.

3.3.6.1 Sample Preparation

Prepare as directed in *Preparation of Extracts*, using ether sodium chloride injection (0.9% NcCl) or serum-free mammalian cell culture media as *Extraction Solvents*. If the size of the *Sample* cannot be readily measured, a mass of not less

Table 3.5 Reactivity grades for elution test

Grade	Reactivity	Conditions of all cultures
0	None	Discrete intracytoplasmic granules; no cell lysis
1	Slight	More than 20% of the cells are round, loosely attached, and without intracytoplasmic granules; occasional lysed cells are present
2	Mild	More than 50% of the cells are round and devoid of intracytoplasmic granules; extensive cell lysis and empty areas between cells
3	Moderate	Greater than 70% of the cell layers contain rounded cells and/or are lysed
4	Severe	Nearly complete destruction of the cell layers

than 0.1 g of elastomeric material or 0.2 g of plastic or polymeric material per mL of extraction medium may be used. Alternatively, use of serum-supplemented mammalian cell culture media as the extracting medium acts to simulate more closely physiological conditions. Prepare the extracts by heating for 24 hours in an incubator preferably containing 5% ± 1% of carbon dioxide. Maintain the extraction temperature at 37° ± 1°, because higher temperatures may cause denaturation of serum proteins.

3.3.6.2 Procedure

Prepare the monolayers in 35-mm-diameter plates using 2 mL of the above cell culture preparation. Aspirate the culture medium from the monolayers, and replace it with either extracts of the sample, USP negative control plastic RS (to provide a Negative Control), or USP Positive Bioreaction Extract RS (to provide a Positive Control). The serum-supplemented and serum-free cell culture media extracts are tested in duplicate without dilution (100%). The sodium chloride injection extract is diluted with serum-supplemented cell culture medium and tested in duplicate at 25% extract concentration. Incubate all cultures for 48 hours at 37 ± 1° in an incubator preferably containing 5 ± 1% of carbon dioxide. Examine each culture at 48 hours, under a microscope, using cytochemical stains, if desired.

3.3.6.3 Interpretation of Results

Proceed as directed for *Interpretation of Results* under *Agar Diffusion Test* but using Table 3.5. Repeat the test if the suitability of the system is not confirmed. The Sample meets the requirements of the test if the cultures treated with the Samples show not greater than a mild reactivity (Grade 2). If the cultures treated with the *Sample* show a significantly greater reaction than the cultures treated with the negative control, repeat the test with several quantitative dilutions of the extracts.

For each of these three procedures, it should be kept in mind that while USP requires that tests be performed in duplicate, ISO requires that they be done in triplicate.

3.3.7 Correlation with In Vivo Results

Cytotoxicity testing for medical devices is a very useful screening tool, but it must be kept firmly in mind that the correlation of results from these assays with intact animal tests (and with observed effects in humans) is poor at best. This issue was researched more than 20 years ago (Wilsnack et al. 1973; Wilsnack 1976) with

side-by-side comparisons of cytotoxicity results with those of animal tests conducted on the same samples. The results demonstrated limited correlation and "the ellipsoid effect," the best correlation being between results at the extremes of concentration and response (Gad 2000).

The author has also tried to correlate results of cytotoxicity and concurrent animal tests (particularly subcutaneous injection and implantation tests, where one would expect the best case) only to find that there were high levels of false negatives and positives, though predominantly the latter. Therefore, investigators are cautioned to not place too much faith and weight on the results of these assays. As the ISO 10993-5 guidance states, positive results in cytotoxicity studies must be considered in light of concurrent in vivo studies (irritation, implantation, and acute systemic toxicity).

3.3.8 Conclusion

Cytotoxicity tests, as an initial screen for toxicities of both plastics and elastomers and of leachates from them, have been in use since the 1960s (Rosenbluth et al. 1965) but have also been recognized for a long time to be limited to effectively serving only as screens that in effect say, "look at this and evaluate/consider further." This limitation to use as screens as opposed to as definitive tests is due to the, at best, moderate correlation of their results with in vivo findings (Wilsnack 1976). Regulatory agencies recognize the limitations of these test systems, and users should bear in mind the categorical (as opposed to truly quantitative) nature of the scoring systems.

3.4 Sensitization (ISO 10993-10)

The evaluation of the immunotoxicity of medical devices as part of their biocompatibility assessment has traditionally, and still the case for most skin only contact devices, been limited to delayed contact dermal sensitization. Later in this chapter, the more recent 10,993–20 evaluation will be presented.

The specific, or adaptive, immune system is characterized by memory, specificity, and the ability to distinguish "self" from "nonself" (Battisto et al. 1983), though an alternative to this basic self-nonself paradigm has been proposed that rather the immune system actually responds to some form of "danger" manager (Pennisi 1996). This is the portion of the immune system involved in delayed contact hypersensitivity.

The important cells of the adaptive immune system are the lymphocytes and antigen-presenting cells that are part of nonspecific immunity. The lymphocytes, which originate from pluripotent stem cells located in the hematopoietic tissues of the liver

(fetal) and bone marrow, are composed of two general cell types: T and B cells. The T cells differentiate in the thymus and are made up of three subsets: helper, suppressor, and cytotoxic. The B cells, which have the capacity to produce antibodies, differentiate in the bone marrow or fetal liver. The various functions of the T cells include presenting antigen to B cells, helping B cells to make antibody, killing infected cells, regulating the level of the immune response, and stimulating cytotoxic activity of other cells such as macrophages (Male et al. 1982).

3.4.1 Hypersensitivity

The four types of hypersensitivity reactions ("sensitization") as classified by Coombs and Gell (1975) are outlined in Table 3.6. The first three types are immediate antibody-mediated reactions, whereas the fourth type is a cellular-mediated delayed-type response that may require 1–2 days to occur after a secondary exposure. Type I reactions are characterized by an anaphylaxis response to a variety of compounds, including proteinaceous materials and pharmaceuticals such as penicillin. Various target organs may be involved depending on the route of exposure. For example, the gastrointestinal tract is usually involved with food allergies, the respiratory system with inhaled allergens, the skin with dermal exposure, and smooth muscle vasculature with systemic exposure. The type of response elicited often depends on the site of exposure and includes dermatitis and urticaria (dermal), rhinitis and asthma (inhalation), increased gastrointestinal emptying (ingestion), and systemic anaphylactic shock (parenteral).

Table 3.6 Types of hypersensitivity responses

Type and designation	Components	Effects	Mechanism
I, immediate	Mast cells; IgE	Anaphylaxis, asthma, urticar uriticaria, rhinitis, dermatdermatitis	IgE binds to mast cells to stimulate release of humoral factors
II, cytoxic	Antibodies	Hemolytic anemia, Goodpasture's disease	IgG and IgM bind to cells (e.g., RBCs), fix complement (opsonization), then lyse cells
III, immune complex (Arthus)	Antigen-antibody complexes (Ag-Ab)	SLE rheumatoid arthritis, glomerular nephritis, serum sickness, vasculitis	Ag-Ab complexes deposit in tissues and may fix complement
IV, delayed hypersensitivity	T_D cells; macrophages	Contact dermatitis, tuberculosis	Sensitized T cells induce a delayed hypersensitivity response upon challenge

Source: Based on classification system of Gell and Coombs (1975)

3.4.1.1 Type I Hypersensitivity

During an initial exposure, IgE antibodies are produced and bind to the cell surface of mast cells and basophils. Upon subsequent exposures to the antigen, reaginic IgE antibodies bound to the surface of target cells at the F_c region (mast cells and basophils) become cross-linked (at the F_{ab} regions) by the antigen.

3.4.1.2 Type IV Delayed-Type Hypersensitivity (DTH)

Delayed-type hypersensitivity reactions are T-cell mediated with no involvement of antibodies. However, these reactions are controlled through accessory cells, suppressor T cells, and monokine-secreting macrophages, which regulate the proliferation and differentiation of T cells. The most frequent form of DTH manifests itself as contact dermatitis. The drug or metabolite binds to a protein in the skin or the Langerhans cell membrane (class II MHC molecules) where it is recognized as an antigen and triggers cell proliferation. After a sufficient period of time for migration of the antigen and clonal expansion (latency period), a subsequent exposure will elicit a dermatitis reaction. A 24–48 h delay often occurs between the time of exposure and onset of symptoms to allow time for infiltration of lymphocytes to the site of exposure. The T cells ($CD4^+$) that react with the antigen are activated and release lymphokines that are chemotactic for monocytes and macrophages. Although these cells infiltrate to the site via the circulatory vessels, an intact lymphatic drainage system from the site is necessary since the reaction is initiated in drainage lymph nodes proximal to the site (Clark 1983). The release (degranulation) of enzymes and histamines from the macrophages may then result in tissue damage. Clinical symptoms of local dermal reactions may include a rash (not limited to sites of exposure), itching, and/or burning sensations. Erythema is generally observed in the area around the site, which may become thickened and hard to the touch. In severe cases, necrosis may appear in the center of the site followed by desquamation during the healing process. The immune-enhancing drugs isoprinosine and avridine have been shown to induce a delayed-type hypersensitivity reaction in rats (Exon et al. 1984).

A second form of delayed-type hypersensitivity response is similar to that of contact dermatitis in that macrophages are the primary effector cells responsible for stimulating CD4+ T cells; however, this response is not necessarily localized to the epidermis. A classical example of this type of response is demonstrated by the tuberculin diagnostic tests. To determine if an individual has been exposed to tuberculosis, a small amount of fluid from tubercle bacilli cultures is injected subcutaneously. The development of induration after 48 h at the site of injection is diagnostic of prior exposure.

Shock, similar to that of anaphylaxis, may occur as a third form of a delayed systemic hypersensitivity response. However, unlike anaphylaxis, IgE antibodies are not involved. This type of response may occur 5–8 h after systemic exposure

and can result in fatality within 24 h following intravenous or intraperitoneal injection.

A fourth form of delayed hypersensitivity results in the formation of granulomas. If the antigen is allowed to persist unchecked, macrophages and fibroblasts are recruited to the site to proliferate, produce collagen, and effectively "wall off" the antigen. A granuloma requires a minimum of 1–2 weeks to form.

In the ISO 10993-20 immunotoxicity assessment scheme, the delayed-type hypersensitivity (DTH) response assay is considered to be a comprehensive Tier II assay for cell-mediated immunity by the NTP.

To express a DTH inflammatory response, the immune system must be capable of recognizing and processing antigen, blastogenesis and proliferation of T cells, migration of memory T cells to the challenge site of exposure to antigen, and subsequent production of inflammatory mediators and lymphokines that elicit the inflammatory response. Thus, by measuring a DTH response to an antigen, these assays assess the functional status of both the afferent (antigen recognition and processing) and efferent (lymphokine production) arms of cellular immunity. Various antigens have been used for assessing DTH, including keyhole limpet hemocyanin (KLH), oxazolone, dinitrochlorobenzene, and sheep red blood cells (SRBCs) (Vos 1977; Godfrey and Gell 1978; Luster et al. 1988).

There are several well-established preclinical models for assessing Type IV (delayed-type) hypersensitivity reactions following dermal exposure, but not for predicting this response after systemic exposure. The dermal exposure mode is the only currently required and widely performed immunotoxicity assay on devices.

Type IV hypersensitivity responses are elicited by T lymphocytes and are controlled by accessory cells and suppressor T cells. Macrophages are also involved in that they secrete several monokines, which results in proliferation and differentiation of T cells. Thus, there are numerous points along this intricate pathway in which drugs may modulate the final response. To achieve a Type IV response, an initial high-dose exposure or repeated lower-dose exposures are applied to the skin; the antigen is carried from the skin by Langerhans cells and presented to cells in the thymus to initiate T-cell proliferation and sensitization. Once sensitized, a second "challenge" dose will elicit an inflammatory response. Thus, before sensitivity can be assessed, each of the models used to evaluate dermal hypersensitivity requires as a minimum:

- An initial induction exposure
- A latency period for expression
- A challenge exposure

A preliminary test for acute irritancy is also required to ensure that the initial dose is sufficient to stimulate sensitization and that the challenge dose is sufficient to ensure expression of the response without producing irritation, which would confound the response. To confirm suspected sensitization or determine a threshold dose, each assay may also include a second challenge dose 1–2 weeks after the first challenge, at the same or lower concentrations. To increase penetration of the test article, various methods of abrasion (e.g., tape stripping) and occlusive coverings

may also be used. Assessing materials to determine if they can act as delayed contact dermal sensitizers in humans is different on a number of grounds from the other tests we have looked at so far and, indeed, from most of the other test systems presented later in this book. These differences all stem from how the immune system, which is the mechanistic basis for this set of adverse responses, functions.

Bringing about this Coombs Type IV hypersensitivity response (which is commonly called "sensitization," for short) requires more than a single exposure to the causative material, both in humans and in test animals. Unlike irritation responses, sensitization occurs in individuals in an extremely variable manner. A portion of the human population is considerably more liable to be sensitized, while others are infrequently affected. And the response, once sensitization is achieved, becomes progressively more severe with each additional exposure. All three of these characteristics are due to the underlying mechanism for the response and influenced the manner in which we conduct tests. These factors mean that in vivo test systems require multiple exposures of animals and tend to underpredict the potential for an adverse response in those individuals who are most susceptible to sensitization. But because the response to repeated exposures of even minimal amounts of material in these susceptible individuals can lead to such striking adverse responses, we must be concerned about them.

A number of factors influence the potential for a chemical to be a sensitizer in humans and, in turn, also influence the performance of test systems. These are summarized in Table 3.7. Various test systems manipulate these in different ways.

There are a number of references which explore and discuss the underlying immune system mechanisms and operation in greater detail. Particularly recommended is Gibson et al. (1983).

3.4.2 Objectives and General Features

Given the considerations of mechanism, degree of concern about protecting people, and practicality, the desired characteristics of a sensitization test include the following:

Table 3.7 Factors influencing delayed type sensitization responses

1. Percutaneous absorption of agent
2. Genetic status of host
3. Immunological status of host
4. Host nutrition
5. Chemical and physical nature of potential sensitizing agent
6. Number, frequency, and degree of exposures of immune system to potential antigen
7. Concurrent immunological stimuli (such as adjuvants, inoculations, and infections). System can be "up-modulated" by mild stimuli or overburdened by excessive stimulation
8. Age, sex, and pregnancy (by influencing factors 1, 3, and 4 above)

Figure 3.2. There are other versions which also comply

1. Be reproducible.
2. Involve fairly low technical skills so that it may be performed as a general laboratory test.
3. Do not involve the use of exotic animals or equipment.
4. Use relatively small amounts of test material.
5. Be capable of evaluating almost any material of interest.
6. Be sensitive enough to detect weak sensitizers (i.e., those which would require extensive exposure to sensitize other than the most sensitive individuals).
7. Predict the relative potency of sensitizing agents accurately.

Several of these desired characteristics are mutually contradictory; as with most other test systems, the methods for detecting dermal sensitization each incorporate a set of compromises.

All the in vivo tests have some common features, however. The most striking is that they involve at least three (and frequently four) different phases—they are multiphasic. These phases are, in order, the irritation/toxicity screen, the induction phase, the challenge phase, and (often) the rechallenge phase. Irritation/toxicity screen: All assays require knowledge of the dermal irritancy and systemic toxicity of the test material(s) to be used in the induction, challenge, and rechallenge. These properties are defined in this pretest phase. Most tests desire (or will allow) mild irritation in the induction phase. Most tests desire (or will allow) mild irritation in the induction phase, but no systemic toxicity. Generally, a nonirritating concentration is required for the challenge and for any rechallenge, as having irritation present either confounds the results or precludes having a valid test. As will be discussed in the sections on the individual tests, even a carefully designed screen does not necessarily provide the desired guidance in selecting usable concentrations. During this phase, solvent systems are also selected.

3.4.2.1 Induction Phase

This requires exposing the test animals to the test material several times over a period of days or weeks. A number of events must be accomplished during this phase if a sensitization response is to be elicited. The test material must penetrate through the epidermis and into the dermis. There, it must interact with dermis protein. The protein-test material complex must be perceived by the immune system as an allergen. Finally, the production of sensitized T cells must be accomplished. Some assays enhance the sensitivity of the induction phase by compromising the natural ability of the epidermis to act as a barrier. These enhancement techniques include irritation of the induction site, intradermal injection, tape stripping, and occlusive dressings. In contrast, events such as the development of a scab over the induction site may reduce percutaneous absorption. The attention of the immune system can be drawn to the induction site by the intradermal injection of oil-coated bacteria (Freund's complete adjuvant, which serves as a mild immunological stimulant).

3.4.2.2 Challenge Phase

This consists of exposing the animals to a concentration of the test material which would normally not be expected to cause a response (usually an erythema type response). The responses in the test animals and of the control animals are then scored or measured.

3.4.2.3 Rechallenge Phase

This is a repeat of the challenge phase and can be a very valuable tool if used properly. Sensitized animals can be rechallenged with the same test material at the same concentration used in the challenge in order to assist in confirming sensitization. Sensitized animals can be rechallenged with different concentrations of the allergen to evaluate dose-response relationships. Animals sensitized to an ingredient to evaluate can be challenged to a formulation containing the ingredient to evaluate the potential of the formulated product to elicit a sensitization response under adverse conditions. Conversely, animals which responded (sometimes unexpectedly) to a final formulation can be challenged with formulation without the suspected sensitizer or to the ingredient which is suspected to be the allergen. A well-designed rechallenge is important and should be considered at the same time that the sensitization evaluation is being designed since the rechallenge must be run within 1–2 weeks after the primary challenge. Unless plans have been made for a possible rechallenge, one may have to reformulate a test material or obtain additional pure ingredient and perhaps run additional irritation/toxicity screens before the rechallenge can be run. The ability of the sensitized animals to respond at a rechallenge being run shortly after the challenge, serves to confirm that a sensitization response has been enabled. In addition, some assays use sham-treated controls, and these must be procured, while the induction phase is in progress. One additional piece of information must be kept in mind when evaluating a rechallenge. The animal does not differentiate between an induction exposure and a challenge exposure. If one is using an assay which involves three induction exposures and one challenge exposure, then at the rechallenge, the animal has received four induction exposures. This “extra” induction may serve to strengthen a sensitization response.

After the study is done, one must evaluate the data and decide how to translate it to human relevance. We will look at this problem toward the end of this chapter.

The basis of modern predictive tests is the Draize tests, as established by Landsteiner and Draize et al. in 1944. It consists of ten intradermal injections of the test compound into the skin of albino guinea pigs during the 3-week induction period and a single intracutaneous challenge application 14 days after the last induction injection. A standardized 0.1% test concentration is used for induction and challenge. This method was widely used and recommended until the end of the 1960s. Its disadvantage is that only strong allergens are detected, while well-known moderate allergens fail to sensitize the animals at all.

Starting in 1964, however, a wide variety of new test designs started to be proposed. Buehler (1964 and 1965) proposed what is now considered the first modern test (described in detail in this chapter), which used an occlusive patch to increase test sensitivity. The Buehler test is the primary example of the so-called "epidermal" methods, which have been criticized for giving false-negative results for moderate to weak sensitizers such as nickel.

A new generation of tests was established by using Freund complete adjuvant (FCA) during the induction process to stimulate the immune system, independent of the type of hapten and independent of the method or application, that is, whether or not the substance is incorporated in the adjuvant mixture. It is claimed that this family of tests display the same level of susceptibility to sensitization in guinea pigs as is normally observed in humans (Cronin and Agrup 1970). The adjuvant tests include the guinea pig maximization test (Maurer et al. 1975, 1980), split adjuvant test (Maguire and Chase 1967), and the epicutaneous maximization test (EMP, Guillot and Gonnet 1985).

3.4.3 Modified Buehler Procedure

This is a closed patch procedure for evaluating test substances for potential delayed contact dermal sensitization in guinea pigs. The procedure, based on that described by Buehler (1965), is practical for test substances that cannot be evaluated by the traditional intradermal injection procedure of Landsteiner and Jacobs (1935) or by the GPMT for skin sensitization testing. The closed patch procedure is performed when a test substance either is highly irritating to the skin by the intradermal injection route of exposure or cannot be dissolved or suspended in a form allowing injection. It is also the method of choice for some companies. This procedure, which is one version of the Buehler test, complies with the test standards set forth in the Toxic Substances Control Act (TSCA 1979) and other regulatory test rules and is presented diagrammatically in Fig. 3.1. There are other versions which also comply.

3.4.3.1 Animals

- Young albino female guinea pigs, weighing between 300 and 400 g are used.
- Although several proposed test rules suggest the use of male guinea pigs, the female sex is preferred because the aggressive social behavior of males may result in considerable skin damage that might interfere with the interpretation of challenge reactions. This concern occurs because animals are group housed (Marzulli and Maibach 1996).
- Animals that show poor growth or are ill in any way are not used, since illness markedly decreases the response. Animals with skin marked or scarred from fighting are avoided. The guinea pigs are observed for at least 2 weeks to detect any illness before starting a study.

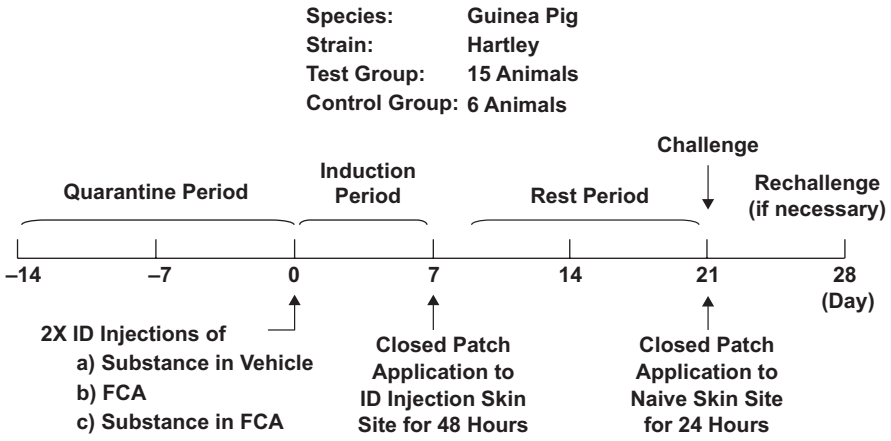


Fig. 3.1 Illustrative figures for injection and patching of animals in GPMT

- The guinea pigs are identified by a cage card and marking pen or any other suitable method. There is no regulatory requirements, however, for the identification of individual animals.
- The guinea pigs are randomly assigned to test and negative control group consisting of at least 15 and at least 6 animals each, respectively. If a pretest group is necessary, as many animals as needed for that group are randomized also.

3.4.3.2 Pretest Screen

- If practical, the dermal irritation threshold concentration should be established for the test substance prior to the first induction application. A concentration of the test substance that produces minimal or no irritation (erythema and/or edema formation) is determined. The highest concentration that produces no irritation is preferred for the dermal sensitization study challenge dose.
- Those animals randomly assigned to the pretest group are used.
- Each animal is prepared by clipping a 1-inch square area of hair from the left upper flank using a small animal clipper with a size no. 40 blade.
- The test substance is diluted, emulsified, or suspended in a suitable vehicle. Vehicles are selected on the basis of their solubilizing capacity for the test substance and on their ability to penetrate the skin.
- Different concentrations of the test substance are tested on the pretest group of guinea pigs; a few animals are used for each concentration tested.
- A volume of 0.l5 ml is applied to a patch consisting of a cotton pad (2.5 × 2.5 cm) occluded with impermeable surgical tape or placed in a Hilltop-style occlusive "chamber."
- The patch is applied to the shaved left flank of a guinea pig. The patch is held firmly in place for 24 hours by wrapping the trunk of the animal with a 3-inch-

wide elastic bandage. A 2-inch-wide strip of tape is used to line the center adhesive side of the bandage in order to prevent skin damage from the adhesive.
- After 24 hours of exposure, the wrappings and patches are removed.
- Observations of skin reactions (erythema and/or edema formation) are recorded 48 hours after application.
- A judgment is made as to which concentration will be used for the dermal sensitization study, based on the dermal irritation data which has been collected. The highest concentration that produces minimal or no dermal irritation is selected.

3.4.3.3 Induction Phase

- Test group and control group guinea pigs are weighed at the beginning of the study and weekly thereafter.
- Test control group guinea pigs are clipped as described earlier in this procedure.
- If the test substance is a liquid solution, suspension, or emulsion, a volume of l5 ml of the highest concentration found to be nonirritating in a suitable vehicle (as determined in the pretest portion of this procedure) is applied to a patch consisting of a cotton pad (1″ × 1″) occluded with impermeable surgical tape. If the test substance is a solid or semisolid, 0.5 g[1] is applied. If the test substance is a fabric, a 1-inch square is moistened with 0.5 ml of physiological saline before application.
- The first induction patch is applied to the clipped left flank of each test group guinea pig. The patch is held firmly in place for 24 hours by wrapping the trunk of each animal with a 3-inch-wide elastic bandage. A 2-inch-wide strip of tape is used to line the center adhesive side of the bandage in order to prevent skin damage from the adhesive. A 2-inch length of athletic adhesive tape is placed over the bandage wrap as a precautionary measure to prevent unraveling.
- After 24 hours of exposure, the wrappings and patches are removed and disposed of in a plastic bag.
- Each dermal reaction, if any, is scored on the basis of the designated values for erythema and edema formation presented in Table 3.8. Observations are made 48 hours after initiation of the first induction application. Resulting dermal irritation scores are recorded.
- After the initial induction application, subsequent induction applications (2–10) are made on alternate days (3 times weekly) until a total of 10 treatments is administered. Each of these patches is removed after 6 hours of exposure. It should be noted that some use a modification which calls for one application per week for 3 weeks.

[1] When the test substance is in flake, granule, powder, or other particulates form, the weight of the test substance that has a volume of 0.5 ml (after compacting as much as possible without crushing or altering the individual particles, such as by tapping the measuring container) is used whenever this volume weighs less than 0.5 g.

Table 3.8 Evaluation of local tissue reactions in tissue irritation studies

Method	Basis/end point	Reference
Erythema and eschar formation		
No erythema	0	
Very slight erythema (barely perceptible)	1	
Well-defined erythema	2	
Moderate to severe erythema	3	
Severe erythema (beet redness) to slight eschar formation (injuries in depth)	4	
Necrosis (death of tissue)	+N	
Eschar (sloughing and scar formation)	+E	
Edema formation		
No edema	0	
Very slight edema (barely perceptible)	1	
Slight edema (edges of area well-defined by definite raising)	2	
Moderate edema (raised approximately 1 mm)	3	
Severe edema (raised more than 1 mm and extending beyond the area of exposure)	4	
Total possible score for primary irritation	8	

Draize et al. (1944)

- Observations are made 24 and 48 hours after initiation of each subsequent induction application. Dermal scores of the remaining nine induction applications are recorded.
- Clipping the hair from the left flank induction sites of test group animals and corresponding sites on negative control group animals is performed just prior to each subsequent induction application. Only the test group guinea pigs receive the induction applications.

3.4.3.4 Challenge Phase

- Fourteen days after the tenth induction application, all ten test groups, and three of five control groups, of guinea pigs are prepared for challenge application by clipping a 1-inch square of hair from the right side, the side opposite that which was clipped during the induction phase.
- A challenge dose, using freshly prepared test substance (solution, suspension, emulsion, semisolid, solid, or fabric), is applied topically to the right side (which had remained untreated during the induction application) of test group animals. The left side, which had previously received induction applications, is not challenge dosed.
- The concentrations of the challenge dose are the same as that used for the first induction application. (It must be a concentration that does not produce dermal irritation after one 24 h application.)

- Each of three negative control group guinea pigs is challenge dosed on the right flank at approximately the same time that the test group guinea pigs are challenge dosed.
- All patches are held in contact with the skin for 24 hours before removal.
- The skin sites are evaluated using the scoring system for erythema and edema formation presented in Table 3.8. Observations are made 48, 72, and 96 hours after initiation of the challenge application. Skin reactions are recorded.

3.4.3.5 Rechallenge Phase

- If the test substance is judged a nonsensitizing agent after the first challenge application or causes dermal sensitization in only a few animals or causes dermal reactions that are weak or questionable, then a second and final challenge application will be performed on each test animal 7 days after the initiation of the first challenge dose.
- Controls from the first challenge application are not rechallenged because they have been exposed to the test substances and are no longer true negative controls. The three remaining naive control group animals (not used for the first challenge) are challenged for comparison to the test group animals.
- The procedure used for the first challenge application will be used for the second challenge application (including reclipping, patching method, and duration of exposure). Either the same concentration or a new concentration (higher or lower) of test substances may be sued, depending on the results of the first challenge. Observations are made 48, 72, and 96 hours after initiation of the rechallenge application and skin reactions are recorded.
- When a rechallenge application is performed, the data from both challenges are compared. If neither challenge produces a positive dermal reaction, the classification of the test substance is based on both challenge applications. If one challenge application (whether it is the first or second) produces a greater number of positive dermal reactions than the other, the classification of the test substance is based on the challenge with the most positive responses.
- Two or more unequivocally positive responses in a group of 15 animals should be considered significant. A negative, equivocal, or single response probably assures that a substance is not a strong sensitizer, although this is best confirmed by further testing with human subjects (NAS 1977).

3.4.3.6 Interpretation of Results

- Judgment concerning the presence or absence of sensitization is made for each animal. The judgment is made by comparing the test animal's challenge responses to its first induction treatment response, as well as to those challenge responses of negative control animals.

- Challenge reactions to the test substance that are stronger than challenge reactions to negative controls or to those seen after the initial induction application should be suspected as results of sensitization (NAS 1977). A reaction that occurs at 48 hours, but resolves by 72 hours or 96 hours, should be considered a positive response as long as it is stronger than that which is displayed by controls at the same time interval.

3.4.3.7 Strengths and Weaknesses

There are a number of both advantages and disadvantages to the Buehler methodology, which has been in use for 20 years. The relative importance and merits of each depend on the intended use of the material. The four advantages are:

1. Virtually no false positives (in fact, in the experience of the author when the pretest is properly conducted, there are no false positives), compared to human experience, are generated by this test.
2. The techniques involved are easy to learn and very reproducible.
3. The Buehler-style test does not overpredict the potency of sensitizers. That is, materials which are identified as sensitizers are truly classified as very strong, weak, or in-between—not all (or nearly all) as very strong.
4. There is a large database in existence for the Buehler-style test. Unfortunately, the vast majority is not in the published literature.

Likewise, there are three disadvantages associated with the Buehler-style test.

1. The test gives a high rate of false negatives for weak sensitizers and a detectable rate of false negatives for moderate sensitizers. That is, the method is somewhat insensitive—particularly if techniques for occlusive wrapping are inadequate.
2. The test takes a long time to complete. If animals are on-hand when started, the test is 5–6 weeks long. As few laboratories keep a "pool" of guinea pigs on-hand (especially as they are the most expensive of the common lab species), the usual case is that 8–10 weeks is the minimum time required to get an answer from this test.
3. The test uses a relatively large amount of test material. In the normal acute "battery," the guinea pig test systems use more material than any other test systems unless an acute inhalation study is included. With ten induction applications, this is particularly true for the Buehler-style test.

3.4.4 Guinea Pig Maximization Test

The guinea pig maximization test (GPMT) was developed by Magnusson and Kligman (1969, 1970) and Magnusson (1975) and is considered a highly sensitive procedure for evaluating test substances for potential dermal sensitization. The procedure presented here is illustrated diagrammatically in Figs. 3.2 and 3.3. is one common version of the test.

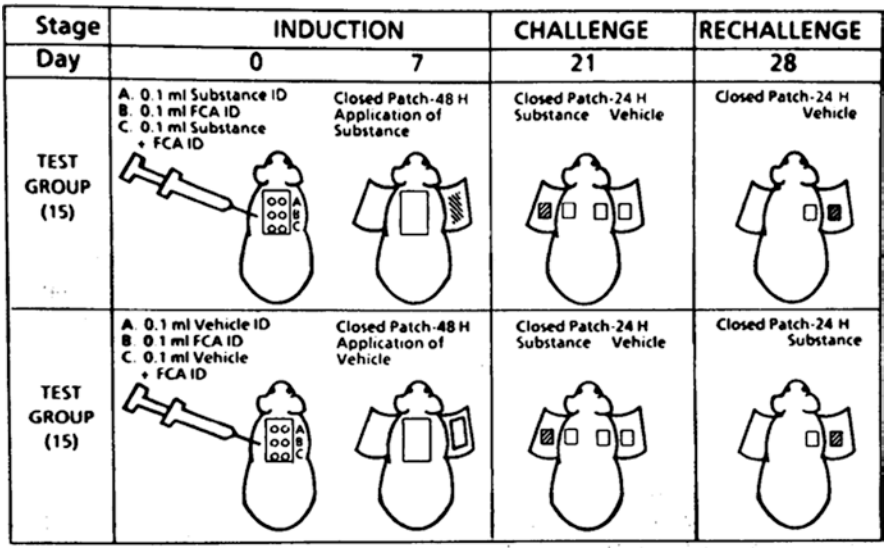


Fig. 3.2 Line chart of study design for guinea pig maximization test (GMPT) for predicting delayed dermal sensitization

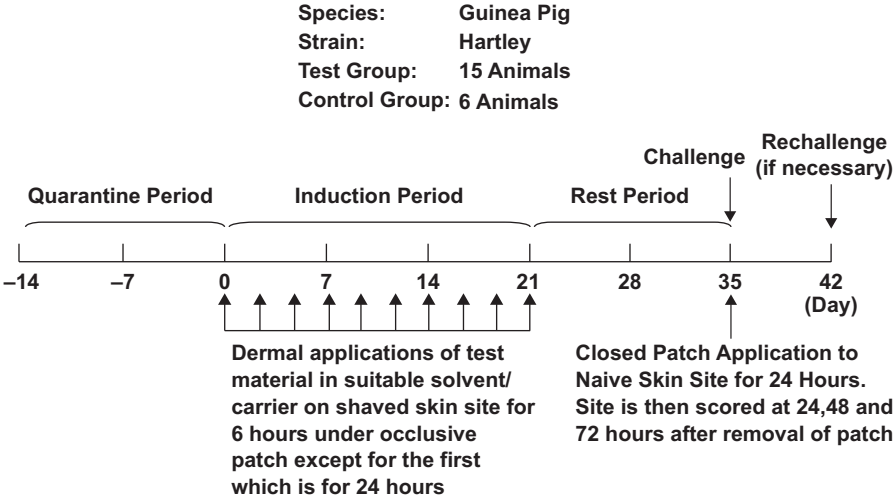


Fig. 3.3 Line chart of study design for modified Buehler test for delayed contact dermal sensitization in the guinea pig

3.4.4.1 Animals

- Young adult female guinea pigs, weighing between 250 and 350 g at the initiation of the study, are used.
- Although several proposed test rules suggest the use of male guinea pigs, the female sex is preferred because the aggressive social behavior of males may result in considerable skin damage that might interfere with the interpretation of challenge reactions.
- Animals that show poor growth or are ill in any way are not used, since illness markedly decreases the response. Animals with skin marked or scarred from fighting are avoided. The guinea pigs are observed for at least 2 weeks to detect any illness before starting the study.
- The guinea pigs are randomly assigned to two groups: (1) a test group consisting of 15 animals and (2) a control group consisting of 6 animals. If a pretest group is necessary, as many animals as needed for that group are randomized also.
- Test and control group guinea pigs are weighed one week prior to dosing (day 7), on the day of dosing (day 0), and weekly thereafter.

3.4.4.2 Pretest

- Several animals are used to pretest the test substance and vehicles to determine the topical dermal irritation threshold concentration.
- These animals are shaved on the left flank, to which is applied a 2 × 2 cm filter paper patch which contains 0.1 ml of test concentration.
- The trunks of the animals are wrapped for 24 hours with a 3-inch-wide elastic bandage to hold the patch in contact with the skin.
- Wrappings are removed after the 24 h exposure, and, based on skin reactions at 48 h, a concentration of the test substance to be used on test is determined. Dermal irritation values are recorded for future reference.
- In addition, several guinea pigs are utilized to determine a concentration (generally, between 1% and 5%) of test substance in vehicle and in FCA emulsion that can be injected id without eliciting a strong local or systemic toxic reaction.
- The hair is clipped in an area of approximately 4 × 6 cm from the upper shoulder region of these animals.
- Several concentrations of test substances (ranging between 1% and 5%) can be injected in the same animal to compare local dermal reactions produced by the different concentrations.
- However, if systemic toxicity is suspected, then each concentration should be tested in separate animals to determine local and systemic effects.
- The dermal reactions (erythema, edema, and diameter) are recorded 24 hours after the id injections.

Induction Stage 1 (Day 0)

- The hair in an area of 4 × 6 cm is clipped from the shoulder region of each test and control group guinea pig on day 0.
- Three pairs of intradermal (id) injections are made with a glass 1-ml tuberculin syringe with a 26-gauge needle, each pair flanking the dorsal midline.
- The three pairs of id injections for *test group* animals are as follows:

  1. ml test substance in appropriate vehicle,
  2. ml Freund's complete adjuvant (FCA) emulsion alone,
  3. ml test substance in FCA emulsion.

- The three pairs of id injection for *control group* animals are as follows:

  1. 0.1 ml vehicle alone
  2. 0.1 FCA emulsion alone
  3. 0.1 ml vehicle in FCA emulsion.

- Injections (1) and (2) in the above two steps are given close to each other and nearest the head; injection (3) is given most posteriorly.
- The date, time, and initials of those individuals performing the id injections are recorded.
- Immediately before injection, an emulsion is prepared by blending commercial FCA with an equal volume of house distilled water or other solvent as appropriate.

  (a) Water-soluble test materials are dissolved in the water phase prior to emulsification.
  (b) Oil-soluble or water-insoluble materials are dissolved or suspended in FCA prior to adding water.
  (c) Paraffin oil, peanut oil, or propylene glycol can be used for dissolving or suspending water-insoluble materials.
  (d) A homogenizer is used to emulsify the FCA alone and the test substance in other in either FCA or vehicle prior to the id injections.
  (e) The concentration of the test substance for id injections is adjusted to the highest level that can be well tolerated locally and generally.

- The adjuvant injection infiltration sometimes causes ulceration, especially when it is superficial, which lasts several weeks. These lesions are undesirable but do not invalidate the test results except for lowering the threshold level for skin irritation.

3.4.4.3 Induction Stage 2 (Day 7)

- Test Substance Preparation.

  (a) The concentration of the test substance is adjusted to the highest level that can be well tolerated.

(b) If the test substance is an irritant, a concentration is chosen that causes a weak to moderate inflammation (as determined by the pretest).
(c) Solids are micronized or reduced to a fine powder and then suspended in a vehicle, such as petrolatum or propylene glycol.
(d) Water- and oil-soluble test substances are dissolved in an appropriate vehicle.
(e) The concentration of the test substance for id injections is adjusted to the highest level that can be well tolerated locally and generally.

- The same area over the shoulder region that received id injections on day 0 is again shaved on both test and control guinea pigs.
- A volume of 0.3 ml of a mildly irritating concentration (if possible) of the test substance (determined by the pretest) is spread over a 1 × 2 inch filter of each test group animal.
- The control group animals are exposed to 0.3 ml of 100% vehicle using the same procedure.
- The date, time, and initials of those individuals performing the second induction are recorded.
- The dressings of both groups are left in place for 48 hours before removal.

3.4.4.4 Challenge Stage (Day 21)

1. An area of hair (1.5 × 1.5 in) on both flanks of the guinea pigs (15 test and 3 controls) is shaved.
2. A 1 × 1-inch patch with a nonirritating concentration of test substance in vehicle (as determined by the pretest) is applied to the left flank and a 1 × 1 inch patch with 100% vehicle is applied to the right flank.
3. The torso of each guinea pig is wrapped in an elastic bandage to secure the patches for 24 hours.
4. The date, time, and initials of those individuals performing the challenge dose are recorded.
5. The patches are removed 24 hours after application.

3.4.4.5 Rechallenge (Day 28)

1. If the first challenge application of test substance does not cause dermal sensitization, causes dermal sensitization in only a few animals, or causes dermal reactions that are weak or questionable, then a second challenge application of test substance to the 15 test group guinea pigs will be conducted on day 28 (1 week after the first challenge). The three remaining naive control group animals (not used for the first challenge) are challenged for comparison to the test group animals.

2. The three negative control group animals used on day 21 will not be rechallenged. These animals will be terminated because they were exposed to the test substance during the first challenge and are no longer negative controls.
3. A 1 × 1-inch patch with a nonirritating concentration of test substance in vehicle is applied to the right flank of test and control group animals. The left flanks are not dosed.
4. The date, time, and initials of those individuals performing the rechallenge dose are recorded.
5. Steps 3 and 5 are followed as for Challenge State (Day 21).

3.4.4.6 Observations: Challenge and/or Rechallenge Readings

1. Twenty-one hours after removing the patch, the challenge area on each flank is cleaned and clipped, if necessary.
2. Twenty-four hours after removing the patch, the first reading of dermal reactions is taken.
3. The dermal reactions are scored on a four-point scale (as below):
4. No reaction.
5. Scattered mild redness.
6. Moderate and diffuse redness.
7. Intense redness and swelling.
8. Forty-eight hours after removing the patch, the second reading is taken and the scores are recorded.

3.4.4.7 Interpretation of Results

1. Both the intensity and duration of the test responses to the test substance and the vehicle are evaluated.
2. The important statistic in the GPMT is the frequency of sensitization and not the intensity of challenge responses. A value of 1 is considered just as positive as a value of 3 (as long as the values for controls are zero).
3. The test agent is a sensitizer if the challenge reactions in the test group clearly outweigh those in the control group. A reaction that occurs at 24 hours, but resolves by 48 hours after removal of patches, should be considered a positive response, as long as it is stronger than that which is displayed by controls. The sensitization rate (% of positive responders) is based on the greatest number of animals showing a positive response, whether it is from the 24-hour data or the 48-hour data after removal of patches.
4. When a second challenge application is performed, the data from both challenges are compared. If neither challenge produces a positive dermal reaction, the classification of the test substance is based on both challenge applications.

Table 3.9 Sensitization severity grading based on incidence of positive responses

Sensitization rate %	Grade	Classification
0–8	I	Weak
9–28	II	Mild
29–64	III	Moderate
65–80	IV	Strong
81–100	V	Extreme

Kligman (1966)

If one challenge application (whether it is the first or second) produces a greater number of positive dermal reactions than the other, the classification of the test substance is based on the challenge with the most positive responses.

5. Under the classification scheme of Kligman (1966, shown in Table 3.9), the test substance is assigned to one of five classes, according to the percentage of animals sensitized, ranging from a week grade I to an extreme grade V.

The advantages and disadvantages of the GPMT can be summarized as follows. First, the advantages:

1. The test system is sensitive and effectively detects weak sensitizers. It has a low false-negative rate.
2. If properly conducted, there are no false positives—that is, materials which are identified as potential sensitizers will act as such at some incidence level in humans.
3. There is a large database available on the evaluation of compounds in this test system, and many people are familiar with the test system.

The disadvantages, meanwhile, are:

1. The test system is sensitive; it overpredicts potency for many sensitizers. There is no real differentiation between weak, moderate, and strong sensitizers; virtually all positive test results identify a material as strong.
2. The techniques involved (particularly the intradermal injections) are not easy. Some regulatory officials have estimated that as many as 35% of the laboratories which try cannot master the system to get it to work reproducibly.
3. The test, though not as long as the Buehler, still takes a minimum of 4 weeks to produce an answer.
4. The test uses a significant amount of test material.
5. One cannot evaluate fibers or other materials which cannot be injected (such as either solids which cannot be finely ground and/or suspended or which are highly irritating or toxic by the iv route).
6. The irritation pretest is critical. Failure to detect irritation in this small group of animals does not guarantee against irritation in test animals at challenge.

3.4.5 *Local Lymph Node Assay (LLNA)*

This method has developed out of the work of Ian Kimber and associates (Kimber et al. 1986, 1994; Kimber and Weisenberger 1989). It has the advantage over the other methods discussed in this chapter in that it provides an objective and quantifiable end point. The method is based on the fact that dermal sensitization requires the elicitation of an immune response. This immune response requires proliferation of a lymphocyte subpopulation. The local lymph node assay (LLNA) relies on the detection of increased DNA synthesis via titriated thymidine incorporation. Sensitization is measured as a function of lymph node cell proliferative responses induced in a draining lymph node following repeated topical exposure of the test animal to the test article. Unlike the other tests discussed in this chapter, this assay looks only at induction because there is no challenge phase.

The typical test (illustrated in Fig. 3.4) is performed using mice—normally female CBA mice 6–10 weeks of age. Female BALB/c and ICR mice have also been used. After animal receipt, they are typically acclimated to standard laboratory husbandry conditions for 7–10 days. The usual protocol will consist of at least two groups (vehicle control and test article treated) of five mice each. They are treated on the dorsal surface of both ears with 25 μl (on each ear) of test article solution for three consecutive days. 24 to 48 hours after the last test article exposure, the animals are given a bolus (0.25 ml) dose of [^{3}H]thymidine (20 μCi with a specific activity of 5.0–7.0 Ci/mmol) in phosphate buffered saline via a tail vein. Five hours after the injection, the animals are euthanized by CO_2 asphyxiation and the auricular lymph nodes removed.

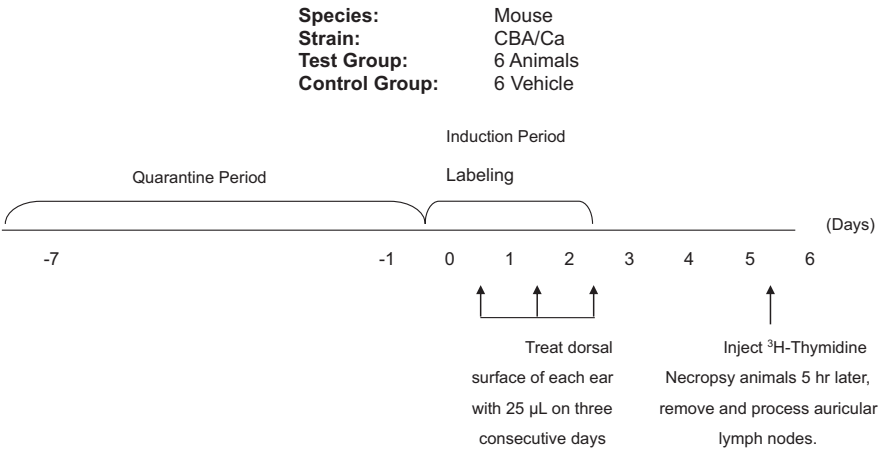


Fig. 3.4 Mouse local lymph node assay (LLNA) (ICVAM protocol)

After removal, the lymph nodes can either be pooled by group or processed individually. Single cell suspensions are prepared by gentle mechanical disaggregation through a nylon (100 μm) mesh. Cells are washed twice by centrifugation in an excess of PBS. After the final supernatant wash is removed, the cells are precipitated with cold 5% trichloroacetic acid (TCA) and kept at 4 °C for 12–18 hours. The precipitate is then pelleted by centrifugation and resuspended in 1 ml 5% TCA, and the amount of radioactivity is determined by liquid scintillation counting, using established techniques for tritium.

The data are reduced to the stimulation index (SI):

$$\mathrm{SI} = \frac{H\left(\mathrm{dpm}\right)\text{treated group}}{H\left(\mathrm{dpm}\right)\text{control group}}$$

An SI of 3 or greater is considered a positive response, i.e., the data support the hypothesis that the test material is a sensitizer.

The test article concentration is normally the highest nonirritating concentration. Several concentrations could be tested at the same time should one wish to establish a dose-response curve for induction. The test is easiest to perform if the vehicle is a standard nonirritating organic, such as acetone, ethanol, or dimethylformamide, or a solvent-olive old blend. Until a laboratory develops its own historical control base, it is also preferable to include a positive control group. Either 0.25% dinitrochlorobenzene or 0.05% oxazalone are recommended for positive controls. If the vehicle for the positive control is different than the vehicle for the test material, then two vehicle control groups may be necessary.

This method has been extensively validated in two international laboratory exercises (Basketter et al. 1992; Loveless et al. 1996). In the earlier work (Basketter et al. 1991), there was good correlation between the results obtained with guinea pig tests and those obtained with the LLNA. In the recent report, for example, five laboratories correctly identified dinitrochlorobenzene and oxazalone as sensitizers and the fact that *p*-aminobenzoic acid was not (Loveless et al. 1996). Arts and colleagues (Arts et al. 1996) demonstrated that rats could be used as well as mice. Interestingly, they validated their assay (for both rats and mice) using BrDU uptake and immunohistochemical staining (rather than [^{3}H] thymidine) to quantitated lymph node cell proliferation.

This method is relatively quick and inexpensive because it uses relatively few mice (which are much less expensive than guinea pigs) and takes considerably less time than traditional guinea pig assays. It has an advantage over other methods in that it does not depend on a somewhat subjective scoring system and produces an objective and quantifiable end point. It does require a radiochemistry laboratory. Unless one already has an appropriately equipped laboratory used for other purposes (most likely metabolism studies), setting one up for the sole purpose of running the LLNA does not make economic sense. The standard version of the test has been adopted by OECD, ISO, ICH, and ICVAM, but also has been shown to have a modest false-positive rate (misidentifying strong irritants as sensitizers). There is a modified version which (using flow cytometry and/or measurement of cytokine levels) is believed to solve this false-positive rate problem.

3.4.5.1 Test System Manipulation (For All In Vivo Test Systems)

Increasing percutaneous absorption will increase test sensitivity. Factors which will increase absorption (and techniques for achieving them) include the following:

1. Increase surface area of solids.
2. Hydrate region of skin exposed to chemical. This can be done by wetting solids and using very occlusive wrapping of application.
3. Irritate application site.
4. Abrade application site.
5. Injection of test material (if possible).
6. Proper selection of solvent or suspending system. (See Christensen et al. (1984) for a discussion of the effect of vehicle in the case of even a strong sensitizer.)
7. Remove part or all of the "barrier layer" (stratum corneum) by tape stripping the application site.
8. Increase the number of induction applications.

Though it is not a factor which increases percutaneous absorption, mildly stimulating the immune system of test animals (by such means as injecting FCA (or some other adjuvant) alone or FCA blended with the test material) also increases responsiveness to the test system.

Also, it is generally believed that using the highest possible test material concentrations (mildly irritating for induction, just below irritating for challenge) will guarantee the greatest possible sensitization response and will therefore also serve to universally increase sensitivity. There are reports, however (Gad et al. 1985 for croton oil and Thorne et al. 1986 for isocyanates), that this is not true for all compounds and that a multiple-dose (i.e., two or more concentrations) study design would increase sensitivity. Such designs, however, would also significantly increase cost.

Concurrent or frequent positive and negative controls are essential to guard against test system failure. Any of these test systems should show 0.05% dinitrochlorobenzene (DNCB) in 70% ethanol to be a strong sensitizer.

In Vitro Methods: There are actually several approaches available to in vitro evaluation of materials for sensitizing potential. These use cultured cells from various sources and, as end points, look at either biochemical factors (such as production of MIF-migration inhibition factor) or cellular events (such as cell migration or cell "transformation").

Milner (1970) reported that lymphocytes from guinea pigs sensitized to dinitrofluorobenzene (DNFB) would transform in culture, as measured by the incorporation of tritiated thymidine, when exposed to epidermal proteins conjugated with DNFB. This work was later extended to guinea pigs sensitized to p-phenylenediamine. He later (Milner 1971) reported that his method was capable of detecting allergic contact hypersensitivity to DNFB in humans when he used human lymphocytes from sensitized donors and human epidermal extracts conjugated with DNFB.

Miller and Levis (1973) reported the in vitro detection of allergic contact hypersensitivity to DNCB conjugated to leukocyte and erythrocyte cellular membranes.

Table 3.10 Requested reference compounds for skin sensitization studies (US Consumer Product Safety Commission)

Tribromophylophosphate	Formalin
Ditallow dimethyl ammonium methyl sulfate	Turpentine
Hydroxylamine sulfate	Potassium dichromate
Ethyl amino benzoate	Penicillin G
Todochlorohydroxy quinoline	*p*-Phenylenediamine
(Clioquinol, chinoform)	Epoxy systems (ethylenediamine,
Nickel sulfate	diethylenetriamine, and diglycidyl ethers)
Monomethyl methacrylate	Toluene 2,4-diisocyanate
Mercaptobenzothiazole	Oil of bergamot

This indicated that reaction was not specifically directed toward epidermal cell conjugates.

Thulin and Zacharian (1972) extended others' earlier work on MIF-induced migration of human peripheral blood lymphocytes to a test for delayed contact hypersensitivity.

None of these approaches has yet been developed as an in vitro predictive test, but work is progressing. Milner (1983) has published a review of the history and state of this field.

Any alternative (in vitro or in vivo) test for sensitization will need to be evaluated against a battery of "known" compounds. The Consumer Product Safety Commission in 1977 proposed such a battery, which is shown in Table 3.10.

3.5 Irritation (Local Tissue Tolerance) (ISO 10993-10)

Local tissue tolerance or irritation studies assess the short-term and generally localized hazards of medical devices in the immediate region of their tissue contact. Topical local (tissue) tolerance effects are almost entirely limited to the inflammatory aspects of the innate immune response, primary irritation. Though this usually means dermal irritation, it can also be vaginal, muscular, vascular, mucous membrane, rectal, nasal, intracutaneous, or ocular. All of these but ocular irritation utilize some version of a common subjective rating scale derived from the original Draize scale (see Table 3.11) to evaluate responses. These are the second of the universally required tests for all devices, irregardless of route or duration of patient exposure.

Most commonly recognized is the use of this scale in the primary dermal irritation (PDI) test, which is performed for those agents that are to be administered to patients by application to the skin. As with all local tolerance tests, it is essential that the material be evaluated in "condition of use"—that is, in the final product ready for human use, applied to test animals in the same manner that the device or bioma-

Table 3.11 Evaluation of skin reactions

Skin reaction	Value
Erythema and eschar formation:	
No erythema	0
Very slight erythema (barely perceptible)	1
Well-defined erythema	2
Moderate to severe erythema	3
Severe erythema (beef redness) to slight eschar formation (injuries in-depth)	4
Necrosis (death of tissue)	+N
Eschar (sloughing)	+E
Edema formation:	
No edema	
Very slight edema (barely perceptible)	
Slight edema (edges of area well-defined by definite raising)	
Moderate edema (raised approximately 1 millimeter)	
Severe edema (raised more than 1 millimeter and extending beyond the area of exposure)	

Draize (1959)

terial is to be used clinically. If appropriate (under applicable regulations) or necessary due to the nature or mode of use of the device, an extract can be evaluated. Such extracts are generally evaluated in the intracutaneous reactivity test.

3.5.1 Dermal Irritation

Skin irritation testing is performed to demonstrate the irritation potential of the device, i.e., for initiating or aggravating damage through its contact with the skin (Draize 1955, 1959). Primary skin irritation is usually done according to the regulations of the Consumer Product Safety Commission, Title 16, Chap. II, Part 1500, or some variation thereof (such as ISO and ASTM 1991). The purpose of the study is to determine the dermal irritation potential of the test article to the intact and abraded skin of the rabbit (the latter to stimulate wound tissue).

Skin absorption occurs through a process of binding, partitioning, and diffusion or active transport of test materials on and into the skin. Penetration has been assessed in vivo by measuring at different times the amount of test substances at different layers of the skin. Blood levels of the test sample have been measured in this test.

A complicated series of chemical and physiological responses result in primary skin irritation. When the skin is exposed to toxic substances, a modified form or the Draize rabbit skin test (fewer animals are now used), first outlined by John Draize in 1944, remains an important source of safety information for government and industry (Draize et al. 1944). In this test, the dermal irritation caused by a substance is investigated by observing changes ranging from erythema and edema to ulceration

produced in rabbit skin when irritants are applied. These skin reactions are produced by diverse physiologic mechanisms, although they are easily observed visually and by palpitation. In recent years, it has become more common to evaluate irritation in/on the skin of pigs/minipigs that are being utilized primarily for evaluating systemic topical effects of repeat-dose dermal application.

Evaluation of materials for their potential to cause dermal irritation and corrosion due to acute contact has been common for industrial chemicals, cosmetics, agricultural chemicals, and consumer products since at least the 1930s (generally, pharmaceuticals are only evaluated for dermal effects if they are to be administered topically—and then by repeat exposure tests, which will not be addressed here). As with acute eye irritation tests, one of the earliest formal publications of a test method (though others were used) was that of Draize et al. in 1944 (Geller et al. 1985; SOT 1989). The methods currently used are fundamentally still those proposed by Draize et al. and, to date, have changed very little since 1944. Efforts have been underway for some 20 years to develop alternatives that either don't use animals or are performed in a more humane and relevant (to human exposure) manner.

Among the most fundamental assessments of the safety of a device or, indeed, of any material that has the potential to be in contact with a significant number of people in our society are tests in animals which seek to predict potential skin irritation or corrosion. Like all the other tests in what is classically called range-finding, Tier I, or acute battery, the tests used here both are among the oldest designs and are currently undergoing the greatest degree of scrutiny or change. Currently, the most common test methods for these end points use the same animal model, the rabbit (almost exclusively the New Zealand white), though some other animal models have been proposed.

Testing is performed to evaluate the potential occurrence of two different, yet related, end points. The broadest application of these is an evaluation of the potential to cause skin irritation, characterized by erythema (redness) and edema (swelling). Severity of irritation is measured in terms of both the degree of these two parameters and how long they persist. There are two types of irritation tests, each designed to address a different concern:

- Primary (or acute) irritation, a localized reversible dermal exposure response resulting from a single application of, or exposure to, a chemical without the involvement of the immune system.
- Cumulative irritation, a reversible dermal response which results from repeated exposure to a device (each individual exposure possessing no or limited potential to causing acute irritation). However, it should be noted that repeated acute or continuous inflammation due to irritation may trigger more complex and less readily reversible responses.

Most regulations and common practice characterize an irritation that persists 14 days past the end of exposure as other than reversible. The adult human has 1.8m^2 of skin, varying in thickness from 0.02 inches on the eyelids to 0.12 to 0.16 inches on the back, palms, and soles of the feet (Hipp 1978). The epidermis, the outer portion

of the skin, is several layers thick, covers the entire surface of the body, and is referred to as the horny layer or *stratum corneum*. It is the first line of defense against physical, chemical, and thermal exposure. The skin is host to normal bacterial flora consisting of *Micrococcus* and *Corynebacterium*, which play an important role in the protection against infection. The melanocyte system, responsible for skin colonization, is located at the interface of the epidermis and the dermis. New cells are constantly being formed from the basal layer and slowly migrate to the surface, replenishing themselves approximately every 2 weeks (Monash and Blank 1958; Matoltry et al. 1968).

Irritation is generally a localized reaction resulting from either a single or multiple exposure to a physical or chemical entity at the same site. It is characterized by the presence of erythema (redness) and edema and may or may not result in cell death. The observed signs are heat (caused by vessel dilation and the presence of large amounts of warm blood in the affected area), redness (due to capillary dilation), and pain (due to pressure on the sensory nerves). The edema often observed is largely due to plasma, which coagulates in the injured area, precipitating a fibrous network to screen off the area, thereby permitting leukocytes to destroy exogenous materials by phagocytosis. If the severity of injury is sufficient, cell death may occur, thereby negating the possibility of cellular regeneration. Necrosis is a term often used in conjunction with cell death and is the degeneration of the dead cell into component molecules which approach equilibrium with surrounding tissue (Montagna 1961).

3.5.1.1 Primary Dermal Irritation Test

Rabbit Testing Procedure

- A group of at least five New Zealand white rabbits are screened for the study.
- All rabbits selected for the study must be in good health; any rabbit exhibiting sniffles, hair loss, loose stools, or apparent weight loss is rejected and replaced.
- One day (at least 18 h) prior to application of the test substance, each rabbit is prepared by clipping the hair from the back and sides using a small animal clipper. A size No. 10 blade is used to remove long hair and then a size No. 40 blade is used to remove the remaining hair.
- Three animals with skin sites that are free from hyperemia or abrasion (due to shaving) are selected. Skin sites that are in the telogen phase (resting stage of hair growth) are used; those skin sites that are in the anagen phase (stage of active growth, indicated by the presence of a thick undercoat of hair) or not used.

Study Procedure

- As many as four areas of skin, two on each side of the rabbit's back, can be utilized for sites of administration.

- Separate animals are not required for an untreated control group. Each animal serves as its own control; Indeed, up to eight separate sites may be used for a single rabbit.
- The intact (free of abrasion) sites of administration are assigned a code number. Typically, a suitably prepared extract from a device or biomaterial and an extraction solution ("vehicle" control) are tested on each animal.
- Application sites should be rotated from one animal to the next to ensure that the test substance and controls are applied to each position at least once.
- Each test or control substance is held in place with a 1″ × 1″ 12-ply surgical gauze patch. The gauze patch is applied to the appropriate skin site and secured with 1″-wide strips of surgical tape at the four edges, leaving the center of the gauze patch nonoccluded.
- If an extraction solution is being evaluated, a patch is applied and secured to the appropriate skin site. A 1-ml tuberculin syringe is used to measure and apply 0.5 ml of test substance to the patch.
- The negative control site is covered with an untreated 12-ply surgical gauze patch (1″ × 1″).
- The entire trunk of the animal may be covered with an impervious material (such as Saran Wrap®) for a 24-h period of exposure. The Saran Wrap® is secured by wrapping several long strips of athletic adhesive tape around the trunk of the animal. The impervious material aids in maintaining the position of the patches and retards the evaporation of volatile test substances. Alternatively, a Hilltop chamber-type self-enclosed applicator may be utilized.
- An Elizabethan collar is fitted and fastened around the neck of each test animal. The collar remains in place for the 24-h exposure period. The collars are utilized to prevent removal of wrappings and patches by the animals while allowing the animals food and water ad libitum.
- The wrapping is removed at the end of the 24-h exposure period. The test substance skin site is wiped to remove any test substance still remaining. When colored test substances (such as dyes) are used, it may be necessary to wash the test substance from the test site with appropriate solvent or vehicle (one that is suitable for the substance being tested). This is done to facilitate accurate evaluation for skin irritation.
- Immediately after the removal of the patches, each 1″ × 1″ test or control site is outlined with an indelible marker by dotting each of the four corners. This procedure delineates the site for identification.

Observations

- Observations are made of the test and control skin sites 1 h after removal of the patches (25 h postinitiation of application). Erythremia and edema are evaluated and scored on the basis of the designated values presented earlier in Table 3.11.
- Observations are again performed 46 and 72 h after application and scores are recorded.

- If necrosis is present or the dermal reaction is unusual, the reaction should be described. Severe erythema should receive the maximum score (4), and + N should be used to designate the presence of necrosis and + E the presence of eschar.
- When the test substance produces dermal irritation that persists 72 h postapplication, daily observations of test and control sites are continued on all animals until all irritation caused by the test substance resolves until Day 14 postapplication.

Evaluation of Results

- A *subtotal irritation value* for erythema and eschar formation is determined for each rabbit by adding the values observed at 25, 48, and 72 h postapplication.
- A *subtotal irritation value* for edema formation is determined for each rabbit by adding the values observed at 25, 48, and 72 h postapplication.
- A *total irritation score* is calculated for each rabbit by adding the subtotal value for erythema or eschar formation to the subtotal irritation value for edema formation.
- The *primary dermal irritation index* is calculated for the test substance or control substance by diving the sum of the total irritation scores by the number of observations (three days × three animals = nine observations).
- The categories of the primary dermal irritation index (PDII) are as follows [this categorization of dermal irritation is a modification of the original classification described by Draize et al. (1944)]:

PDII = 0.0	Nonirritant
>0.0–0.5	Negligible irritant
>0.5–2.0	Mild irritant
>2.0–5.0	Moderate irritant
>5.0–8.0	Severe irritant

Other abnormalities, such as atonia or desquamation, should be noted and recorded.

3.6 Intracutaneous Reactivity

Essentially this is a modification of the primary dermal irritation study except that the test device or extracts from it are inserted/injected into the subcutaneous tissue of the test animal. The animal model used is the albino rabbit; the test object or injection is allowed to stay in place for 72 hours after insertion. The (primarily innate immune) response is score at 24, 48, and 72 hours using an eight-point edema and erythema scale. Three albino New Zealand rabbits are used, and up to five separate site can be used and evaluated on each animal, with one or two being vehicle controls.

The intracutaneous irritation test is a sensitive acute toxicity screening test and is generally accepted for detecting potential local irritation by extracts from a biomaterial. Extracts of material obtained with nonirritation polar and nonpolar extraction media are suitable, and sterile extracts are desirable.

3.6.1 Intracutaneous Test

This test is designed to evaluate local responses to the extracts of materials under test following intracutaneous injection into rabbits.

3.6.1.1 Test Animal

Select healthy, thin-skinned albino rabbits whose fur can be clipped closely and whose skin is free from mechanical irritation or trauma. In handling the animals, avoid touching the injection sites during observation periods, except to discriminate between edema and an oil residue. Rabbits previously used in unrelated tests, such as the *pyrogen test*, and that have received the prescribed rest period, may be used for this test provided that they have clean unblemished skin.

3.6.1.2 Procedure

Agitate each extract vigorously prior to withdrawal of injection doses to ensure even distribution of the extracted matter. On the day of the test, closely clip the fur on the animal's back on both sides of the spinal column over a sufficiently large test area. Avoid mechanical irritation and trauma. Remove loose hair by means of vacuum. If necessary, swab the skin lightly with diluted alcohol, and dry the skin prior to injection. More than one extract from a given material can be used per rabbit, if you have determined that the test results will not be affected. For each *Sample* use two animals and inject each intracutaneously, suing one side of the animal for the *Sample* and the other side for the *Blank*, as outlined in Table 3.12. [*Note*—Dilute each gram of the extract of the *Sample* prepared with *polyethylene glycol 400*, and the corresponding *Blank*, with 7.5 volumes of *sodium chloride injection* to obtain a solution having a concentration of about 120 mg of polyethylene glycol per mL.]

Examine injection sites for evidence of any tissue reaction such as erythema, edema, and necrosis. Swab the skin lightly, if necessary, with diluted alcohol to

Table 3.12 Intracutaneous test

Extract or blank	Number of sites (per animal)	Dose, *μL per site*
Sample	5	200
Blank	5	200

facilitate reading of injection sites. Observe all animals at 24, 48, and 72 hours after injection. Rate the observations on a numerical scale for the extract of the *Sample* and for the *Blank*, using Table 3.8. Reclip the fur as necessary during the observation period.

If each animal at any observation period shows an average reaction to the *Sample* that is not significantly greater than to the *Blank*, the *Sample* meets the requirements of this test. If at any observation period the average reaction to the *Sample* is questionably greater than the average reaction to the *Blank*, repeat the test using three additional rabbits. On the repeat test, the average reaction to the *Sample* in any of the three animals is not significantly greater than the *Blank*.

3.7 Ocular Irritation Testing

Unless the device is specifically intended for ocular contact, in vivo eye irritation testing is no longer conducted.

Ocular irritation is significantly different from the other local tissue irritation tests on a number of grounds (Grant 1993). For the medical device industry, eye irritation testing is performed when the device is intended to be put into the eye as a means or route of application for ocular therapy. There are a number of special tests applicable to medical devices that are beyond the scope of this chapter, since they are intended to assess potential effects or irritation of a specific device. These are addressed later in the chapter on special cases. In general, however, it is desired that an eye irritation test that is utilized by this group be both sensitive and accurate in predicting the potential to cause irritation in humans. Failing to identify human ocular irritants (lack of sensitivity) is to be avoided, but of equal concern is the occurrence of false positives.

3.7.1 Primary Eye Irritation Test

The primary eye irritation test was originally intended to predict the potential for a single splash of chemical into the eye of a human being to cause reversible and/or permanent damage. Since the introduction of the original Draize test 50 years ago (Draize et al. 1944), ocular irritation testing in rabbits has both developed and diverged. There is no longer a single test design that is used and different objectives are pursued by different groups using the same test. This lack of standardization has been recognized for some time, and attempts have been made to address standardization of at least the methodological aspects of the test, if not the design aspects.

One widely used study design, which begins with a screening procedure as an attempt to avoid testing severe irritants or corrosives in animals, goes as follows:

3.7.1.1 Test Article Screening Procedure

- Each test substance will be screened in order to eliminate potentially corrosive or severely irritating materials from being studied for eye irritation in the rabbit.
- The pH of the test substance (ISO compliant extract or ocular lens solution) measured.
- A primary dermal irritation test will be performed prior to the study.
- The test substance will not be studied for eye irritation if it is a strong acid (pH of 2.0 or less) or strong alkali (pH of 11.0 or greater) and/or if the test substance is a severe dermal irritant (with a PDII of 5 to 8) or causes corrosion of the skin.
- If it is predicted that the test substance does not have the potential to be severely irritating or corrosive to the eye, continue to Section B, Rabbit Screen Procedure.

3.7.1.2 Rabbit Screening Procedure

- A group of at least six New Zealand white rabbits of either sex are screened for the study. The animals are removed from their cages and placed in rabbit restraints. Care should be taken to prevent mechanical damage to the eye during this procedure.
- All rabbits selected for the study must be in good health; any rabbit exhibiting sniffles, hair loss, loose stools, or apparent weight loss is rejected and replaced.
- One hour prior to instillation of the test substance, both eyes of each rabbit are examined for signs of irritation and corneal defects with a handheld slit lamp. All eyes are stained with 2.0% sodium fluorescein and examined to confirm the absence of corneal lesions.
- *Fluorescein Staining:* Cup the lower lid of the eye to be tested and instill one drop of a 2% (in water) sodium fluorescein solution onto the surface of the cornea. After 15 sec, thoroughly rinse the eye with physiological saline. Examine the eye, employing a handheld long-wave ultraviolent illuminator in a darkened room. Corneal lesions, if present, appear as bright yellowish-green fluorescent areas.
- Only three of the six animals are selected for the study. The three rabbits must not show any signs of eye irritation and must show either a negative or minimum fluorescein reaction (due to normal epithelial desquamation).

3.7.1.3 Study Procedure

- At least 1 h after fluorescein staining, the test substance is placed in one eye of each animal by gently pulling the lower lid away from the eyeball to form a cup (conjunctival cul-de-sac) into which the test material is dropped. The upper and lower lids are then gently head together for 1 sec to prevent immediate loss of material.
- The other eye remains untreated (if an ocular lens solution is being tested) or receives just extraction solution and serves as a control.
- For testing liquids, 0.01 ml of the test substance is used.

- The treated eyes of the three rabbits are not washed following instillation of the test substance.
- To prevent self-inflicted trauma by the animals immediately after instillation of the test substance, the animals are not immediately returned to their cages. After the test and control eyes are examined and graded at 1-h postexposure, the animals are returned carefully to their respective cages.

3.7.1.4 Observations

- The eyes are observed for any immediate signs of discomfort after instilling the test substance. Blepharospasm and/or excessive tearing are indicative of irritating sensations caused by the test substance, and their duration should be noted. Blepharospasm does not necessarily indicate that the eye will show signs of ocular irritation.
- Grading and scoring of ocular irritation are performed in accordance with Table 3.13. The eyes are examined and grades of ocular reactions are recorded.
- If signs of irritation persist at Day 7, readings are continued on Days 10 and 14 after exposure or until all signs of reversible toxicity are resolved.
- In addition to the required observation of the cornea, iris, and conjunctiva, serious effects (such as pannus, rupture of the globe, or blistering of the conjunctivae) indicative of a corrosive action are reported.
- Whether or not toxic effects are reversible depends on the nature, extent, and intensity of damage. Most lesions, if reversible, will heal or clear within 21 days. Therefore, if ocular irritation is present at the 14-day reading, a 21-day reading is required to determine whether the ocular damage is reversible or nonreversible.

3.7.2 Alternatives

Testing for potential to cause irritation or damage to the eyes remains the most active area for the development (and validation) of alternatives and the most sensitive area of animal testing in biomedical research. This has been true since the beginning of the 1980s. Table 3.14 presents an overview of the reasons for pursuing such alternatives. The major reason, of course, has been the pressure from public opinion.

Indeed, many of the in vitro tests now being evaluated for other end points (such as skin irritation and lethality) are adaptations of test systems first developed for eye irritation uses. A detailed review of the underlying theory of each test system is beyond the scope of this chapter. Frazier et al. (1987a) performed such a review, and Table 3.15 presents an updated version of the list of test systems overviewed in that volume.

Table 3.13 Scale of weighted scores for grading the severity of ocular lesions[a]

Reaction criteria	Score
I. Cornea	
A. Opacity degree of density (area that is most dense is taken for reading)	
1. Scattered or diffuse area, details of iris clearly visible	1
2. Easily discernible translucent area, details of iris slightly obscured	2
3. Opalescent areas, no details of iris visible, size of pupil barely discernable	3
B. Area of cornea involved	
1. One-quarter (or less) but not zero	1
2. Greater than one-quarter, less than one-half	2
3. Greater than one-half/less than whole area	3
4. Greater than three-quarter up to whole area	4
Scoring equals A × B × 5; total maximum = 80[b]	
II. Iris	
A. Values	
1. Folds above normal, congestion, swelling, circumcorneal ingestion (any one or all of these or combination of any thereof), iris still reacting to light (sluggish reaction is possible)	1
2. No reaction to light, hemorrhage, gross destruction (any one or all of these)	2
Scoring equals A × B (where B is the area of the iris involved, graded as "under cornea"); total maximum = 10	
III. Conjunctivae	
A. Redness (refers to palpebral conjunctivae only)	1
1. Vessels definitely injected above normal	2
2. More diffuse, deeper crimson red, individual vessels not easily discernible	3
3. Diffuse beefy red	
B. Chemosis	
1. Any swelling above normal (include initiating membrane)	1
2. Obvious swelling with partial eversion of the lids	2
3. Swelling with lids about half closed	3
4. Swelling with lids about half closed to completely closed	4
C. Discharge	
1. Any amount different from normal (does not include small amount observed in inner canthus of normal animals)	1
2. Discharge with moistening of the lids and hair just adjacent to the lids	2
3. Discharge with moistening of the lids and considerable area around the eye	3
Scoring = (A + B + C) × 2; total maximum = 20	

[a]The maximum total score is the sum of all scores obtained for the cornea, iris, and conjunctivae
[b]All A × B = Σ (1−3) × Σ (1−4) for three animals

There are six major categories of approach to replacing in vivo systems (such as the rabbit), and these almost certainly would require some form of battery of such test systems. Many individual systems, however, might constitute effective screens in defined situations. The first five of these aim at assessing portions of the irritation response, including alterations in tissue morphology, toxicity to individual complete

Table 3.14 Rationales for seeking in vitro alternatives for eye irritancy tests

1. Avoid whole-animal and organ in vivo evaluation
2. Strict Draize scale testing in the rabbit assesses only three eye structures (conjunctiva, cornea, and iris), and traditional rabbit eye irritancy tests do not assess cataracts, pain, discomfort, or clouding of the lens
3. In vivo tests assess only inflammation and immediate structural alterations produced by irritants (not sensitizers, photoirritants, or photoallergens). Note, however, that the test was (and generally is) intended to evaluate any pain or discomfort
4. Technician training and monitoring are critical (particularly in view of the subjective nature of evaluation)
5. Rabbit eye tests do not perfectly predict results in humans, if the objective is either the total exclusion of irritants or the identification of truly severe irritants on an absolute basis (i.e., without false positives or negatives). Some (such as Reinhardt et al. 1985) have claimed that these tests are too sensitive for such uses
6. There are structural and biochemical differences between rabbit and human eyes which make extrapolation from one to the other difficult. For example, Bowman's membrane is present and well developed in man (8–12 μm thick) but not in the rabbit, possibly giving the cornea greater protection
7. Lack of standardization
8. Variable correlation with human results
9. Large biological variability between experimental units
10. Large, diverse, and fragmented databases which are not readily comparable

cells or tissue physiology, inflammation or immune modulation, and alterations in repair and/or recovery processes. These methods have the limitation that they assume that one of the component parts can or will predict effects in the complete organ system. While each component may serve well to predict the effects of a set of chemical structures which determine part of the ocular irritation response, a valid assessment across a broad range of structures will require the use of a collection or battery of such tests.

The sixth category contains tests that have little or no empirical basis, such as computer-assisted structure-activity relationship models. These approaches can only be assessed in terms of how well or poorly they perform. Table 3.15 presents an overview of all six categories and some of the component tests within them, updated from the assessment by Frazier et al. (1987b), along with references for each test.

3.8 Other Nonparenteral Route Irritation Tests

Mucosal irritation may be evaluated by a number of tests; each of them has serious limitations. In the cheek pouch mucosal test, intact samples or sample extracts are inserted into the cheek pouches of Chinese hamsters. In the vaginal mucosal tests, sample extracts are injected into the vagina of albino rabbits. Rabbits in estrous may give false-positive results. In the penile mucosal tests, sample extracts are dripped

Table 3.15 In vitro alternatives for eye irritation tests

Morphology
Enucleated superfused rabbit eye system (Burton et al. 1981)
Balb/c 3 T3 cells/morphological assays (HTS) (Borenfreund and Puerner 1984)
Cell toxicity
Adhesion/cell proliferation
BHK cells/growth inhibition (Reinhardt et al. 1985)
BHK cells/colony formation efficiency (Reinhardt et al. 1985)
BHK cells/detachment (Reinhardt et al. 1985)
SIRC cells/colony-forming assay (North-Root et al. 1982)
Balb/c 3 T3 cells/total protein (Shopsis and Eng 1985)
BCL/D1 cells/total protein (Balls and Horner 1985)
Primary rabbit corneal cells/colony0forming assay (Watanabe et al. 1988)
Membrane integrity
LS cells/dual dye staining (Scaife 1982)
Thymocytes/dual fluorescent dye staining (Aeschbacher et al. 1986)
LS cells/dual dye staining (Kemp et al. 1983)
RCE-SIRC-P815-YAC-1/Cr release (Shadduck et al. 1985)
L929 cells/cell viability (Simons 1981)
Bovine red blood cell/hemolysis (Shadduck et al. 1985)
Mouse L929/fibroblasts/erythrocin C staining (Frzaier 1988)
Rabbit corneal epithelial and endothelial cells/membrane leakage (Meyer and McCulley 1988)
Agarose diffusion (Barnard 1989)
Corneal protein profiles (Eurell and Meachum 1994)
Cell metabolism
Rabbit corneal cell cultures/plasminogen activator (Chan 1985)
LS cells/ATP assay (Kemp et al. 1985)
Balb/c 3 T3 cells/neutral red uptake (Borenfreund and Puerner 1984)
Balb/c 3 T3 cells/uridine uptake inhibition assay (Shopsis and Sathe 1984)
HeLa cells/metabolic inhibition test (MIT-24) (selling and Ekwall 1985)
MDCK cells/dye diffusion (Tchao 1988)
Cell and tissue physiology
Epidermal slice/electrical conductivity (Oliver and Pemberton 1985)
Rabbit ileum/contraction inhibition (Muir et al. 1983)
Bovine cornea/corneal opacity (Muir 1984)
Proposed mouse eye/permeability test (Maurice and Singh 1986)
Inflammation/immunity
Chlorioallantonic membrane (CAM)
CAM (Leighton et al. 1983)
HET-CAM (Luepke 1985)
Bovine corneal cup model/leukocyte chemotactic factors (Elgebaly et al. 1985)
Rat peritoneal cells/histamine release (Jacaruso et al. 1985)
Rat peritoneal mast cells/serotonin release (Dubin et al. 1984)

(continued)

Table 3.15 (continued)

Rat vaginal explant/prostaglandin release (Dubin et al. 1984)
Bovine eye cup/histamine (Hm) and leukotriene C4 (Lt/C4) release (Benassi et al. 1986)
Recovery/repair
Rabbit corneal epithelial cells/wound healing (Jumblatt and Neufeld 1985)
Others
EYTEX assay (Gordon et al. 1986; Soto et al. 1988)
Computer-based structure-activity relationship (SAR) (Enslein 1984)
Tetrahymena motility (Silverman 1983)

onto the expressed penises of albino rabbits. Most of the sample is removed when the penis is withdrawn into the body. The oral mucosa and rectal mucosa may also be evaluated. Methods for these tests are set forth in ISO 10993 Part 10.

The design of vaginal, rectal, penile, and nasal irritation studies is less formalized, but follows the same basic pattern as the primary dermal irritation test. The rabbit is the preferred species for vaginal and rectal irritation studies, but the monkey and dog have also been used for these (Chvapil 1979; Eckstein et al. 1969; Lilly et al. 1972; Lindhe et al. 1970; Muller et al. 1988; Nixon et al. 1972; Bernstein and Carlish 1979; Kaminsky and Willigan 1982; Davidson et al. 1982; Haugen 1980). Both the rabbit and rat have commonly seen use for nasal irritation evaluations. Defined quantities (typically 1.0 ml) of test solutions or suspensions are instilled into the orifice in question. For the vagina or rectum, inert bungs are usually installed immediately thereafter to continue exposure for a defined period of time (usually the same period of hours as future human exposure). The orifice is then flushed clean, and 24 h after exposure, it is examined and evaluated (graded) for irritation using the scale in Table 3.8.

3.8.1 Pyrogenicity

Pyrogenicity is the induction of a febrile (fever) response by the parenteral (usually IV or IM) administration of exogenous material, usually (but not always) bacterial endotoxins. Pyrogenicity is usually associated with microbiological contamination of a final formulation or product but is now increasingly of concern because of the increase in interest in biosynthetically produced materials. Generally, ensuring the sterility of product and process will guard against pyrogenicity. Pyrogenicity testing is performed extensively in the medical device industry. If a device is to be introduced directly or indirectly into the fluid path, it is required that it be evaluated for pyrogenic potential (USP 2013; European Pharmacopoeia 1990).

The USP Pyrogen Test suing rabbit or the limulus amebocyte lysate (LAL) test can be used to support pyrogen-free claims. If the LAL test is used, the LAL test method must either meet the FDA's documented titled *Guideline on Validation of*

the Limulus Amebocyte Lysate Test as an End-Product Endotoxin Test for Human and Animal Parental Drugs, Biological Products, and Medical Devices, or a 510(k) or PMA application must be submitted for the LAL test.

The bacterial endotoxin limit for medical devices is 0.5 EU/ml. Manufacturers may retest LAL test failures with another LAL test or the USP rabbit pyrogen test. Medical devices that contact cerebrospinal fluid should have less than 0.06 EU/ml of endotoxin.

In vitro pyrogenicity testing (or bacterial endotoxin testing) is one of the great success stories for in vitro testing. Some 30 years ago, the limulus amebocyte lysate (LAL) test was developed, validated, and accepted as an in vitro test for estimating the concentration of bacterial endotoxins that may be present in or on a sample of the article(s) to which the test is applied using LAL that has been obtained from aqueous extracts of the circulating amebocytes of the horseshoe crab, *Limulus polyphemus*, and that has been prepared and characterized for use as an LAL reagent for gel-clot formation (Cooper 1975; Weary and Baker 1977). The test's limitation is that it detects only the pyrogens of gram-negative bacteria. This is generally not significant (at least for use in lot release assays) since most environmental contaminants that gain entrance to sterile products are gram-negative (Bulich et al. 1990; Devleeschouwer et al. 1985).

Where the test is conducted as a limit test, the specimen is determined to be positive or negative to the test judged against the endotoxin concentration specified in the individual monograph (USP 2007). Where the test is conducted as an assay of the concentration of endotoxin, with calculation of confidence limits of the result obtained, the specimen is judged to comply with the requirements if the result does not exceed (1) the concentration limit specified in the individual monograph and (2) the specified confidence limits for the assay. In either case the determination of the reaction end point is made with parallel dilutions of redefined endotoxin units.

Since LAL reagents have also been formulated to be used for turbidimetric (including kinetic) assays or colorimetric readings, such tests may be used if shown to comply with the requirements for alternative methods. These tests require the establishment of a standard regression curve, and the endotoxin content of the test material is determined by interpolation from the curve. The procedures include incubation for a preselected time of reacting endotoxin and control solutions with LAL reagent and reading the spectrophotometric light absorbance at suitable wavelengths. In the case of the turbidimetric procedure, the reading is made immediately at the end of the incubation period. In the kinetic assays, the absorbance is measured throughout the reaction period, and rate values are determined from those readings. In the colorimetric procedure, the reaction is arrested at the end of the preselected time by the addition of an appropriate amount of acetic acid solution prior to the readings. A possible advantage in the mathematical treatment of results, if the test is otherwise validated and the assay suitable designed, could be the confidence interval and limits of potency from the internal evidence of each assay itself.

3.8.2 Reference Standard and Control Standard Endotoxins

The reference standard endotoxin (RSE) is the USP Endotoxin Reference Standard, which has a defined potency of 10,000 USP endotoxin units (EU) per vial. Constitute the entire contents of one vial of the RSE with 5 ml of LAL reagent water, vortex for not less than 20 min., and use this concentrate for making appropriate serial dilutions. Preserve the concentrate in a refrigerator, for making subsequent dilutions, for not more than 14 days. Allow it to reach room temperature, if applicable, and vortex it vigorously for not less than 5 min. Before use. Vortex each dilution for not less than 1 min. Before proceeding to make the next dilution. Do not use stored dilutions. A control standard endotoxin (CSE) is an endotoxin preparation other than the RSE that has been standardized against the RSE. If a CSE is a preparation not already adequately characterized, its evaluation should include characterizing parameters both for endotoxin quality and performance (such as reaction in the rabbit) and for suitability of the material to serve as a reference (such as uniformity and stability). Detailed procedures for its weighing and/or constitution and use to ensure consistency in performance should also be included. Standardization of CSE against the RSE using an LAL reagent for the gel-clot procedure may be effected by assaying a minimum of four vials of the CSE or four corresponding aliquots, where applicable, of the bulk CSE and one vial of the RSE as directed under Test Procedure, but using four replicate reaction tubes at each level of the dilution series for the RSE and four replicate reaction tubes similarly for each vial or aliquot of the CSE. If the dilutions for the four vials or aliquots of the CSE cannot all be accommodated with the dilutions for the one vial of the RSE on the same rack for incubation, additional racks may be used for accommodating some of the replicate dilutions for the CSE, but all of the racks containing the dilutions of the RSE and CSE are incubated as a block. However, in such cases, the replicate dilution series from the one vial of the RSE are accommodated together on a single rack and the replicate dilution series from any one of the four vials or aliquots of the CSE are not divided between racks. The antilog of the difference between the mean log 10 end point of the RSE and the mean log 10 end point of the CSE is the standardized potency of the CSE, which is then converted to and expressed in units/ng under stated drying conditions for the CSE, or units per container, whichever is appropriate. Standardize each new lot of CSE prior to use in the test. Calibration of a CSE in terms of the RSE must be with the specific lot of LAL reagent and the test procedure with which it is to be used. Subsequent lots of LAL reagent from the same source and with similar characteristics need only checking of the potency ratio. The inclusion of one or more dilution series made from the RSE when the CSE is used for testing will enable observation of whether or not the relative potency shown by the latter remains within the determined confidence limits. A large lot of a CSE may, however, be characterized by a collaborative assay of a suitable design to provide a representative relative potency and the within-laboratory and between-laboratory variance.

A suitable CSE has a potency of not less than 2 EU/ng and not more than 50 EN/ng, where in bulk form, under adopted uniform drying conditions, e.g., to a particular low moisture content and other specified conditions of use, and a potency within a corresponding range where filled in vials of a homogeneous lot.

3.8.3 Preparatory Testing

Use an LAL agent of confirmed label or determined sensitivity. In addition, where there is to be a change in lot of CSE, LAL reagent, or another reagent, conducts tests of a prior satisfactory lot of CSE, LAL, and/or other reagents in parallel on changeover. Treat any containers or utensils employed so as to destroy extraneous surface endotoxins that may be present, such as by heating in an oven at 250 ° F or above for sufficient time.

The validity of test results for bacterial endotoxins requires an adequate demonstration that specimens of the article, or of solutions, washings, or extracts thereof to which the test is to be applied, do not of themselves inhibit or enhance the reaction or otherwise interfere with the test. Validation is accomplished by testing untreated specimens or appropriate dilutions thereof, concomitantly with and without known and demonstrable added amounts of RSE or a CSE, and comparing the results obtained. Appropriate negative controls are included. Validation must be repeated if the LAL reagent source or the method of manufacture or formulations of the article is changed.

Test for confirmation of labeled LAL reagent sensitivity: Confirm the labeled sensitivity of the particular LAL reagent with the RSE (or CSE) using not less than four replicate vials, under conditions shown to achieve an acceptable variability of the test, viz., the antilog of the geometric mean log 10 lystate gel-clot sensitivity is within 0.5 to 2.0, where the labeled sensitivity is in EU/ml. The RSE (or CSE) concentrations selected to confirm to LAL reagent label potency should bracket the stated sensitivity of the LAL reagent. Confirm the labeled sensitivity of each new lot of LAL reagent prior to use in the test.

3.8.4 Inhibitions or Enhancement Test

Conduct assays, with standard endotoxin, or untreated specimens in which there is no endogenous endotoxin detectable, and of the same specimens to which endotoxin has been added, as directed under Test Procedures, but use not less than four replicate reaction tubes at each level of the dilution series for each untreated specimen and for each specimen to which endotoxin has been added. Record the end points (E, in units/ml) observed in the replicates. Take the logarithms (e) of the end points, and compute the geometric means of the log end points for the RSE (or CSE) for the untreated specimens and for specimens containing endotoxin by the formula antilog,

elf, where *e* is the sum of the log end points of the dilution series used and *f* is the number of replicate end points in each case. Compute the amount of endotoxin in the specimen to which endotoxin has been added. The test is valid for the article if this result is within twofold of the known added amount of endotoxin. Alternatively, if the test has been appropriately set up, the test is valid for the article if the geometric mean end point dilution for the specimen to which endotoxin has been added is within one twofold dilution of the corresponding geometric mean end point dilution of the standard endotoxin.

Repeat the test for inhibition or enhancement using specimens diluted by a factor not exceeding that given by the formula *x/y* (see Maximum Valid Dilution). Use the least dilution sufficient to overcome the inhibition or enhancement of the known endotoxin for subsequent assays of endotoxin in test specimens.

If endogenous endotoxin is detectable in the untreated specimens under the conditions of the test, the article is unsuitable for the inhibition or enhancement test, or it may be rendered suitable by removing the endotoxin present by ultrafiltration or by appropriate dilution. Dilute the untreated specimen (as constituted, where applicable, for administration or use) to a level not exceeding the maximum valid dilution, at which no endotoxin is detectable. Repeat the test for inhibition or enhancement using the specimens at those dilutions.

3.8.4.1 Test Procedure

In preparing for and applying the test, observe precautions in handling the specimens in order to avoid gross microbial contamination. Washings or rinsings of devices must be with LAL reagent water in volumes appropriate to their use and, where applicable, of the surface area which comes into contact with body tissues or fluids. Use such washings or rinsings if the extracting fluid has been in contact with the relevant pathway or surface for not less than 1 hr. at controlled room temperature (15–30 °C). Such extracts may be combined, where appropriate.

For validating the test for an article, for endotoxin limit tests or assays, or for special purposes where so specified, testing of specimens is conducted quantitatively to determine response end points for gel-clot readings. Usually graded strengths of the specimen and standard endotoxin are made by multifold dilutions. Select dilutions so that they correspond to an geometric series in which each step is greater than the next lower by a constant ratio. Do not store diluted endotoxin, because of loss of activity by absorption. In the absence of supporting data to the contrary, negative and positive controls are incorporated into the test.

Use not less than two replicate reactions tubes at each level of the dilution series for each specimen under test. Whether the test is employed as a limit test or as a quantitative assay, a standard endotoxin dilution series involving not less than two replicate reaction tubes is conducted in parallel. A set of standard endotoxin dilution series is included for each block of tubes, which may consist of a number of racks for incubation together, provided the environmental conditions within blocks are uniform.

3.8.4.2 Preparation

Since the form and amount per container of standard endotoxin and of LAL reagent may vary, constitution and/or dilution of contents should be as directed in the labeling. The pH of the test mixture of the specimen and the LAL reagent is in the range of 6.0–7.5 unless specifically directed otherwise in the individual monograph. The pH may be adjusted by the addition of sterile, endotoxin-free sodium hydroxide or hydrochloric acid or suitable buffers to the specimen prior to testing.

3.8.4.3 Maximum Valid Dilution

The maximum valid dilution (MVD) is appropriate to injections or to solutions for parenteral administration in the form constituted or diluted for administration, or, where applicable, to the amount of drug by weight if the volume of the dosage form for administration could be varied. Where the endotoxin limit concentration is specified in the individual monograph in terms of volume (in EU/ml), divide the limit by γ, which is the labeled sensitivity (in EU/ml) of the lysate employed in the assay, to obtain the MVD factor. Where the endotoxin limit concentration is specified in the individual monograph in terms of weight or of units of active drug (in EU/mg or in EU/unit), multiply the limit by the concentration (in mg/ml or in units/ml of the drug in the solution tested or of the drug constituted according to the label instructions, whichever is applicable), and divide the product of the multiplication by γ to obtain the MVD factor. The MVD factor so obtained is the limit dilution factor for the preparation for the test to be valid.

3.8.4.4 Procedure

To 10 × 75-mm test tubes, add aliquots of the appropriately constituted LAL reagent and the specified volumes of specimens, endotoxin standard, negative controls, and a positive product control consisting of the article, or of solutions, washings, or extracts thereof, to which the RSE (or a standardized CSE) has been added at a concentration of endotoxin of 2 for LAL reagent (see under Test for confirmation of labeled LAL reagent sensitivity). Swirl each gently to mix and place in an incubating device such as water bath or heating block, accurately recording the time at which the tubes are so placed. Incubate each tube, undisturbed, for 60 ± 2 min at 37 ± 1 °C, and carefully remove it for observation. A positive reaction is characterized by the formation of a firm gel that remains when inverted through 180 degrees. Record such a result as a positive (+). A negative result is characterized by the absence of such a gel or by the formation of a viscous gel that does not maintain its integrity. Record such a result as a negative (−). Handle the tubes with care, and avoid subjecting them to unwanted vibrations, or false-negative observations may result. The test is invalid if the positive product control or the endotoxin standard

does not show the end point concentration to be within ± twofold dilutions from the label claim sensitivity of the LAL reagent or if any negative control shows a gel-clot end point.

3.8.4.5 Calculation and Interpretation

Calculate the concentration of endotoxin (in units/ml or in units/g or mg) in or on the article under test by the formula *pS/U*, where *S* is the antilog of the geometric mean log 10 of the end points, expressed in EU/ml for the standard endotoxin; *U* is the antilog of *elf*, where *e* is the log 10 of the end point dilution factors, expressed in decimal fractions, and *f* is the number of replicate reaction tubes read at the end point level for the specimen under test; and *p* is the correction factor for those cases where a specimen of the article cannot be taken directly into test but is processed as an extract, solution, or washing.

Where the test is conducted as an assay with sufficient replication to provide a suitable number of independent results, calculate for each replicate assay the concentration of endotoxin in or on the article under test from the antilog of the geometric mean log end point ratios. Calculate the mean and the confidence limits from the replicate logarithmic values of all the obtained assay results by a suitable statistical method.

3.8.4.6 Interpretation

The article meets the requirements of the test if the concentration of endotoxin does not exceed that specified in the individual monograph, and the confidence limits of the assay do not exceed those specified.

3.8.5 Material-Mediated Pyrogenicity (USP Rabbit Pyrogen Test)

The *United States Pharmacopeia* describes a pyrogen test using rabbits as a model. This test, which is the standard for limiting risks of a febrile reaction to an acceptable level, involves measuring the rise in body temperature in a group of three rabbits for 3 h after injection of 10 ml of test solution.

3.8.5.1 Apparatus and Diluents

Render the syringes, needles, and glassware free of pyrogens by heating at 250 °F for not less than 30 min or by any other suitable method. Treat all diluents and solutions by washing and rinsing of devices or parenteral injection assemblies in a man-

ner that will ensure that they are sterile and pyrogen-free. Periodically perform control pyrogen tests on representative portions of the diluents and solutions that are used for washing and rinsing of the apparatus.

3.8.5.2 Temperature Recording

Use an accurate temperature-sensing device, such as a clinical thermometer or thermistor or similar probe, that has been calibrated to ensure an accuracy of ±0.1° and has been tested to determine that a maximum reading is reached in less than 5 min. Insert the temperature-sensing probe into the rectum of the test rabbit to a depth of not less than 7.5 cm, and, after a period of time not less than that previously determined as sufficient, record the rabbit's temperature.

3.8.5.3 Test Animals

Use healthy, mature rabbits. House the rabbits individually in an area of uniform temperature (between 20 °C and 23 °C) free from disturbances likely to excite them. The temperature should vary no more than ±3 °C from the selected temperature. Before using a rabbit for the first time in a pyrogen test, condition it for not more than 7 days before use by a sham test that includes all of the steps as directed under Procedure, except injection. Do not use a rabbit for pyrogen testing more frequently than once every 48 h, nor prior to 2 weeks following a maximum rise in its temperature of 0.6° or more while being subjected to the pyrogen test, or following its having been given a test specimen that was found to be pyrogenic.

3.8.5.4 Procedure

Perform the test in a separate area designated solely for pyrogen testing and under environmental conditions similar to those under which the animals are housed. Withhold all food from the test rabbits during the period of the test. Access to water is allowed at all times, but may be restricted during the test. A probe measuring rectal temperature remain inserted throughout the testing period, restrain the rabbits with loose-fitting Elizabethan collars that allow the rabbits to assume a natural resting posture. Not more than 30 min prior to the injection of the test dose, determine the "control temperature" of each rabbit; this is the base for the determination of any temperature increase resulting from the injection of a test solution. In any one group of test rabbits, use only those rabbits whose control temperatures do not vary by more than 1 °C from each other and do not use any rabbit having a temperature exceeding 39.8 °C.

Unless otherwise specified in the individual protocol, inject 10 ml of the test solution per kilogram of body weight into an ear vein of each of three rabbits, completing each injection within 10 min after the start of administration. The test solu-

tion is either the product, constituted if necessary as directed in the labeling, or the material under test. For pyrogen testing of devices or injection assemblies, use washings or rinsings of the surfaces that come in contact with the parenterally administered material or with the injection site or internal tissues of the patient. Ensure that all test solutions are protected from contamination. Perform the injection after warming the test solution to a temperature of 37 °C ± 2°. Record the temperature at 1, 2, and 3 h subsequent to the injection.

3.8.5.5 Test Interpretation and Continuation

Consider any temperature decreases as zero rise. If no rabbit shows an individual rise in temperature of 0.6° or more above its respective control temperature, and if the sum of the three individual maximum temperature rises does not exceed 1.4°, the product meets the requirements for the absence of pyrogens. If any rabbit shows an individual temperature rise of 0.6° or more, or if the sum of the three individual maximum temperature rises exceeds 1.4°, continue the test using five other rabbits. If not more than three of the eight rabbits show individual rises in temperature of 0.6° or more, and if the sum of the eight individual maximum temperature rises does not exceed 3.7°, the material under examination meets the requirements for the absence of pyrogens.

3.9 Acute Systemic Toxicity Testing (ISO 10993-11)

3.9.1 Introduction

Acute toxicity testing is the defining and evaluation of the toxic syndrome (if any) produced by a single dosage of extracts from a device or biomaterials. In the initial testing program promulgated for “medical closures” (and therefore medical devices), this was meant to serve the purpose of screening for rapidly toxic constituents in a device or contaminants on it. It sought (and seeks) to have increased sensitivity by using a broader range of extraction vehicles and using dose levels (in terms of high volumes of extraction fluids—up to 50 ml/kg). The current version of the test is a holdover from these earlier days.

Historically, the main focus of these tests has been lethality determinations and the identification of overt signs and symptoms of overdosage. For a complete historical perspective, see Rhodes (2000), Gad and Chengelis (1999), Auletta (1998), or Piegorsh (1989). A more enlightened and modern view holds that, especially for pharmaceutical agents, lethality in animals is a relatively poor predictor of hazard in man (Gad and Chengelis 1999). The current trend is toward gaining increasing amounts of more sophisticated data from these tests, such as in the expanded acute studies done to enable so called “Phase 0” clinical trials. The various types of acute study designs, their utility in pharmaceutical product testing, and the resultant sample data are discussed in this chapter.

In the pharmaceutical industry, acute toxicity testing has uses other than for product safety determinations. First, as in other industries, acute toxicity determinations are part of industrial hygiene or occupational health environmental impact assessments. These requirements demand testing not only for finished products but frequently of intermediates as well.

For medical devices, acute systemic toxicity typically means the ISO-10993 or pharmacopeial (such as USP and other pharmacopeias) "acute systemic toxicity" testing is actually lethality testing using one or more extractants ("eluants") from the material of interest. It is designed to screen broadly for the presence of any extractable materials which may be lethal to the intact organism. The "model" employed is the albino mouse (17–23 grams in weight). USP section <88> provides guidance on preparation of materials and test conduct. The following is an example protocol:

Extractions media area:

- Sodium chloride 0.9% (physiological saline solution).
- Polyethylene glycol 400.
- 5% ethanol in 0.9% aqueous saline
- Vegetable (cottonseed or sesame) oil.
- If not covered by any of the above, device product vehicle should be employed, as seen in Table 3.1.

3.9.2 *Acute Systemic Toxicity Characterization*

The acute systemic toxicity study as performed for medical devices is a true screen—the parenteral routes of administration and the number and volumes of solvents used are all intended to maximize the sensitivity of a test system which truly has only a limited set of parameters of interest, primarily lethality (Fig. 3.5; Table 3.16).

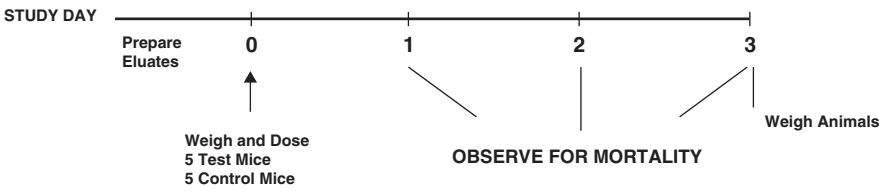


Fig. 3.5 USP acute systemic toxicity test- live chart

Table 3.16 USP acute systemic toxicity test

Extract or blank	Dose (/kg)	Route	Injection rate (μL/second)
SAL	50 mL	IV	100
5% ETOH	50 mL	IV	100
PEG 400	10 g	IP	NA
CSO	50 mL	IP	NA

3.9.2.1 Body Weight Considerations

Body weight and feed consumption are frequently determined parameters in toxicity testing. To an extent, the ability of an animal to gain or maintain weight may be considered a sensitive, but nonspecific, indicator of health. While this is true in subchronic or chronic studies, its relevance in acute studies must be carefully considered. In most protocols, body weights are determined on Day 1 (prior to dosing), Day 7, and Day 14, which are the days mandated by most regulatory guidelines. Despite being common, the design is not well founded: if an animal has not died within 7 days postdosing, it has probably recovered, and its body weight may not be noticeably different from controls by Day 14. A complete protocol addresses this problem by specifying more frequent body weight determinations (daily for the first 3 to 5 days of the observation period) so that not only can initial decreases (if they occur) be detected, but recovery can also be charted. Feed consumption measurements should be made at the same times, because it is difficult to determine the causes behind body weight changes in the absence of feed consumption data. Body weight loss accompanied by normal feed consumption implies something very different than body weight loss (or lack of gain) accompanied by lack of feed consumption. In the absence of feed consumption data, however, changes in body weight should still be considered indicative of a change in an animal's health status.

Yet another reason why body weight determinations are of questionable value in acute studies has to do with the statistical analysis of the data. Deaths may substantially alter group size and complicate analysis. The death of two of five animals causes a 40% decrease in group size and a substantial diminution of the power of any statistical test. In addition, the resulting data sets are censored: comparisons will often be between the control group, a dosage group where all the animals survive, and a high-dosage group where less than 50% of the animals survive to the end of the observation period. One has to question the utility of body weight changes if they occur at dosages that are acutely lethal. The data in Table 3.17 illustrate this point. Body weight changes tended to occur only at dosages that were acutely lethal. Additionally, one would suspect that the censoring of body weights in groups where death occurs is not random; that is, the animals that die are most likely those that are most sensitive, while those that survive are the most resistant or robust. This problem can be addressed by building exclusionary criteria into a protocol. For example, one could statistically analyze body weight data in groups that only had less than 50% mortality.

3.9.2.2 Factors that Can Affect Acute Tests

Many investigations into the sources of variability in acute toxicity testing have been conducted, and these have been reviewed by Elsberry (1986). The factors causing the greatest interstudy variation included lack of specifications for sex, strain, age, and weight range. When clearly defined, detailed protocols were used, interlaboratory

Table 3.17 Information, including lethality, that can be gained in acute toxicity testing

Lethality/mortality
LD_{50} with confidence limits
Shape and slope of lethality curves
Estimation of maximum nonlethal dose or minimum lethal dose (LD_{01})
Time to dose estimates
Clinical signs
Times of onset and recovery
Thresholds
Agonal versus nonagonal (i.e., do signs occur only in animals that die?)
Specific versus general responses
Separation of dose-response curves from lethality curves
Body weight changes
Actual loss versus decreased gain
Recovery
Accompanied by changes in feed consumption
Changes in animals that die versus those that survive
Target organ identification
Gross examinations
Histological examinations
Clinical chemical changes
Hematological changes
Specialized function test
Immunocompetency
Neuromuscular screening
Behavioral screening
Pharmacokinetic considerations
Different routes of administration yielding difference in toxicity
Plasma levels of test article
Areas under the curves, volume of distribution, half-life
Metabolic pattern of test article
Distribution to key organs
Relationship between plasma levels and occurrence of clinical signs

variation was found to be minimal. Hence, it is equally important that the details of the protocol be well described and followed. It is not appropriate to draw dosage-response conclusion by comparing groups that differ substantially in age or that have been fed, fasted, or otherwise manipulated differently.

3.10 Implantation Studies and Implantation Biology (ISO 10993-6)

Implantation studies are a type of assay unique to medical device and biomaterials, having been specifically devised for those situations where an exogenous (and usually man-made) construct or material is enclosed in the body or partially entered into it by a breached surface. It is intended to assess the effects of devices (usually polymers or elastomers) which are in direct contact with living tissue (not including the skin). The effects of concern may be either short or long term, with a range of responses over the course of the host body and the device interacting with each other. Longer-term studies are conducted for long-term implants and focus more on broader systemic effects and potential carcinogenicity. The whole field of implantation biology (which studies these interactions) is an extremely complete and active one (Greco 1994; Black 2000; Grethcer and Hollinger 2006).

3.10.1 ISO 10993 Implantation Test

The ISO 10993 mandated test is covered by Part 6 of the test guidelines and is specifically intended to test for local effects after implantation. It can be performed for either short term (from 1 up to 12 weeks) or long term (from 12 to 104 weeks).

The test specimen is implanted into a site and tissue appropriate for evaluation of the biological safety of the material. The implant is not intended to be subjected to mechanical or functional loading. The local effects are evaluated by a comparison of the tissue response caused by a test specimen to that caused by materials used in medical devices whose clinical acceptability has been established.

3.10.1.1 Preparation of Specimens for Implantation

Solid Specimens (Excluding Powders)

Physical characteristics (i.e., form, density, hardness, surface finish) can influence the character of the tissue response to the test material. Each implant shall be manufactured, processed, cleaned of contaminants, and sterilized by the method intended for the final product. After final preparation and sterilization, the implant specimens shall be handled in such a way as to ensure that they are not scratched, damaged, or contaminated in any way prior to or during insertion.

Nonsolid Specimens (Including Powders)

Nonsolid specimens may be liquids, pastes, and particulates, as distinct from the materials covered otherwise. The components may be mixed before use (e.g., bone cements, dental materials) and set after varying time periods. The materials may be contained in tubes for the purpose of testing for local effects after implantation. Polyethylene (PE), polypropylene (PP), or polytetrafluoroethylene (PTFE) tubes are commonly used for this purpose.

Prior to test the tubes shall be rinsed with 70% (V/V) ethanol and distilled water and sterilized by autoclaving or other appropriate methods relevant for clinical applications. Materials tested in their freshly mixed state shall be tested for microbiological contamination.

Prepare the test material according to the manufacturer's instructions and insert the material into the tube until level with the top. Exercise the utmost care to prevent contamination of the outer surface of the tube by the test material. Avoid entrapment of air in the tube and ensure that the end surfaces of the inserted material in the tube and the tube ends are smooth.

Note—PE tubes may be deformed by autoclaving. It is difficult to section PTFE tubes in the microtome, and substitution by PE or PP tubes of the same dimensions may be preferable when the tubes are to remain in the tissue blocks during sectioning.

Control Specimens

The size, shape, and especially the surface condition of the control(s) shall be as similar to that of the implant test specimens as is practically possible. When the test material is contained in a tube, the control shall be a rod of the same material as the tube and with the same diameter as the outer diameter of the tube. The control specimens shall be handled, cleaned, and sterilized in such a manner as to maintain them as acceptable and well-characterized controls.

Selection of control material(s) should be based on their established use in clinical applications similar to those proposed for the candidate test material and is not otherwise restricted.

Animals and Tissues

Select an animal species with due consideration of the size of the implant test specimens, the intended duration of the test in relation to the expected life-span of the animals, as well as the recognized species differences in biological response in both hard and soft tissues. For short-term testing in subcutaneous tissue and muscle, animals such as mice, rats, guinea pigs, and rabbits are commonly used (Gad 2006). Select one species among these. For long-term testing in subcutaneous tissue, muscle, and bone, animals such as rats, guinea pigs, rabbits, dogs, sheep, goats, pigs,

and other animals with a relatively long life expectancy are suitable. Select one species among these.

The specimens of test and control materials shall be implanted under the same conditions in the same species of the same age, sex, and strain in corresponding anatomical sites. The number and size of implants inserted in an animal depends on the size of the species and the anatomical location of the implantation.

3.10.1.2 Test Periods

The local tissue response to implanted materials is assessed in short-term tests up to 12 weeks and in long-term tests exceeding 12 weeks.

Test periods are chosen to ascertain that a steady state has been reached with respect to biological response. The local biological response to implanted materials depends both on the properties of the materials and on the trauma of surgery. The tissue configuration found in the vicinity of an implant changes with the time elapsed after surgery. Usually, at 1-week observation periods, a high cell activity is found, followed by a transitional stage. In muscle and connective tissue, depending on the species, a steady state is seen in the cell population after 9–12 weeks. Implantation in bone tissues may need longer observation periods.

Test periods shall be selected from those specified in Table 3.18 for short-term implantation or from Table 3.19 for long-term implantation.

Table 3.18 Selection of test periods for short-term implantation in subcutaneous tissue and muscle

	Implantation period (weeks)					
Species	1	2	3	4	9	12
Mice	X		X		X	
Rats	X	X		X		X
Guinea pigs	X	X		X		X
Rabbits	X	X		X		X

Table 3.19 Selection of test periods for long-term implantation in subcutaneous tissue, muscle, and bond

	Implantation period (weeks)				
Species	12	26	52	78	104
Rats	X	X	X		
Guinea pigs	X	X	X		
Rabbits	X	X	X	X	
Dogs	X	X	X	X	X
Sheep	X	X	X	X	X
Goats	X	X	X	X	X
Pigs	X	X	X	X	X

Depending on the intended use of the test material, not all implantation periods may be necessary (see ISO 10993-1). An observation period of 104 weeks may be of interest in selected instances. The number of implants per animal and the number of animals per observation period are described in the appropriate sections below. A sufficient number of implants shall be inserted to ensure that the final number of specimens to be evaluated will give valid results.

3.10.1.3 Surgery

Anesthetize the animals. Remove hair from the surgical area by clipping, shaving, or other mechanical means. Wash the area with an antiseptic solution. Ensure that hair does not come in contact with the implants or the wound surfaces.

The surgical technique may profoundly influence the result of any implantation procedure. The surgery shall be carried out under aseptic conditions and in a manner that minimizes trauma at the implant site. After surgery close the wound, using either wound clips or sutures, taking precautions to maintain aseptic conditions.

3.10.1.4 Postoperative Assessment

Observe each animal at appropriate intervals during the test period and record any abnormal findings, including local, systemic, and behavioral abnormalities.

3.10.1.5 Euthanasia

At the termination of the experimental period, euthanize the animals with an overdose of anesthetic or by some other acceptable humane method (see ISO 10993-2).

3.10.1.6 Evaluation of Biological Response

Evaluate the biological response by grading and documenting the macroscopic and histopathological test responses as a function of time. Compare the responses to the test material and control material.

Carry out comparison of the control and the test implants at equivalent locations relative to each implant so that the effect of relative motion between tissue and implant is at a minimum. For a cylindrical specimen, the region is midway between ends. With grooved cylindrical implants, the center portions between the grooves as well as the flat top end surfaces of the implant are suitable for evaluation.

For a nonsolid or particulate material incorporated into a tube, the area at the end of the tube is the only available area for evaluation.

3.10.1.7 Macroscopic Assessment

Examine each implant site with the aid of a low magnification lens. Record the nature and extent of any tissue reaction observed.

3.10.1.8 Preparation for Histology: Implant Retrieval and Specimen Preparation

Excise the implant together with sufficient unaffected surrounding tissue to enable evaluation of the local biological response. Process the excised tissue blocks containing test or control implants for histopathological and other studies as appropriate.

When conventional techniques are used, the tissue envelope may be opened before or after exposure to a fixative and the condition of the implant surface and tissue bed shall be reported. However, with this technique the tissue layers closest to the implant are usually destroyed.

When the implant/tissue surface is to be studied, embedding of the intact tissue envelope with the implant in situ using hard plastics is preferred. Appropriate sectioning or grinding techniques are employed for the preparation of histological sections. It shall be demonstrated that the technique of embedding in plastics does not markedly alter the interface tissue.

3.10.1.9 Histological Assessment

The extent of response may be determined by measurement of the distance from the implant/tissue interface to unaffected areas with the characteristics of normal tissue and of normal vascularity. Record the section orientation in relation to the implant dimensions. Record the implant orientation, number of sections, and cutting geometry.

The biological response parameters which shall be assessed and recorded include:

- Extent of fibrosis/fibrous capsule and inflammation.
- Degeneration as determined by changes in tissue morphology.
- Number and distribution as a function of distance from the material/tissue interface of the inflammatory cell types, namely, polymorphonuclear leukocytes, lymphocytes, plasma cells, eosinophils, macrophages, and multinucleated cells.
- Presence of necrosis as determined by nuclear debris and/or capillary wall breakdown.
- Other parameters such as material debris, fatty infiltrations, granuloma.
- For porous implant materials, the quality and quantity of tissue ingrowth.

In the case of bone, the interface between the tissue and the material is of special interest. Evaluate the area of bone contact and the amount of bone in the vicinity of the implant as well as the presence of intervening non-calcified tissues. Note the presence of bone resorption and bone formation.

3.10.1.10 Implant Specimens

Description of test and control materials, material condition, fabrication, surface condition, and the shape and size of implants.

Remember to specify the rationale for selection of control material(s).

The surface preparation of the specimens can affect the tissue reaction. Therefore, the preparation procedure should be noted in the report.

Report cleaning, handling, and sterilization techniques employed. If not done in-house, this information should be supplied by the manufacturer before the investigation commences.

3.10.1.11 Animals and Implantation

Report in origin, age, sex, and strain of animals. Report housing conditions, diet, and mass of animals during the study period. The health of the animals shall be evaluated during the study. All observations, including unexpected death, shall be reported.

Report insertion techniques. Report number of implants inserted per animal, per site, and per observation period.

3.10.1.12 Retrieval and Histological Procedure

The report shall include a description of the retrieval technique. The number of implants retrieved per animal and per observation period shall be recorded. All specimens shall be accounted for and considered as part of the test. The techniques for taking histological sections shall be described.

3.10.1.13 Evaluation

Macroscopic observations shall include the observations made on implant as well as the macroscopic appearance of the tissue surrounding the implant. The report shall include the results obtained from each histological examination.

The report shall include a comparative evaluation of the biological responses to test and control materials, as well as a descriptive narrative of the biological response.

3.10.1.14 Test Method for Implantation in Subcutaneous Tissue

Field of Application

This test material is used for assessing the biological response of subcutaneous tissue to an implanted material.

The study may be used to compare the effect of different surface textures or conditions of the same material or to assess the effect of various treatments or medications or a material.

Principle

Insertion of the implants in the subcutaneous tissue of test animals. The method compares the biological response to implants of test specimens with the biological response to implants of control specimens made of materials which are established in clinical use.

Test Specimens

Common provisions for preparation of test and control specimens as previously described. Implant sizes are based on the size of the test animal.

Specimens made of sheet material shall be 10 mm to 12 mm in diameter and from 0.3 mm to 1 mm in thickness.

Note—The subcutaneous site, deep to the panniculus carnosus muscle, is particularly suitable for the evaluation of polymeric sheet material. In an intramuscular site, sheet material may become folded, which makes it difficult to assess the effect of the material per se.

Bulk materials shall be fabricated into specimens 1.5 mm in diameter and 5 mm in length and have radiused ends.

Grooved specimens shall be 4 mm in diameter and 7 mm in length (see annex B).

Note—Tissue ingrowth into the grooves minimizes tissue irritation caused by interface motion.

Nonsolid specimens (including powders) shall be prepared in tubes 1.5 mm in diameter and 5 mm in length.

Test Animals and Implant Sites

The implants shall be inserted in the dorsal subcutaneous tissue of adult mice, rats, guinea pigs, or rabbits. Select one species among these.

Use at least three animals and sufficient sites to yield 10 specimens for each material and implantation period.

3.10.1.15 Implantation Procedure

Select one of the procedures described below.

Implantation Along Dorsal Midline

Make an incision of the skin and make one or more subcutaneous pockets by blunt dissection. The base of the pocket shall be more than 10 mm from the line of incision. Place one implant in each pocket. The implants shall not be able to touch one another.

Note—Alternatively, the implants may be delivered by a trocar to the desired site.

Implantation in Neck

In mice, make a 10-mm-long incision above the sacrum and prepare a subcutaneous tunnel by blunt dissection toward the neck. Push one implant (for design see annex B) through the tunnel to position it at the neck.

In rats, insert one implant of each of the control and candidate materials separately on each side of the neck. The implants shall not be able to touch one another.

At some distance from the implant, close the tunnel with stitches of appropriate suture material to prevent the implant from moving.

Implantation Period

To ensure a steady state of biological tissue response, the implantation period(s) shall be as specified in Tables 3.18 and 3.19.

3.10.1.16 Evaluation of Biological Response

The evaluation shall take into account the items specified earlier.

3.10.1.17 Test Method for Implantation in Muscle

Field of Application

This test is used for assessing the biological response of muscle tissue to an implanted material.

Principle

Insertion of the implant in the muscle of a test animal. The method compares the biological response to implants of test specimens with the biological response to implants of control specimens made of materials which are established in clinical use.

Test Specimens

Common provisions for preparation of test and control specimens are described earlier. Implant sizes are based on the size of the muscle group chosen.

For rabbit paravertebral muscles, implants of a width of 1 mm to 3 mm with a length of approximately 10 mm shall be used.

The specimens shall have rounded edges and the ends finished to a full radius.

Test Animals and Implant Sites

Insert the implants in the muscle tissue of rabbits or other animals. Ensure that the muscles are of sufficient size to accommodate the implant specimens. Use only one species per test.

Note—The paravertebral muscles of rabbits are the preferred implant sites. Alternatively, the gluteal muscles of rats, or the thigh muscles of rabbits, may be used.

Use at least three animals and sufficient implant sites to yield eight test specimens and eight control specimens for each implantation period.

In cases where the control material is expected to elicit more than a minimal response, use two specimens of this control. Implant two additional control specimens, composed of a material known to evoke a minimal tissue reaction, in a location opposite to the test materials.

Implantation Procedure

Implantation shall be by hypodermic needle or trocar. For larger implants other appropriate surgical implantation techniques may be used.

Implant test specimens into the body muscle with the long axis parallel to the muscle fibers.

For rabbit paravertebral muscles, implant four specimens of the test materials along one side of the spine, 25 mm to 50 mm from the midline and parallel to the spinal column, and about 25 mm apart from each other. In similar fashion implant four specimens of the control material in the contralateral muscle of each animal.

Implantation Period

To ensure a steady state of biological tissue response, the implantation period(s) shall be as specified in Tables 3.18 and 3.19.

3.10.1.18 Test Method for Implantation in Bone

Field of Application

This test method is used for assessing the biological response of bone tissue to an implanted material.

The study may be used to compare the effect of different surface textures or conditions of the same material or to assess the effect of various treatments or modifications of a material.

Principal

Insertion of the implants into the bone tissue of test animals, serves to evaluate systemic response to the local foreign body insertion. The method compares the biological response to implants of test specimens with the biological response to implants of control specimens made of materials which are established in clinical use.

Shape of Implant Specimens

The specimens may be screw-shaped or threaded to provide initial stability of the implants in the bone. If preparation of a screw shape is impractical, a cylinder shape may be used.

Size of Test Specimens

Implant sizes are based on the size of the test animal and bone chosen. The following dimensions shall be considered:

- Rabbits: cylindrical implants 2 mm in diameter and 6 mm in length.
- Dogs, sheep, and goats: cylindrical implants 4 mm in diameter and 12 mm in length.
- Rabbits, dogs, sheep, goats, and pigs: 2 mm to 4.5 mm orthopedic bone screw-type implants.

Test Animals

The implants shall be inserted into the bone of dogs, sheep, goats, pigs, or rabbits. Select one species among these. Species differences are important in bone physiology and should be assessed before implantation procedures are initiated.

At least four rabbits, or at least two each of other animals, shall be used for each implantation period.

Implant Sites

Equivalent anatomical sites shall be used for test and control specimens. The test implants shall be contralateral to the control implants. Select the implant site to minimize the risk of mobility of the implant.

Note—The femur and tibia are suitable. Other sites may be considered.

The number of implant sites shall be as follows:

(a) In each rabbit there shall be a maximum of six implant sites: three for test specimens and three for control specimens.
(b) In each dog, sheep, goat, or pig, there shall be a maximum of 12 implant sites; 6 for test specimens and 6 for control specimens. Do not insert more than 12 specimens in any one animal.

The size, mass, and age of the animal and the implant site chosen should ensure that the implant placement does not cause significant risk of pathological fracture of the test site. In younger animals it is especially important to ensure that the implants avoid the epiphyseal area or other immature bones.

Implantation Procedure

Perform bone preparation using low drilling speed and intermittent drilling with profuse irrigation with physiological saline solution and suction, because overheating will result in local tissue necrosis.

It is important that the diameter of the implant and the implant bed in the bone match well enough to avoid ingrowth of fibrous tissue.

Expose the cortex of each femur or tibia and drill the appropriate number of holes to receive implants. For rabbits, prepare up to three holes; for larger animals prepare up to six holes. Ream to final diameter or tap screw thread before insertion. Insert cylinders by finger pressure to allow press fit. Tighten screw-shaped implants in place with an instrument capable of delivering a predetermined torque. Record the torque.

Implantation Period

To ensure a steady state of biological tissue response the implantation period(s) shall be as specified in Tables 3.18 and 3.19.

3.10.1.19 Control Materials

Response

The biological response to these materials is not defined as to response, but rather the response is used as a reference against which a reaction to another material is compared.

As a porous control material is not available at present, it is acceptable to use a dense control material for comparative purposes.

If the most appropriate control material is expected to elicit a tissue response greater than that normally observed with the control materials cited in this annex, samples of these latter materials may be implanted as controls to check the surgical technique.

Metallic Control Materials

Stainless steel, cobalt-chromium, titanium, and titanium alloys are used to fabricate control specimens. The biological response to these materials has been well characterized by their extensive use in research and clinical practice. See for further information ISO 5832, Parts 1 to 8.

Polymeric and Ceramic Control Materials

Information on nonmetallic control materials is to be found in ASTM F 748, 763, and 981.

Implantation as a Method for Other End Points

Implantation studies can also be conducted to evaluate the longer-term (subchronic and chronic) potential for devices to elicit systemic toxicity or to evaluate the carcinogenic potential of devices. Uses for these cases are addressed later in this chapter.

3.11 Long-Term Implant Studies

Here we address the issues and considerations involved in evaluating the systemic effects of long-term implant devices, with the case of evaluation of materials for potential carcinogenicity being addressed in a later section. There are also local tissue and body/implant interactions that must be evaluated (Leninger et al. 1964). The spectrum of interactions can be thought of as presented in Table 3.20.

Such interactions are assessed in long-term studies which may or may not include the eventual retrieval of the implant itself from the host. Retrieval studies seek to study the biological and device-related performance characteristics under actual conditions of use and to determine the efficacy, reliability, and biocompatibility (safety) of medical devices.

Table 3.20 Host/implant interactions

Effects of implant on host
Local
Blood material interactions
Protein adsorption
Coagulation
Fibrinolysis
Platelet reactions
Complement activation
Blood
Leukocyte reactions
Hemolysis
Toxicity
Derangements of healing
Encapsulation
Foreign body reaction
Pannus formation
Infection
Tumorigenesis
Systemic
Embolization
Thrombus
Hypersensitivity
Alteration of lymphatic system
Effects of host on implant
Physical
Abrasive wear
Fatigue
Stress corrosion
Degeneration
Dissolution
Biological
Adsorption of tissue substances
Enzymatic degradation
Calcification

As such, retrieval studies have seven objects:

1. Enhanced patient management.
2. Recognition of complications.
3. Device design criteria.
4. Evaluation of patient/prosthesis matching.
5. Elimination of complications.
6. Identification of interactions.
7. Elucidation of mechanisms of interactions.

Implants can fail for any of six different categories of causes:

1. Thrombosis and thromboembolism.
2. Device-assisted infection.
3. Inappropriate healing.
4. Degradation, fracture.
5. Adverse local tissue reaction.
6. Adverse systemic reaction.

ASTM standard practice F981-87 ("Standard Practice for Assessment of Compatibility of Biomaterials [Nonporous] for Surgical Implants with Respect to Effect of Materials on Muscle and Bone") provides a framework for evaluating long-term host/implant interactions.

The practice provides a series of experimental protocols for biological assays of tissue reaction to nonporous, nonabsorbable biomaterials for surgical implants. It assesses the effects of the material on animal tissue in which it is implanted. The specified experimental protocol is not designed to provide a comprehensive assessment of the systemic toxicity, carcinogenicity, teratogenicity, or mutagenicity of the material. It applies only to materials with projected applications in human subjects where the materials will reside in bone or soft tissue in excess of 30 days and will remain unabsorbed. Applications in other organ systems or tissues may be inappropriate and are therefore excluded. Control materials will consist of any one of the metal alloys in ASRM Specifications F67, F75, F90, F136, F138, or F562 or ultra-high-molecular-weight polyethylene as stated in ASTM Specifications F648 or USP polyethylene negative control.

Referenced ASTM Standards include:

F67 Specification for Unalloyed Titanium for Surgical Implant Applications[a]
F75 Specification for Cast Cobalt-Chromium-Molybdenum Alloy for Surgical Implant Applications[b]
F86 Practice for Surface Preparation and Marking of Metallic Surgical Implants[a]
F90 Specification for Wrought Cobalt-Chromium-Tungsten-Nickel Alloy for Surgical Implant Applications[b]
F136 Specification for Wrought Titanium 7A1-4V ELI Alloy for Surgical Implant Applications[a]
F138 Specification for Stainless Steel Bars and Wire for Surgical Implants (Special Quality)[a]

F361 Practice for Assessment of Compatibility of Metallic Materials for Surgical Implants with Respect to Effect of Materials on Tissue[b]
F469 Practice for Assessment of Compatibility of Nonporous Polymeric Materials for Surgical Implants with Regard to Effect of Materials on Tissue[c]
F562 Specification for Wrought Cobalt-Nickel-Chromium-Molybdenum Alloys for Surgical Implant Application[a]
F648 Specification for Ultra-High-Molecular-Weight Polyethylene Powder and Fabricated Form for Surgical Implants[a]
F673 Practice for Short-Term Screening of Implant Materials[a]
[a]*Annual Book of ASTM Standards,* Vo. 13.01
[b]Discontinued. See 1986 *Annual Book of ASTM Standards,* Vo. 13.01
[c]Discontinued. See 1987 *Annual Book of ASTM Standards,* Vo. 13.01

The practice describes the preparation of implants; the number of implants and test hosts, test sites, exposure schedule, implant sterilization techniques; and methods of implant retrieval and tissue examination of each test site. Histological criteria for evaluating tissue reaction are provided. A test protocol for comparing the local tissue response evoked by biomaterials is specified, from which medical implantable devices might ultimately be fabricated, with the local tissue response elicited by control materials currently accepted for the fabrication of surgical devices. Currently accepted materials are the metals, metal alloys, and polyethylene previously specified which are standardized on the basis of acceptable long-term clinical experience. The controls consistently produce cellular reaction and scar to a degree that has been found to be acceptable to the host.

Rats (acceptable strains such as Fischer 344), New Zealand rabbits, and dogs may be used as test hosts for soft tissue implant response. It is suggested that the rats be age and sex matched. Rabbits and dogs may be used as test hosts for bone implants.

The sacrospinalis, paralumbar, gluteal muscles, and the femur or tibia can serve as the test site for implants. However, the same site must be used for test and material implants in all the animal species.

Table 3.21 contains a suggested minimum number of study animals and a suggested schedule for the necropsy of animals.

Each implant shall be made in a cylindrical shape with hemispherical ends (see below for sizes). If the ends are not hemispherical, this must be reported.

Table 3.21 Intervals of sacrifice

	Number of animals to be necropsied		
Necropsy periods (weeks after insertion of implants)	Rat	Rabbit	Dog
12 weeks	4	4	2
26 weeks	4	4	2
52 weeks	4	4	2
104 weeks	–	–	2

Each implant shall be fabricated and finished and its surface cleaned in a manner appropriate for its projected application in human subjects in accordance with ASTM Practice F86.

Reference metallic specimens shall be fabricated from materials such as the metal alloys in ASTM Specifications F67, F75, F90, F138, or F562 or polymeric polyethylene USP-negative control plastic.

Suggested sizes and shapes of implants for insertion in muscle are as follows:

- For rats 1-mm-diameter by 2-cm-long cylindrical implants
- For rabbits 2-mm-diameter by 10–15-mm-long cylindrical implants
- For dogs 6-mm-diameter by 18-mm-long cylindrical implants

If fabrication problems prevent preparing specimens 1 mm in diameter, alternative specimen sizes are 2-mm diameter by 6-mm long for rats and 4-mm diameter by 12-mm long for rabbits. If these alternate dimensions are used, such should be reported and such use justified.

Sizes and shapes of implants for insertion in the bone are as follows:

- For rabbits 2-mm-diameter by 6-mm-long cylindrical implants
- For dogs 4-mm-diameter by 12-mm-long cylindrical implants

If the length of the bone implants needs to be less than that designated because of anatomical constraints, such should be reported.

3.11.1 Number of Test and Control Implants

In each rat, due to size, there should be two implants: one each for test and control material implant. In each rabbit, due to size, there should be six implants: four for test materials and two control material implants. In each dog, there should be twelve implants: eight for test materials and four control material implants.

3.11.2 Conditioning

Remove all surface contaminants with appropriate solvents and rinse all test and control implants in distilled water prior to sterilization. It is recommended that the implant materials be processed and cleaned in the same way the final product will be. That is, clean, package, and sterilize all implants in the same way as used for human implantation.

After final preparation and sterilization, handle the test and control implants with great care to ensure that they are not scratched, damaged, or contaminated in any way prior to insertion.

Report all details of conditioning.

3.11.3 Implantation Period

Insert all implants into each animal at the same surgical session so that implantation periods run concurrently. The implantation period is 52 weeks for rats and rabbits and 104 weeks for dogs, with interim sacrifices at 12, 26, and 52 weeks (see Table 3.21).

3.11.3.1 Implantation (Muscle)

Place material implants in the paravertebral muscles of the adult rats, rabbits, or dogs in such a manner that they are directly in contact with muscle tissue.

Introduce material implants in dogs by the technique of making an implantation site in the muscle by using a hemostat to separate the muscle fibers. Then insert the implant using plastic-tipped forceps or any tool that is nonabrasive to avoid damage to the implant. Do not insert more than 12 implant materials in each dog.

Introduce material implants in rabbits and rats using a sterile technique. Sterile disposable Luer lock needles may be used to implant the material implants into the paravertebral muscles along the spine. In rats insert a negative control implant on one side of the spine and a test material implant on the other side. In rabbits implant one negative control material on each side of the spine and implant two test materials on each side of the spine. If larger diameter specimens are used, an alternative implantation technique such as that described above should be employed.

3.11.3.2 Implantation (Femur)

Expose the lateral cortex of each rabbit femur and drill three holes 1/16 in. (1.6 mm) through the lateral cortex using the technique and instrument appropriate for the procedure. For dogs, make the holes 1/8 inch (3.2 nun) in diameter; make six holes in each femur. Into each one of these holes, insert one of the implants by finger pressure. Then close the wound.

Caution should be taken to minimize the motion of the implant in the tissue on the desired result.

3.11.4 Sacrifice and Implant Retrieval

Euthanize animals by a humane method at the intervals listed in Table 3.21. The necropsy periods start at 12 weeks because it is assumed that acceptable implant data have been received for earlier periods such as 1, 4, and 8 weeks from short-term implant testing.

At necropsy, record any gross abnormalities of color or consistency observed in the tissue surrounding the implant. Remove each implant with an intact envelope of surrounding tissue. Include in the tissue sample a minimum of a 4-mm-thick layer of tissue surrounding the implant. If less than a 4-mm-thick layer is removed, report such a case.

3.11.4.1 Postmortem Observations

Necropsy all animals that are sacrificed for the purposes of the assay or die during the assay period in accordance with standard laboratory practice. Establish the status of the health of the experimental animals during the period of the assay.

3.11.4.2 Histological Procedure

Tissue Sample Preparation

Prepare two blocks from each implantation site. Process the excised tissue block containing either a test implant or control implant for histopathological examination and such other studies as are appropriate. Cut the sample midway from end to end into appropriate size for each study. Record the gross appearance of the implant and the tissue. If special stains are deemed necessary, prepare additional tissue blocks or slides, or both, and make appropriate observations.

Histopathological Observations

Compare the amount of tissue reaction adjacent to the test implant to that adjacent to a similar location on the control implant with respect to thickness of scar, presence of inflammatory or other cell types, presence of particles, and such other indications of interaction of tissue and material as might occur with the actual material under test (Pizzoferrato et al. 1988; Rahn et al. 1982). A suggested method for the evaluation of tissue response after implantation can be found in Turner et al. (1973), as summarized in Table 3.22.

Pathologists may choose to use this scoring system while comparing the negative control to the test material as an aid in their evaluation. The overall toxicity of the test material as compared to the negative control is to be evaluated independently for all time periods. Table 3.23 provides a suggested format for evaluation and scoring.

Porous or porous-coated materials are specifically excluded since the response to such materials includes ingrowth of tissue into the pores. As a result, the method of tissue fixation and sectioning and the evaluation scheme are substantially different.

Table 3.22 Suggested methods for tissue response evaluation

Number of elements[a]	Score
0	0
1–5	0.5
6–15	1
16–25	2
26 or more	3
Degree of necrosis	
Not present	0
Minimal present	0.5
Mild degree of involvement	1
Moderate degree of involvement	2
Marked degree of involvement	3
Overall toxicity rating of test samples	
Nontoxic	
Very slight toxic reaction	0
Mild toxic reaction	1
Moderate toxic reaction	2
Marked toxic reaction	3
	4

[a]Cellular elements to be evaluated based upon the number of elements in high power field (470×), average of five fields

Table 3.23 Suggested evaluation format and scoring range

Animal number				
Duration of implant (weeks)				
Sample description				
Gross response				
Histopath (number)				
Score	0.51	2	3	4
Necrosis				
Degeneration				
Inflammation				
Polymorphonuclear leukocytes				
Lymphocytes				
Eosinophils				
Plasma cells				
Macrophages				
Fibrosis				
Giant cells				
Foreign body debris				
Fatty infiltration				
Relative size of involved area (mm)				
Histopathologic toxicity rating				

Stainless steel, cobalt-chromium, and titanium alloys are sued as reference materials since the biological response to these materials has been well characterized by their extensive use in research. The response to these materials is not defined as compatible, but rather the response is used as a reference against which reactions to other materials is compared.

This practice is a modification of the original Practice F361 in that it only involved long-term test periods. The short-term response to materials is to be evaluated using Practice F763. Special methods exist to reduce the impact of relative motion at the implant/tissue interface (Geret et al. 1980a, b).

This practice was revised in 1987 to allow for alternative specimen dimensions for rats and rabbits for muscle implantation. The original specimen dimensions were intended to be implanted through a needle, which was a change from F361 and F469. The alternate dimensions restore those specified since 1972 which some members felt were more appropriate for some material types.

3.11.5 Considerations

One problem of the implantation tests is a tendency for the strips or prototype implant devices to migrate from their implantation sites, even to subcutaneous positions, and this often prolongs the search for them. Nevertheless, the test is an effective detection system for toxic ingredients of solid materials which leach in contact with tissue fluid. It is important to recognize the microscopic effects of the standard negative control strips (such as the USP additive-free polyethylene). These reactions are typical of skeletal muscle in contact for a week or more with a foreign body and comprise mild mononuclear cell infiltration, multinucleated giant cell formation, fibroplasia, slight dystrophic calcification, muscle fiber atrophy, and centripetal migration of sarcolemmal nuclei. Traumatic hemorrhage is also common. Positive reactions are similar but more pronounced and additionally include focal necrosis and exudation, particularly of heterophils. These are useful tests, not only for finished products but also to identify unacceptable changes in formulation or manufacturing processes such as the introduction of chlorinating cycles to remove bloom on latex catheters. For materials to be utilized in long-term implants, these tests are the only means of accurately predicting long-term tissue and systemic interactions.

The finished state and handling of all components, but particularly metals, in an implant are critical variables, as the primary tissue/implant interactions revolve around surface effects such as the ionization of metals (Ferguson et al. 1960). Such interactions are so predominantly surface interactions (Kordan 1967) that the surface conditions of an implant, such as porosity and pore size, are critical (Goldhaber 1961, 1962). The actual site of implantation in the body also influences the nature of interaction significantly (Kaminski et al. 1968).

3.12 Genotoxicity (ISO 10993-3)

Genotoxicity encompasses all the potential means by which the genetic material of higher organisms may be damaged, with the potential for resulting serious consequences. Most forms of genotoxicity are expressions of mutagenicity—the induction of DNA damage and other genetic alterations, with binding to and changes in one or a few of DNA base pairs (gene mutations). Elastogenicity that is gross changes in chromosomal structure (i.e., chromosomal aberrations) or in chromosome numbers represents a different class of genotoxic changes, for which practical thresholds of effect are identifiable.

It has been known for several hundred years that exposure to particular chemicals or complex mixtures can lead to cancer in later life (Doll 1977), and it has been postulated more recently that chemicals can also induce heritable changes in man, leading to diseases in the next generation (ICEMC 1983). There has been accumulating evidence that such changes can arise following damage to DNA and resulting mutations (see, e.g., Bridges 1976). Therefore, it has become necessary to determine whether widely used chemicals or potentially useful new chemicals possess the ability to damage DNA. In industry, such information may be used to discard a new chemical drug or biomaterial if a safer alternative can be found, to control or eliminate human exposure for a genotoxic industrial compound or, for a drug, to proceed with development if benefits clearly outweigh risks. Data concerning the genotoxicity of a new material have become part of the basic biocompatibility information package. They are needed for decision-making and to reduce risks that might otherwise be unforeseen.

ISO 10993-3 sets forth clear guidance on testing requirements or summarized in Table 3.24.

3.12.1 Test Systems

In vivo and in vitro techniques are available to test mutagenic properties to demonstrate presence or lack of ability of the test material to cause mutation or chromosomal damage or cause cancer, as summarized in Table 3.25. The material intended for intimate contact and long exposure should not have any genotoxic properties. The presence of unpolymerized materials and traces of monomers, oligomers, additives or biodegration products can cause mutations. Mutation can be a point mutation or chromosomal rearrangement caused by DNA damage. Therefore, the material's ability to cause point mutation, chromosomal change, or evidence of DNA damage are tested. As we have seen, correlations exist between mutagenic and carcinogenic properties. Most carcinogens are mutagens, but not all mutagens are human carcinogens.

The Ames *Salmonella*/microsome test is a principal sensitive mutagen screening test. Compounds are tested on the mutants of *Salmonella typhimurium* for reversion

Table 3.24 ISO genotoxicity guidance

Genotoxic effect to be assessed for conformance with ISO 1099-3	Significance of test	Tests meeting requirements
DNA effects	Damage to DNA (deoxyribonucleic acid) by a chemical or material may result in genotoxic effects such as mutations, which in turn may lead to carcinogenicity Damage to DNA causes the cell to manufacture new DNA to compensate for the loss or damage. This can be assessed by evaluating the formation of newly synthesized DNA	Unscheduled DNA synthesis
Gene mutations	Gene mutations (changes in the sequences of DNA that code for critical proteins or functions) have been correlated to carcinogenicity and tumorigenicity	Ames assay (4 *Salmonella* typhimurium bacterial strains and *Escherichia coli*) is a reverse mutation assay. A bacterial mutation event causes the bacteria to become histidine (a vital amino acid) independent. Normal bacteria will not survive in the absence of histidine Hypoxanthine guanine phosphoribosyl transferase (HGPRT) is a forward mutation assay. Mammalian cells that have been exposed to a mutagen will survive in the presence of a toxic substance (6-thioguanine)
Chromosomal aberrations	Physical damage to chromosomes (large ordered stretches of DNA in the nuclei of cells) or clastogenicity can lead to DNA damage, in turn leading to abnormal and/or carcinogenic growth of cells	Chromosomal aberration assay assesses the potential for physical damage to the chromosomes of mammalian cells by a biomaterial

ISO 10993-3 (1993) Biological Evaluation of Medical Devices—Part 3: Tests for Genotoxicity, Carcinogenicity, and Reproductive Toxicity states that at least three in vitro tests, two of which use mammalian cells, should be used to test for three levels of genotoxic effects: DNA effects, gene mutations, and chromosomal aberrations

ANSI/AAMI/ISO 10993-3 (1993) Biological Evaluation of Medical Devices—Part 3: Tests for Genotoxicity, Carcinogenicity, and Reproductive Toxicity states that "Suitable cell transformation systems may be used for carcinogenicity prescreening"

from a histidine requirement back to prototrophy. A positive result is seen by the growth of revertant bacteria A microsomal activation system should be included in this assay. The use of all five bacterial test strains are generally required.

Table 3.25 Common genotoxicity tests

Assays for gene mutations	In vitro	In vivo
Salmonella typhimurium reverse mutation assay (Ames test, bacteria) [OECD 471]	✓	
Escherichia coli reverse mutation assay (bacteria) [OECD472]	✓	
Gene mutation in mammalian cells in culture [OECD 476]	✓	
Drosophila sex-linked recessive lethal assay (fruit fly) [OECD 477]		✓
Gene mutation in *Saccharomyces cerevisiae* (yeast) [OECD 480]	✓	
Mouse spot test [OECD 484]		✓
Assays for chromosomal and genomic mutations		
In vitro cytogenetic assay [OECD 473]	✓	
In vivo cytogenetic assay [OICD 475]		✓
Micronucleus test [OECD 474]		✓
Dominant lethal assay [OECD 478]		✓
Heritable translocation assay [OECD 485]		✓
Mammalian germ cell cytogenetic assay [OECD 483]		✓
Assays for DNA effects		
DNA damage and repair: Unscheduled DNA synthesis in vitro [OECD 482]	✓	
Mitotic recombination in *Saccharomyces cerevisiae* (yeast) [OECD 481]	✓	
In vitro sister chromatid exchange assay [OECD 479]	✓	

A non-bacterial mutagenicity tests are generally required to support the lack of mutagenic or carcinogenic potential. Some well-known tests are:

- The L5178Y mouse lymphoma test for mutants at the TK locus
- Sister chromatid exchange assay

ISO 10993 specifically requires three genotoxicity assays for all devices. The assays should preferably evaluate DNA effects, gene mutations, and chromosomal aberrations; and two of the assays should preferably use mammalian cells. Guidances for providing tests for selection to meet these needs are the OECD guidelines, which include eight in vitro and seven in vivo assays.

3.12.2 ISO Test Profile

ISO 10993 Part 3: Tests for Genotoxicity, Carcinogenicity and Reproductive Toxicity suggests that a series of three in vitro assays be conducted, at least two of which should use mammalian cells as a target (Table 3.26). The tests should address the three types of genotoxic effects: (1) gene mutations, (2) chromosomal and genomic aberrations, and (3) DNA effects. Three tests that are recommended are shown below (note that none of the three recommended tests assays for DNA effects!). In ISO's opinion, a profile of three in vitro genotoxicity tests is considered

Table 3.26 Genotoxicity tests recommended by ISO

Test	Mutation	Cell type	Method
Salmonella reverse mutation assay [OECD 471]	Gene	Bacterial	In vitro
In vitro cytogenetic assay [OECD 473]	Chromosome	Mammalian	In vitro
Gene mutation in mammalian cells [OECD 476]	Gene	Mammalian	In vitro

Table 3.27 Genotoxicity tests recommended by ICH

Genotoxicity test-ICH	Mutation	Cell type	Method
A test for gene mutation in bacteria	Gene	Bacterial	In vitro
In vitro cytogenetic assay using mouse lymphomas tk cells	Chromosome	Mammalian	In vitro
In vivo test for chromosomal damage using rodent hematopoietic cells	Gene	Mammalian	In vivo

sufficient to establish safety for most medical devices; in vivo testing need only be done if in vitro tests are positive.

3.12.3 ICH Test Profile

The International Conference on Harmonization recommends a rather different profile of genotoxicity tests for drugs (Table 3.27). There is a requirement to have an in vivo test conducted, but such is not needed prior to human exposure unless one or more of the in vitro studies yields a positive response.

3.12.4 In Vitro Test Systems

The principal tests can be broadly categorized into microbial and mammalian cell assays. In both cases the tests are carried out in the presence and absence of in vitro metabolic activation enzymes, usually derived from rodent liver.

3.12.4.1 In Vitro Metabolic Activation

The target cells for in vitro mutagenicity tests often possess a limited (often overlooked) capacity for endogenous metabolism of xenobiotics. However, to simulate the complexity of metabolic events that occur in the whole animal, there is a critical need to supplement this activity.

3.12.5 Bacterial Mutation Tests

The study of mutation in bacteria (and bacterial viruses) has had a fundamental role in the science of genetics in the twentieth century. In particular, the unraveling of biochemical anabolic and catabolic pathways, the identification of DNA as the hereditary material, the fine structure of the gene, the nature of gene regulation, etc. have all been aided by bacterial mutants.

As an offshoot of studies of genes concerned with the biosynthesis of amino acids, a range of *E. coli* (see, e.g., Yanofsky 1971) and *Salmonella typhimurium* strains (see, e.g., Ames 1971) with relatively well-defined mutations in known genes became available. Thus, bacteria already mutant at an easily detectable locus are treated with a range of doses of the test material to determine whether the compound can induce a second mutation that directly reverses or suppresses the original mutations. Thus, for amino acid auxotrophs, the original mutation has resulted in loss of ability to grow in the absence of the required amino acid. The second mutation restores prototrophy, i.e., the affected cell is now able to grow in the absence of the relevant amino acid, if provided with inorganic salts and a carbon source. This simple concept, in fact, underlines the great strength of these assays, for it provides enormous selective power which can identify a small number off the chosen mutants from a population of millions of unmutated cells and cells mutated in other genes. The genetic target, i.e., the mutated DNA bases in the gene in question (or bases in the relevant tRNA genes; see the discussion of suppressor mutations), can thus be very small, just one or a few bases in length.

An alternative approach is to use bacteria to detect "forward mutations." Genetic systems which detect forward mutations have an apparent advantage, in that a wide variety of genetic changes may lead to a forward mutation., e.g., point mutation, deletions, insertions, etc. In addition, forward mutations in a number of different genes may lead to the same change in phenotype; thus, the genetic target is much larger than that seen in most reverse mutation assays. However, if a particular mutagen causes rare specific changes, these changes may be lost against the background of more common events (Gatehouse et al. 1990). Spontaneous mutation rates tend to be relatively high in forward mutation systems. Acquisition of resistance to a toxic chemical (e.g., an amino acid analogue or antibiotic) is a frequently used genetic marker in these systems. For instance, the use of resistance to the antibiotic streptomycin preceded the reversion assays in common use today.

3.12.5.1 Reversion Test Background

There are several excellent references describing the background and use of bacteria for reversion tests (Brusick 1987a, b; Gatehouse et al. 1990). Three different protocols have been widely used: plate incorporation assays, treat and plate tests, and fluctuation tests. These methods are described in detail in the following sections.

Fundamental to the operation of these tests are the genetic compositions of the tester strains selected for use.

3.12.5.2 Genetic Makeup of Tester Strains

The most widely used strains are those developed by Bruce Ames and colleagues which are mutant derivatives of the organism *Salmonella typhimurium*. Each strain carries one of a number of mutations in the operon coding for histidine biosynthesis. In each case the mutation can be reverted either by base change or by frameshift mutations. The genotype of the commonly used strains is shown in Table 3.28.

3.12.5.3 Protocol for Dose Ranging and Selection

Before carrying out the main tests, it is necessary to carry out a preliminary toxicity dose ranging test. This should be carried out following the same basic protocol as the mutation test, except that instead of scoring the number of mutants on, for example, minimal media plates with limiting amounts of a required amino acid, the number of survivors is scored on fully supplemented minimal media. A typical protocol is outlined below:

- Prepare a stock solution of the test compound at a concentration of 50 mg ml^{-1} in an appropriate solvent. It may be necessary to prepare a lower concentration of stock solution, depending on the solubility of the test compound.
- Make dilutions of the stock solution.
- To 2.0 ml aliquots of soft agar overlay medium (0.6% agar and 0.5% sodium chloride in distilled water) containing a trace of histidine and excess biotin and maintained at 45 °C in a dry block, add 100 μl at a solution of the test article. Use only one plate per dilution.

Table 3.28 Ames test-test strains

Strain	Genotype	Reversion events
TA1535	$hisG_{46}$ rfa *f* gal chlD bio uvrB	Subset of base-pair substitution events
TA100	$hisG_{46}$ *f*rfa gal chlD bio uvrB (pKM101)	Subset of base-pair substitution events
TA1537	$hisC_{3076}$ *f*rfa gal chlD bio uvrB	Frameshifts
TA1538	$hisD_{3052}$ *f*rfa gal chlD bio uvrB	Frameshifts
TA98	$hisD_{3052}$ *f*rfa gal chlD bio uvrB pKM101)	Frameshifts
TA97	$hisD_{6610}$ $hisO_{1242}$ rfa *f* gal chl^{D} bio uvrB (pKM101)	Frameshifts
TA102	his*f* $(G)_{8476}$ rfa galE (pAQ1) (pKM101)	All possible transitions and transversions; small deletions

- Mix and pour onto dried Vogel and Bonner minimal medium plates as in an Ames test, including an untreated control and a solvent control, if necessary. The final concentrations of test compound will be 5000, 1500, 500, 150, and 50 μg plate^{-1}.
- Repeat step (3), using 0.5 ml of 8 percent S9 mix per 2.0 ml aliquot of soft agar in addition to the test compound and tester strain. The S9 mix is kept on ice during the experiment.
- Incubate the plates for 2 days at 37 °C and examine the background lawn of growth with a microscope (58 eyepiece lens, 510 objective lens). The lowest concentration giving a depleted background lawn is regarded as a toxic dose.

This test will also demonstrate excess growth, which may indicate the presence of histidine or tryptophan or their precursors in the test material, which could make testing for mutagenicity impracticable by this method.

When setting the maximum test concentration, it is important to test into the mg plate^{-1} range where possible (Gatehouse et al. 1990), as some mutagens are only detectable when tested at high concentrations. However, for nontoxic, soluble mutagens an upper limit of 5 mg plate^{-1} is recommended (DeSerres and Shelby 1979). For less soluble compounds, at least one dose exhibiting precipitation should be included.

3.12.5.4 Ames Salmonella/Plate Incorporation Method

The following procedure is based on that described by Ames and colleagues (Maron and Ames 1983), with additional modifications:

- Each selected test strain is grown for 10 h at 37 °C in nutrient broth (Oxoid No. 2) or supplemented minimal media (Vogel-Bonner) on an orbital shaker. A timing device can be used to ensure that cultures are ready at the beginning of the working day.
- 2.0 ml aliquots of soft agar overlay medium are melted just prior to use and cooled to 50 °C and relevant supplements added—i.e., L-histidine, final concentration 9.55 μg ml^{-1}, and D-biotin, 12 μg ml^{-1}. (N.B.: If *E. coli* WP2 tester strains are used, the only supplement required is tryptophan 3.6 μg ml^{-1}.) The medium is kept semi-molten by holding the tubes containing the medium in a hot aluminum dry block, held at 45 °C. It is best to avoid water baths as microbial contamination can cause problems.
- The following additions are made to each tube of top agar: the test article (or solvent control) in solution (10–200 μl), the test strain (100 μl), and, where necessary, S9 mix (500 μl). The test is carried out in the presence and absence of S9 mix. The exact volume of test article or solvent may depend on toxicity or solubility, as described in the preceding section.
- There should be at least three replicate plates per treatment with at least five test doses plus untreated controls. Duplicate plates are sufficient for the positive and sterility control treatments. The use of twice as many negative control plates as used in each treatment group will lead to more powerful tests from a statistical standpoint (Mahon et al. 1989).

- Each tube of top agar is mixed and quickly poured onto dried prelabeled Vogel-Bonner basal agar plates.
- The soft agar is allowed to set at room temperature and the plates are inverted and incubated (within 1 h of pouring) at 37 °C in the dark. Incubation is continued for 2–3 days.
- Before scoring the plates for revertant colonies, the presence of a light background lawn of growth (due to limited growth of non-revertant colonies before the trace of histidine or tryptophan is exhausted) should be confirmed for each concentration of test article by examination of the plate under low power of a light microscope. At concentrations that are toxic to the test strains, such a lawn will be depleted and colonies may appear that are not true revertants but surviving, non-prototrophic cells. If necessary, the phenotype of any questionable colonies (pseudo-revertants) should be checked by plating on histidine or tryptophan-free medium.
- Revertant colonies can be counted by hand or with an automatic colony counter. Such machines are relatively accurate in the range of colonies normally observed (although calibration against manual counts is a wise precaution). Where accurate quantitative counts of plates with large numbers of colonies are required, only manual counts will give accurate results.

3.12.5.5 Controls

Positive Controls

Where possible, positive controls should be chosen that are structurally related to the test article. This increases the confidence in the results. In the absence of structurally related mutagens, each strain has one or more standard positive control designated. The use of such controls validates each test run and helps to confirm the nature of each strain. Pagano and Zeger (1985) have shown that it is possible to store stock solutions of most routinely used positive controls (sodium azide, 2-aminoanthracene, benzo[*a*]phyene, 4-nitroquinoline oxide) at −20 °C to −80 °C for several months, without loss of activity. This measure can help reduce potential exposure to laboratory personnel.

Untreated/Vehicle Controls

Untreated controls omit the test article, but are made up to volume with buffer. The vehicle control is made up to volume with the solvent used to dissolve the test substance. It is preferable to ensure that each of the treated plates contains the same volume of vehicle throughout.

As detailed by Gatehouse and Tweats (1988), the nature and concentration of solvent may have a marked effect on the test result. Dimethysolphoxide is often used as the solvent of choice for hydrophobic compounds. However, there may be

unforeseen effects, such as an increase in mutagenicity of come compounds—e.g., *p*-phenylenediamine (Burnett et al., 1982)—or a decrease in mutagenicity of others, such as simple aliphatic nitrosamines (Yahagi et al. 1977). It is essential to use fresh batches of the highest purity grade available and to prevent decomposition/oxidation on storage. The products after oxidation, etc. are both toxic and can induce base-pair substitutions in both bacterial and mammalian assays. Finally, DMSO and other organic solvents can inhibit the oxidation of different premutagens by microsomal monooxygenases (Wolff 1977a, b). To reduce the risk of artifactual results, it is essential to use the minimum amount of organic solvent (e.g., <2% w/w/) compatible with adequate testing of the test chemical.

It is important to keep a careful check of the number of mutant colonies present on untreated or vehicle control plates. These numbers depend on the following factors:

- The repair status of the cell—i.e., excision repair-deficient strains tend to have more "spontaneous mutants" than repair-proficient cells.
- The presence of mutator plasmids. Excision-deficient strains containing pKM101 have a higher spontaneous mutation rate at both base substitution and frameshift loci than excision-proficient strains.
- The total number of cell divisions that take place of the cells in the supplemented top agar. This is controlled by the supply of nutrients—in particular, histidine. Rat liver extracts may also supply trace amounts of limiting nutrients, resulting in a slight increase in the spontaneous yield of mutants in the presence of S9 mix.
- The size of the initial inoculum. During growth of the starting culture, mutants will arise. Thus, if a larger starting inoculum is used, more of these "preexisting" mutants will be present per plate. In fact, the "plate mutants" arising as described in point (3) predominate.
- The intrinsic mutability of the mutation in question. In practice the control mutation values tend to fall within in relatively precise range for each strain. Each laboratory should determine the normal range of revertant colonies per plate for each strain (Table 3.29).

Deviations in background reversion counts from the normal range should be investigated. It is possible that cross-contamination, variations in media quality, etc. have occurred that may invalidate particular experiments.

Frequent checks should also be made on the sterility of S9 preparations, media, and test articles. These simple precautions can prevent loss of valuable time and resources.

Evaluation of Results

At least two independent assays are carried out for each test article. The criterion for positive response is a reproducible and statistically significant result at any concentration for any strain. When positive results are obtained, the test is repeated, using the strain(s) and concentration range with which the initial positive results were observed. This range may be quite narrow for toxic agents.

Table 3.29 Positive controls for use in plate incorporation assays

Species	Strain	Mutagen	Conc. (μg plate^{-1})[a]
(a) In the absence of S9 mix			
S. Typhimurium	TA1535 TA100 TA1538 TA98 TA1537	Sodium azide Hycanthone methane sulphonate ICR 191	1–5 5–20 1
E. coli	WP2 uvrA	Nifuroxime	5–15
(b) in the presence of S9 mix			
E. coli	WP2 uvraA (pKM101)		
S. Typhimurium	TA1538 TA1535 TA100 TA90 TA1537	2-Aminoanthracene Neutral red	1–10 10–20

3.12.5.6 Preincubation Tests

Some mutagens are poorly detected in the standard plate incorporation assay, particularly those that are metabolized to short-lived reactive electrophiles—e.g., short-chain aliphatic *N*-nitroso compounds (Bartsch et al. 1976). It is also possible that some metabolites may bind to components within the agar. Such compounds can be detected by using a preincubation method first described by Yahagi et al. (1975) in which the bacteria, test compound, and S9 mix are incubated together in a small volume at 37 °C for a short period (30–60 min) before adding the soft agar and pouring as for the standard assay. In this variation of the test, during the preincubation step, the test compound, S9 mix, and bacteria are incubated in liquid at higher concentrations than in the standard test, and this may account for the increased sensitivity with relevant mutagens. In the standard method, the soluble enzymes in the S9 mix, cofactors, and the test agent may diffuse into the bottom agar. This can interfere with the detection of some mutagens—a problem that is overcome in the preincubation method (Forster et al. 1980; Gatehouse and Wedd 1984).

The test is carried out as follows:

- The strains are cultured overnight, and the inocula and S9 mix are prepared as in the standard Ames test.
- The soft agar overlays are prepared and maintained at 45 °C prior to use.
- To each of 3–5 tubes maintained at 37 °C in a Driblock are added 0.5 ml of S9 mix, 0.1 ml of the tester strain (10–18 h culture), and a suitable volume of the test compound, to yield the desired range of concentrations. The S9 mix is kept on ice prior to use.
- The reaction mixtures are incubated for us to 1 h at 37 °C.

- 2.0 ml of soft agar is added to each tube. After mixing, the agar and reaction mixture are poured onto previously labeled, dried Vogel-Bonner plates.
- Once the agar has set, the plates are incubated for 2–3 days before revertant colonies are scored.

The use of controls is as described for the plate incorporation assay. It is crucial to use the minimum amount of organic solvent in this assay, as the total volume of the incubation mixture is small relative to the solvent component.

This procedure can be modified to provide optimum conditions for particular chemical classes. For instance, preincubation times greater than 60 min plus aeration have been found necessary in the detection of allyl compounds (Neudecker and Henschler 1985).

3.12.6 Forward Mutation Tests

Forward mutation is an end point that may arise from various events, including base substitutions, frameshifts, DNA deletions, etc., as mentioned earlier.

Although bacterial forward mutation systems have not gained the popularity of reverse mutation tests (owing, in part, to lower sensitivity to some mutagens and lack of specificity), they have proved useful on occasion and have their supporters.

Several forward mutation tests have been devised, and a brief mention of two of the more widely used systems is provided below.

3.12.6.1 The Forward Mutation Test

The L-arabinose resistance test with *Salmonella typhimurium* is based on *ara D* mutants of the L-arabinose operon (Hera and Pueyo 1986); *ara D* mutants are unable to use L-arabinose as the sole carbon source. The assay scores a change from L-arabinose sensitivity to L-arabinose resistance, which is defined as the ability to grow in a medium containing L-arabinose plus another carbon source such as glycerol.

This phenotypic change reflects forward mutations in at least three different loci in the arabinose operon (Pueyo and Lopez-Barea 1979).

Strains have been constructed along the same lines as the recommended Ames strains with mutations to remove excision repair and mutations in increase permeability, and including the mutator plasmic pKM101—i.e., *Salmonella typhimurium* BA3 *ara D531, hisG46*, ΔuvrB *bio*, and BA9 *araD531, hisG46* ΔuvrB, *bio*, and *rfa* (p.KM101).

Protocols for the test have included plate incorporation, preincubation, and treat and plate tests (Hera and Pueyo 1986). In the latter tests the assay does not have the problem of "plate mutants" as described for reverse mutation tests in the previous section. The recommended procedure has the following outline protocol:

- Incubate the test strain of bacteria (10^7–10^8 cells per ml) and the test agent at 37 °C in nonselective DM medium with shaking.
- Wash the cells after a 2 h exposure period.
- Plate on selective medium (DM salts, 2 mg ml^{-1} glycerol, 2 ml ml^{-1} L-arabinose, 20 µg ml^{-1} L-histidine, 12 µg ml^{-1} biotin) containing an additional supplement of D-glucose, 0.5 mg per plate.

For metabolic activation 30 µl of S9 fraction and appropriate cofactors are included in the initial incubation mixture as the standard level. Different concentrations of S9 fraction can be used as required.

The group who have developed this test recommend that strain BA9 can replace the four strains used in the standard Ames test and that for the mutagens tested to date, this strain detects the same range of mutagens as the Ames test strains with equal or better sensitivity. The test does seem suitable for testing complex mixtures such as red wine (Dorado et al. 1988). However, the spontaneous background count using the protocol outlined above is over 500 per plate. If fewer cells are used, false-negative results are obtained (Xu et al. 1984).

3.13 In Vitro Cytogenetic Assays

The in vitro cytogenetic assay is a short-term mutagenicity test for detecting chromosomal damage in cultured mammalian cells.

Cultured cells have a limited ability metabolically to activate some potential clastogens. This can be overcome by adding an exogenous metabolic activation system such as S9 mix to the cells (Ames et al. 1975; Natarajan and Obe 1982; Maron and Ames 1983; Madle and Obe 1980).

Observations are made in metaphase cells arrested with a spindle inhibitor such as colchicine or colcemid to accumulate cells in a metaphase-like stage of mitosis (c-metaphase) before hypotonic treatment to enlarge cells and fixation with alcohol/acetic acid solution. Cells are then dispersed onto microscope slides and stained and slides are randomized, coded, and analyzed for chromosome aberrations with high-power light microscopy. Details of the procedure are given in Dean and Danford (1984) and Preston et al. (1981, 1987). The UKEMS guidelines (Scott et al. 1990) recommend that all tests be repeated regardless of the outcome of the first test and that, if a negative or equivocal result is obtained in the first test, the repeat should include an additional sampling time. In the earlier version of the guidelines (Scott et al. 1983), a single sampling at approximately 1.5 normal cycle times (−24 h for a 1.5 cell cycle) from the beginning of treatment was recommended, provided that a range of concentrations was used which induced marginal to substantial reductions in mitotic index, usually an indicator of mitotic delay. However, Ishidate (1988a) reported a number of chemicals which gave negative responses with a fixation time of 24 h but which were positive at 48 h. This was when a Chinese hamster fibroblast line (CHO) with a doubling time of 15 h was used. It would appear, therefore, that

there are chemicals which can induce extensive mitotic delay at clastogenic doses and may be clastogenic only when cells have passed through more than one cell cycle since treatment (Thust et al. 1980). A repeat test should include an additional sample at approximately 24 h later, but it may only be necessary to score cells from the highest dose at this later fixation time. When the first test gives a clearly positive result, the repeat test need only utilize the same fixation time. The use of other sampling times is in agreement with other guidelines (European Community EEC Directive—OECD 1983; American Society for Testing and Materials—Preston et al. 1987; Japanese Guidelines—JMHW 1984, 1987; Ishidate 1988b).

3.13.1 Cell Types

Established cell lines, cell strains, or primary cell cultures may be used. The most often used are Chinese hamster cell lines and human peripheral blood lymphocytes. The merits of these two cell lines have been reported (Ishidate and Harnois 1987; Kirkland and Garner 1987). The cell system must be validated and consistently sensitive to known clastogens.

3.13.2 Chinese Hamster Cell Lines

Chinese hamster cell lines have a small number of large chromosomes (11 pairs). Chinese hamster ovary cells, in which there has been an extensive rearrangement of chromosome material and the chromosome number may not be constant from cell to cell, are frequently used. Polyploidy, endoreduplication, and high spontaneous chromosome aberration frequencies can sometimes be found in these established cell lines, but careful cell culture techniques should minimize such effects. Cells should be treated in exponential growth when cells are in all stages of the cell cycle.

3.13.3 Human Peripheral Blood Lymphocytes

Blood should be taken from healthy donors not known to be suffering from viral infections or receiving medication. Staff handling blood should be immunized against hepatitis B, and regular donors should be shown to be hepatitis B antigen negative. Donors and staff should be aware of AIDS implications, and blood and cultures should be handled at containment level 2 (Advisory Committee on Dangerous Pathogens 1984).

Peripheral blood cultures are stimulated to divide by the addition of a T cell mitogen such as phytohemagglutinin (PHA) to the culture medium. Mitotic activity is at a maximum at about 3 days but begins at about 40 h after PHA stimulation, and

the chromosome constitution remains diploid during short-term culture (Evans and O'Riordan 1975). Treatments should commence at about 44 h after culture initiation. This is when cells are actively proliferating and cells are in all stages of the cell cycle. They should be sampled about 20 h later. In a repeat study, the second sample time should be about 92 h after culture initiation. Morimoto et al. (1983) report that the cycle time for lymphocytes averages about 12–14 h except for the first cycle.

Female donors can give higher yields of chromosome damage (Anderson 1988).

3.13.4 Positive and Negative Controls

When the solvent is not the culture medium or water, the solvent, liver enzyme activation mixture and solvent with untreated controls are used as negative controls.

Since cultured cells are normally treated in their usual growth medium, the solubility of the test material in the medium should be ascertained before testing. Extremes of pH can be clastogenic (Cifone et al. 1987), so the effect of the test material on pH should also be determined, but buffers can be utilized.

Various organic solvents are used, such as dimethyl sulfoxide (DMSO), dimethylformamide, ethanol, and acetone. The volume added must not be toxic to cells. Greater than 10% water v/v can be toxic because of nutrient dilution and osmolality changes.

A known clastogen should always be included as a positive control. When metabolic activation is used, a positive control chemical known to require metabolic activation should also be used to ensure that the system is functioning properly. Without metabolic activation, a direct-acting positive control chemical should be used. A structurally related positive control can also be used. Appropriate safety precautions must be taken in handling clastogens (IARC 1979; MRC 1981).

Positive control chemicals should be used to produce relatively low frequencies of aberrations so that the sensitivity of the assay for detecting weak clastogens can be established (Preston et al. 1987).

Aberration yields in negative and positive controls should be used to provide a historical database.

3.13.5 Treatment of Cells

When an exogenous activation system is employed, short treatments (about 2 h) are usually necessary because S9 mix is often cytotoxic when used for extended lengths of time. However, cells may be treated with chemicals either continuously up to harvest time or for a short time followed by washing and addition of fresh medium to allow cell cycle progression. Continuous treatment avoids centrifugation steps required with washing of cells and optimizes the endogenous metabolic capacity of the lymphocytes.

When metabolic activation is used, S9 mix should not exceed 1–10 percent of the culture medium by volume. It has been shown that the S9 mix is clastogenic in CHO cells and mouse lymphoma cells (Cifone et al. 1987; Kirkland et al. 1989) but not in human lymphocytes, where blood components can inactivate active oxygen species which could cause chromosome damage. When S9 mix from animals treated with other enzyme-inducing agents such as phenobarbitone/beta-naphthoflavone is used, clastogenesis may be minimized (Kirkland et al. 1989).

Prior to testing, it is necessary to determine the cytotoxicity of the test material, in order to select a suitable dose range for the chromosome assay both with and without metabolic activation. The range most commonly used determines the effect of the agent on the mitotic index (MI), i.e., the percentage of cells in mitoses at the time of cell harvest. The highest dose should inhibit mitotic activity by approximately 50% (EEC Annex V) and 75% (UKEMS: Scott et al. 1990) or exhibit some other indication of cytotoxicity. If the reduction in MI is too great, insufficient cells can be found for chromosome analysis. Cytotoxicity can also be assessed by making cell counts in the chromosome aberration test when using cell lines. In the lymphocyte assay, total white cell counts can be used in addition to MI. A dose which induces 50–75% toxicity in these assays should be accompanied by a suitable reduction in mitotic index.

If the test material is not toxic, it is recommended by, for example, the EEC (Annex V) that it be tested up to 5 mg ml^{-1}. The UKEMS recommends that chemicals be tested up to their maximum solubility in the treatment medium and not just their maximum solubility in stock solutions.

For highly soluble nontoxic agents, concentrations above 10 mM may produce substantial increases in the osmolality of the culture medium which could be clastogenic by causing ionic imbalance within the cells (Ishidate et al. 1984; Brusick 1987a, b). At concentrations exceeding 10 mM, the osmolality of the treatment media should be measured, and if the increase exceeds 50 mmol kg^{-1}, clastogenicity resulting from high osmolality should be suspected and, according to the UKEMS, is unlikely to be of relevance to human risk. The UKEMS also does not recommend the testing of chemicals at concentrations exceeding their solubility limits as suspensions or precipitate.

A minimum of three doses of the test material should be used—the highest chosen as described above, the lowest on the borderline of toxicity and an intermediate one. Up to six doses can be managed satisfactorily, and this ensures the detection of any dose response and that a toxic range is covered. MIs are as required for the preliminary study (at least 1000 cells per culture). It is also useful to score endoreduplication and polyploidy for historical data. Cells from only three doses need to be analyzed.

The range of doses used at the repeat fixation time can be those which induce a suitable degree of mitotic inhibition at the earlier fixation time, but if the highest dose reduces the MI to an unacceptably low level at the second sampling time, the next highest dose should be chosen for screening.

A complete assay requires the test material to be investigated at a minimum of three doses together with a positive (untreated) and solvent-only control which can

be omitted if tissue culture medium is used as a solvent. When two fixation times are used in repeat tests, the positive control is necessary at only one time but the negative or solvent control is necessary at both times.

Duplicates of each test group and quadruplicates of solvent or negative controls should be set up. The sensitivity of the assay is improved with larger numbers scored in the negative controls (Richardson et al. 1989).

3.13.6 Scoring Procedures

Prior to scoring, slides should be coded, randomized, and then scored "blind." Metaphase analysis should only be carried out by an experienced observer. Metaphase cells should be sought under low-power magnification and those with well-spread, i.e., nonoverlapping, clearly defined non-fuzzy chromosomes examined under high power with oil immersion. It is acceptable to analyze cells with total chromosome numbers or that have lost one or two chromosomes during processing. In human lymphocytes ($2n$–46), 44 or more centromeres and in CHO cells ($2n = 22$; range 21–24) 20 or more centromeres can be scored. Chromosome numbers can be recorded for each cell, to give an indication of aneuploidy. Only cells with increases in numbers (above 46 in human lymphocytes and 24 in CHO cells) should be considered in this category, since decreases can occur through processing.

Recording microscope coordinates of cells is necessary and allows verification of abnormal cells. A photographic record is also useful of cells with aberrations. Two hundred cells (100 from each of two replicates) should be scored per treatment group. When ambiguous results are obtained, there may be further "blind" reading of these samples.

3.13.7 Presentation of Results

The test material, test cells, method of treatment, harvesting of cells, cytotoxicity assay, etc. should be clearly stated as well as the statistical methods used. Richardson et al. (1989) recommend that comparison be made between the frequencies in control cells and at each dose level using Fisher's exact test.

In cytogenetic assays, the absence of a clear positive dose-response relationship at a particular time frequently arises. This is because a single common sampling time may be used for all doses of a test compound. Chromosome aberration yields can vary markedly with posttreatment sampling time of an asynchronous population, and increasing doses of clastogens can induce increasing degrees of mitotic delay (Scott et al. 1990). Additional fixation times should clarify the relationship between dose and aberration yield.

Gaps are by tradition excluded from quantification of chromosome aberration yields. Some gaps have been shown to be real discontinuities in DNA (e.g., Heddle

and Bodycote 1970). Where chromosome aberration yields are on the borderline of statistical significance above control values, the inclusion of gaps could be useful. Further details on this approach may be found in the UKEMS guidelines (Scott et al. 1990).

Since chromosome exchanges are relatively rare events, greater biological significance should be attached to their presence than to gaps and breaks.

Chemicals which are clastogenic in vitro at low doses are more likely to be clastogenic in vivo than those where clastogenicity is detected only at high concentrations (Ishidate et al. 1988a, b). Negative results in well-conducted in vitro tests are a good indication of a lack of potential for in vivo clastogenesis, since almost all in vivo clastogens have given positive results in vitro when adequately tested (Thompson 1986; Ishidate et al. 1988a, b).

3.14 In Vivo Cytogenetics Assays

Damage induced in whole animals can be detected in in vivo chromosome assays in either somatic or germinal cells by examination of metaphases or the formation of micronuclei. The micronucleus test can also detect whole chromosome loss or aneuploidy in the absence of clastogenic activity and is considered comparable in sensitivity to chromosome analysis (Tsuchimoto and Matter 1979).

Rats and mice are generally used for in vivo studies, with the mouse being employed for bone marrow micronucleus analysis and the rat for metaphase analysis, but both can be used for either. Mice are cheaper and easier to handle than rats, and only a qualitative difference in response has been found between the species (Albanese 1987). Chinese hamsters are also widely used for metaphase analysis because of their low diploid chromosome number of 22. However, there are few other historical toxicological data for this species.

3.14.1 Somatic Cell Assays

3.14.1.1 Metaphase Analysis

Metaphase analysis can be performed in any tissue with actively dividing cells, but bone marrow is the tissue most often examined. Cells are treated with a test compound and are arrested in metaphase by the administration of colcemid or colchicine at various sampling times after treatment. Preparations are examined for structural chromosome damage. Because the bone marrow has a good blood supply, the cells should be exposed to the test compound or its metabolites in the peripheral blood supply, and the cells are sensitive to S-dependent and S-independent mutagens (Topham et al. 1983).

Peripheral blood cells can be stimulated to divide even though the target cell is relatively insensitive (Newton and Lilly 1986). It is necessary to stimulate them with a mitogen since the number of lymphocytes which are dividing at any one time is very low. Cells are in G_0 when exposure is taking place, so they may not be sensitive to cell cycle stage-specific mutagens, and any damage might be repaired before sampling.

3.14.1.2 Micronuclei

The assessment of micronuclei is considered simpler than the assessment of metaphase analysis. This assay is most often carried out in bone marrow cells, where polychromatic erythrocytes are examined. Damage is induced in the immature erythroblast and results in a micronucleus outside the main nucleus, which is easily detected after staining as a chromatid-containing body. When the erythroblast matures, the micronucleus, whose formation results from chromosome loss during cell division or from chromosome breakage forming centric and acentric fragments, is not extruded with the nucleus. Micronuclei can also be detected in peripheral blood cells (MacGregor et al. 1980). In addition, they can be detected in liver (Tates et al. 1980; Braithwaite and Ashby 1988) after partial hepatectomy or stimulation with 4-acetylaminofluorene, or they can be detected in any proliferating cells.

The concentration given above will give relatively small increases in revertant count above the spontaneous level. There is little point in using large concentrations of reference mutagens which invariably give huge increases in revertant counts. This would give little information on the day-to-day performance of the assay.

3.15 Subacute, Subchronic and Chronic Toxicity (ISO 10993-11)

These studies can be 2 weeks long (what used to be called “subacute” studies because they were conducted at dose levels below those employed for single-dose or acute studies) or last up to a year. Another name for these studies is repeat-dose studies (Ballantyne 2009)—that is, those studies whereby animals have exposure to a device or extract from a device over a period of 1 year or less (at 9 months to a year, such studies become “chronic” – that is, for a majority of the lifetime on an animal. Currently, the duration of such general repeat-dose toxicity studies ranges from 14 days (long enough for the drug levels in the body to reach steady state and for the adaptive immune response to begin to be active) to 9 months (in nonrodents).

3.15.1 Objectives

As with any scientific study or experiment (but especially for those in safety assessment), the essential first step is to define and understand the reason(s) for the conduct of the study—that is, its objectives. There are three major (scientific) reasons for conducting subchronic and chronic studies, but a basic characteristic of all but a few subchronic studies needs to be understood. The subchronic study is (as are most other studies in whole animal toxicology) a broad screen. It is not focused on a specific end point; rather, it is a broad exploration of the cumulative biological effects of the administered agent over a range of doses. So broad an exploration, in fact, that it can be called a "shotgun" study.

The objectives of the typical subchronic and chronic studies fall into three categories. The first is to broadly identify the toxicity and other aspects of device/host interaction over a protracted period of time. Unlike the ISO 10933-6 implantation studies, effects and interactions distal from the immediate site of device implantation are of interest and concern (Traina 1983). This definition is both qualitative (what are the target organs and the nature of the effects seen) and quantitative (at what exposure levels are effects definitely seen and not seen). Unlike pharmaceutical and other types of repeat-dose studies, there is no intent nor effort to evaluate dose-response relationships.

The second is to provide support for the initiation of and/or continued conduct of one or more clinical trials in man. The duration of exposure is driven by a compromise between meeting regulatorily established guidelines and the economic pressure to initiate clinical trials as soon as possible. These studies are most often the longest to conduct prior to initiation of any clinical trial.

Chronic studies (those that last 6 or 9 months or a year) may also be conducted for implanted or other systemic contact devices, but are in reality uncommon.

3.15.2 Regulatory Considerations

Much of what is done (and how it is done) in longer than acute systemic biocompatibility studies is a response to a number of regulations. Three of these have very broad impact. These are the Good Laboratory Practices requirements, Animal Welfare Act requirements, and regulatory requirements that actually govern study design.

3.15.3 Regulatory Requirements for Study Design

The first consideration in the construction of a study is a clear statement of its objectives, which are almost always headed by meeting regulatory requirements to support drug development and registration. Accordingly, the relevant regulatory requirements must be analyzed, which is complicated by the fact that requirements

for devices, though in general governed by ISO 10993-1, may be slightly different for some regulatory bodies.

New devices are frequently not developed for registration and sale in a single-market country. The expense is too great and the potential for broad international sales too appealing. While each major country has its own requirements as to study designs and studies required (with most of the smaller countries adhering to the regulations of one of the major players), harmonization has done much to smooth these differences (Gad 2010). Agents intended to treat or arrest the progress of rapidly spreading life-threatening diseases (such as AIDS) are subject to less stringent safety assessment requirements prior to initial clinical evaluations than are other drugs. However, even though approval (if clinical efficacy is established) for marketing can be granted with preclinical testing still under way, all applicable safety assessments (as with any other class of drugs) must still be completed (FDA 1988).

3.15.4 Study Design and Conduct

3.15.4.1 Animals

Unlike pharmaceutical safety assessment, studies in only a single suitable species is required. If a rodent, most commonly the species employed is the rat. There is considerably more variability in the nonrodent species, with a range of factors determining whether the minipig, rabbit, or dog is employed. The factors that should and do govern species selection are presented in detail in Gad (2015). Numbers of animals to be used in each dose group of a study are presented in Table 3.30.

Animals are assigned to groups (test and control) by one or another form of statistical randomization. Prior to assignment, animals are evaluated for some period of time after being received in house (usually at least 1 week for rodents and 2 for nonrodents) to ensure that they are healthy and have no discernible abnormalities. The randomization is rarely pure; it is often "blocked" in some form or another (by initial body weight, at least) so that each group is not (statistically) significantly different from the others in terms of the "blocked" parameters.

Proper facilities and care for test animals are not only a matter of regulatory compliance (and a legal requirement) but also essential for a scientifically sound and valid study.

Table 3.30 Number of animals for chronic and subchronic study per test group

Study length	Rats per sex	Dogs per sex	Rabbits per sex
2–4 weeks	5	3	4
3 months[a]	20	6	8
6 months	30	8	8
1 year	50	10	10

[a]Starting with 13-week studies, one should consider adding animals (particularly to the high-dose group) to allow evaluation of reversal (or progression) of effects

Husbandry requires clean cages of sufficient size and continuous availability of clean water and food (unless the protocol requires some restriction on their availability). Environmental conditions (temperature, humidity, and light-dark cycle) must be kept within specified limits. All of these must, in turn, be detailed in the protocols of studies. The limits for these conditions are set forth in relevant NIH and USDA publications.

3.15.4.2 Parameters to Measure

As was stated earlier, repeat-dose general (systemic) toxicity studies are "shotgun" in nature; that is, they are designed to look at a very broad range of end points with the intention of screening as broadly as indications of toxicity. Meaningful findings are rarely limited to a single end point—rather, what typically emerges is a pattern of findings. This broad search for components of toxicity profile is not just a response to regulatory guidelines intended to identify potentially unsafe drugs. An understanding of all the indicators of biological effect can also frequently help one to understand the relevance of findings, to establish some as unrepresentative of a risk to humans, and even to identify new therapeutic uses of an agent.

Parameters of interest in the repeat-dose study can be considered as sets of measures, each with its own history, rationale, and requirements. Chapter 6 sought to present an overview of such parameters. It is critical to remember, however, that the strength of the study design as a scientific evaluation lies in the relationships and patterns of effects that are seen not in simply looking at each of these measures (or groups) as independent findings, but rather as integrated profiles of biological effects.

3.15.4.3 Study Designs

The traditional design for a repeat-dose toxicity study is very straightforward. The appropriate numbers of animals of each sex are assigned to each of the designated dose and control groups. Unfortunately, this basic design is taken by many to be dogma, even when it does not suit the purposes of the investigator. There are many possible variations to study design, but four basic factors should be considered: controls, the use of interval and satellite groups, balanced and unbalanced designs, and staggered starts.

Classically, a single control group of the same size as each of the dose groups is incorporated into each study. Some studies incorporate two control groups (each the same size as the experimental groups) to guard against having a statistically significant effect due to one control group being abnormal for one or more parameters (a much more likely event when laboratory animals were less genetically homogeneous than they are now). The belief is that a "significant" finding that differs from one (but not both) of the concurrent control groups, and does not differ from historical control data, can be considered as not biologically significant. This is, however,

an indefensible approach. Historical controls have value, but it is the concurrent control group(s) in a study that is of concern.

Interval or satellite groups have been discussed at two earlier points in this chapter. They allow measurement of termination parameters at intervals other than at termination of the study. They are also useful when the manipulation involved in making a measurement (such as the collection of an extensive blood sample), while not terminal, may compromise (relative to other animals) the subject animals. Another common use of such groups is to evaluate recovery from some observed effect at study termination.

Usually, each of the groups in a study is the same size, with each of the sexes being equally represented. The result is called a balanced design, with statistical power for detection of effects optimized for each of the treatment groups. If one knows little about the dose-toxicity profile, this is an entirely sound and rational approach. However, there are situations when one may wish to utilize an unbalanced design—that is, to have one or more dose groups larger than the others. This is usually the case when either greater sensitivity is desired (typically in a low-dose group), or an unusual degree of attrition of test animals is expected (usually due to mortality in a high-dose group), or as a guard against a single animal's idiopathic response being sufficient to cause "statistical significance."

As it is, the normal practice to have a balanced design, it is also traditional to initiate treatment of all animals at the same time. This may lead to problems at study termination, however. It is a very uncommon toxicology laboratory that can "bring a study down" on a single day. In fact, there are no labs that can collect blood and perform necropsies in a single day on even the 48 to 80 dogs involved in a study, much less the 160 to 400+ rats in the rodent version. Starting all animals on study the same day presents a number of less than desirable options. The first is to terminate as many animals as can be done each day, continuing to dose (and therefore, further affect) the remaining test animals. Assuming that the animals are being terminated in a random, balanced manner, this means that the last animals terminated will have received from three to ten additional days of treatment. At the least, this is likely to cause some variance inflation (and therefore both decrease the power of the study design and possibly confound interpretation). If the difference in the length of treatment of test animals is greater than 3% of the intended length of the study, one should consider alternative designs.

3.15.5 Study Interpretation and Reporting

For a successful repeat-dose study, the bottom line is the clear demonstration of a no-effect level, characterization of a toxicity profile (providing guidance for any clinical studies), enough information on pharmacokinetics and metabolism to scale dosages to human applications, and at least a basic understanding of the mechanisms involved in any identified pathogenesis. The report that is produced as a result of the study should clearly communicate these points—along with the study design

and experimental procedures, summarized data, and their statistical analysis—and it should be GLP compliant, suitable for FDA submission format.

Interpretation of the results of a study should be truly scientific and integrative. It is elementary to have the report state only each statistically and biologically significant finding in an orderly manner and not just a recitation of all observations. The meaning and significance of each in relation to other findings, as well as the relevance to potential human effects, must be evaluated and addressed.

The author of the report should ensure that it is accurate and complete, but also that it clearly tells a story and concludes with the relevant (to clinical development) findings. A useful approach is to construct a summary table (such as illustrated in Table 3.30) which gives an overview by dose group, gender, and grouping of observation. The initial use of such a table should be as soon as the "in-life" data (all but the histopathology) from main groups in studies is available, as it can serve as both a tool for early understanding of findings and a guide to what examination may be added or modified in recovery group animals.

There are some common problems encountered in general toxicity studies. The most common of these and their usual causality are presented in Table 3.31 (which table are you referring to??).

Table 3.31 Problems in biocompatibility study design and conduct

Study	Issue	Solution
All Using Extractions	Particles in extraction fluids	Let settle or centrifuge then decant and use clean fluid
All In Vitro Using Extractions	pH ≤ 2 or ≥ 11	Test solution would need to be buffered to between 5 and 9
Cytotoxicity	Tests performed with hypo- or hyperosmolarity extractant solutions	Perform tests with serial dilutions of extract solution due to level of monotensive osmolarity
All Using Extractions	Very small and expensive devices	Customized studies to minimize total amounts of extract solution required AND base volume on device weight
Implantation Studies	Device larger than can be accommodated in rats	1. Do not implant subcutaneously – rather, IM 2. Use rabbits, or if needed, dogs
Cytotoxicity	Score of 3 or 4 – that is, test failure	Perform two in vitro studies at least one of which has parenteral exposure (such as intracutaneous reactivity or implantation)
Compliment Activation	Failure	Check material components – use of surfactants in hydrophilic coating will cause failure
Ames (mutagenicity)	Device has an antimicrobial component	Use a mammalian mutagenicity assay, such as a mouse lymphoma assay
Sensitization	For US FDA	Must be GP assay – if for internal exposure device, use GPMT
Subchronic	Study not accepted by FDA due to length (less than 30 days) or incomplete data set	See section in Chapter 3 for description of complete requirements

The usual case is that over the course of drug development, we go from shorter (14 or 28 day) studies in progressive steps to longer studies (90 day/13 week) than chronic studies of 6 or 7 months. As we progress through this sequence, the results of earlier studies should modify the design of longer studies.

This chapter addresses a group of studies that have in common the facts that (1) they are intended to predict longer-term effects that occur after repeated exposure to an agent and (2) they are tests that have historically only been performed on a small subset of devices and the materials used to make them, but for which requirements have been significantly increased.

Subchronic and chronic studies for medical devices are generally in the range of being hybrids between what we are used to regarding as subchronic studies and the simple implant studies. The studies historically have been performed using only one route—implantation—with "dose" being determined in terms of how many devices or much material are implanted. And in their simplest forms, these subchronic and chronic studies are conducted as primarily modified forms of longer implantation studies with only the limited set of local issue tolerance indicators in the region of the implants being evaluated. It should also be kept in mind that it is frequent practice to combine such biocompatibility studies with evaluations of efficacy and/or device performance.

3.15.6 Objectives

As with any scientific study or experiment (but especially for those in safety assessment), the essential first step is to define and understand the reason(s) for the conduct of the study—that is, its objectives. There are three major (scientific) reasons for conducting subchronic studies, but a basic characteristic of all but a few subchronic studies needs to be understood. The subchronic study is (as are most other studies in whole animal toxicology) a broad screen. It is not focused on a specific end point; rather it is a broad exploration of the cumulative biological effects off the administered agent over a range of doses, so broad an exploration, in fact, that it can be called a "shotgun" study. The objectives of the typical subchronic device study fall into two categories. The first is to broadly define the toxicity of prolonged exposure to a medical device or medical device material in an animal model (most commonly, the rabbit). The second objective is one of looking forward to later studies. The subchronic study must provide sufficient information to allow a prudent setting of doses for later, longer studies (including, ultimately, carcinogenicity studies). At the same time, the subchronic study must also provide guidance for the other (than dose) design features of longer-term studies (such as what parameters to measure and when to measure them, how many animals to use, and how long to conduct the study). These objectives are addressed by the usual subchronic study.

Chronic studies (those that last 6 months or a year) may also be conducted for the above purposes but are primarily done to fulfill registration requirements for drugs that are intended for continuous long-term (lifetime) use or frequent intermittent use.

3.15.7 Regulatory Requirements for Study Design

The first consideration in the construction of a study is a clear statement of its objectives, which are almost always headed by meeting regulatory requirements to support device development and registration. Accordingly, the relevant regulatory requirement must be analyzed, which is complicated by the fact that new drugs are no longer developed for registration and sale in a single-market country. The expense is too great and the potential for broad international sales too appealing. Chapter 2 should be consulted for the broad overview of such regulation.

3.15.8 Study Design and Conduct

3.15.8.1 Animals

In all but a few rare cases, medical devices are evaluated for subchronic and chronic biocompatibility in only a single species. This is most often the rabbit, though the rat, dog, and hamster have also been used. The factors that should and do govern species selection are reviewed in detail in Gad and Chengelis (1998) and Gad (2015).

Except in rare cases, the animals used are young, healthy adults in the logarithmic phase of their growth curve. (The FDA specifies that rodents be less than 6 weeks of age at the initiation of dosing.)

Numbers of animals to be used in each dose group of a study are presented in Table 3.32. Though the usual practice is to use three different dose groups and at least one equal-sized control group, this number is not fixed and should be viewed as a minimum (see the section on study design later in this chapter). There must be as many control animals as are in the largest-size test group.

Animals are assigned to groups (test and control) by one or another form of statistical randomization. Prior to assignment, animals are evaluated for some period of time after being received in-house (usually at least 1 week for rodents and two for nonrodents) to ensure that they are healthy and have no discernible abnormalities. The randomization is never pure; it is always "blocked" in some form or another (by initial body weight, at least) so that each group is not (statistically) significantly different from the others in terms of the "blocked" parameters (usually initial body weight).

Proper facilities and care for test animals are not only a matter of regulatory compliance (and a legal requirement) but also essential for a scientifically sound and valid study. Husbandry requires clean cages of sufficient size and continuous availability of clean water and food (unless the protocol requires some restriction on their availability). Environmental conditions (temperature, humidity, and light-dark cycle) must be kept within specified limits. All of these must, in turn, be detailed in the protocols of studies. The limits for these conditions are set forth in relevant NIH and USDA publications.

Table 3.32 Number of animals for chronic and subchronic study per test group

Clinical chemistry	Hematology	Urinalysis
Albumin	Erythrocyte count (RBC)	Chloride
Alkaline phosphatase (ALP)	Hemoglobin (HGB)	Bilirubin
Blood urea nitrogen (BUN)	Hematocrit (HCT)	Glucose
Calcium Chloride	Mean corpuscular Hemoglobin (MCH)	Ketone Osmolality
Creatine	Mean corpuscular	Occult blood
Creatine phosphokinase (CPK)	Volume (MCV)	pH
Direct bilirubin	Platelet count	Phosphorus
Gamma-glutamyltransferase (GGT)	Prothrombin time	Potassium
Globulin	Reticulocyte count	Protein
Glucose	White cell count (WBC)	Sodium
Lactic dehydrogenase (LDH)	White cell differential count	Specific gravity
Phosphorus		Volume
Potassium		
Serum glutamic-oxaloacetic transaminase (SGOT)		
Serum glutamic-pyruvic transaminase (SGPT)		
Sodium		
Total bilirubin		
Total cholesterol		
Total protein		
Triglycerides		

3.15.9 Parameters to Measure

As was stated earlier, subchronic studies are usually "shotgun" in nature, that is, they are designed to look at a very broad range of end points with the intention of screening as broadly as indications of toxicity. Meaningful findings are rarely limited to a single end point—rather, what typically emerges is a pattern of findings. This broad search for components of a toxicity profile is not just a response to regulatory guidelines intended to identify potentially unsafe drugs. An understanding of all indicators of biological effect can also frequently help one to understand the relevance of findings, to establish some as unrepresentative of a risk to humans, and even to identify new therapeutic uses of an agent.

Parameters of interest in the repeat-dose study can be considered as sets of measure, each with its own history, rationale, and requirements. It is critical to remember, however, that the strength of the study design as a scientific evaluation lies in the relationships and patterns of effects that are seen not in simply looking at each of these measures (or groups) as independent findings, but rather as integrated profiles of biological effects.

3.15.9.1 Body Weight

Body weight (and the associated calculated parameter of body weight gain) is a nonspecific, broad screen for adverse systemic toxicity. Animals are initially assigned to groups based on a randomization scheme that includes having each group vary insignificantly from one another in terms of body weight. Weights are measured prior to the initial dose, then typically 1–3, 5, 7, 11, and 14 days thereafter. The frequency of measurement of weights goes down as the study proceeds after 2 weeks, and weighting is typically weekly through 6 weeks, then every other week through 3 months, and monthly thereafter. Because the animals used in these studies are young adults in the early log phase of their growth, decreases in the rate of gain relative to control animals are a very sensitive (albeit nonspecific) indicator of systemic toxicity.

3.15.9.2 Food Consumption

Food consumption is typically measured with one or two uses in mind. First, it may be explanatory in the interpretation of reductions (either absolute or relative) in body weight. In cases where administration of the test compound is via diet, it is essential to be able to adjust dietary content so as to accurately maintain dose levels. Additionally, the actual parameter itself is a broad and nonspecific indicator of systemic toxicity. Food consumption is usually measured over a period of several days, first weekly and then on a once-a-month basis. Water consumption, which is also sometimes measured, is similar in interpretation and use.

3.15.9.3 Clinical Signs

Clinical signs are generally vastly underrated in value, probably because insufficient attention is paid to care in their collection. Two separate levels of data collection are actually involved here. The first is the morbidity and mortality observation, which is made twice a day. This generally consists of a simple cage-side visual assessment of each animal to determine if it is still alive and, if so, whether it appears in good (or at least stable) health. Historically, this regulatorily required observation was intended to ensure that tissues from intoxicated animals were not lost for meaningful histopathologic evaluation due to autolysis (Arnold et al. 1990).

The second level of clinical observation is the detailed hands-on examination analogous to the human physical examination. It is usually performed against a checklist (see Gad and Chengelis 1998, for an example), and evaluation is of the incidence of observations of a particular type in a group of treated animals compared to controls. Observations range from being indicative of nonspecific systemic toxicity to fairly specific indicators of target organ toxicity. These more detailed observations are typically taken after the first week of a study and on a monthly basis thereafter.

Ophthalmologic examinations are typically made immediately prior to initiation of a study (and thus serve to screen out animals with preexisting conditions) and toward the end of a study.

Particularly when the agent under investigation either targets or acts via a mechanism likely to have a primary effect on a certain organ for which functional measures are available, an extra set of measurements of functional performance should be considered. The organs or organ systems that are usually of particular concern are the kidneys, liver, and cardiovascular, nervous, and immune system. Special measures (such as creatinine clearance as a measure of renal function) are combined with other data already collected (organ weights, histopathology, clinical pathology, etc.) to provide a focused "special" investigation or evaluation of adverse effects on the target organ system of concern.

3.15.9.4 Clinical Pathology

Clinical pathology covers a number of biochemical and morphological evaluations based on invasive and noninvasive sampling of fluids from animals that are made periodically during the course of a subchronic study. These evaluations are sometimes labeled as clinical (as opposed to anatomical) pathology determinations. Table 3.32 presents a summary of the parameters measured under the headings of clinical chemistry, hematology, and urinalysis, using samples of blood and urine collected at predetermined intervals during the study. Conventionally, these intervals are typically at three points evenly spaced over the course of the study, with the first being 1 month after study initiation and the last being immediately prior to termination of the test animals. For a 3-month study, this means that samples of blood and urine would be collected at 1, 2, and 3 months after study initiation (i.e., after the first day of dosing of the animals). There are some implications of these sampling plans that should be considered when the data are being interpreted. Many of the clinical chemistry (and some of the hematologic) markers are really the result of organ system damage that may be transient in nature (see Table 3.35 for a summary of interpretations of clinical chemistry findings and Table 3.33 for a similar summary for hematological findings). The samples on which analysis is performed are from fixed points in time, which may miss transient changes (typically, increases) in some enzyme levels.

3.15.10 Histopathology

Histopathology is generally considered the single most significant portion of data to come out of a repeat-dose toxicity study. It actually consists of three related sets of data (gross pathology observations, organ weights, and microscopic pathology) that are collected during the termination of the study animals. At the end of the study, a number of tissues are collected during termination of all surviving animals

Table 3.33 Association of changes in biochemical parameters with actions at particular target organs

Organ system									
Parameter	Blood	Heart	Lung	Kidney	Liver	*Bone*	Intestine	Pancreas	Notes
Albumin				↓	↓				Produced by the liver. Very significant reductions indicate extensive liver damage3
ALP (alkaline phosphatase)					↑	↑	↑		Elevations usually associated with cholestasis. Bone alkaline phosphatase tends to be higher in young animas
Bilirubin (total)	↑				↑				Usually elevated due to cholestasis either due to obstruction or hepatopathy
BUN (blood urea nitrogen)				↑	↓				Estimates blood-filtering capacity of the kidneys. Doesn't become significantly elevated until kidney function is reduced 60–75%
Calcium				↑					Can be life-threatening and result in acute death
Cholinesterase				↑	↓				Found in plasma, brain, and RBC
CPK (creatinine phosphokinase)		↑							Most often elevated due to skeletal muscle damage but can also be produced by cardiac muscle damage. Can be more sensitive than histopathology
Creatine				↑					Also estimates blood-filtering capacity of kidney as BUN does. More specific than BUN

(continued)

Table 3.33 (continued)

Organ system									
Parameter	Blood	Heart	Lung	Kidney	Liver	*Bone*	Intestine	Pancreas	Notes
Glucose								↑	Alterations other than those associated with stress are uncommon and reflect an effect on the pancreatic islets or anorexia
GGT (gamma-glutamyltransferase)					↑				Elevated in cholestasis. This is a microsomal enzyme and levels often increase in response to microsomal enzyme induction
HBDH (hydroxybutyric dehydrogenase)		↑			↑				Most prominent in cardiac muscle tissue
LDH (lactic dehydrogenase)		↑	↑	↑	↑				Increase usually due to skeletal muscle, cardiac muscle, and liver damage. Not very specific unless isozymes are evaluated
Protein (total)				↑	↑				Absolute alterations are usually associated with decreased production (liver) or increased loss (kidney)
SGOT (serum glutamic-oxaloacetic transaminase); also called AST (aspirate amino transferase)		↑		↑	↑			↑	Present in skeletal muscle and heart and most commonly associated with damage to these
SGPT (serum glutamic-pyruvic transaminase; also called ALT (alanine aminotransferase)					↑				Elevations usually associated with hepatic damage or disease
SDH (sorbitol dehydrogenase)					↑ or ↓				Liver enzyme which can be quite sensitive but is fairly unstable
									Samples should be processed as soon as possible

Table 3.34 Tissues for histopathology

Adrenals[a]	Mainstream bronchi
Body and cervix	Major salivary gland
Brain, all three levels[a]	Mesenteric lymph nodes
Cervical lymph nodes	Ovaries and tubes
Cervical spinal cord	Pancreas
Duodenum	Pituitary
Esophagogastric junction	Prostate
Esophagus	Skeletal muscle from proximal hind limb
Eyes with optic nerves	Spleen[a]
Femur with marrow	Sternbrae with marrow
Heart	Stomach
Ileum	Testes with epididymides[a]
Kidneys[a]	Thymus and mediastinal contents[a]
Large bowel	Thyroid with parathyroid[a]
Larynx with thyroid and parathyroid	Trachea
Liver[a]	Urinary bladder
Lungs[a]	Uterus including horns

[a]Organs to be weighed

(test and control). Organ weights and terminal body weights are recorded at study termination, so that absolute and relative (to body weight) values can be statistically evaluated.

These tissues, along with the organs for which weights are determined, are listed in Table 3.34. All tissues collected are typically processed for microscopic observation, but only those from the high-dose and control groups are necessarily evaluated microscopically. If a target organ is discovered in the high-dose group, then successively lower-dose groups are examined until a "clean" (devoid of effect) level is discovered (Haschek and Rousseaux 1991) (Table 3.35).

In theory, all microscopic evaluations should be performed blind (without the pathologist knowing from which dose group a particular animal came), but this is difficult to do in practice, and such an approach frequently degrades the quality of the evaluation. Like all the other portions of data in the study, proper evaluation benefits from having access to all data that addresses the relevance, severity, timing, and potential mechanisms of a specific toxicity. Blind examination is best applied in peer review or consultations on specific findings.

In addition to the "standard" set of tissues specified in Table 3.33, observations during the course of the study or in other or previous studies may dictate that additional tissues be collected or special examinations (e.g., special stains, polarized light or electron microscopy, immunocytochemistry, or quantitative morphometry) be undertaken to evaluate the relevance of, or understand the mechanisms underlying, certain observations.

Histopathology testing is a terminal procedure, and, therefore, sampling of any single animal is a one-time event (except in the case of a tissue collected by biopsy).

Table 3.35 Some probable conditions affecting hematological changes

Parameter	Elevation	Depression
Red blood cells	1. Vascular shock 2. Excessive diuresis 3. Chronic hypoxia 4. Hyperadrenocorticism	1. Anemia 2. Blood loss 3. Hemolysis 4. Low RBC production
Hematocrit	1. Increased RBC 2. Stress 3. Shock (a) Trauma (b) Surgery 4. Polycythemia	1. Anemias 2. Pregnancy 3. Excessive hydration
Hemoglobin	1. Polycythemia (increase in production of RBC)	1. Anemias 2. Lead poisonings
Mean cell volume	1. Anemias 2. B-12 deficiency	1. Iron deficiency
Mean corpuscular hemoglobin	1. Reticulocytosis	1. Iron deficiency
White blood cells	1. Bacterial infections 2. Bone marrow stimulation	1. Bone marrow depression 2. Cancer chemotherapy 3. Chemical intoxication 4. Splenic disorders
Platelets		1. Bone marrow depression 2. Immune disorder
Neutrophilis	1. Acute bacterial infections 2. Tissue necrosis 3. Strenuous exercise 4. Convulsions 5. Tachycardia 6. Acute hemorrhage	1. Viral infections
Lymphocytes	1. Leukemia 2. Malnutrition 3. Viral infections	
Monocytes	1. Protozoal infections	
Eosinophils	1. Allergy 2. Irradiation 3. Pernicious anemia 4. Parasitism	
Basophils	1. Lead poisoning	

Because it is a regulatory requirement that the tissues from a basic number of animals be examined at the stated end of the study, an assessment of effects at any other time course (most commonly, to investigate recovery from an effect found at study termination) requires that satellite groups of animals be incorporated into the study at start-up. Such animals are randomly assigned at the beginning of the study and otherwise treated exactly the same as the equivalent treatment (or control) animals.

Components of subacute or chronic response:

- Mononuclear inflammatory cells (lymphocytes, macrophages, plasma cells)
- Epithelioid or giant cells
- Fibroplasia or fibrosis
- Measure width of reactive zone
- Score on scale of 0 (not present) to 5 (extreme)

Components of acute inflammatory response:

- Inflammatory cells (polymorphonuclear leukocytes)
- Necrosis
- Hemorrhage
- Fibrin/serum

3.15.11 Study Interpretation and Reporting

For a successful repeat-dose study, the bottom line is the clear demonstration of a no-effect level, characterization of a toxicity profile (providing guidance for any clinical studies), and at least a basic understanding of the mechanisms involved in any identified pathogenesis. The report that is produced as a result of the study should clearly communicate these points—along with the study design and experimental procedures, summarized data, and their statistical analysis—and it should be GLP compliant, suitable for FDA submission format.

Interpretation of the results of a study should be truly scientific and integrative. It is elementary to have the report state only each statistically and biologically significant finding in an orderly manner. The meaning and significance of each in relation to other findings, as well as the relevance to potential human effects, must be evaluated and addressed.

The author of the report should insure that it is accurate and complete but also that it clearly tells a story and concludes with the relevant (to clinical development) findings.

3.16 Hemocompatibility (ISO-10993-4)

Hemocompatibility—a lack of significant adverse interactions of a device with the formed elements of the blood—can be one of the most complex of the standard safety concerns for devices to be evaluated. Properly done for a long-term cardiovascular implant device (such as a stent) as an independent entity, it could also be the most expensive of the standard, short-term responses end point to evaluate. ISO 10993 Part 4 (*Selection of Tests for Interactions with Blood*) presents 25 different categories of assays for such evaluations, and FDA expectations have become yet

more extensive. While in vitro methods have been used for screening materials (Motlagh et al. 2006), they have severe limitations.

The composition of blood is:

- 55% fluid elements
  (a) Plasma (91% H_2O)
  (b) 7% dissolved proteins
  (c) 45% globulins
  (d) 7% fibrinogen, trace proteins
  (e) 2% other stuffs
- 45% formed elements
  (a) Red blood cells 5,000,000/μL
  (b) Platelets 300,000/μL
  (c) White blood cells 7000/μL

Few materials have consistently shown good hemocompatibility in both arterial and venous blood flow environments. Because of its complexity (Beutler et al. 1982; Collman et al. 2006) results obtained from laboratory animals may not apply to man, and results from one test system may not necessarily be correlated to those obtained from a different test system. In vitro results may not predict well what happens in vivo (Didisheim et al. 1984; Lindon et al. 1978). Any hemocompatibility statement must be linked to the intended use and conditions for which the statement is valid.

Blood-material interaction can range from transient hemolysis and minimal protein adsorption to activation of coagulation, complement, and significant destruction of cells. Complicated mechanisms exist in the cardiovascular system which may interact with medical devices. Devices vary enormously in type, function, and duration of blood contact (Cooper et al. 1987; Dewangee 1987), particularly now that combination and nanotechnology devices are increasingly moving to market. Therefore, a multidisciplinary approach to hemocompatibility testing is important. This includes in vitro static and dynamic tests, acute extracorporeal tests, tests of cardiovascular devices in appropriate animal models, and clinical studies. For most devices only the in vitro static tests are performed. Complex interactions are operative between the surfaces of devices/materials and the blood, based on both chemical and physical parameters (Zaslavsky et al. 1978), such as the rate of release of chemical moieties from a device and the nature of the blood contacting device surface.

A thorough review of the normal function and structure of the hematopoietic system is a discipline in itself and far beyond the range of this chapter. Irons (1985), Brown (1993), and Williams et al. (1995) should be consulted by those interested in the background. Figures 3.6 and 3.7 provide a very rudimentary overview of the pathways involved in the generation and differentiation of the formed elements of the blood.

The FDA and ISO requirements for hematocompatability evaluation for most devices that were not implants in the vascular system (but rather have limited duration on contact) are frequently met in the part by performing a simple in vitro hemolysis test. The ASTM guidelines (ASTM 2010) and NIH (1985) called for a more

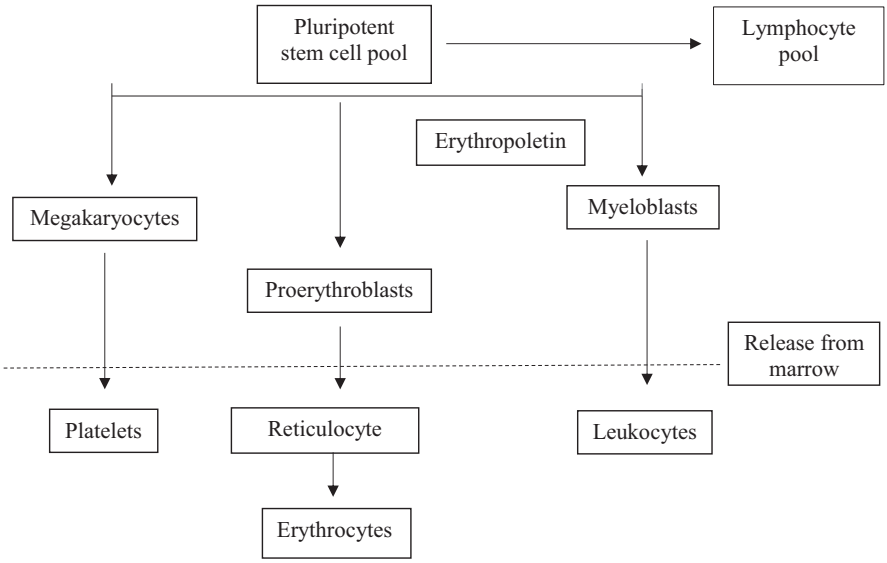


Fig. 3.6 Differentiation of formed blood elements

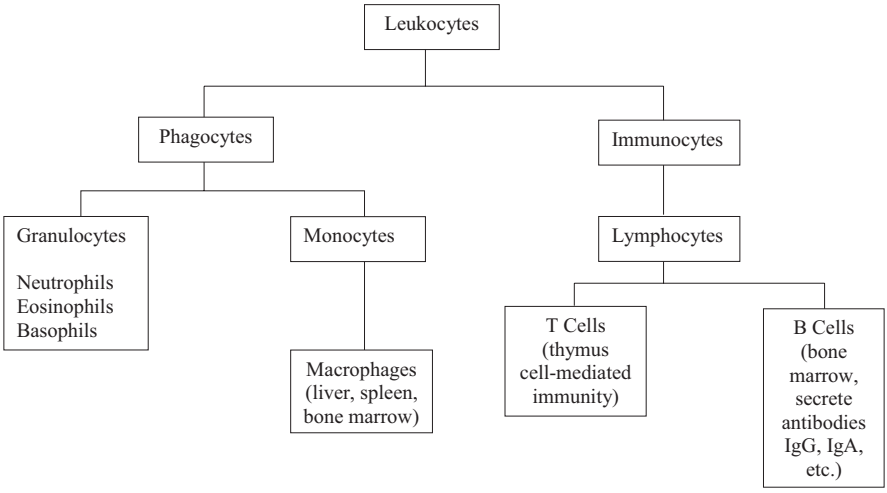


Fig. 3.7 Development of formed blood cells from plaripotent stem cells

stringent approach. Under ISO-10993-4, externally communicating devices, with indirect or circulating contact with the bloodstream, or implant devices in the vascular system must be evaluated. It is recommended that this evaluation look at five different end points (test categories: thrombosis, coagulation, platelets, hematology, and immunology). Tables 3.36 and 3.37 summarize representation tests available to evaluate each of these end points.

Table 3.36 Abbreviations

Abbreviation	Meaning
Bb	Product of alternate pathway complement activation
β-TG	Beta-thromboglobulin
C-4d	Product of classical pathway complement activation
C-3a, C5a	(Active) complement split products from C3 to C5
D-dimer	Specific fibrin degradation products (F XIII cross-linked fibrin)
ECMO	Extracorporeal membrane oxygenator
EM	Electron microscopy
FDP	Fibrin/fibrinogen degradation products
FPA	Fibrinopeptide A
F_{1+2}	Prothrombin activation fragment 1 + 2
iC3b	Product of central C complement activation
IL-1	Interleukin-1
IVC	Inferior vena cava
MRI	Magnetic resonance imaging
PAC-1	Monoclonal antibody which recognizes the activated form of platelet surface glycoprotein IIb/IIIa
PET	Positron emission topography
PF-4	Platelet factor 4
PT	Prothrombin time
PTT	Partial thromboplastin time
RIA	Radioimmunoassay
S-12	Monoclonal antibody which recognizes the alpha granule membrane component GMP140 exposed during the platelet release reaction
SC5b-9	Product of terminal pathway complement activation
TAT	Thrombin-antithrombin complex
TCC	Terminal complement complex
TT	Thrombin time
VWF	von Willebrand factor

Each of these categories, of course, is a potential type of interaction between the blood and materials used in devices.

Devices contacting and therefore potentially having an interaction with the blood are categorized by ISO as follows:

3.16.1 Non-contact Devices

An example is in vitro diagnostic devices, which have no biocompatibility testing requirements.

Table 3.37 ISO 10993-1 (2018) Test Requirement Chart

Medical device categorization by			End points of biological evaluation															
Nature of body contact		Contact duration																
Category	Contact	A—Limited (≤24 h) B—Prolonged (>24 h to 30 d) C—Long term (>30 d)	Physical and/or chemical information	Cytotoxicity	Sensitization	Irritation or intracutaneous reactivity	Material-mediated pyrogenicity[a]	Acute systemic toxicity[b]	Subacute toxicity[b]	Subchronic toxicity[b]	Chronic toxicity[b]	Implantation effects[b,c]	Hemocompatibility	Genotoxicity[d]	Carcinogenicity[d]	Reproductive/developmental toxicity[d,e]	Degradation[f]
Surface medical device	Intact skin	A	Xg	E[h]	E	E											
		B	X	E	E	E											
		C	X	E	E	E											
	Mucosal membrane	A	X	E	E	E											
		B	X	E	E	E		E	E			E					
		C	X	E	E	E		E	E	E	E	E		E			
	Breached or compromised surface	A	X	E	E	E	E	E									
		B	X	E	E	E	E	E	E			E					
		C	X	E	E	E	E	E	E	E	E	E		E	E		
Externally communicating medical device	Blood path, indirect	A	X	E	E	E	E	E					E				
		B	X	E	E	E	E	E	E				E				
		C	X	E	E	E	E	E	E	E	E	E	E	E	E		
	Tissue/bone/dentin	A	X	E	E	E	E	E									
		B	X	E	E	E	E	E	E			E		E			
		C	X	E	E	E	E	E	E	E	E	E		E	E		
	Circulating blood	A	X	E	E	E	E	E					E	E[j]			
		B	X	E	E	E	E	E	E			E	E	E			
		C	X	E	E	E	E	E	E	E	E	E	E	E	E		

Implant medical device	Tissue/bone[i]	A	X	E	E	E	E	E									
		B	X	E	E	E	E	E	E			E		E			
		C	X	E	E	E	E	E	E	E	E	E		E	E		
	Blood	A	X	E	E	E	E	E				E	E	E			
		B	X	E	E	E	E	E	E			E	E	E			
		C	X	E	E	E	E	E	E	E	E	E	E	E	E		

[a]Refer to ISO 10993-ww: 2017, Annex F

[b]Information obtained from comprehensive implantation assessments that include acute systemic toxicity, subacute toxicity, subchronic toxicity, and/or chronic toxicity may be appropriate if sufficient animals and timepoints are included and assessed. It is not always necessary to perform separate studies for acute, subacute, subchronic, and chronic toxicity

[c]Relevant implantation sites should be considered. For instance, medical devices in contact with intact mucosal membranes should ideally be studied/considered in contact with intact mucosal membranes

[d]If the medical device can contain substances known to be carcinogenic, mutagenic, and/or toxic to reproduction, this should be considered in the risk assessment

[e]Reproductive and developmental toxicity should be addressed for novel materials, materials with a known reproductive or developmental toxicity, medical devices with relevant target populations (e.g. pregnant women), and/or medical devices where there is the potential for local presence of deice materials in the reproductive organs

[f]Degradation information should be provided for any medical devices, medical device components, or materials remaining within the patient that have the potential for degradation

[g]X means prerequisite information needed for a risk assessment

[h]E means end points to be evaluated in the risk assessment (either through the use of existing data, additional end point-specific testing, or a rationale for why assessment of the end point does not require an additional data set). If a medical device is manufactured from novel materials, not previously used in medical device applications, and no toxicology data exists in the literature, additional end points beyond those marked “E” in this table should be considered. For particular medical devices, there is a possibility that it will be appropriate to include additional or fewer end points than indicated

[i]Tissue includes fluids and subcutaneous spaces. For gas pathway devices or components with only indirect tissue contact, see device-specific standards for biocompatibility information relevant to these medical devices

[j]For all medical devices used in extracorporeal circuits

3.16.2 External Communicating Devices

These are devices that contact the circulating blood and serve as a conduit into the vascular system. Examples include but are not limited to those below.

External communicating devices that serve as an indirect blood path include but are not limited to:

- Cannulae
- Extension sets
- Devices for the collection of blood
- Devices for the storage and administration of blood and blood products (e.g., tubing, needles, and bags)

Indirect blood path devices are assigned the simplest testing strategy by ISO. A profile of six, relatively inexpensive, in vitro tests is recommended, one test for thrombosis, one for coagulation, one for platelet count, two hematology tests, and a complement activation panel for immunology. Optional tests may also be required by the regulatory authority and this point should be clarified with them before a testing program is initiated.

External communicating devices in contact with circulating blood include but are not limited to:

- Cardiopulmonary bypass
- Extracorporeal membrane oxygenators
- Hemodialysis equipment
- Donor and therapeutic apheresis equipment
- Devices for absorption of specific substances from blood
- Interventional cardiology and vascular devices
- Percutaneous circulatory support systems
- Temporary pacemaker electrodes

Circulating blood devices are assigned a somewhat more complex testing strategy, reflecting the fact that circulating blood must blow through the device; hence device patency becomes an issue. Again, tests from the five basic categories are recommended by ISO. Additional, optional tests are also listed.

3.16.3 Implant Devices

These are devices that are placed largely or entirely within the vascular system. Examples include but are not limited to:

- Mechanical or tissue heart valves
- Prosthetic or tissue vascular grafts
- Circulatory support devices (ventricular assist devices, artificial hearts, intra-aortic balloon pumps)

- Inferior vena cava filters
- Stents
- Arteriovenous shunts
- Blood monitors
- Internal drug delivery catheters
- Pacemaker electrodes
- Intravascular membrane oxygenators (artificial lungs)

The ISO guidelines also provide detailed guidance as to tests to be performed for each of these types of devices. This guidance is summarized in Tables 3.38, 3.39, 3.40, 3.41, and 3.42. Table 3.36 provides a codex for the significant aberrations utilized in Tables 3.38, 3.39, 3.40, 3.41, and 3.42. Not covered in these tables are the specialized cases associated with cardiovascular devices.

3.16.4 Standard Tests

Among the wide range of tests described in Tables 3.38, 3.39, 3.40, 3.41, and 3.42, there are a number which are most commonly performed. These are available at most contract research organizations and hospitals and include (besides the simple

Table 3.38 External communicating devices—level 1—blood path, indirect

Test category	**Method**	**Comments**
Thrombosis	Light microscopy (adhered platelets, leukocytes, aggregates, erythrocytes, fibrin, etc.)	Light microscopy can be replaced by scanning EM if the nature of the material presents technical problems for light microscopy
Coagulation	PTT (non-activated)	
Platelets	Platelet count	
Hematology	Leukocyte count and differential: hemolysis (plasma hemoglobin)	Hemolysis is regarded as an especially significant screening test to perform in this category because of its measurement of red blood cell membrane fragility in contact with materials and devices. The method used should be one of the normative standard test methods for hemolysis
Immunology	C3a, C5a, TCC, Bb, iC3b, C4d, SC5b-9	A panel including the last four tests encompasses the various complement activation pathways
Level 2 (optional)		
Thrombosis	**Scanning EM**	
Coagulation	Coagulation factor assays, includingFPA, D-dimer, F_{1+2}, PAC-1, s-12, TAT	
Platelets	PF-4, β-TG, thromboxane B2, ^{111}In-labelled platelet survival	

Table 3.39 External communicating devices level 1: circulating blood

Test category	Method	Comments
Thrombosis	Percent occlusion	Light microscopy can be replaced by scanning EM if the nature of the material presents technical problems for light microscopy
	Flow reduction	
	Gravimetric analysis (thrombus mass)	
	Light microscopy (adhered platelets, leukocytes, aggregates, erythrocytes, fibrin, etc.)	Pressure drop not recommended for devices intended for PR
	Pressure drop across device	
Coagulation	PTT (nonactivated)	
Platelets	Platelet count	
	Platelet aggregation	
	Template bleeding time	
Hematology	Leukocyte count and differential: hemolysis (plasma hemoglobin)	Hemolysis is regarded as an especially significant screening test to perform in this category because of its measurement of red blood cell membrane fragility in contact with materials and devices. The method used should be one of the normative standard test methods for hemolysis
Immunology	C3a, C5a, TCC, Bb, iC3b, C4d, SC5b-9	A panel including the last four tests encompasses the various complement activation pathways

Table 3.40 External communicating devices level 2: optional

Test category	Method	Comments
Thrombosis	Scanning EM (platelet adhesion and aggregation; platelet and leukocyte morphology; fibrin)	
Coagulation	Specific coagulation factor assays; FPA, D-dimer, F_{1+2}, PAC-1, S-12, TAT	
Platelets	PF-4, βthromboxane B2; gamma imaging of radiolabeled platelets^{111}In-labeled platelet survival	^{111}In-labeling is recommended for PR only
Hematology	Reticulocyte count; activation-specific release products of peripheral blood cells (i.e., granulocytes)	
Immunology	C3a, C5a, TCC, bb, iC3b, C4d, SC5b-9	A panel including the last four tests encompasses the various complement activation pathways

tests presented here) determination of the numbers and types of formed elements of the blood (Lewis et al. 1990) and other end 3.39.

The surfaces of polymers and ceramics (Yokoyama et al. 1986) may require more specialized equipment and approaches than are generally available.

Table 3.41 Implant devices: level 1

Test	Method	Comments
Thrombosis	Percent occulsion	
	Autopsy of device (gross and microscopic)	
	Autopsy of distal organs (gross and microscopic)	
Coagulation	PTT (nonactivated), PT, TT	
	Plasma fibrinogen, FDP	
Platelets	Platelet count	
	Platelet aggregation	
Hematology	Leukocyte count and differential	Hemolysis is regarded as an especially significant screening test to perform in this category because of its measurement of red blood cell membrane fragility in contact with materials and devices. The method used should be one of the normative standard test methods for hemolysis
	Hemolysis (plasma hemoglobin)	
Immunology	C3a, C5a, TCC, Bb, iC3b, C4d, SC5b-9	A panel including the four tests encompasses the various complement activation pathways

Table 3.42 Implant devices level 2: optional

Test category	Method	Comments
Thrombosis	Scanning EM	
	Angiography	
Coagulation	Specific coagulation factor assays	
	FPA, D-dimer, F_{1+2}, PAC-1, S-12, TAT	
Platelets	^{111}In = labeled platelet survival PF-$, βthromboxane B2	
	Gamma imaging of radiolabeled platelets	
Hematology	Reticulocyte count; activation specific release products of peripheral blood cells (i.e., granulocytes)	
Immunology	IL-1 and other cytokines; detection of messenger RNA-specific for cytokines	

3.16.5 Hemolysis Tests

The simplest and most commonly conducted hemocompatibility test is the in vitro homolysis test. In the direct contact hemolysis test, the intact test article is placed in a solution of saline, and a small amount of whole red blood is added. After a period of incubation (one hour) at 37 ° C, the supernate is decanted and assayed for hemoglobin. The concentration of hemoglobin is proportional to the number of red cells that were lysed.

The percent hemolysis is calculated by the equation below:

$$\text{Percent Hemolysis} = \frac{A - B}{C} \times 100$$

where *A* is the absorbance of the test sample, *B* is the absorbance of a negative control, and *C* is the absorbance of a positive control.

Hemolysis tests evaluate the acute in vitro hemolytic properties of materials, especially those intended for use in contact with blood. The concentration of substances which produces hemolysis is generally higher than that needed to produce a cytotoxic effect. The result of hemolysis testing can be correlated with acute in vivo toxicity tests. A hemolysis test is rapid, requires simple equipment, gives easily interpretable quantitative results, and can be performed in the presence of the material or on the extract. The results are compared to the controls and expressed as percent hemolysis.

The average life-span of human red cells is 120 days. When the life-span is shortened, whatever the cause, there is said to be a hemolytic process, and when the marrow fails to replace the lost cells quickly enough, then a hemolytic anemia reduction in red blood cell count develops.

The term "hemolytic" is rather misleading, as it implies actual lysis or bursting of the red cells in the circulation. Sometimes this does occur, and then it is known as intravascular hemolysis. More often, however, the cells are damaged or in some way inadequate and are then removed from the circulating blood by macrophages in the spleen in the usual way. This process is known as extra vascular hemolysis. One of the most common tests for hemolysis is the osmotic fragility test.

The membranes surrounding most cells in animal tissues are semipermeable, which means that they allow the passage of water, but prevent the passage of dissolved substances. When two solutions of different concentration are separated by such a membrane, water passes from the more dilute solution to the more concentrated one, until the concentration on both sides is equal. This tendency for water to flow in one direction is called osmosis, and the pressure exerted as it does so is known as the osmotic pressure.

Two solutions of equal concentration are known as "isotonic." For human blood this is equivalent to 0.9% saline solution, when their concentrations are unequal, the more dilute solution is "hypotonic" (hypo = low) and the more concentrated solution is "hypertonic" (hyper = over, above).

If red cells are to retain their shape and function properly, the hemoglobin solution inside the cell and the plasma outside the cell must be isotonic. Similarly, when dealing with red cells in the laboratory, solutions must be isotonic with the contents of the cell.

3.16.6 The Osmotic Fragility Test

The degree of hemolysis of cells in hypotonic solutions depends largely on their shape. Cells which are already spherocytic (spherical) are easily lysed, that is, they are more fragile. Flattened cells on the other hand are more resistant to lysis than normal cells.

A series of solutions are prepared containing salt concentrations from 0.3 g to 0.6 g per 100 ml. A volume of blood is then added to each of these, and after half an hour, the degree of hemolysis is found by spinning down the intact red cells and measuring the color intensity of the supernatant. A normal control blood must always be treated in the same way for the purpose of comparison and to check the quality of the reagents (Kirk et al. 1975). The process proceeds as follows:

3.16.6.1 Method

- Place two rows of seven 10-ml tubes in a rack, and label one row T1–T7 (test) and the other row C1–C7 (control).
- Label another four tubes T. Std., T. Blank, C. Std., and C. Blank. Place these in the appropriate places in the rack.
- Using one row only, set up tubes as below.
- Mix the solutions carefully, then transfer 5 ml from each tube to the corresponding tube of the second row. This ensures that each concentration is the same for test and control.
- To the test row, add 0.05 ml of control blood in the same way.
- Mix all the tubes and allow them to stand at room temperature for 30 min.
- Mix them again, and then centrifuge all the tubes at 3000 rev/min for 5 min.
- Using the appropriate blank, read the color intensities of the test row in a colorimeter, using an Ilford 625 green filter. Take care not to disturb the red cell layer when transferring the clear supernatants to the cuvettes for reading.
- Repeat the procedure with the control row, taking care to change the blank.

3.16.6.2 Calculation and Results

One now has two sets of eight readings, including the standards. The standards, which are cells in water, represent 100% hemolysis. They are different from each other, because the hemoglobin levels of the two bloods are different.

Using one set of readings, calculate the percentage hemolysis in each tube.

$$\text{Percentage hemolysis} = \frac{\text{Test reading}}{\text{Standard reading}} \times 100$$

Calculate the second set of results in the same way. The results are best expressed in the form of a graph showing percentage hemolysis against the NaCl concentrations.

3.16.6.3 Factors Influencing the Results

- The blood must be as fresh as possible, preferably less than 4 hours old. Cells deteriorate on standing and begin to lyse spontaneously.

- Defibrinated or heparinized blood is most suitable for this test, as such blood would not contain any extra salts. Blood anticoagulated with EDTA is used also.
- The ratio of blood to saline affects the results; a ratio of 1:100 is usually used.
- The pH of the saline also affects the results. In order to standardize this, a stock solution of buffered saline may be prepared which is osmotically equivalent to 10% NaCl:

(a) Sodium chloride (NaCl)	178 g
(b) Disodium hydrogen phosphate ($Na_2PO_4.2H_2O$)	27.21 g
(c) Sodium dihydrogen phosphate ($NaH_2PO_4.2H_2$)	4.86 g
(d) Distilled water to 2 liters	

This solution is diluted 1 in 10 just before use.

3.16.6.4 Normal Range

Slight hemolysis:	0.45–0.4 g% NaCl
Complete hemolysis:	0.35–0.3 g% NaCl

3.16.7 Erythrocyte Stability

The erythrocyte stability test provides a sensitive measure of the interaction of extractable or leachable substances with the plasma membrane of erythrocytes and is reflected as changes in the osmotic fragility of the erythrocytes. This test can detect leachables at concentrations slightly below the sensitive levels of many cytotoxicity systems.

Hypotonic saline or distilled water (as described earlier under hemolysis) can be adjusted to the required tonicity. Extractors are adjusted to give osmolarity appropriate to hemolyze about 50% of the erythrocytes. Usually rabbit blood is used (though human blood is preferred), diluted with isotonic saline to about 1% hematocrit. One-tenth of a millimeter of this stock erythrocyte solution is added to 5 mL of hypotonic extract, and the surviving cells are counted. The relative hemolysis, number of cells lysed in the extract versus number of cells lysed in the control, is reported. By performing the tests with a series of dilutions of the extract, the concentration of extract at which no detectable change occurs can be established and compared with data from other materials or from extracts prepared under different conditions. Cell size distribution profiles can also be obtained in this test, giving an indication of the degree of swelling or morphologic changes.

3.16.8 Whole Blood Clotting Time

Whole blood clotting time may be measured by modifying the Lee-White method or other relevant tests. The Lee-White test measures the recalcification as an indicator of coagulability (Kretschner et al. 2004). Such measures, however, are influenced by blood or sample dilution. Thrombin time, prothrombin time, and/or platelet counts should be included.

- Partial thromboplastin time (PTI)
  (a) Shortened time = activation of intrinsic pathway
  (b) Sensitive to all know clotting factors except VII and XII
- Prothrombin time (PT)
  (a) Measures activation of extrinsic pathway
  (b) Measures time required for recalcified plasma to clot in presence of thromboplastin
- Thrombin time (TT)
  (a) Useful in detecting inhibitors of the thrombin-fibrinogen reaction
  (b) Measure the availability of functional fibrinogen

3.16.9 Thrombogenicity

Testing for thrombogenicity is normally done by examining platelet and fibrinogen turnover and observing thrombus formation and resulting emboli. Because thrombogenicity tests are usually difficult, controversial, and expensive, manufacturers should consult with the FDA to verify the proper model and test protocol (as briefly described below).

Thrombosis is the formation or existence of a blood clot within the vascular system. When associated with a device, it can be a life-threatening event because the clot, called a thrombus, can occlude a vessel and stop the blood supply to an organ or body part. If detached, the thrombus becomes an embolus and may occlude a vessel at a distance from the original site. When measuring thrombosis, the test end point is the size of clot formation or the adherence of platelets, leukocytes, erythrocytes, or other aggregates on the test device.

In the *light microscopy* method, an intact sample may be exposed to whole blood ex vivo. Ex vivo means away from or outside of the body, and, in ex vivo experiments, some of an animal's blood is caused to bypass the normal circulatory system and pass through or across a device and then flow back into the animal's body. Light microscopy is used to scan the material for evidence of thrombus formation. Alternatively, the material or device may be excised after a suitable period of exposure in vivo and then scanned for thrombus formation using a light microscope.

Table 3.43 NAVI model-scoring

Score	Thrombus formation score description
0	No significant thrombosis (very small clot acceptable at insertion)
1	Minimal thrombosis, one location
2	Minimal thrombosis, multiple locations
3	Significant thrombosis, ≤1/2 the length of the implant, vessel patient
4	Significant thrombosis, ≥1/2 the length of the implant, vessel patient
5	Vessel completely occluded

Thrombosis "tests" give a yes or no answer. There is either thrombus formation or there is not. This is the NAVI (non-anticoagulated venous implant) model, which can use dogs, pigs, or sheep. It is scored as per Table 3.43, but only as utility for assessing short term (up to 5 days) effect. Its limits are:

- The implant position.
- The implant technique.
- The extent of device-vessel wall contact.
- Time/incubation period.
- The explant technique.
- The material or the material surface.
- Nonthromboadherent materials get labeled nonthrombogenic.
- The recipient/subject thrombotic potential.
- Statistical power.
- Evaluator expertise.

3.16.9.1 Percent Occlusion, Flow Reduction and Gravimetric Analysis

These are all attempts to quantify the amount of thrombus formation. Percent occlusion is visually assessed after a device has been in use and has been removed. Percent occlusion is a measure of the severity of the thrombotic process in a conduit. Flow reduction is a measure of the drop-in rate or volume of blood flow through a device after a period of implantation. Gravimetric analysis is a weight measure of thrombus mass after removal of the mass from a device after a period of use.

3.16.9.2 Scanning Electron Microscopy

This is a method of visually assessing a device on a micron scale. When used on explanted materials or devices, it may give a closer visual assessment of thrombus formation, capsular formation, or device performance.

3.16.9.3 Angiography

This is an X-ray of blood vessels (which have been made rediopaque by the injection of a dye) and is a method of taking an X-ray of the vasculature following injection of a radiopaque substance to obtain a description of the blood vessels or the arterial pulse.

3.16.10 Complement Activation

Inappropriate of excess complement activation may lead to unwanted tissue damages or cause cardiopulmonary distress in patients (Henderson 1989). Complement activation is usually measured by the conversion of C_3 to C_{3a} and/or C_5 to C_{5a}. The hemolytic complement expressed in CH_{50} is generally not sensitive enough to detect complement activation caused by biomaterials and is not acceptable as a complement activation assay.

The classical complement system consists of nine separate protein components (numbered C_1 through C_9) acting in sequence. When activated, complement components interact sequentially with one another in a cascade. Activation of some complement components results in the cleavage of a component into two fragments. In some cases, the larger fragments join other activated fragments, and the smaller fragments, such as C_{3a} and C_{5a}, have inflammatory properties. The C_{3a} and C_{5a} cause vasodilation and increase capillary permeability.

3.16.11 Protein Adsorption

The adsorption of plasma protein is generally the first event that occurs when blood contacts a foreign surface (Lemm and Unger 1980). This protein layer has a great influence on the thrombogenicity of a material. One of the more commonly used techniques is the radiolabeling of protein with ^{125}I. The measurements consist of three steps:

- The exposure of a solid surface to a solution containing the radiolabeled proteins
- Rinsing to remove all but the adsorbed protein
- Measurement of the radioactivity retained by the surface

This technique provides a direct measurement of the amount of protein adsorbed on a surface. Protein adsorption can also be studied from flowing solutions in specially designed flow chambers. Recently, real-time spectrophotometric measurements of dynamic protein adsorption have been done by Fourier and transformed into infrared-attenuated total reflectance.

3.16.12 Coagulation

Coagulation refers to the process of blood clotting, which results from the initiation of a cascading enzymatic pathway where the product on one reaction is an enzyme which catalyzes another, subsequent, reaction. The outcome of coagulation is the formation of a clot or thrombus. When measuring coagulation, the test end point is enzyme activation or suppression (not thrombus formation).

3.16.12.1 Clotting Time

As designed by Lee and White, is probably the earliest coagulation test developed. It is not discussed in ISO 10993 Part 4, although it is still frequently used to screen materials for blood compatibility. A sample of blood is removed from an animal exposed to material or device. The time at which the blood is withdrawn is noted as accurately as possible. The syringe is emptied into a small glass tube, which is rotated endwise every 30 seconds. The point at which the blood no longer flows from its position but maintains its surface contour when inverted is taken as the end point. Normal clotting time in humans is about 6.5 minutes.

3.16.12.2 Thromboplastin

The third blood coagulation factor (Factor III). Partial thromboplastin time (PTT) is the clotting time of recalcified citrated plasma upon the addition of partial thromboplastin obtained from mammalian brain or lung. Shortening of the PTT following contact with a material indicates activation of coagulation factors; a prolonged PTT suggests a deficiency. A blood sample, as citrated plasma, is obtained from an animal that has been exposed to the intact material. An excess of calcium ions and thromboplastin are added and the time to clotting measured.

3.16.12.3 Prothrombin

A circulating protein which, when acted upon by thrombokinase, forms thrombin. Prothrombin time (PT) is related to prothrombin concentration and the accessory factors, Factor V, Factor VII, and Factor X. In the presence of thromboplastin, clotting time depends on the concentrations of these four factors. A blood sample is obtained from an animal that has been exposed to the intact material. An excess of calcium ions and thromboplastin is added and the time to clotting measured. A prolonged prothrombin time indicated a deficiency of prothrombin, Factors V, VII, and X or fibrinogen, indicating the implant has inactivated, absorbed, or otherwise interfered with the concentration of these proteins.

3.16.12.4 Thrombin

It is a protein found in shed blood. Formed from prothrombin, it reacts with soluble fibrinogen, converting it to fibrin which forms the basis of blood clots. Thrombin time (TT) is the time required for plasma to clot when a solution of thrombin is added. A blood sample as plasma is obtained from an animal that has been exposed to the intact material. A solution of thrombin is added to the plasma and the time to clotting measured. A prolonged thrombin time indicates a deficiency in fibrinogen.

3.16.12.5 Plasma Fibrinogen

A protein in the blood which, when acted upon by thrombin and calcium, forms fibrin. A sample of plasma is obtained from an animal that has been exposed to the intact material. Active fibrinogen is measured indirectly by using a commercially available thrombin time assay. Thrombin time is dependent on fibrinogen and can be an accurate measure of its active concentration.

3.16.12.6 Fibrin/Fibrinogen Degradation Products

By-products of degraded fibrin and/or fibrinogen. A sample of plasma is obtained from an animal that has been exposed to the intact material. An immunoassay is performed by exposing the plasma to fibrin/fibrinogen antibodies per the instructions in commercially available tests.

3.16.12.7 Specific Coagulation Factor Assays

Fibrinopeptide A, D dimer (a fibrin degradation product), F_{1+2} (prothrombin activation fragment 1 + 2), PAC-1 (monoclonal antibody which recognizes the activated form of platelet surface glycoprotein IIb/IIIa), S-12 (monoclonal antibody which recognizes the alpha granule membrane component 9GMP140 exposed during the platelet release reaction), or TAT (thrombin-antithrombin complex) may be performed on blood samples taken from animals exposed to intact, implanted material.

3.16.13 Platelets

3.16.13.1 Platelet Count

Platelets are flat, round cells found in the circulating blood. They play an important role in blood coagulation, hemostasis, and thrombus formation. When a small vessel is injured, platelets adhere to each other and the edges of the injury and form a plug.

The plug or blood clot soon retracts and stops the loss of blood. A blood sample is obtained from an animal that has been exposed to the intact material and the number of platelets per mm^3 determined. Normal human values are 200,000 to 300,000.

3.16.13.2 Platelet Aggregation

Induced when cells at the site of injury secrete epinephrine, or when collagen, thrombin or other agents are produced at the site. Platelet aggregation can by induced in vitro by the addition of these agents exogenously. To evaluate the ability of platelets to aggregate, plasma is placed in a beaker and the exogenous agents added with constant stirring. As the platelets aggregate, the plasma becomes progressively clearer. An optical system (aggregometer) is used to detect the change in light transmission. Delayed or reduced platelet aggregation, or spontaneous aggregation, is a sign of platelet activation.

Assays for PF-4 (platelet factor 4), β-TG (beta-thromboglobulin), or thromboxane B2 may be performed on blood samples taken from animals exposed to intact, implanted material.

Gamma imaging of radiolabeled platelets may be performed on ^{111}Indium-labelled platelet survival times may be determined in situ in animals exposed to intact, implanted material.

3.16.14 Conclusion

Though hematocompatability has long been identified as a concern for medical devices and biomaterials (Mason 1972; Autian 1977; Wilsnack and Bernadyn 1979), it is only recently that the standards for evaluation of the relevant end points ave come to utilize available technology (ISO-10993-4).

Currently, for materials and for devices with limited (in either extent of duration) exposure to the circulated blood, a limited battery of in vitro evaluations as described in Table 7.1 should be adequate to ensure hematocompatability. For devices with extended contact with circulating blood (systemic circulation), however, a much more extensive evaluation, including at least an in vivo study in a suitable model species should be considered.

3.17 Carcinogenicity (ISO 10993-3)

This section studies the potential tumorigenicity and carcinogenicity of devices and biomaterials with prolonged human exposure via implantation.

Carcinogenicity studies are infrequently required for medical devices. The tables in Chap. 1 cite each of the general cases though these do not catch all of the nuances.

Under ISO 10993-11, for example, the need to perform carcinogenicity tests may be triggered by:

- Devices introduced in the body>30 days cumulative contact
- Devices or materials with positive genotoxicity tests
- Resorbable materials and devices
- Note that in those cases where carcinogenicity testing is required but no effects have occurred in genotoxicity tests, clinical testing may be performed concurrently with carcinogenicity testing where implantation does not represent the most appropriate route of exposure, scientifically justified alternative should be considered.

Even then, the ISO standard states that "carcinogenicity should be conducted only if there are suggestive data from other sources." Where implantation does not represent the most appropriate or there is a more practical route of exposure, scientifically justified alternative routes should be considered (Henry 1985). The intent of such testing is to determine the carcinogenic ("tumorigenic") potential of devices, materials, and/or multiple exposures over a period of the total life-span of the test animal. Such tests are frequently designed to evaluate both the chronic toxicity and the tumorigenicity in a single study as well as device efficacy. These studies are the longest and most expensive of the nonclinical studies typically conducted on any new device or device material. These studies are important because, as noted by the International Agency for Research on Cancer (1987), "in the absence of adequate data on humans, it is biologically plausible and prudent to regard agents for which there is sufficient evidence of carcinogenicity in experimental animals as if they presented a carcinogenic risk to human." The best established risks of carcinogenicity have to do with the effects of metals leading from long-term implants.

3.17.1 Animal Model

Unlike for pharmaceuticals and agrichemicals, only one species is required to be evaluated in a carcinogenicity assay for a device or device material. The Sprague-Dawley-derived rat is by far the most commonly used with the other strains of rats (Wistar, Long-Evans, CFE, and Fischer 344 seeing only rare use). On very rare occasions, dogs have been used (besides other concerns, a dog tumorigenial study is required to run 7 years (as opposed to two for a rat, to be valid)), but this is now so infrequent; we will concentrate on the case of the rat study in this section. The use of a single species is unlikely to adversely affect overall ability to detect potential risks (Zbinden 1993).

The choice of species and strain to be used in a carcinogenicity study is based on various criteria including susceptibility to tumor induction, incidence of spontaneous tumors survival, existence of an adequate historical data base, and availability (Cameron et al. 1985; Arnold et al. 1990; Gad 2002).

Susceptibility to tumor induction is an important criterion. There would be little justification for doing carcinogenicity studies in an animal model that did not respond when treated with a "true" carcinogen. Ideally, the perfect species/strain would have the same susceptibility to tumor induction as the human. Unfortunately, this information is usually unavailable, and the tendency has been to choose animal models that are highly sensitive to tumor induction to minimize the probability of false negatives.

The incidence of spontaneous tumors is also an important issue. Rodent species and strains differ greatly in the incidence of various types of spontaneous tumors. The Sprague-Dawley stock, although preferred, has a very high incidence of mammary tumors in aging females, which results in substantial morbidity during the second year of a carcinogenicity study. If one chooses the Fischer 344 (F344) strain, the female mammary tumor incidence will be lower, but the incidence of testicular tumors will be higher (close to 100%), than that in Sprague-Dawley rats.

A high spontaneous tumor incidence can compromise the results of a carcinogenicity study in two ways. If a compound induces tumors at a site that already has a high spontaneous tumor incidence, it may be impossible to detect an increase above the high background "noise." Conversely, if a significant increase above levels is demonstrated, one may question the relevance of this finding to humans on the basis that the species is "highly susceptible" to tumors of this type (Hajian 1983).

Such considerations are further compounded by the "Oppenheimer effect" (Turner 1941; Oppenheimer et al. 1948, 1952, 1953, 1955, 1958, 1961, 1964). This is the occurrence of parenchymal tumors produced after long-lasting periods when smooth-surfaced solids are implanted. Such solids have included everything from plastics to marble chips. This is a well-established phenomenon in rodents which has not been demonstrated in nonrodents or humans. These tumors are thought to be due to a epigenetic mechanism, and no sex differences in response have been seen. About 80% of the resulting tumors are fibrosarcomas (Alexander and Horning 1959; Brand et al. 1975, 1976; Ecanow et al. 1977; Brand and Brand 1982; Memol 1986). Particulate "generation" by the degradation of device components increases the degree of problems.

From these early investigations, one can derive a number of characteristics for the phenomenon termed "solid-state carcinogenesis." The major ones are:

- Composition of the material per se appears to be of little importance (unless it contains leachable carcinogens) because a wide variety of materials elicits a similar response.
- A continuous, impermeable surface is important since perforations, weaves, or powders tend to reduce or abolish tumorigenicity of the material (Bates and Klein 1966; Bischoff and Bryson 1964; Dukes and Mitchley 1962; Goldhaber 1961, 1962).
- The implant must be of at least a minimum ("critical" size).
- The implant must remain in situ for a minimum period of time. The studies of Oppenheimer et al. (1958) found the presarcomatous changes occurred when the material was in place for about 6 months, although tumors may not appear for many more months.

The ability of a species/strain to survive for an adequate period is essential for a valid assessment of carcinogenicity. Poor survival has caused regulatory problems for pharmaceutical companies (PMA 1988) and is, therefore, an important issue for medical devises. The underlying concept is that animals should be exposed to the drug for the greater part of their normal life-span to make a valid assessment of carcinogenicity. If animals on study die from causes other than drug-induced tumors, they may not have been at risk long enough for tumors to have developed. The sensitivity of the bioassay would be reduced, and the probability of a false negative result would be increased.

The availability of an adequate historical data base if often cited as an important criterion for species/strain selection. Historical control data can sometimes be useful in evaluating the results of a study. Although such data are not considered equal in value to concurrent control data, they can be helpful if there is reason to believe that the concurrent control data are "atypical" for the species/strain.

Advantages of the Sprague-Dawley rat are (1) a large historical data base including various routes of exposure, (2) demonstrated susceptibility to known carcinogens, (3) generally good survival until recently (see below), and (4) ease of handling compared with certain other stocks. Disadvantages include (1) moderate to high incidence of spontaneous tumors, especially mammary and pituitary, (2) old rat nephropathy, and (3) marked genetic variability in stocks obtained from different suppliers (Chu et al. 1981; Sher et al. 1982).

There has recently been a reduction in survival of Sprague-Dawley rats and rats of other strains [Food and Drug Administration (FDA) 1993]. This reduction may be the result of ad libitum feeding, as preliminary results suggest that caloric restriction may improve survival. Leukemia appears to be the major cause of decreasing survival in the F344 rat. The problem of reduced survival may necessitate a reevaluation of the survival requirements for carcinogenicity studies by regulatory agencies. There is also now a significant body of data that suggest that switching from the long favored ad libitum feeding of animals in bioassays can both extend their life-span and decrease the incidences of some background tumors (Rao and Huff 1990).

3.17.2 Dose Selection

3.17.2.1 Number of Dose Levels

Unlike for drugs or agricultural chemicals, there will ordinarily be two dose levels, the maximum implantable dose (MID) and a fraction thereof (usually one-half of the MID). The controls will generally include polyethylene implants or other materials whose lack of carcinogenic potential is documented in a comparable form and shape.

In carcinogenicity testing on rodents, the maximum implantable dose (MID) of a material or device should be applied. Where possible, this dose should be expressed as multiple of the worst-case human exposure in milligrams per kilogram.

3.17.3 Group Size

The minimum number of animals assigned to each dose group in implant carcinogenicity studies is 50 of each sex. Most companies, however, use more than the minimum number, and some use up to 80 animals per sex per group. The most important factor in determining group size is the need to have an adequate number of animals for a valid assessment of carcinogenic activity at the end of the study. Larger group sizes are also used when the carcinogenicity study is combined with a chronic toxicity study in the rat. In this case, serial sacrifices are performed at 6 and 12 months to evaluate potential toxic effects of the device.

In the final analysis, the sensitivity of the bioassay for detecting carcinogens is directly related to the sample size. Use of the MTD has often been justified based on the small number of animals at risk compared to the potential human population, in spite of the difficulties inherent in extrapolating effects at high doses to those expected at much lower clinical doses. A reasonable compromise may be the use of doses lower than the MTD combined with a larger group size than the 50 per sex minimum accepted by regulatory agencies.

3.17.4 Route of Administration

Device carcinogenicity studies are conducted with the device or material being implanted into the test animals. Prior to implantation the samples are prepared. Whenever possible, the device shall be tested in its "ready-to-use" form. Otherwise a suitably formed implant shall be made of the test material, with appropriate consideration of potential solid-state carcinogenicity.

Treated animals typically receive single implants in a flank by making an incision, opening a pouch, inserting the sample, and closing the pouch. Dose groups are achieved by implanting variable numbers of devices in multiple flanks. Controls are generally untreated in the sense that no device is implanted—only the surgical procedure is performed.

3.17.5 Study Duration

The duration of carcinogenicity studies for rats is 2 years. Occasionally, rat studies are extended to 30 months. When hamsters are used, the study duration is limited to 18 months, a time period that is consistent with the shorter survival characteristics of this species.

Irrespective of the intended duration of the study, the most important consideration is that adequate numbers of animals survive long enough to allow for a valid assessment of carcinogenic activity. When survival is problematic, the duration of the study may be modified accordingly. The effect of survival on study duration is discussed in the next section.

3.17.6 Survival

As stated earlier, adequate survival is of primary importance in carcinogenicity studies because animals must be exposed to a drug for the greater part of their life-span to increase the probability that late-occurring tumors can be detected. Early mortality, resulting from causes other than tumors, can jeopardize the validity of a study because dead animals cannot get tumors.

In general, the sensitivity of a carcinogenicity bioassay is increased when animals survive to the end of their natural life-span, because weak carcinogens may induce late-occurring tumors. The potency of a carcinogen is often inversely related to the time to tumor development. By analogy, as the dose of a carcinogen is reduced, the time to tumor occurrence is increased (Littlefield et al. 1979; DePass et al. 1986).

Why do we not allow all animals on a carcinogenicity study to live until they die a natural death if by so doing we could identify more drugs as carcinogens? In fact, the sensitivity of a bioassay may not be improved by allowing the animals to live out their natural life-span because the incidence of spontaneous tumors tends to increase with age. Thus, depending on the tumor type, the ability of the bioassay to detect a device-related increase in tumor incidence may actually decrease, rather than increase, with time. Therefore, the optimum duration of a carcinogenicity study is that which allows late-occurring tumors to be detected but does not allow the incidence of spontaneous tumors to become excessive.

Reduced survival in a carcinogenicity study may or may not be device-related. Sometimes, the MTD is exceeded and increased mortality occurs at the highest dose level and, occasionally, at the mid-dose level as well. This situation may not necessarily invalidate a study; in fact, the protocol may be amended to minimize the impact of the device-induced mortality. For example, cessation of drug treatment may enhance the survival of the animals in the affected groups and allow previously initiated tumors to develop. As shown by Littlefield et al. (1979) in the NCTR ED01 study, liver tumors induced by 2-acetylaminofluorene, which appeared very late in the study, were shown to have been induced much earlier and not to require the continuous presence of the carcinogen to develop. By contrast, bladder tumors that occurred in the same study were dependent on the continued presence of the carcinogen.

Whether drug treatment is terminated or not, device-related toxicity may also be managed by performing complete histopathology on animals in the lower-dose groups rather than on high-dose and control animals only. If there is no increase in tumor incidence at a lower-dose level that is not compromised by reduced survival, the study may still be considered valid as an assessment of carcinogenicity.

When reduced survival is related to factors other than excessive toxicity, the number of animals at risk for tumor development may be inadequate, and the validity of the study may be compromised even in the absence of a device effect on survival. Obviously, the adjustments described above for excessive, drug-related toxicity are not relevant to this situation.

There is no unanimity of opinion among regulatory agencies as to the minimum survival required to produce a valid carcinogenicity study or as to the best approach for dealing with survival problems. Even with a single agency such as the FDA, different opinions exist on these issues. For example, the recently issued *FDA Redbook II Draft Guidelines* requires that rats, mice, or hamsters be treated for 24 months. Early termination due to decreased survival is not recommended. The EEC guidelines differ in that they suggest termination of the study when survival in the control group reaches 20%, while Japanese guidelines suggest termination at 25% survival in the control or low-dose groups (Speid et al. 1990). These provisions make good sense in that they do not request termination of the study when device-related mortality may be present only at the highest dose.

3.17.7 Parameters Evaluated

In a pure carcinogenicity study the chief parameters measured are survival and occurrence of tumors (Table 3.44).

Also measured are typically urinalysis parameters on samples collected prior to study start, at 6-month intervals during the study, and just prior to the final sacrifice, as presented in Table 3.45.

Clinical pathology and hematology measurements are made on blood samples collected at the same intervals with parameters measured in Table 3.44.

3.17.8 Statistical Analysis

Irrespective of the specific protocols used, all carcinogenicity studies end with a statistical comparison of tumor proportions between treated and control groups. This analysis is necessary because the control incidence of most tumor types is rarely zero. In the unlikely case that a type of tumor is found in treated animals but not in concurrent or appropriate historical controls, it is reasonable to conclude that the tumor is treatment-related without statistical analysis (Haschek and Rousseaux 1991).

Most companies analyze tumor data using mortality-adjusted methods (PMA 1988). Peto/International Agency for Research on Cancer (IARC) methodology is most commonly used, perhaps because this method is currently favored by the FDA (Peto et al. 1980). The use of life-table methods is most appropriate for “lethal” tumors, that is, those that cause the death of the animals. Various statistical methods are available for analyzing the incidence of lethal and nonlethal tumors (e.g., Gart

Table 3.44 Lifetime carcinogenicity study (implant) organs and sites to be examined

Adrenals (2)	Harderian glands (2)	Spleen
Brain (3 levels: forebrain, midbrain, and hindbrain including brainstem)	Heart	Spinal cord (cervical)
Eyes (2)	Kidneys (2)	Skin (dorsal)
Gastrointestinal tract:	Larynx	Sternum/bone marrow
Esophagus	Liver (2 lobes)	Thymic region
Stomach (glandular and nonglandular)	Lung (2 coronal sectionsincluding all lobes and mainstem bronchi)	Thyroid/parathyroid
Duodenum	Lymph node (mesenteric)	Trachea
Jejunum	Mammary region (males and females)	Urinary bladder
Ileum	Pancreas	Uterus
Cecum	Pituitary	Implant site (4 sections ofsubcutaneous site and contiguous performal region)
Colon	Prostate	Any other grossly abnormaltissues or organs
Rectum	Salivary gland (submaxillary)	
Gonads: Ovaries with oviducts (2) testes with epididymids (2)	Sciatic nerve	
	Seminal vesicles	
	Skeletal muscle (thigh)	

Table 3.45 Urinalysis parameters measured

Appearance (color)
pH
Ketones
Urobilinogen
Specific gravity (refractive index)
Albumin
Glucose
Occult blood
Urinary sediment
Volume
Bilirubin

et al. 1979, 1986; Dinse and Lagokos 1983; McKnight 1988; Portier and Bailer 1989; Gad 2004). These methods are especially useful when there are drug-related differences in mortality rates. When there is no drug effect of survival, unadjusted methods will generally give the same results.

As a general approach, most pharmaceutical statisticians begin by testing for the presence of a dose-related trend in tumor proportions. If the trend test is significant, that is, the p value is less than or equal to 0.05, pairwise comparisons are performed between the treated and control groups. Trend and pairwise analyses may be adjusted for mortality as stated earlier or performed without mortality adjustment using such simple methods as chi-square or Fisher's exact tests.

Although in most cases the use of trend tests is appropriate since most biological responses are dose-related, there are exceptions to this rule. Certain drugs, especially those with hormonal activity, may not produce classical dose responses and may even induce inverse dose-response phenomena. In these cases, a pairwise comparison may be appropriate in the absence of a significant positive trend.

Most companies use one-tailed comparisons, and a substantial number use two-tailed methods. Since regulatory agencies are primarily interested in identifying carcinogenic drugs, as opposed to those that inhibit carcinogenesis, the use of one-tailed tests is generally considered more appropriate. Some companies prefer two-tailed comparisons because, in the absence of a true carcinogenic effect, there is an equal probability of seeing significant decreases as well as significant increases by chance alone.

One of the most important statistical issues in the analysis of carcinogenicity data is the frequency of "false positives" or Type I errors. Because of the multiplicity of tumor sites examined and the number of tests employed, there is concern that noncarcinogenic devices may be erroneously declared carcinogens. If any $p < 0.05$ increase in tumor incidence is automatically regarded as a biologically meaningful result, then the false-positive rate may be as high as 47–50% (Haseman et al. 1986).

Several statistical procedures designed to correct for the multiplicity of significance tests have been published (Haseman 1990). One approach to the problem of multiple tumor site/type testing is a procedure attributed to Tukey by Mantel (1980). This method is used to adjust a calculated p value based on the number of tumor types/sites for which there are a minimum number of tumors in the particular study. The reasoning here is that, for most tumor sites, the number of tumors found is so small that it is impossible to obtain a significant result for that tumor site no matter how the tumors might have been distributed among the dose groups. Only those sites for which a minimum number of tumors is present can contribute to the false-positive rate for a particular study.

A method proposed by Schweder and Spjotvoll (1982) is based on a plot of the cumulative distribution of observed p values. Farrar and Crump (1988) have published a statistical procedure designed not only to control the probability of false-positive findings but also to combine the probabilities of a carcinogenic effect across tumor sites, sexes, and species.

Another approach to controlling the false-positive rate in carcinogenicity studies was proposed by Haseman (1983). Under this "rule," a compound would be declared a carcinogen if it produced an increase significant at the 1% level in a common tumor or an increase significant at the 5% level in a rare tumor. A rare neoplasm was

defined as a neoplasm that occurred with a frequency of less than 1% in control animals. The overall false-positive rate associated with the decision rule was found to be no more than 7–8%, based on control tumor incidences from NTP studies in rats and mice. This false-positive rate compares favorably with the expected rate of 5%, which is the probability at which one would erroneously conclude that a compound was a carcinogen. This method is notable for its simplicity and deserves serious consideration by pharmaceutical statisticians and toxicologists. Without resorting to sophisticated mathematics, this method recognizes the fact that tumors differ in their spontaneous frequencies and, therefore, in their contribution to the overall false-positive rates in carcinogenicity studies. False-positive results are much less likely to occur at tissue sites with low spontaneous tumor incidences than at those with high frequencies.

As a final point that has special relevance to pharmaceutical carcinogenicity studies, one may question whether the corrections for multiple comparisons and their effect on the overall false-positive rate are appropriate for all tumor types. For example, if a compound is known to bind to receptors and produce pharmacological effects in a certain organ, is it justified to arbitrarily correct the calculated p value for the incidence of tumors in that organ, using the methods described above? It is difficult to justify such a correction considering that the basis for correcting the calculated p value is that the true probability of observing an increased incidence of tumors at any site by chance alone may be much higher than the nominal alpha level (usually 0.05). It is reasonable to expect that, when a drug has known pharmacological effects on a given organ, the probability of observing an increased tumor incidence in that organ by chance alone is unlikely to be higher than the nominal 5% alpha level.

Although most pharmaceutical statisticians and toxicologists agree on the need to control the probability of false-positive results, there is no consensus as to which method is most appropriate or most acceptable to regulatory agencies. The FDA and other such agencies will accept a variety of statistical procedures but will often reanalyze the data and draw their own conclusions based on their analyses.

3.17.9 Interpretation of Results

3.17.9.1 Criteria for a Positive Result

There are three generally accepted criteria for a positive result in a carcinogenicity study. The first two are derived directly from the results of the statistical analysis: (1) a statistically significant increase in the incidence of a common tumor and (2) a statistically significant reduction in the time to tumor development. The third criterion is the occurrence of very rare tumors, that is, those not normally seen in control animals, even if the incidence is not statistically significant.

3.17.9.2 Use of Historical Controls

When the study is over, the data analyzed, and the *p* values corrected, as appropriate, one may find that one or more tumor types increased in drug-treated groups relative to concurrent control group(s), but comparable to or lower than the historical incidence. Occasionally, a small number of tumors may be found in a treated group, and the incidence may be significant because of the absence of this tumor in the concurrent controls. Review of appropriate historical control data may reveal that the low tumor incidence in the treated group is within the "expected" range for this tumor.

The role of historical control data in interpreting carcinogenicity findings depends on the "quality" of the historical data. Ideally, the data should be derived from animals of the same age, sex, strain, and supplier, housed in the same facility, and the pathology examinations should have been performed by the same pathologist or using the same pathological criteria for diagnosis. Since genetic drift occurs even in animals of a given strain and supplier, recent data are more useful than older data. The value of historical control data is directly proportional to the extent to which these conditions are fulfilled.

Although methods are available for including historical control data in the formal statistical analysis (Tarone 1982; Dempster et al. 1983), this is usually not done and for good reason. The heterogeneity of historical data requires that they be used qualitatively and selectively to aid in the final interpretation of the data, after completion of the formal statistical analysis.

3.18 Immunotoxicology (ISO 10993-20)

The evaluation of the immunotoxicity of medical devices as part of their biocompatibility assessment is the subject of the most recent ISO-10993 guidance (Part 20). Traditionally and still one of the three tests required for all medical devices is dermal delayed contact sensitization (covered in an earlier portion of this chapter). More extensive evaluation of immune system interactions with devices is both a much more recent requirement and not expected for all devices. This is in spite of the fact that the association between implanted or indwelling devices and granuloma formation has been known for some time (Adams 1983; Anderson 1988; Black 1981; Burkett et al. 1986; Woodward and Salthouse 1986; Unanue 1994; Salthouse 1982; Marchant et al. 1985). However, improved science and the record of immune-based device problem since 1980 has brought the adequacy of this approach into question.

The immune system is a highly complex system of cells tissues and mediators involved in a multitude of functions including antigen presentation and recognition, amplification, and cell proliferation with subsequent differentiation and secretion of lymphokines and antibodies (Bick 1985). The end result is an integrated and highly

interdependent system responsible for defense against foreign pathogens and spontaneously occurring neoplasms that, if left unchecked, may result in infection and malignancy. We also now speak of the system being composed of innate (present in both invertebrates and vertebrates) and adaptive (present only in vertebrates) components. To be effective, the immune system must be able to both recognize and destroy foreign antigens. To accomplish this, cellular and soluble components of the adaptive immune system of diverse function and specificity circulate through blood and lymphatic vessels, thus allowing them to act at remote sites and tissues. For this system to function properly requires regulation through cell-to-cell communications and precise recognition of self versus nonself (or threat/nonthreat) . Immunotoxicants can upset this balance if they are lethal to one of more of the cell types or alter membrane morphology and receptors. There are several undesired immune system responses that may potentially occur upon repeated exposure to a medical device material that may ultimately present barriers to its development, including:

- Down-modulation of the immune response (immunosuppression or hypoimmunity), which may result in an impaired ability to deal with neoplasia and infections. This is of particular concern if the devise is intended or likely to be used in patients with preexisting conditions such as cancer, severe infection, or immunodeficiency diseases.
- Up-modulation of the immune system (or hyperimmunity, such as autoimmunity).
- Direct adverse immune responses to the agent itself in the form of hypersensitivity responses (anaphylaxis and delayed contact hypersensitivity).
- Direct immune responses to the device that limit or nullify its utility (i.e., the development of neutralizing antibodies to be a delivered agent).

Immunotoxicology has evolved since the late 1970s as a specialty within toxicology that brings together knowledge from basic immunology, molecular biology, microbiology, pharmacology, and physiology. As a discipline, immunotoxicology involves the study of the adverse effects that xenobiotics have on the immune system. As listed above, several different types of adverse immunological effects may occur, including immunosuppression, autoimmunity, and hypersensitivity. Although these effects are clearly distinct, they are not mutually exclusive. For example, immunosuppressive drugs that suppress suppressor-cell activity can also induce autoimmunity (Hutchings et al. 1985), and agents that are immunoenhancing at low doses may be immunotoxic at high doses. Chemical xenobiotics may be in the form of natural or man-made environmental chemicals—pharmaceuticals and biologicals that are pharmacologically, endocrinologically, or toxicologically active. Although, in general, xenobiotics are not endogenously produced, immunologically active biological response modifiers that naturally occur in the body should also be included, since many are now known to compromise immune function when present in pharmacologically effective doses (Koller 1987). The success in development of immunomodulatory protein therapeutics likewise has had to both promise and challenge.

Although the types immunological responses to various xenobiotics may be similar, the approach taken for screening potential immunological activity will vary depending on the application of the compound. In contrast to potential environmental exposures, medical devices and the materials they are composed of are developed with intentional but restricted human exposure, and their biological effects are extensively studied in surveillance.

In 1993, the FDA issued draft guidelines for immunotoxicity testing in the revision of the "Redbook" (FDA 1993). Although these guidelines have been established through the Center for Food Safety and Applied Nutrition, other centers within the FDA extended the usage of these guidelines to cover testing for human and veterinary pharmaceuticals. Both ICH (for pharmaceuticals) and ISO (for drugs) have promulgated guidances for the evaluation of subject products for adverse immune effects.

Unanticipated immunotoxicity is infrequently observed with drugs that have been approved for marketing. With the exception of drugs that are intended to be immunomodulatory or immunosuppressive as part of their therapeutic mode of action, there is little evidence that drugs or devices cause unintended functional immunosuppression in man (Gleichman et al. 1989). However, hypersensitivity (allergy) and autoimmunity are frequently observed and are serious consequences of some therapies (DeSwarte 1986). An adverse immune response in the form of hypersensitivity is one of the most frequent safety causes for withdrawal of drugs that have already made it to market (see Table 3.46) and accounts for approximately 15% of adverse reactions to xenobiotics (de Weck 1983). In addition, adverse immune responses such as this (usually urticaria and frank rashes) are the chief "unexpected" finding in clinical studies. These findings are unexpected in that they are not predicted by preclinical studies. These findings are unexpected in that they are not predicted by preclinical studies because there is a lack of good preclinical models for predicting systemic hypersensitivity responses, especially to orally administered agents. As a consequence, the unexpected occurrence of hypersensitivity in the clinic may delay, or even preclude, further development and commercialization. Thus, a primary purpose for preclinical immunotoxicology testing is to help us detect these adverse effects earlier in development, before they are found in clinical trials.

3.18.1 Immunotoxic Effects

The immune system is a highly integrated and regulated network of cell types that requires continual renewal to achieve balance and immunocompetence. Fortunately, the multiple components of the system serve to overlap and respond to challenges in a manner which provides redundancy. The delicacy of this balance makes specific components of the immune system a natural target for cytotoxic drugs or their metabolites. Since renewal is dependent on the ability of cells to proliferate and dif-

Table 3.46 Drugs withdrawn from the market due to dose- and time-unrelated toxicity not identified in animal experiments

Compound	Adverse reaction	Year of introduction	Years on the market
Aminopyrine	Agranulocytosis	Approx 1900	75
Phenacetin	Interstitial nephritis	Approx 1900	83
Dipyrone	Agranulocytosis	Approx 1930	47
Clioquinol	Subacute myelo-optic neuropathy	Approx 1930	51
Oxyphenisatin	Chronic active hepatitis	Approx 1955	23
Nialamide	Liver damage	1959	19
Phenoxyorioazine	Liver damage	1961	5
Mebanazine	Liver damage	1963	3
Ibufenac	Hepatotoxicity	1966	2
Practolol	Oculo-mucocutaneous syndrome	1970	6
Alclofenace	Hypersensitivity	1972	7
Azaribine	Thrombosis	1975	1
Ticynafen	Nephropathy	1979	1
Benoxaprofen	Photosensitivity, hematoxicity	1980	2
Zomepirac	Urticaria, anaphylactic shock	1980	3
Zimelidine	Hepatotoxicity	1982	2

Source: Adapted from Bakke et al. (1984)

ferentiate, exposure to agents that arrest cell division can subsequently lead to reduced immune function or immunosuppression. This concept has been exploited in the development of therapeutic drugs intended to treat leukemias, autoimmune disease, and chronic inflammatory diseases and to prevent transplant rejection. However, some drugs adversely modulate (or overly stimulate) the immune system secondarily to their therapeutic effects. (Simply consider the case of the monoclonal antibody TGN-412).

Two broad categories of immunotoxicity have been defined on the basis of suppression or stimulation of normal immune function. Immunosuppression is a down-modulation of the immune system characterized by cell depletion, dysfunction, or dysregulation that may subsequently result in increased susceptibility to infection and tumors. By contrast, immunostimulation is an increased or exaggerated immune responsiveness that may be apparent in the form of a tissue-damaging allergic hypersensitivity response or pathological autoimmunity. However, as knowledge of the mechanisms involved in each of these conditions has expanded, the distinction between them has become less clear. Some agents can cause immunosuppression at one dose or duration of exposure and immunostimulation at others. For instance, the chemotherapeutic drug cyclophosphamide is in most cases immunosuppressive; however, it can also induce autoimmunity (Hutchings et al. 1985). Likewise, dimethylnitrosamine, a nitrosamine detected in some foods, has been shown to have both suppressing and enhancing effects on the immune system (Yoshida et al. 1989).

3.18.2 Immunosuppression

The various cells of the immune system may differ in their sensitivity to a given xenobiotic. Thus, immunosuppression may be expressed as varying degrees of reduced activity of a single cell type or multiple populations of immunocytes. Several lymphoid organs such as the bone marrow, spleen, thymus, and lymph nodes may be affected simultaneously, or the immunodeficiency may be isolated to a single tissue, such as the Peyer's patches of the intestines. The resulting deficiency may in turn lead to an array of clinical outcomes of varying ranges of severity. These outcomes include increased susceptibility to infections, increased severity or persistence of infections, or infections with unusual organisms (e.g., system fungal infections). Immunosuppression can be induced in a dose-related manner by a variety of therapeutic agents at dose levels lower than those required to produce overt clinical signs of general toxicity. In addition, immunosuppression can occur without regard to genetic predisposition, given that a sufficient dose level and duration of exposure has been achieved.

Humoral immunity is characterized by the production of antigen-specific antibodies that enhance phagocytosis and destruction of microorganisms through opsonization. Thus, deficiencies of humoral immunity (B lymphocytes) may lead to reduced antibody titers and are typically associated with acute gram-positive bacterial infections (i.e., *Streptococcus*). Although chronic infection is usually associated with dysfunction of some aspect of cellular immunity, chronic infections can also occur when facultative intracellular organisms such as *Listeria* or *Mycobacterium* evade antibodies and multiply within phagocytic cells.

Since cellular immunity results in the release of chemotactic lymphocytes that in turn enhance phagocytosis, a deficiency in cellular immunity may also result in chronic infections. Cellular immunity is mediated by T cells, macrophages, and NK cells involved in complex compensatory networks and secondary changes. Immunosuppressive agents may act directly by cytotoxicity to T cells or indirectly by blocking mitosis, lymphokine synthesis, lymphokine release, or membrane receptors to lymphokines. In addition, cellular immunity is involved in the production and release of interferon, a lymphokine that ultimately results in blockage of viral replication (Table 3.47). Viruses are particularly susceptible to cytolysis by T cells since they often attach to the surface of infected cells. Thus, immunosuppression of any of the components of cellular immunity may result in an increase in protozoan, fungal, and viral infections as well as opportunistic bacterial infections.

Immune depression may result unintentionally as a side effect of cancer chemotherapy or intentionally from therapeutics administered to prevent graft rejection. In fact, both transplant patients administered immunosuppressive drugs and cancer patients treated with chemotherapeutic agents have been shown to be at high risk of developing secondary cancers, particularly of lymphoreticular etiology (Penn 1977). Most of these drugs are alkylating or cross-linking agents that by their chemical nature are electrophilic and highly reactive with nucleophilic macromolecules (protein and nucleic acids). Nucleophilic sites are quite ubiquitous and include amino, hydroxyl, mercapto, and histidine functional groups. Thus, immunotoxic

Table 3.47 Growth and differentiation factors of the immune system

Factors	Cell of origin	Primary immune functions
Interleukins[a]		
IL-1	Macrophage, B and T cells	Lymphocyte-activating factor; enhances activation of T and B cells, NK cells, and macrophages
IL-2	T cells (T_h)	T-cell growth factor; stimulates T-cell growth and effector differentiation; stimulates B-cell proliferation/ differentiation
IL-3	T cells (T_h)	Mast-cell growth factor; stimulates proliferation/ differentiation of mast cells, neutrophils, and macrophages
IL-4	T cells (T_h), mast cells, B cells	B-cell growth factor; induces proliferation/differentiation of B cells and secretion of IgA, IgG_1, and IgE; promotes T-cell growth; activates macrophages
IL-6	T cells, fibroblasts, monocytes	Stimulates growth/differentiation of B cells and secretion of IgG; promotes IL-2-induced growth of T cells
IL-7	Bone marrow stromal cells	Stimulates pre-B- and pre-T-cell growth/differentiation; enhances thymocyte Proliferation
IL-8	Monocytes, fibroblasts	Neutrophils chemotaxis
IL-9	T cells	Stimulates T cells and mast cells
IL-10	T cells	Stimulates mast cells and thymocytes; induction of class II MHC
Interferons (INF)		
α-INF	Leukocytes and mast cells	Antiviral; increases NK-cell function, B-cell differentiation, potentiates macrophage production of IL-1
β-INF	Fibroblasts, epithelial cells	Antiviral; potentiates macrophage production of IL-1; increases NK-cell function
γ-INF	T cells (T_h), cytotoxic T cells	Antiviral; activates macrophages; induces MHC class II expression on macrophages, epithelial, and endothelial cells
Tumor necrosis factors (TNF)		
TNFα	Macrophage, B and T cells	Catectin; promotes tumor cytotoxicity; activates macrophages and neutrophils; enhances IL-2 receptor expression on T cells; inhibits antibody secretion
TNFβ	T cells (T_h)	Lymphotoxin; promotes T-cell-mediated cytotoxicity
	NK cells	B-cell activation
Colony-stimulating factors (CSF)		
	Stem cells:	*Promotes growth and differentiation of:*
Granulocyte CSF	Myeloid	Granulocytes and macrophages
Macrophage CSF	Myeloid	Macrophages and granulocytes
Granulocytes-macrophage CSF	Myeloid	Granulocytes, macrophages, eosinophils, mast cells, and pluripotent progenitor cells

Source: Extracted and modified from Golub and Green (1991)
[a]Includes lymphokines, monokines, and cytokines produced by T cells, macrophages, and other cells, respectively

agents used in chemotherapy may induce secondary tumors through direct genotoxic mechanisms (i.e., DNA alkylation).

Reduced cellular immunity may result in increased malignancy and decreased viral resistance through indirect mechanisms as well, by modulating immune surveillance of aberrant cells. T lymphocytes, macrophage cells, and NK cells are all involved in immunosurveillance through cytolysis of virally inflected cells or tumor cells, each by a different mechanism (Table 3.47) (Burnet 1970). In addition to the common cell types described in Table 3.47, at least two other types of cytotoxic effector cells of T-cell origin have been identified, each of which has a unique lytic specificity phenotype and activity profile (Merluzzi 1985). Of these, both LAK and TIL cells have been shown to lyse a variety of different tumor cells. However, TIL cells have 50–100 times more lytic activity than LAK cells. Most tumor cells express unique surface antigens that render them different from normal cells. Once detected as foreign, they are presented to the T helper cells in association with MHC molecules to form an antigen-MHC complex. This association elicits a genetic component to the immunospecificity reaction. T helper cells subsequently direct the antigen complex toward the cytotoxic T lymphocytes, which possess receptors for antigen-MHC complexes. These cells can then proliferate, respond to specific viral antigens or antigens on the membranes of tumor cells, and destroy them (Yoshida et al. 1989).

In contrast, the macrophages and natural killer (NK) cells are involved in nonspecific immunosurveillance in that they do not require prior sensitization with a foreign antigen as a prerequisite for lysis and are not involved with MHC molecules. The enhancement of either NK cell function or macrophage function has been shown to reduce metastasis of some types of tumors. Macrophage cells accumulate at the tumor site and have been shown to lyse a variety of transformed tumor cells (Volkman 1984). Natural killer cells are involved in the lysis of primary autochthonous tumor cells. Migration of NK cells to tumor sites has been well documented. Although not clearly defined, it appears that they can recognize certain proteinaceous structures on tumor cells and lyse them with cytolysin.

3.18.3 Immunostimulation

A variety of drugs as well as environmental chemicals have been shown to have immunostimulatory or sensitizing effects on the immune system and these effects are well documented in humans exposed to drugs (DeSwarte 1986). The drug or metabolite can act as a hapten and covalently bind to a protein or other cellular constituent of the host to appear foreign and become antigenic. Haptens are low molecular weight substances that are not in themselves immunogenic but will induce an immune response if conjugated with nucleophilic groups on proteins or other macromolecular carriers. In both allergy and autoimmunity, the immune system is stimulated or sensitized by the drug conjugate to produce specific pathological responses. An allergic hypersensitivity reaction may vary from one which results

in an immediate anaphylactic response to one which produces a delayed hypersensitivity reaction or immune complex reaction. Allergic hypersensitivity reactions result in a heightened sensitivity to nonself antigens, whereas autoimmunity results in an altered response to self-antigens. Unlike immunosuppression, which nonspecifically affects all individuals in a dose-related manner, both allergy and autoimmunity have a genetic component that creates susceptibility in those individuals with a genetic predisposition. Susceptible individuals, once sensitized, can respond to genetic predisposition. Susceptible individuals, once sensitized, can respond to even minute quantities of the antigen.

3.18.4 Autoimmunity

In autoimmunity, as with hypersensitivity, the immune system is stimulated by specific responses that are pathogenic, and both tend to have a genetic component that predisposes some individuals more than others. However, as is the case with hypersensitivity, the adverse immune response of drug-induced autoimmunity is not restricted to the drug itself, but also involves a response to self-antigens.

Autoimmune responses directed against normal components of the body may consist of antibody-driven humoral responses and/or cell-mediated, delayed-type hypersensitivity responses. T cells can react directly against specific target organs, or B cells can secrete autoantibodies that target "self." Autoimmunity may occur spontaneously as the result of a loss of regulatory controls that initiate or suppress normal immunity causing the immune system to produce lymphocytes reactive against its own cells and macromolecules such as DNA, RNA, or erythrocytes.

Although autoantibodies are often associated with autoimmune reactions, they are not necessarily indicative of autoimmunity (Russell 1981). Antinuclear antibodies can occur normally with aging in some healthy women without autoimmune disease, and all individuals have B cells with the potential of reacting with self-antigens through Ig receptors (Dighiero et al. 1983). The presence of an antibody titer to a particular immunogen indicates that haptenization of serum albumin has occurred as part of a normal immune response. However, if cells are stimulated to proliferate and secrete autoantibodies directed against a specific cell or cellular component, a pathological response may result. The tissue damage associated with autoimmune disease is usually a consequence of Type II or III hypersensitivity reactions that result in the deposition of antibody-antigen complexes.

Several diseases have been associated with the production of autoantibodies against various tissues. For example, an autoimmune form of hemolytic anemia can occur if the antibodies are directed against erythrocytes. Similarly, antibodies that react with acetylcholine receptors may cause myasthenia gravis, those directed against glomerular basement membranes may cause Goodpasture's syndrome, and those that target the liver may cause hepatitis. Other forms of organ-specific autoimmunity include autoimmune thyroiditis (as seen with amiodarone) and juvenile diabetes mellitus, which result from autoantibodies directed against the tissue-specific

antigens thyroglobulin and cytoplasmic components of pancreatic islet cells, respectively. In contrast, systemic autoimmune diseases may occur if the autoantibodies are directed against an antigen that is ubiquitous throughout the body, such as DNA or RNA. For example, systemic lupus erythematosus (SLE) occurs as the result of autoimmunity to nuclear antigens that form immune complexes in the walls of blood vessels and basement membranes of tissues throughout the body.

The etiology of renal autoimmunity is not well established and is confounded by factors such as age, sex, and nutritional state, as well as genetic influences on pharmacological and immune susceptibility. Unlike idiopathic autoimmunity, which is progressive or characterized by an alternating series of relapses and remissions, drug-induced autoimmunity is thought to subside after the drug is discontinued. However, this is not certain since a major determining factor for diagnosis of a drug-related disorder is dependent on the observation of remission upon withdrawal of the drug (Bigazzi 1988).

One possible mechanism for xenobiotic-induced autoimmunity involves xenobiotic binding to autologous molecules, which then appear foreign to the immunosurveillance system. If a self-antigen is chemically altered, a specific T helper (T_h) cell may see it as foreign and react to the altered antigenic determinant portion, allowing an autoreactive B cell to react to the unaltered hapten. This interaction results in a carrier-hapten bridge between the specific T_h and autoreactive B cell, bringing them together for subsequent production of auto-antibodies specific to the self-antigen that was chemically altered (Weigle 1980). Conversely, a xenobiotic may alter B cells directly, including those that are autoreactive. Thus, the altered B cells may react to self-antigens independent from T_h-cell recognition and in a nontissue-specific manner.

Another possible mechanism is that the xenobiotic may stimulate nonspecific mitogenicity of B cells. This could result in a polyclonal activation of B cells with subsequent production of autoantibodies. Alternatively, the xenobiotic may stimulate mitogenicity of T cells that recognize self, which in turn activate B-cell production of antibodies in response to "self" molecules. There is also evidence to suggest that anti-DNA autoantibodies may originate from somatic mutations in lymphocyte precursors with antibacterial or antiviral specificity. For example, a single amino acid substitution resulting from a mutation in a monoclonal antibody to polyphorlcholine was shown to result in a loss of the original specificity and an acquisition of DNA reactivity similar to that observed for anti-DNA antibodies in SLE (Talal 1987).

The mechanism of autoimmunity may also entail interaction with MHC structures determined by the HLA alleles. Individuals carrying certain HLA alleles have been shown to be predisposed to certain autoimmune diseases, which may account in part for the genetic variability of autoimmunity. In addition, metabolites of a particular drug my vary between individuals to confound the development of drug-induced autoimmunity. Dendritic cells, such as the Langerhans cells of the skin and B lymphocytes that function to present antigens to T_h cells, express class II MHC structures. Although the exact involvement of these MHC structures is unknown, Gleichmann et al. (1989) have theorized that self-antigens rendered foreign by

drugs such as D-penicillamine may be presented to T_h cells by MHC class II structures. An alternate hypothesis is that the drug or a metabolite may alter MHC class II structures on B cells, making them appear foreign to T_h cells.

A number of different drugs have been shown to induce autoimmunity in susceptible individuals. A syndrome similar to that of SLE was described in a patient administered sulfadiazine in 1945 by Hoffman (see Bigazzi 1985). Sulfonamides were one of the first classes of drugs identified to induce an autoimmune response, while to date, more than 50 other drugs have been associated with a similar syndrome.

Autoantibodies to red blood cells and autoimmune hemolytic anemia have been observed in patients treated with numerous drugs, including procainamide, chlorpropamide, captopril, cefalexin, penicillin, and methyldopa (Logue et al. 1970; Kleinman et al. 1984). Hydralazine- and procainamide-induced autoantibodies may also result in SLE. Approximately 20% of patients administered methyldopa for several weeks for the treatment of essential hypertension developed a dose-related titer and incidence of autoantibodies to erythrocytes, 1% of which presented with hemolytic anemia. Methyldopa does not appear to act as a hapten but appears to act by modifying erythrocyte surface antigens. IgG autoantibodies then develop against the modified erythrocytes.

Some metals that are used in devices have also been shown to induce autoimmune responses. Gold salts used to treat arthritis may induce formation of antiglomerular basement membrane antibodies, which may lead to glomerulonephritis similar to that seen in Goodpasture's disease (see Type II hypersensitivity). Since gold is not observed at the site of the lesions (Druet et al. 1982), it has been hypothesized that the metal elicits an antiself response. Lithium, used to treat manic-depression, is thought to induce autoantibodies against thyroglobulin, which in some patient results in hypothyroidism. In studies with rats, levels of antibodies to thyroglobulin were shown to increase significantly in lithium-treated rats compared to controls immediately after immunization with thyroglobulin; however, rats that were not immunized with thyroglobulin did not produce circulation antithyroglobulin antibodies upon receiving lithium, and there was no effect of lithium on lymphocytic infiltration of the thyroid in either group (Hassman et al. 1985).

In addition, silicone-containing medical devices, particularly breast prostheses, have been reported to cause serum-sickness-like reactions, scleroderma-like lesions, and an SLE-like disease termed human adjuvant disease (Kumagai et al. 1984; Guillaume et al. 1984). Some patients may also present with granulomas and autoantibodies. Human adjuvant disease is a connective tissue or autoimmune disease similar to that of adjuvant arthritis in rats and rheumatoid arthritis in humans. Autoimmune disease-like symptoms usually develop 2–5 years after implantation in a small percentage of people that receive implants, which may indicate that there is a genetic predisposition similar to that for SLE. An early hypothesis is that the prosthesis or injected silicone plays an adjuvant role by enhancing the immune response through increased macrophage and T-cell helper function. There is currently controversy as to whether silicone, as a foreign body, induces a nonspecific inflammation reaction, a specific cell-mediated immunological reaction, or no reaction at all. However, there is strong support to indicate that silicone microparticles

can act as haptens to produce a delayed hypersensitivity reaction in a genetically susceptible population of people.

3.19 Evaluation of the Immune System

3.19.1 Regulatory Positions

Since the last edition of this book, the pharmaceutical and medical device industries have come to have specific regulatory guidances requiring routine evaluation of the toxicological profile of the xenobiotic (drugs and devices) in routine preclinical safety testing. The chemical industry has been a proponent of using a battery of assays to assess chemical-induced immunotoxicity, hence guidelines for a two-tiered screen approach have been proposed by the National Toxicology Program (NTP) (Luster et al. 1988). This strategy, which was developed for nontherapeutic chemicals and environmental contaminants that have different safety standards, does not address some of the safety issues and test strategy issues that are unique to pharmaceuticals. ICH and ISO (and by reference the FDA) have promulgated drugs (ICH) and devices (ISO 10993-20) a similar two-leveled approach (Hinton 1992) for assessing immunotoxicity of food colors and additives. In all of these testing schemes, the initial tier generally includes a fundamental histopathologic assessment of the major components of the immune system. Additional tiers are then added to more precisely evaluate the functionality of the components that appeared to be adversely affected in the first tier of tests. These test strategies are primarily geared toward the detection of chemical-induced immunosuppression, thus the effectiveness of these test schemes for detecting immunostimulation has not yet been determined (Spreafico 1988) (Table 3.48).

ISO defines the first tier of assays (Table 3.49) to include an assessment of immunopathology: humoral, cell-mediated, and nonspecific immunity such as natural killer cell activity. The second tier includes a more comprehensive battery that should be used once functional changes are observed in the Tier I assays. The Tier II assays focus on mechanisms of immunotoxicity such as depletion of specific cell subsets by flow cytometry analysis or evaluation of secondary immune responses by examining IgG response. Cell-mediated immunity is assessed through a functional assay that looks at the ability of cytotoxic T cells to kill target cells, and nonspecific immunity is evaluated by examining various function of macrophages: (1) the ability to phagocytize inert fluorescent beads or radiolabeled chicken erythrocytes and (2) the ability to produce cytokines such as IL-1 or macrophage activation factor. The ultimate immune test would be to examine the effects of xenobiotics on the intact animal's response to challenge by viral, bacterial, or parasitic pathogens or neoplastic cells. The ability of the immune system to compensate or, conversely, its inability to compensate for loss or inhibition of its components is fully examined through host resistance mechanisms. This tiered test approach has been validated

Table 3.48 Cellular components of the immune system and their functions

Cell subpopulations	Markers[a]	Functions
Nonspecific immunity		
Granulocytes		Degranulate to release mediators
Neutrophils (blood)		
Basophils (blood)		
Eosinophils (blood)		
Mast cells (connective tissue)		
Natural killer cells (NK)		Nonsensitized lymphocytes; directly kill target cells
Reticuloendothelial	CD14; HLA-DR	Antigen processing, presentation, and phagocytosis (humoral and some cell-mediated responses)
Macrophage (peritoneal, pleural, alveolar spaces)		
Histiocytes (tissues)		
Monocytes (blood)		
Specific immunity		
Humoral immunity		
Activated B cells	CD19; CD23	Proliferate; form plasma cells
Plasma cells		Secrete antibody; terminally differentiated
Resting		Secrete I_gM antibodies (primary response)
Memory		Secrete I_gG antibodies (secondary response)
Cell-mediated immunity		
T-cell types:		
Helper (T_k)	CD4; CD25	Assists in humoral immunity; required for antibody production
Cytotoxic (T_k)	CD8; CD25	Targets lysis
Suppressor (T_s)	CD8; CD25	Suppresses/regulates humoral and cell-medicated responses

[a]Activation surface markers detected by specific monoclonal antibodies; can be assayed with flow cytometry

Table 3.49 Tier I screen

Parameter	Procedures
Immunopathology	Routine hematology-complete and differential count; routine toxicology information-weights of body, immune organs (spleen and thymus), liver, and kidney; histopathology of immune organs
Humoral-mediated immunity	LPS (lipopolysaccharide) mitogen response or $F(ab)_2$ mitogenic response; enumeration of plaques by IgM antibody-forming cells to a T-dependent antigen (sheep red blood cells; serum IgM concentration
Cell-mediated immunity	Lymphocyte mitogenic response to concanavalin A and mixed lymphocyte response to allogeneic lymphocytes; local lymph node assay
Nonspecific immunity	Natural killer cell activity

Source: Adapted from Luster et al. (1988) and Vos et al. (1989)

with 50 selected compounds, and results from these studies have shown that the use of only 2 or 3 immune tests are sufficient to predict known immunotoxic compounds in rodents with a >90% concordance (Luster et al. 1992a, b). Specifically the use of either a humoral response assay for plaque-forming colonies (PFC response) or determination of surface marker expression in combination with almost any other parameter significantly increased the ability to predict immunotoxicity when compared to the predictivity of any assay alone.

The FDA guidelines for immunotoxicity testing of food additives start with a Type I battery of tests. Type I tests can be derived from the routine measurements and examinations performed in short-term and subchronic rodent toxicity studies, since they do not require any perturbation of the test animals (immunization or challenge with infectious agents). These measurements include hematology and serum chemistry profiles, routine histopathologic examinations of immune-associated organs and tissues, and organ and body weight measurements including thymus and spleen. If a compound produces any primary indicators of immunotoxicity from these measurements, more definitive immunotoxicity tests, such as those indicated in the preceding paragraph, may be recommended on a case-by-case basis.

The following is a brief explanation of some of the indicators that may be used to trigger additional definitive testing and a description of some of the most commonly used assays to assess humoral, cell-mediated, or nonspecific immune dysfunction, which are common to most immunotoxicology test strategies.

3.19.2 Immunopathologic Assessments

Various general toxicological and histopathologic evaluations of the immune system can be made as part of routine preclinical safety testing to obtain a preliminary assessment of potential drug-related effects on the immune system. At necropsy, various immunological organs of the immune system such as thymus, spleen, and lymph nodes are typically observed for gross abnormalities and weighed in order to detect decreased or increased cellularity. Bone marrow and peripheral blood samples are also taken to evaluate abnormal types and/or frequencies of the various cellular components.

3.19.2.1 Organ and Body Weights

Changes in absolute weight, organ-to-body weight ratios, and organ-to-brain weight ratios of tissues such as thymus and spleen are useful general indicators of potential immunotoxicity. However, these measures are nonspecific for immunotoxicity since they may also reflect general toxicity and effects on endocrine function that can indirectly affect the immune system.

3.19.2.2 Hematology

Hemacytometers or electronic cell counters can be used to assess the numbers of lymphocytes, neutrophils, monocytes, basophils, and eosinophils in the peripheral blood, while changes in relative ratios of the various cell types can be assessed by microscopic differential evaluation. Similar evaluations can be performed with bone marrow aspirates, where changes may reflect immunotoxicity to the pluripotent stem cells and newly developing lymphoid precursor cells. Potential hematological indicators of immunotoxicity include altered white blood cell counts or differential ratios, lymphocytosis, lymphopenia, or eosinophilia. Changes in any of these parameters can be followed up with more sophisticated flow cytometric analyses or immunostaining techniques that are useful for phenotyping the various types of lymphocytes (B cell, T cell) and the T-cell subsets (CD4+ and CD8+) on the basis of unique surface markers. Decreases or increases in the percentages of any of the cell populations relative to controls, or in the ratios of B cells/T cells, or CD4+/CD8+ cells may be indicators of immunotoxicity.

3.19.2.3 Clinical Chemistry

Nonspecific clinical chemistry indicators of potential immune dysfunction include changes in serum protein levels in conjunction with changes in the albumin-to-globulin (A/G) ratio. Immunoelectrophoretic analysis of serum proteins can then be performed to quantify the relative percentages of albumin and the α-, β-, and γ-globulin fractions. To perform these assays, a drop of serum (antigen) is placed into a well cut in a gel, and then the gel is subjected to electrophoresis so that each molecule in the serum moves in the electric field according to its charge. This separation is then exposed to specific antiserum, which is placed in a trough cut parallel to the direction in which the components have moved. By passive diffusion, the antibody reaches the electrophoretically separated antigen and reacts to form Ag-Ab complexes. The γ-globulin fractions can be separated and further quantified for the relative proportions of IgG, IgM, IgA, and IgE using similar techniques.

Serum concentrations of immunoglobulin classes and subclasses can also be measured using various techniques such as radioimmunoassays (RIAs) or enzyme-linked immunosorbent assays (ELISAs). In the ELISA, antigens specific for each class of immunoglobulin can be adsorbed onto the surfaces for microtiter plates. To determine the quantity of each antibody in a test sample, an aliquot of antiserum is allowed to react with the adsorbed antigens. Unreacted molecules are rinsed off and an enzyme-linked anti-Ig is then added to each well. Next, substrate is added, and the amount of color that develops is quantified using a spectrophotometric device. The amount of antibody can then be extrapolated from standard curves since the amount of color is proportional to the amount of enzyme-linked antibody that reacts. Variations in levels of a given antibody may indicate the decreased ability of B cells or decreased numbers of B cells producing that antibody. In addition, serum autoantibodies to DNA, mitochondria, and parietal cells can be used to assess autoimmunity. Serum cytokines

(IL-1, IL-2, and γ-interferon) can also be evaluated using immunochemical assays to evaluate macrophage, lymphocyte, and lymphokine activity; prostaglandin E_2 can also be measured to evaluate macrophage function.

CH50 determinations can be used to analyze the total serum complement and are useful for monitoring immune complex diseases (Sullivan 1989); activation of complement (Table 3.50) in the presence of autoantibodies is indicative of immune complex diseases and autoimmunity. The various components of the complement system (C3, C4) can also be measured to assess the integrity of the system. For instance, low serum concentrations of C3 and C4, with a concomitant decrease in CH50, may indicate activation of complement, while a low C4 alone is a sensitive indicator of reduced activation of the complement system. Since C3 is used as an alternate complement pathway, it usually measures high. Therefore, a low C3 with a normal C4 may indicate an alternate pathway of activation.

3.19.2.4 Histopathology

Histopathologic abnormalities can be found in lymphoid tissues during gross and routine microscopic evaluations of the spleen, lymph nodes, thymus, bone marrow, and gut-associated lymphoid tissues such as Peyer's patches and mesenteric

Table 3.50 Examples of tests for and indicators of the evaluation of immune responses

		Nonfunctional assays		
Immune responses	Functional assays	Soluble mediators	Phenotyping	Others[a]
Tissue/inflammatory	Implant/systemic ISO 10993-6 and ISO 10993-11	NA	Cell surface markers	Organ weight analysis
Humoral response	Immunoassays (e.g., ELISA) for antibody responses to antigen plus adjuvant[b]	Complement (including C3a and C5a anaphylatoxins)	Cell surface markers	
	Plaque-forming cells	Immune complexes		
	Lymphocyte proliferation			
	Antibody-dependent cellular cytotoxicity			
	Passive cutaneous anaphylaxis			
	Direct anaphylaxis			
		Cellular responses		

(continued)

Table 3.50 (continued)

		Nonfunctional assays		
Immune responses	Functional assays	Soluble mediators	Phenotyping	Others[a]
T cells	Guinea pig maximization test	Cytokine patterns indicative of T-cell subset (Th1, Th2)	Cell surface markers (helper and cytotoxic T cells)	
	Mouse local lymph node assay			
	Mouse ear swelling test			
	Lymphocyte proliferation			
	Mixed lymphocyte reaction			
NK cells	Tumor cytotoxicity	NA	Cell surface markers	
Macrophages and other monocytes	Phagocytosis	Cytokines (IL-1, TNFα, IL-6, TGFβ, IL-10, ϒ-interferon)	MHC markers	
	Antigen presentation			
Dendritic cells	Antigen presentation to T cells	NA	Cell surface markers	
Vascular endothelial cells	Activation			
Granulocytes (basophils, eosinophils, neutrophils)	Degranulation	Chemokines, bioactive amines, inflammatory cytokines, enzymes	NA	Cytochemistry
	Phagocytosis			
Host resistance	NA	NA	NA	
Clinical symptoms	NA	NA	NA	Allergy, skin rash, urticaria, edema, lymphadenopathy, inflammation

[a]Animal models of some human autoimmune diseases are available. However, routine testing for induction of autoimmune diseases by materials/devices is not recommended
[b]Most commonly used tests. Functional assays are generally more important than tests for soluble mediators or phenotyping

lymph nodes. Microscopic evaluations should include descriptive qualitative changes such as types of cells, densities of cell populations, proliferation in known T- and B-cells areas (e.g., germinal centers), relative numbers of follicles and germinal centers (immune activation), and the appearance of atrophy or necrosis. In addition, unusual findings such as granulomas and scattered, focal mononuclear cell infiltrates in nonlymphoid tissues may be observed as indica-

tors of chronic hypersensitivity or autoimmunity. A complete histopathologic evaluation should also include a quantitative assessment of cellularity through direct counts of each cell type in the various lymphoid tissues. In addition, changes in cellularity of the spleen can be more precisely quantitated from routine H&E (hematoxylin and eosin) sections using morphometric analysis of the germinal centers (B cells) and periarteriolar lymphocyte sheath (T cells). Similar morphometric measurements can be made of the relative areas of the cortex and medulla of the thymus. If changes in cellularity are apparent from routinely stained histopathology sections, special immunostaining (immunoperoxidase or immunofluorescence) of B cells in the spleen and lymph nodes using polyclonal antibodies to IgG, or immunostaining of the T cells and their subsets in the spleen using mono-polyclonal antibodies to their specific surface markers, can be used to further characterize changes in cellularity.

Numerous physiological and environmental factors such as age, stress, nutritional deficiency, and infections may affect the immune system (Sullivan 1989). Thus, adverse findings in animal studies may reflect these indirect immunotoxic effects rather than the direct immunotoxic potential of a chemical or drug. Indirect immunotoxic effects may be assessed through histopathologic evaluations of endocrine organs such as the adrenals and pituitary.

It is also well known that the functional reserves of the immune system can allow biologically significant, immunotoxic insults to occur without the appearance of morphological changes. In addition, there is some built-in redundancy in the system in that several mechanisms may produce the same outcome. For instance, cytotoxic T cells may alone be sufficient to protect the organism against a bacterial infection; however, the body will also produce antibodies for future protection. Thus, if one mechanism is insufficient to fight off infection, the second mechanism can serve as a backup. Because of this functional reserve, adverse effects may remain subclinical until the organism is subjected to undue stress or subsequent challenge (Bloom et al. 1987). Therefore, routine immunopathologic assessments as part of standard preclinical toxicity tests may not be sufficient to detect all immunotoxins. Although changes detected in routine toxicological and pathological evaluations are nonspecific, and of undetermined biological significance to the test animal, they can be invaluable as flags for triggering additional testing.

3.19.3 Humoral Immunity

As described previously, the humoral immune response results in the proliferation, activation, and subsequent production of antibodies by B cells following antigenic exposure and stimulation. The functionality and interplay between the three primary types of immune cells (macrophage, B cells, and T cells) required to elicit a humoral response can be assessed through various in vitro assays using cells from the peripheral blood or lymphoid tissues.

3.19.3.1 Antibody Plaque-Forming Cell (PFC) Assay

The number of B cells producing antibody (PFC) to a T-dependent antigen such as sheep red blood cells (SRBCs) can be assessed in vitro following in vivo exposure to the test article and antigen (ex vivo tests). The PFC response to a T-dependent antigen is included as a Tier I test by the NTP since it appears to be the most commonly affected functional parameter of exposure to immunosuppressants. However, this test is designated as a Type II test in the FDA Redbook since it requires an in vivo immunization of the animals with antigen and, thus, cannot be evaluated as part of an initial toxicity screen.

Although this assay requires that B cells be fully competent in secreting antibodies, T cells and macrophage cells are also essential for the proper functioning of humoral immunity. However, this assay is nonspecific in that it cannot determine which cell type(s) is responsible for dysfunction. Macrophage cells are needed to process antigen and produce IL-1. T cells are needed for several functions including antigen recognition of surface membrane proteins and B-cell maturation through the production of various lymphokines that stimulate growth and differentiation. SRBCs are most commonly used as the T-dependent antigen, although T-cell independent antigens may also be useful to rule out T helper dysfunction as a cause of immunodysfunction.

The PFC assay has evolved from methodology originally developed as a hemolytic plaque assay (or Jerne plaque assay) by Nils Jerne to quantitate the number of antibody-forming cells in a cell suspension plated with RBCs onto agar plates (Jerne and Nordin 1963). In its present form, animals are treated in vivo with the test compound, immunized with approximately 5×10^8 SRBCs administered intravenously within 2 to 3 days posttreatment, and then sacrificed 4 days (IgM) or 6 days (IgG) later. Antibody-producing spleen cell suspensions are then mixed in vitro with SRBCs, placed onto covered slides, and incubated for a few hours in the presence of complement. During incubation, antibody diffuses from the anti-SRBC-producing cells and forms Ag-Ab complexes on the surfaces of nearby SRBCs, resulting in the formation of small clear plaques on the slide. Plaques are then counted and expressed as PFCs/10^6 spleen cells. A dose-related reduction in PFCs is indicative of immunosuppression.

3.19.3.2 B-Cell Lymphoproliferation Response

The NTP has classified this assay as a Tier I test since mitogenesis can be performed easily in tandem with other tests to provide an assessment of the proliferative capacity of the cells (Luster et al. 1988). Since this assay is performed ex vivo with peripheral blood (or spleen) and is well characterized for use in various animal species, it has also been included as an Expanded Type I test in the revised Redbook.

The proliferation of peripheral blood or splenic B cells following stimulation with lipopolysaccharide (LPS) or other mitogens (pokeweed mitogen extract) is another measurement of humoral immunity. LPS (a bacterial lipopolysaccharide) is

a B-cell-specific mitogen that stimulates polyclonal proliferation (mitosis) as part of the natural sequence of antigen recognition, activation, and clonal expansion. The mitogen does not interact with just one particular antigen-specific clone, but with all cells bearing the carbohydrate surface marker for which it is specific. Since mitogens are both polyclonal and polyfunctional, they can stimulate a wider spectrum of antigenic determinants than antigens, which can only stimulate a low number (10^{-6}) of specific cells.

In this assay, lymphocytes from animals treated in vivo are cultured in vitro in microtiter plates in the presence of tritiated [^{3}H]thymidine (or uridine) using a range of at least three concentrations of mitogen to optimize the response. Lymphocytes can be obtained aseptically from peripheral blood r from single cell suspensions of spleen cells that are prepared by pushing the tissue through sterile gauze or 60-mesh wire screens. A decrease in DNA synthesis (incorporation of ^{3}H) as compared to the unexposed cells of control animals may indicate that the B cells were unable to respond to antigenic stimulation. Alternative methodology employs an 18–20 h incubation with ^{125}I-labeled iododeoxyuridine ([^{125}I]IUdR) and fluorodeoxyuridine (FUdR) (White et al. 1985). After incubation, the cells are collected onto filter disks and then counted with a gamma counter.

Assays such as this that use polyclonal mitogens for activation may not be as sensitive as specific antigen-driven systems (Luster et al. 1988). In addition, suppression of the mitogen response does not always correlate with the PFC response. Since mitogenesis represents only a small aspect of B-cell function and maturation, this end point is not sensitive to early events that may affect activation or later events that may affect differentiation of B cells into antibody-secreting cells (Klaus and Hawrylowicz 1984).

3.19.4 Cell-Mediated Immunity

3.19.4.1 T-Cell Lymphoproliferation Response

This assay is analogous to the B-cell lymphoproliferative response assay described above. Thus, this assay is also classified as a Tier I test by the NTP and as an Expanded Type I test in the revised draft of the Redbook.

T cells from the peripheral blood or spleen undergo blastogenesis and proliferation in response to specific antigens that evoke a cell-mediated immune response. T-cell proliferation is assessed using T-cell-specific mitogens such as the plant lectins, concanavalin A (Con A), and phytohemagglutinin (PHA) or T-cell-specific antigens (i.e., tuberin, *Listeria*). Uptake of ^{3}H as an indicator of DNA synthesis is used as described above for evaluating B-cell proliferation. T-cell mitogens do not just stimulate synthesis of DNA, but, in fact, they also stimulate the expression of cell-specific function. For instance, Con A can trigger the expression of T helper, suppressor, and cytotoxic effector cells, and either mitogen may induce the expression (or reexpression of memory cells) of the differentiated function (Clark 1983).

Since cell populations responsive to Con A are thought to be relatively immature compared to those that are stimulated with PHA, the parallel usage of both mitogens may be useful for distinguishing the affected subset (Tabo and Paul 1973). A secondary response to T-cell antigens such as purified protein derivative of tuberculin (PPD) or tetanus toxoid can also be assessed.

3.19.4.2 Mixed Lymphocyte Response (MLR) Assay

This assay has been shown to be sensitive for the detection of chemical-induced immunosuppression and is a recommended Tier I assay by the NTP (Luster et al. 1988). In addition, it has been shown to be predictive of host response to transplantation and of general immunocompetence (Harmon et al. 1982).

The mixed lymphocyte response assay assesses the ability of T cells to recognize foreign antigens on allogenic lymphocytes and, thus, is an indirect measure of the cell-mediated ability to recognize graft or tumor cells as foreign. Responder lymphocytes from animals treated in vivo with the test compound are mixed with allogeneic stimulator lymphocytes that have been treated in vitro with mitomycin C or irradiated to render them unable to respond (Bach and Voynow 1966). Both cell types are cultured in vitro for 3–5 days and then incubated with ^{3}H for an additional 6 h. Once the radiolabel is incorporated into the DNA of the responding cells, the DNA is extracted and the amount of radioactive label is measured to quantitate proliferation of the responder cells of drug-treated animals compared to those of the controls.

3.19.4.3 Cytotoxic T Lymphocyte (CTL)-Mediated Assay

This assay is similar to the MLR assay and can be performed in parallel or as a Tier II follow-up to the MLR assay.

The CTL assay ascertains the ability of cytotoxic T cells to lyse an allogeneic target cell or the specific target cell type with which they were immunized. In general, the cytolytic response of activated effector cells is assessed by measuring the amount of radioactivity (^{51}Cr) that is released from the target cell. When performed in conjunction with the MLR assay, lymphoid cells of the two strains are cultured together in vitro as described above; however, ^{51}Cr is added to the culture after 4–5 days (instead of ^{3}H). Both responder and target cells are labeled with the ^{51}Cr, which is taken up rapidly by the cells through passive diffusion but is released slowly as long as the cell membrane is intact. Furthermore, since chromium is reduced from Cr^{6+} to Cr^{3+} which enters the cells at a much slower rate than Cr^{6+}, the ^{51}Cr released from the damaged target cells is not significantly reincorporated into undamaged cells (Clark 1983), which would reduce the sensitivity of the assay. Thus, the amount of chromium released into the medium and recovered in the supernatant of the mixture of the cells is directly proportionate to the extent of lysis of the target cells by the sensitized responder cells.

In a capillary tube assay developed in 1962 by George and Vaughan, the inhibition of migration of macrophage cells can be used to access normal T-cell function (see Clark 1983). T cells are obtained from the peripheral blood of animals treated in vivo with a test article and injected with an antigen (e.g., tuberculin). These T cells are functioning normally; they should release migration inhibition factor (MIF). As a consequence, the macrophages, which generally show a propensity for migration upon stimulation with the antigen, should show a MIG-induced reduction in migratory behavior.

3.19.4.4 Delayed-Type Hypersensitivity (DTH) Response

The DTH response assay is considered to be a comprehensive Tier II assay for cell-mediated immunity by the NTP.

3.19.5 Nonspecific Immunity

3.19.5.1 Natural Killer Cell Assays

This assay is a Tier I test for nonspecific immunity in the NTP testing scheme (Luster et al. 1988) and is proposed as an additional Type I test in the draft Redbook.

Natural killer (NK) cells, like cytotoxic T cells, have the ability to attack and destroy tumor cells or virus-infected cells. However, unlike T cells, they are not antigen specific; do not have unique, clonally distributed receptors; and do not undergo clonal selection. In in vitro or ex vivo tests, target cells (e.g., YAC-1 tumor cells) are radiolabeled in vitro or in vivo with ^{51}Cr and incubated in vitro with effector NK cells from the spleens of animals that had been treated with a xenobiotic. This assay can be run in microtiter plates over the range of various ratios of effector/target cells. Cytotoxic activity is then measured by the amount of radioactivity released from the damaged tumor cells as was previously described for cytotoxic T cells. This assay can also be performed in vivo, where YAC-1 cells labeled with [^{125}I]IUdR are injected directly into mice and NK cell activity is correlated with its level of radioactivity (Riccardi et al. 1979). Immunotoxicity observed as reduced NK cell activity is correlated with increased tumorigenesis and infectivity.

3.19.5.2 Macrophage Function

Several assays are available to measure various aspects of macrophage function, including quantitation of resident peritoneal cells, antigen presentation, cytokine production, phagocytosis, intracellular production of oxygen free radicals (used to kill foreign bodies), and direct tumor-killing potential. Techniques for quantitation of peritoneal cells and functional assays for phagocytic ability are classified as

comprehensive Tier II tests by the NTP and as additional Type I tests in the draft Redbook.

Macrophage cells and other polymorphonuclear cells (PMNs) contribute to the first-line defense of nonspecific immunity through their ability to phagocytize foreign materials, including pathogens, tumor cells, and fibers (e.g., silica, asbestos). Xenobiotics can affect macrophage function direct toxicity to macrophages or by modulating their ability to become activated. Differential counts of resident peritoneal cells can be made as rapid, preliminary assessment of macrophage function for xenobiotics that are not administered parenterally.

Numerous in vitro assays can be employed to assess common function of macrophages and PMNs including adherence to glass, migration inhibition, phagocytosis, respiratory activity (chemiluminescent assays or nitroblue tetrazolium), and target cell killing. In one such assay, the chemotactic response to soluble attractants is evaluated using a Boyden chamber with two compartments that are separated by a filter. Macrophage cells or PMNs from treated animals are place in one side and a chemotactic agent in the other. Chemotaxis is then quantified by counting the number of cells that pass through the filter. In another assay, the ability of the macrophages to phagocytize foreign materials can be evaluated by adding fluorescent latex beads to cultures containing macrophage cells, then determining the proportion of cells that have phagocytized the beads using a fluorescent microscope or by flow cytometry (Duke et al. 1985). Similar functions can be evaluated by incubating the cells with known amounts of bacteria. The cells are then removed by filtration or centrifugation, the remaining fluid is plated onto bacterial nutrient agar, and, after a few days of incubation, the bacterial colonies are counted. Furthermore, the efficiency of the cells to kill the bacteria once phagocytized can be assayed by lysing the cells and plating the lysate onto bacterial agar.

Various in vivo assessments of macrophage function have also been used. For example, peritoneal exudate cell (PEC) recruitment can be assessed using eliciting agents such as *Corynebacterium parvum*, MVE-2, or thioglycolate (Dean et al. 1984). In one such assay (White et al. 1985), mice are injected intraperitoneally with thioglycolate, sacrificed 5 days later, and the peritoneal cavity is flushed with culture medium. The cell suspension is then counted, the cell concentration is adjusted to a known density (2×10^5 ml^{-1}), and the cells are cultured for 1 h in 24-well culture dishes. Adherent cells are then washed with medium, and aliquots of ^{51}Cr-labeled SRBCs that were opsonized with mouse IgG are added to each well and incubated for various times. This same system can be used to assess adherence to and chemotaxis of the PECs (Laskin et al. 1981). Phagocytosis can also be evaluated in vivo by measuring the clearance of injected particles from the circulation and the accumulation of the particles in lymphatic tissues such as the spleen.

3.19.5.3 Mast Cell/Basophil Function

The function of mast cells and basophils to degranulate can be evaluated using a passive cutaneous anaphylaxis test (Cromwell et al. 1986). Serum containing specific anaphylactic (IgE) antibodies from donor animals previously exposed to a

known antigen is first administered by intradermal (or subcutaneous) injection into unexposed host animals. After a sufficient latency period to allow binding of the donor IgE to the host tissue mast cells, the animals are administered a second intravenous injection of the antigen. The anaphylactic antibodies present in the serum will stimulate normally functioning mast cells to degranulate (release histamines) and produce a marked inflammatory response. Using similar in vitro assays with mast cells and basophils, the quantities of histamines that are released from the cells can be measured directly in the culture medium.

3.19.6 Host-Resistance Assays

Host-resistance assays can be used to assess the overall immunocompetence of the humoral or cell-mediated immune systems of the test animal (host) to fend of infection with pathogenic microbes or to resist tumorigenesis and metastasis. These assays are performed entirely in vivo and are dependent on all of the various components of the immune system to be functioning properly. Thus, these assays may be considered to be more biologically relevant than in vitro tests that only assess the function of cells from one source and of one type. Since these assays require that the animal be inoculated with a pathogen or exogenous tumor cell, they cannot be performed as part of a general preclinical toxicity assessment and are thus classified as Type II tests in the revised Redbook. These assays are also included as Tier II tests by the NTP.

Similar host-resistance assays are used to evaluate the immunosurveillance of spontaneous tumors, which is assessed as the capacity of the organism to reject grafted syngeneic tumors. Various animal-bearing tumor models (Pastan et al. 1986) and host-resistance models have been used to assess immunotoxicity. Several of the host-resistance assays utilize cultured tumor cell lines such as PYB6 sarcoma and B16F10 melanoma cells that are used with Fischer 344 rats. For example, the PYB6 sarcoma model uses death as an end point. In this assay, syngeneic mice are injected with the PYB6 sarcoma cells, and death due to tumor is recorded daily. In another routinely used assay, animals that have been treated with a xenobiotic are injected with either B16F10 melanoma cells or Lewis lung carcinoma cells, and then approximately 20 days later, they are sacrificed and pulmonary tumors are measured and counted.

3.19.7 Hypersensitivity

3.19.7.1 Type I Hypersensitivity

Although there are acceptable systems for evaluating Type I (immediate) reactions following systemic exposure, there are no reliable animal models for predicting Type I reactions following dermal applications or oral administration of drug.

Repeated exposure of a xenobiotic is required to produce a Type I response. A drug in the form of a hapten must covalently bind to macromolecules (proteins, nucleic acids) before it can initiate a primary antibody response. Once sensitized, even the smallest exposure to the xenobiotic can elicit a rapid, intensive IgE antibody-mediated inflammatory response. With the exception of antivirals and chemotherapeutic drugs, most drugs should not be reactive with biological nucleophiles since these drugs are usually screened out as mutagens or carcinogens in preclinical safety studies. However, Type I hypersensitivity is a particular problem with biotechnology products themselves (e.g., insulin, growth hormones, interleukins), trace impurities from the producing organisms (e.g., *E. coli* proteins, mycelium), or the vehicles used to form emulsions (Matory et al. 1985).

The production of neutralizing antibodies to recombinant DNA protein products or their contaminants may be assayed using ELISAs or IRAs. A suitable animal model used to evaluate the potential for a Type I response to protein hydrolysates is detailed in the *United States Pharmacopeia*. This test is very sensitive for testing proteins administered by the parenteral route, but is of little value for low molecular weight drugs and those that are administered orally (Descotes and Mazue 1987). Active systemic anaphylaxis can be assessed in guinea pigs following systemic exposure to the test compound. For dermal exposures, however, rabbits or guinea pigs must be exposed to the test article by intradermal injections and then evaluated for their ability to mount a systemic anaphylactic response. The passive cutaneous anaphylaxis test (as described above for mast cells) can also be used to assess a potential anaphylactic response to a test compound. The serum containing potential anaphylactic (IgE) antibodies from donor animals previously exposed to the test compound is first administered by intradermal (or subcutaneous) injection into unexposed host animals. After a latency period, the animals are administered an intravenous injection of the test compound together with a dye. If anaphylactic antibodies are present in the serum, the subsequent exposure to the test compound will cause a release of vasoactive amines (degranulation of mast cells), ultimately resulting in the migration of the dye to the sites of the intradermal serum injection.

3.19.7.2 Types II and III Hypersensitivity

No simple animal models are currently available to assess Type II (antibody-mediated cytotoxicity) hypersensitivity reactions. IgE antibodies and immune complexes in the sera of exposed animals can be assayed using ELISA or RIA techniques that require the use of specific antibodies to the drug.

Type III (immune complex-related disease) reactions have been demonstrated by the presence of proteinuria and immune complex deposits in the kidneys of the Brown-Norway, Lewis, and PVG/C rat strains. However, susceptibility to the deposition and the subsequent lesions (glomerulonephritis) are often variable and dependent on the strain (Bigazzi 1985). For example, despite the appearance of clinical signs and proteinuria, after 2-month administration of mercuric chloride, detectable levels of circulating antinuclear autoantibodies can no longer be observed in the

Brown-Norway strain (Bellon et al. 1982). By contrast, in PVG/C rats administered mercuric chloride, immune complex deposition and antinuclear autoantibodies are present for longer periods of time; however, proteinuria is not observed (Weening et al. 1978).

3.19.8 Approaches

3.19.8.1 Suggested Approaches to Testing

As outlined above, there are numerous assays available to assess the various end points that are relevant to immunotoxicity. Early in the development process, a new compound should be evaluated with regard to various factors that may flag it as a potential immunotoxin, including chemical, structural, or physicochemical properties (e.g., photoallergin) and therapeutic class (i.e., immunomodulators, anti-inflammatories, and antimetabolites). Compounds from therapeutic or structural classes that are known to be potential immunotoxins or immunomodulators should be evaluated for the effects in question on a case-by-case basis. With the exception of immunomodulators, protein products, and products of biotechnology, the majority of pharmaceuticals can be assessed for most forms of immunotoxicity during routine preclinical toxicity tests. In general, a well-conducted preclinical toxicity study can detect most serious immunotoxins in the form of altered clinical, hematologic, or histological end points. For example, possible effects on humoral immunity may be indicated from clinical observation of gastrointestinal or respiratory pathology and changes in serum total protein and globulin and by histological changes in lymphoid cellularity. Likewise, effects on the cell-mediated response may be observed as increases in infections and tumor incidences and by changes in the T-cell compartments of lymphoid tissues. In the case of immunosuppressive drugs such as cyclophosphamide and cyclosporin A, the immune effects seen in rodents are similar to those observed in the clinic (Dean et al. 1987).

If perturbations are observed in any hematologic or histopathologic indicators of immunotoxicity, it is then prudent to follow up these findings with one or more of the following:

- Use of special immunochemical and cytological assays that can be performed retrospectively on samples taken from the animals in question
- Use of more specific in vitro assays to further assess effects on the pertinent target system and potential mechanism of activity
- Use of more specific in vivo and ex vivo assays to determine toxicological significance.
- Inclusion of additional nonroutine parameters for immunotoxicity assessment in subsequent (longer-term) toxicity assays. Can also include additional satellite groups for functional tests that may require coadministration of adjuvants, pathogens, or tumor cells

3.19.9 Suggested Approaches to Evaluation of Results

Several rodent toxicity studies have shown impaired host resistance to infectious agents or tumor cells at exposure levels of drugs that did not cause overt signs of toxicity (Vos 1977; Dean et al. 1982). One serious limitation to the incorporation of specific immunotoxicological evaluations into general use in safety assessment for pharmaceuticals is a lack of clarity in how to evaluate and use such findings. This problem is true for all new diagnostic techniques in medicine and for all the new and more sensitive tools designed to evaluate specific target organ toxicities. Ultimately, as we have more experience and a reliable data base that allows us to correlate laboratory findings with clinical experience, the required course of action will become clearer. However, some general suggestions and guidance can be offered.

- First, it is generally agreed that adverse effects observed above a certain level of severity should be given the same importance as any other life-threatening events when assessing biological significance. These are effects that are so severe that they are detected as part of the routine evaluations made in safety assessment studies. Such findings may include death, severe weight loss, early appearance of tumors, and the like. Findings such as significantly increased mortalities in a host-resistance assay would also fit into this category.
- Second, there are specific end point assays for which an adverse outcome clearly dictates the action to be taken. These end points include either immediate or delayed hypersensitivity reactions, because once the individual is sensitized, a dose-response relationship may not apply.
- Third, as with most toxicological effects, toxic effects to the immune system are dependent upon dose to the target site. The dose-response curve can be used to determine no-effect and low-effect levels for immunotoxicity. These levels can then be compared to the therapeutic levels to assess whether there is an adequate margin of safety for humans.

If we consider both the specific immunotoxicity assays surveyed earlier in the chapter and the arrays of end points evaluated in traditional toxicology studies, which may be indicative of an immune system effect, these guidelines leave many potential questions unanswered. As additional data on individual end points indicative of immune system responses are collected, the pharmaceutical toxicologist is challenged with various issues regarding assay interpretation and relevance to proposed (or future) clinical trials. For example, what do significant, but non-life-threatening, decreases in antibody response, lymphocyte numbers, macrophage functions, or host resistance in an animal mean about the clinic use of a drug in a patient? The intended patient population is clearly relevant here—if the disease is one in which the immune system is already challenged or incorrectly modulated, any immune system effect other than an intended one should be avoided. There are several additional considerations and questions that should be answered when evaluating the biological and clinical significance of a statistically significant immune response:

- *Is there a dose response?* The dose response should be evaluated as a dose-related trend in both incidence and severity of the response. If there is a dose-related response, is the lowest dose (preferably plasma level) at which the effect is seen near or below the target clinical dose (plasma level), and is there an adequate therapeutic margin of safety?
- *Does the finding stand alone?* Is a change observed in only one parameter, or are there correlated findings that suggest a generalized, biologically significant effect? For example, are there changes in lymph node and spleen weights and morphological changes in these tissues to accompany changes in lymphocyte numbers?
- *Is the effect a measure of function or a single end point measurement?* Functional measures such as host resistance or phagocytosis involve multiple cells and immunocomponents and, therefore, are considered to be more biologically relevant than a significant change in a single end point measurement (e.g., T-cell number).
- *Is the effect reversible?* Reversibility of a response is dependent on the drug itself, exposure levels/duration, and factors related to the test animal (metabolic capability, genetic susceptibility, etc.). Most effects produced by immunosuppressive drugs have been shown to be reversible after cessation of therapy, such as those produced during cancer chemotherapy. However, if a tumor develops before the immune system is restored, the effect is not reversible, as is the case of secondary tumors related to chemotherapy.
- *Is there sufficient systematic toxicity data available at levels that demonstrate adequate exposure?* If a study was designed such that there was insufficient exposure or duration of exposure to potential lymphoid target tissues, the test protocol may not be adequate to demonstrate an adverse effect.

In general, a well-conducted long-term study in two species, with no indication of immunotoxicity, based on the considerations outlined above, should be adequate to evaluate the potential for drug-induced immunotoxicity. If the results from these studies do not produce evidence of immune-specific toxicity after examination of standard and/or additional hematologic, serum chemical, and histopathologic parameters, then additional testing should not be indicated. However, if there are structure-activity considerations that may indicate a potential for concern, or if significant abnormalities are observed that cannot be clearly attributed to other toxicities, then it is important to perform additional tests to fully assess the biological significance of the findings.

3.19.10 Problems and Future Directions

There are some very pressing problems for immunotoxicology, particularly in the context of pharmaceuticals and biological therapeutics and the assessment of their safety. Unlike industrial chemicals, environmental agents, or agricultural chemicals,

pharmaceutical products are intended for human exposure, are usually systemically absorbed, and have intentionally biological effects on man—some of which are intentionally immunomodulating (interleukins, growth factors) or immunotoxic (cyclosporin, cyclophosphamide).

3.19.10.1 Data Interpretation

The first major issue was presented and explored in the preceding section. This is how to evaluate and utilize the entire range of data that current immunotoxicological methodologies provide to determine the potential for immunotoxicity and how to interpret the biological significance of minor findings.

3.19.10.2 Appropriate Animal Models

As previously addressed, most routine preclinical toxicology tests are performed with rats and dogs; therefore, toxicity, pharmacokinetic, and pharmacology data are most abundant for these species. However, most immunological parameters are best characterized and validated with mice. In addition, the NTP test battery was developed for the mouse, and some of these assays cannot be readily transferred to the rat. Over the last few years, several laboratories have begun adapting tests to both the rat and dog (Bloom et al. 1985a, b; Thiem et al. 1988); however, efforts need to continue along these lines to further our understanding of the immune responses in these species and how they correlate with other animal models and man.

3.19.10.3 Indirect Immunotoxic Effects

A problem related to data interpretation is how to distinguish secondary effects that may indirectly result in immunotoxicity from the primary effects of immunotoxicity in preclinic toxicity studies. Various factors may produce pathology similar to that of an immunotoxin, including:

- Stress in a chronically ill animal as related to general toxicity, such as lung or liver damage, can result in immune suppression.
- Malnutrition in animals with drug-induced anorexia or malabsorption can trigger immune suppression.
- Infections and/or parasites may also modulate immune parameters.

These indirect factors must be systematically ruled out, and additional mechanistic studies may be necessary to address this problem. The potential for some indirect effects may be assessed through histopathologic evaluation of endocrine organs such as the adrenals and the pituitary.

3.19.10.4 Hypersensitivity Tests

Probably the largest immunotoxicity concern in clinical studies is unexpected hypersensitivity reactions. While the available guinea pig- and mouse-based tests for delayed contact hypersensitivity resulting from dermal exposure are generally good predictors, there are currently no well-validated models for either immediate or delayed hypersensitivity responses resulting from either oral ingestion or parenteral administration. Yet these two situations are the largest single cause for discontinuing clinical trials.

One assay that may hold some promise for delayed hypersensitivity is an adoptive transfer-popliteal lymph node assay (Gleichmann et al. 1989a, b). This assay, based on the techniques previously described for the popliteal lymph node assay, allows assessment of hypersensitivity following systemic exposure of the drug. Donor mice are first injected with drug for five consecutive days. After a 4-week latency period, potentially sensitized T cells obtained from the spleen are injected into the footpad of a syngeneic mouse together with a subcutaneous challenge dose of the drug. Two to five days after the cell transfer, the popliteal lymph nodes are measured and observed for evidence of a response (enlargement). Once this assay is validated, it should allow for a more relevant assessment of hypersensitivity for drugs that are administered systemically (Gleichmann et al. 1989a, b).

3.19.10.5 Autoimmunity

Traditional methods for assessing immunotoxicity as part of routine preclinical toxicity tests are primarily geared toward the detection of immunosuppressive effects. Although it is possible to incorporate clinical methods for detecting immune complexes and autoantibodies into the preclinical test protocols, the significance of adverse findings is ambiguous. Since these effects have a genetic component to their expression, the relevance of findings in animals is of questionable significance, particularly since these findings in the clinic do not always correlate with pathological effects.

3.19.10.6 Functional Reserve Capacity

As previously discussed, the immune system has a tremendous reserve capacity that offers several levels of protection and backups to the primary response. As a consequence, this functional reserve can allow biologically significant, immunotoxic insults to occur without the appearance of morphological changes. Furthermore, adverse effects may remain subclinical until the organism is subjected to undue stress or subsequent challenge. Thus, there is some concern that routine immunopathologic assessments by themselves may not be sufficiently sensitive to detect all immunotoxins, particularly when testing is conducted in a relatively pathogen-free, stress-free laboratory environment.

3.19.10.7 Significance of Minor Perturbations

Although the immune system has a well-developed reserve capacity, some of these systems may act synergistically rather than independently. For instance, a macrophage can recognize and kill bacteria coated with antibodies more effectively than can either the macrophage or antibodies alone. Thus, even minor deficiencies and impairments may have some impact on the organism's ability to fend off infection or tumors, particularly if the organism is very young, old, ill, stressed, genetically predisposed to certain cancers, or otherwise immunocompromised. These considerations lead to some additional questions that must be addressed:

- What level of immunosuppression will predispose healthy or immunocompromised individuals to increased risk of infection or tumors?
- What slight disturbances or immunosuppression lead to a prolonged recovery from viral or bacterial infections?
- Will slight up-modulation for extended periods result in autoimmune diseases or increased susceptibility to allergy?
- Are individuals that are slightly immunosuppressed at higher risk of developing AIDS after exposure to HIV?

The blind men and the elephant, this parable applies all too well to evaluating the relevance of isolated measures of effect on the immune system to truly adverse changes in patient health. We can measure with great sensitivity and precision changes in many components of the system. But we do not yet have sufficient understanding to know where the right line is that says it is an indication of an adverse effect on health.

3.19.11 Challenges and Special Cases

There are a number of issues for which our present testing scheme does not work well or does not address at all.

- Foremost would have to be anaphylaxis—that is, cell-mediated Type I immediate hypersensitivity. While our tests for potential Type IV delayed contact hypersensitivity (DTH) are required for all medical devices with patient contact and do a good job of identifying such potential risks, the potential harm to a patient is limited and transitory. Not so for Type I—here there is a real risk of severe harm and even death. And no predictive testing is performed. Prior to harmonization of nonclinical drug testing under the ICH, the Japanese had and required guinea pig-based tests for both active and passive anaphylaxis (Müller and Kley 1982 and Verdier et al. 1994).
- Inhalation/respiratory devices are not addressed by ISO 10993, nor are there any guidances or requirements for biocompatibility testing. What does exist is ISO (2017). *ISO 18562: Biocompatibility Evaluation of Breathing Gas Pathways in*

Healthcare Application. International Organization for Standardization (ISO), Brussels, Belgium. These provide guidances for physical testing (quantitation of particle sizes below certain sizes) and chemical identification, quantitation, and subsequent risk assessment of extracts from gas pathway components of such devices. Inhalation toxicology is a complex field, with special concerns that are unlikely to be adequately understood by those not in the field.

- A growing number of devices intended for implantation, ranging from cardiovascular stents to repair matrices for bones. With the basis for determining what biocompatibility testing will be required for a device having duration of patient exposure as a fundamental component, and the basis of performing a quantitative risk assessment likewise being dependent on both rate and total duration of patient exposure to extracted chemical moieties, the lack of certainty for the time dimension associated with resorbable devices presents a cause of uncertainty and a problem.
- Ocular devices (contact lenses and their care products, interocular lenses, and specialty devices such as the Vision Care implantable telescope for macular degeneration) represent a special case. They have their own specialized guidances (ISO 9394 and ISO 11979-5) fundamental depend on elements of the ISO-10993 set, but the two specialized guidances also prescribe specialized tests.
- While ISO 10993-1 has situations under which the performance of reproductive toxicity effects should be considered, and a specific guidance document (along with genotoxicity and carcinogenicity), ISO 10993-3, which addresses the design of such studies, it is rare that actual studies are conducted.

References

AAMI. (2014). *Biological evaluation of medical devices (2009, reaffirmed 2014) cytotoxicity ISO10993-5* (pp. 69–82). Arlington: Association for the Advancement of Medical Instrumentation.

Advisory Committee on Dangerous Pathogens. (1984). Acquired immune deficiency syndrome (AIDS)-Interim Guidelines. In *Circular GC(85)2*. Haywood: Health Publications Unit.

Albanese, R. (1987). Mammalian male germ cell cytogenetics. *Mutagenesis, 2*, 79–85.

American Society for Testing and Materials. (2010). *ASTM F 756–87: Standard practice for assessment of hemolytic properties of materials*. Philadelphia: ASTM.

Ames, B. N. (1971). The detection of chemical mutagens with enteric bacteria. In A. Hollaender (Ed.), *Chemical mutagens, principles and methods for their detection* (Vol. 1, pp. 267–282). New York: Plenum Press.

Ames, B. N., McCann, J., & Yamasaki, E. (1975). Methods for detecting carcinogens and mutagens with the Salmonella/mammalian-microsome mutagenicity test. *Mutation Research, 31*, 237–364.

Anderson, J. M. (1988). Inflammatory response to implants. *Transactions of the American Society of Artificial Internal Organs, 34*, 101–107.

Arnold, D. L., Grice, H. C., & Krawksi, D. R. (1990). *Handbook of in vivo toxicity testing*. San Diego: Academic Press.

Arts, J. H. E., Droge, S. C. M., Bloksma, N., & Kuper, C. F. (1996). Local lymph node activation in rats after application of the sensitizers 2,4-dinitrichlorobenzene and trimellic anhydride. *Food and Chemical Toxicology, 34*, 55–62.

ASTM. (1991). *Standard Practice for testing biomaterials in rabbits for primary skin irritation* (Designation F 719–81). Philadelphia: ASTM.

Auletta, C. (1998). Acute systemic toxicity testing. In S. Gad (Ed.), *Handbook of product safety assessment* (2nd ed.). New York: Marcel Dekker.

Autian, J. (1977). Toxicological evaluation of biomaterials: Primary acute toxicity screening program. *Artificial Organs, 1*, 53–60.

Bagley, D. M., Rizvi, P. Y., Kong, B. M., & DeSalva, S. J. (2008). Factors affecting the use of the hens' egg chorioallantoic membrane assay as a model for eye irritation potential. *Toxicology Cutaneous and Ocular Toxicology, 10*(2), 95–104.

Ballantyne, B. (2009). Repeated exposure toxicity. In B. Ballantyne, T. Marrs, & T. Syversen (Eds.), *General and applied toxicology* (3rd ed., pp. 73–88). Hoboken: Wiley.

Barile, F. A. (1994). *Introduction to in vitro cytotoxicity, mechanisms and methods*. Boca Raton: CBC Press.

Bartsch, H., Camus, A. M., & Malaveille, C. (1976). Comparative mutagenicity of N-nitrosamines in a semi-solid and in a liquid incubation system in the presence of rat or human tissue fractions. *Mutation Research, 37*, 149–162.

Basketter, D. A., & Scholes, E. W. (1992). Comparison of the local lymph node assay with the guinea pig maximization test for the detection of a range of contact allergens. *Food and Chemical Toxicology., 30*(1), 65–69.

Basketter, D. A., Scholes, E. W., Kimber, I., Botham, P. A., Hilton, J., Miller, K., Robbins, M. C., Harrison, P. T. C., & Waite, S. J. (1991). Interlaboratory evaluation of the local lymph node assay with 25 chemicals and comparison with guinea pig test data. *Toxicology Methods, 1*(1), 30–43.

Battisto, J. R., Claman, H. N., & Scott, D. W. (1983). Immunological tolerance to self and nonself. *Annals of the New York Academy of Sciences, 392*, 1–433.

Bernstein, M. L., & Carlish, K. (1979). The induction of hyperkeratotic white lesions in hamster cheek pouches with mouthwash. *Oral Surgery, 48*, 517.

Beutler, E., Kuht, W., & Est, C. (1982). The osmostic fragility of ethrycytes after prolonged liquid storage and after reinfusion. *Blood, 59*, 1141.

Black, J. (1981). Rating system for tissues at animal implant sites. In *Biological performance of materials, fundamentals of biocompatibility* (p. 220). New York: Marcel Dekker, Appendix 2.

Black, J. (2000). *Biological performance of materials* (4th ed.). New York: Taylor & Francis Group.

Blein, O., Adolphe, M., Lakhdar, B., et al. (1991). Correlation and validation of alternative methods to the Draize eye irritation test (OPAL project). *Toxicology In Vitro, 5*, 555–557.

Braithwaite, I., & Ashby, J. (1988). A non-invasive micronucleus assay in rat liver. *Mutation Research, 203*, 23–32.

Bridges, B. A. (1976). Short-term screening tests for carcinogens. *Nature, 261*, 195–200.

Brown, B. A. (1993). *Hematology: Principles and procedures* (6th ed.). Philadelphia: Lea and Febiger.

Brusick, D. (1987). *Principles of genetic toxicology* (2nd ed.). New York/London: Plenum Press.

Buehler, E. V. (1964). A new method for detecting potential sensitizers using the guinea pig. *Toxicology and Applied Pharmacology, 6*, 341.

Buehler, E. V. (1965). Delayed contact hypersensitivity in the guinea pig. *Archives of Dermatology, 91*, 171–177.

Bulich, A. A., Tung, K. K., & Scheibner, G. (1990). The luminescent bacteria toxicity test: Its potential as an in vitro alternative. *Journal of Bioluminescence and Chemiluminescence, 5*(2), 71–77.

Burnett, C., Fuchs, C., Corbett, J., & Menkart, J. (1982). The effect of dimethylsulphoxide on the mutagenicity of the hair-dye, p-phenylenediamine. *Mutation Research, 103*, 1–4.

Christensen, O. B., Christnsen, M. B., & Maibach, H. I. (1984). Effect of vehicle on elicitation of DNCB contact allergy in the guinea pig. *Contact Dermatitis, 10*, 166–169.

Chvapil, M., Chvapil, T. A., Owen, J. A., Kantor, M., Ulreich, J. B., & Eskelson, C. (1979). Reaction of vaginal tissue of rabbits to inserted sponges made of various materials. *Journal of Biomedical Materials Research, 13*(1), 1–13.

Ciapetti, G., Granchi, D., Verri, E., Savarino, L., Cavedagna, D., & Pizzoferrato, A. (1996). Application of combination of neutral red and amido black staining for rapid, reliable cytotoxicity testing of biomaterials. *Biomaterials, 17*, 1259–1264.

Cifone, M. A., Myhr, B., Eiche, A., & Bolisfodi, G. (1987). Effect of pH shifts on the mutant frequency at the thymidine kinase locus in mouse lymphoma L5178Y TK+/− cells. *Mutation Research, 189*, 39–46.

Clark, W. R. (1983). *The experimental foundations of modern immunology* (2nd ed., pp. 1–453). New York: John Wiley & Sons.

Coombs, R. R. A., & Gell, P. G. H. (1975). Classification of allergic reactions responsible for clinical hypersensitivity and disease. In P. G. H. Gell, R. A. A. Coombs, & D. J. Lachman (Eds.), *Clinical aspects of immunology* (p. 761). Oxford: Blackwell Scientific Publications.

Cooper, J. F. (1975). Principles and applications of the limulus test for pyrogen in parenteral drugs. *Bulletin of the Parenteral Drug Association, 3*, 122–130.

Cooper, S. L., Fabrizius, D. J., & Grasel, T. G. (1987). Methods of assessment of thrombosis ex vivo. In E. F. Leonard, V. T. Turitto, & L. Vroman (Eds.), *Blood in contact with natural and artificial surfaces* (Vol. 516, pp. 572–585). New York: Ann N.Y. Acad Sciences.

Cronin, E., & Agrup, G. (1970). Contact dermatitis X. *British Journal of Dermatology., 82*, 428–433.

Dean, B. J., & Danford, N. (1984). Assays for the detection of chemically-induced chromosome damage in cultures mammalian cells. In S. Venitt & J. M. Parry (Eds.), *Mutagenicity testing. A practical approach* (pp. 187–232). Oxford: IRL Press.

DeSerres, F. J., & Shelby, M. D. (1979). Recommendations on datas production and analysis using the Salmonella/microsome mutagenicity assay. *Mutation Research, 64*, 159–165.

Devleeschouwer, M. J., Cornil, F. F., & Dony, J. (1985). Studies on the sensitivity and specificity of the limulus amebocyte lysate test and rabbit pyrogen assays. *Applied and Environmental Microbiology, 50*, 1509–1511.

Dewangee, M. K. (1987). Methods of assessment of thrombosis *in vivo*. In E. F. Leonard, V. T. Turitto, & L. Vroman (Eds.), *Blood in contact with natural and artificial surfaces* (Vol. 516, pp. 541–571). New York: Ann N.Y. Acad Sciences.

Di Silvio, L. (2010). *Cellular response to biomaterials*. Cambridge: Woodhead Publishing Ltd.

Didisheim, P., Dewanjee, M. K., Kaye, M. P., Frisk, C. S., Fass, D. N., Tirrell, M. V., & Xillman, P. E. (1984). Nonpredictability of long-term *in vivo* response from short-term *in vitro* or *ex vivo* blood/material interactions. *Transactions of American Society of Artificial Internal Organs, 30*, 370–376.

Doll, R. (1977). Strategy for detection of cancer hazards to man. *Nature, 265*, 589–596.

Dorado, G., Ariza, R. R., & Pueyo, C. (1988). Mutagenicity of red wine in the L-arabinose resistance test with Salmonella typhimurium. *Mutagenesis, 3*, 497–502.

Draize, J. H. (1955). *Dermal Toxicity. Association of Food and Drug Officials of the U.S* (pp. 46–59). Washington, D. C.: FDA.

Draize, J. H. (1959). *Dermal toxicity. appraisal of the safety of chemicals in foods, drugs, and cosmetics*. Austin: Association of Food and Drug Officials of the United States, Texas State Department of Health, Texas.

Draize, J. H., Woodard, G., & Calvery, H. O. (1944). Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. *The Journal of Pharmacology and Experimental Therapeutics, 82*(3), 377–390.

Eckstein, P., Jackson, M. C. N., Millman, N., & Sobrero, A. J. (1969). Comparison of vaginal tolerance tests of spermicidal preparations in rabbits and monkeys. *Journal of Reproduction and Fertility, 20*, 85–93.

Elsberry, D. (1986). Screening approaches for acute ad subacute toxicity studies. In W. Lloyd (Ed.), *Safety evaluation of drugs and chemicals* (pp. 145–151). Washington D.C.: Hemisphere.

European Pharmacopeia. (2019). Ph. *Eur,* 10th Ed.

Evans, H. J., & O'Riordan, M. L. (1975). Human peripheral blood lymphocytes for the analysis of chromosome aberrations in mutagen tests. *Mutation Research, 31*, 135–148.

Exon, J. H., Henningsen, G. M., Osborne, C. A., & Koller, L. D. (1984). Toxicologic, pathologic, and immunotoxic effects of 2,4-dichlorophenol in rats. *Journal of Toxicology and Environmental Health, 14*(5–6), 723–730.

FDA. (1988). Investigational new drug, antibiotic, and biological drug product regulations, procedures for drugs intended to treat life-threatening and severely debilitating illness. *Federal Register, 53*(204), 41516–41524.

Fergeson, A. B., Jr., Laing, P. G., & Hodge, E. S. (1960). The ionoization of metal implants in living tissues. *Journal of Bone and Joint Surgery, 42A*, 77–90.

Forster, R., Green, M. H. L., & Priestley, A. (1980). Optimal Levels of S9 fraction in Ames and fluctuation tests: Apparent importance of diffusion of metabolites from top agar. *Carcinogenesis, 2*, 1081–1085.

Frazier, J. M. (1987). Update: A critical evaluation of alternative to acute ocular irritancy testing. In A. M. Goldberg (Ed.), *Progress in in vitro toxicology* (pp. 67–75). New York: Mary Ann Lierbert.

Frazier, J. M., Gad, S. C., Goldberg, A. M., & McCulley, J. P. (1987). *A critical evaluation of alternative to acute ocular irritancy testing*. New York: Mary Ann Liebert.

Gad, S. C. (2000). *Vitro toxicology* (2nd ed.). Philadelphia: Taylor and Francis.

Gad, S. C. (2006). *Statistics and experimental design for toxicologists* (4th ed.). Boca Raton: CRC Press.

Gad, S. C. (2010). *International regulatory safety evaluation of pharmaceuticals and medical devices*. Berlin: Springer.

Gad, S. C. (2015). *Animal models in toxicology* (3nd ed.). Boca Raton: CRC Press.

Gad, S. C., & Chengelis, C. P. (1998). *Acute toxicology testing* (2nd ed.). San Diego: Academic Press.

Gad, S., & Chengelis, C. P. (1999). *Acute toxicology testing* (2nd ed.). San Diego: Academic Press.

Gad, S. C., & Gad-McDonald, S. E. (2016). *Biomaterials, medical devices, and combination products: Biocompatibility testing and safety assessment*. Boca Raton: CRC Press.

Gad, S. C., Dobbs, D. W., Dunn, B. J., Reilly, C., & Walsh, R. D. (1985). *Elucidation of the delayed contact sensitization response to Croton oil*, presented at the American College of Toxicology in Washington, DC, November 1985.

Gad, S. C., Darr, R. W., Dobbs, D. W., Dunn, B. J., Reilly, C., & Walsh, R. D. (1986). *Comparison of the potency of 52 dermal sensitizers in the Mouse Ear Swelling Test (MEST)*, presented at SOT Meetings, March 1986.

Gatehouse, D. G., & Tweats, D. J. (1988). Letter to the editor. *Mutagenesis, 1*, 307–308.

Gatehouse, D., & Wedd, D. J. (1984). The differential mutagenicity of isoniazid in fluctuation assays and Salmonella plate tests. *Carcinogenesis, 5*, 391–397.

Gatehouse, D. G., Wilcox, P., Forster, R., Rowland, I. R., & Callander, R. D. (1990). Bacterial mutation assays. In D. J. Kirkland (Ed.), *Basic mutagenicity tests: UKEMS. Recommended procedures* (pp. 13–61). Cambridge: Cambridge University Press.

Geller, W., Hobel, W., & Seifert, G. (1985). Overview of animal test methods for skin irritation. *Food and Chemical Toxicology, 23*(2), 165–168.

Geret, V., Rahan, B. A., Mathys, R., Straumann, F., & Perren, S. M. (1980a). In vivo testing of tissue tolerance of implant meterials: Improved quantitative evaluation through reduction of relative motion at the implant tissue interface. In H. K. Uhthoff (Ed.), *Current concepts of internal fixation of fracture* (pp. 108–126). New York: Springer.

Geret, V., Rahan, B. A., Mathys, R., Straumann, F., & Perren, S. M. (1980b). A method for testing tissue tolerance for improved quantitative evaluation through reduction of relative motion at the implant-tissue interace. In G. D. Winter, J. L. Leray, & K. de Groot (Eds.), *Evaluation of biomaterials* (pp. 780–798). London: Wiley.

Gibson, G. G., Hubbard, R., & Parke, D. U. (1983). *Immunotoxicology*. San Diego: Academic.
Godfrey, H. P., & Gell, P. G. H. (1978). Cellular and molecular events in the delayed-onset hypersensitivities. *Reviews in Physiological and Biochemical Pharmacology, 84*, 2.
Goldhaber, P. (1961). The influence of pore size on carcinogenicity of subcutaneously implanted Millipore filters. *Proceeding of American Association for Cancer Research, 3*, 228.
Goldhaber, P. (1962). Further observations concerning the carcinogenicity of Millipore filters. *Proceeding of American Association for Cancer Research, 3*, 323.
Grant, M. W. (1993). *Toxicology of the eye* (4th ed.). Springfield: Thomas.
Greco, R. S. (1994). *Implantation biology*. Boca Raton: CRC Press.
Guelcher, S. A., & Hollinger, J. O. (2006). *An introduction to biomaterial*. Boca Raton: Taylor & Francis Group.
Guillot, J. P., & Gonnet, J. F. (1985). The Epicutaneous maximization test. *Current Problems in Dermatology, 14*, 220–247.
Haschek, W. M., & Rousseaux, C. G. (1991). *Handbook of toxicology pathology*. San Diego: Academic Press.
Haugen, E. (1980). The effect of periodontal dressings on intact mucous membrane and on wound healing. A methodological study. *Acta Odontologica Scandinavica, 8*, 363.
Heddle, J. A., & Bodycote, D. J. (1970). On the formation of chromosomal aberrations. *Mutation Research, 9*, 117–126.
Hera, C., & Pueyo, C. (1986). Conditions for optimal use of the l-arabinose-resistance mutagenesis test with Salmonella typhimurium. *Mutagenesis, 1*, 267–274.
Hipp, I. I. (1978). The skin and industrial dermatosis. *National Safety News,* April:12–16.
IARC. (1979). *Handling chemical carcinogens in the laboratory; problems of safety, scientific publications, no. 33*. International Agency for Research on Cancer, Lyons.
ICEMC. (1983). Committee & Final Report: Screening strategy for chemicals that are potential germ-cell mutagens in mammals. *Mutation Research, 114*, 117–177.
Irons, R. D. (Ed.). (1985). *Toxicology of the blood and bone marrow*. New York: Raven Press.
Ishidate, M., Jr. (1988a). *Data book of chromosomal aberration tests in vitro*. Amsterdam: Elsevier.
Ishidate, M., Jr. (1988b). A proposed battery of tests for the initial evaluation of the mutagenic potential of medicinal and industrial chemicals. *Mutation Research, 205*, 397–407.
Ishidate, M., Jr., & Harnois, M. C. (1987). The clastogenicity of chemicals in mammalian cells. Letter to the editor. *Mutagenesis, 2*, 240–243.
Ishidate, M., Jr., Sofuni, T., Yoshikawa, K., et al. (1984). Primary mutagenicity screening of food additives currently used in Japan. *Food and Chemical Toxicology, 22*, 623–636.
ISO-05832. (2016). *Implants for surgery- Metallic Material*. Silver Springs, MD.
ISO-10993-10. (2010). *Biological evaluation of medical devices*- Part 10 Tests for irritation and skin sensitization. Silver Springs, MD.
ISO 10993-12. (2012). *Biological evaluation of medical devices* – Part 12 Sample Preparation and Reference Materials.
ISO Japan. (2009). *Biocompatibility safety assessment of medical devices: FDA/ISO and Japanese guidelines*. Medical Device & Diagnostic Industry, January.
JMHW. (1984). *Guidelines for testing of drugs for toxicity*. (Pharmaceutical Affairs Bureau, Notice No. 118). Ministry of Health and Welfare, Japan.
Kaminsky, M., & Willigan, D. A. (1982). pH and the potential irritancy of douche formulations to the vaginal mucosa of the albino rabbit and rat. *Food and Chemical Toxicology, 20*, 193.
Kaminski, E. J., Oblesby, R. J., Wood, N. K., & Sandrik, J. (1968). The behavior of biological materials at different sites of implantation. *Journal of Biomedical Materials Research, 2*, 81–86.
Kimber, I., & Dearman, R. J. (1994). Immune responses to contact and respiratory allergens. In J. H. Dean, M. I. Luster, E. A. Munson, & I. Kimber (Eds.), *Immunotoxicology and Immunopharmacology* (pp. 663–679). New York: Raven Press.
Kimber, I., & Weisenberger, C. (1989). A murine local lymph node assay for the identification of contact allergens: Assay development and results of an initial validation study. *Archives of Toxicology, 63*, 274–282.

Kimber, I., Mitchell, J. A., & Griffin, A. C. (1986). Development of an immurine local lymph node assay for the determination of sensitizing potential. *Food and Chemical Toxicology, 24*, 585–586.

Kirkland, D. J., & Garner, R. C. (1987). Testing for genotoxicity-chromosomal aberrations in vitro—CHO cells or human lymphocytes? *Mutation Research, 189*, 186–187.

Kirkland, D. J., Marhsall, R. R., MCEnaney, S., Bidgood, J., Rutter, A., & Mullineux, S. (1989). Aroclor-1254-induced rat-liver S9 causes chromosomal aberrations in CHO cells but not human lymphocytes: A role for active oxygen? *Mutation Research, 214*, 115–122.

Kligman, A. M. (1966). The identification of contact allergens by human assay. III. The maximization test. A procedure for screening and rating contact sensitizers. *The Journal of Investigative Dermatology, 47*, 393–409.

Kordan, H. A. (1967). Localized interfacial forces resulting from implanted plastics as possible physical factors involved in tumor formation. *Journal of Theoretical Biology, 17*, 1–11.

Landsteiner, K., & Jacobs, J. (1935). Studies on sensitization of animals with simple chemical compounds. *The Journal of Experimental Medicine, 61*, 643–656.

Lewis, S. M., Rowan, R. M., & Kubota, F. (1990). Recommended methods for the visual determination of white cell and platelet count, S. M. Lewis, R. M. Rowan and F. Kubota. *Journal of Clinical Pathology, 43*, 932–936.

Lilly, G. E., Cutcher, J. L., & Henderson, M. D. (1972). Reaction of oral and mucous membranes to selected dental materials. *Journal of Biomedical Materials Research, 6*, 545.

Lindhe, J., Heyden, G., Svanberg, G., Loe, H., & Schiott, C. R. (1970). Effect of local applications of chlorhexidine on the oral mucosa of the hamster. *Journal of Periodontal Research, 5*, 177.

Lindon, N. N., Rodvien, R., Brier, D., et al. (1978). In vitro assessment of interaction of blood with model surfaces. *The Journal of Laboratory and Clinical Medicine, 92*, 904–914.

Loveless, S. E., Ladics, G. S., Greberick, G. F., Ryan, C. A., Basketter, D. A., Scholes, E. W., House, R. V., Hilton, J., Dearman, R. J., & Kimber, I. (1996). Further evaluations of the local lymph node assay in the final phase of the international collaborative trial. *Toxicology, 108*, 141–152.

Luster, M. I., Munson, A. E., Thomas, P. T., Holsaple, M. P., Fenters, J. D., White, K. L., Jr., Lauer, L. D., Germolec, D. R., Rosenthal, G. L., & Dean, J. H. (1988). Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program's guidelines for immontoxicity evaluation in mice. *Fundamental and Applied Toxicology, 10*, 2–19.

MacGregor, J. T., Wehr, C. M., & Gould, D. H. (1980). Clastogen-induced micronuclei in peripheral blood erythrocytes: The basis of an improved micronucleus test. *Environmental Mutagenesis, 2*, 509–514.

Madle, S., & Obe, G. (1980). Methods for the analysis of the mutagenicity of indirect mutagens/carcinogens in eukaryotic cells. *Human Genetics, 56*, 7–20.

Magnusson, B. (1975). The relevance of results obtained with the Guinea pig maximization test. In H. Maibach (Ed.), *Animal models in dermatology* (pp. 76–83). Edinburgh: Churchill Livingstone.

Magnusson, B., & Kligman, A. M. (1969). The identification of contact allergens by animal assay. The Guinea pig maximization test. *The Journal of Investigative Dermatology., 52*, 268–276.

Magnusson, B., & Kligman, A. M. (1970). *Allergic contact dermatitis in the Guinea pig: Identification of contact allergens. Chap. 8*. Springfield: Thomas.

Maguire, H. C., & Chase, M. W. (1967). Exaggerated delayed-type hypersensitivity to simple chemical allergens in the Guinea pig. *The Journal of Investigative Dermatology, 49*, 460–468.

Mahon, G. A. T., Green, M. H. L., Middleton, B., Mitchell, I. D. G., Robinson, W. D., & Tweats, D. J. (1989). Analysis of data from microbial colon assays. In D. J. Kirkland (Ed.), *Statistical evaluation of mutagenicity test data* (pp. 26–65). Cambridge: Cambridge University Press.

Male, D., Champion, B., & Cooke, A. (1982). *Advanced immunology*. Philadelphia: J.B. Lippincott.

Maron, D. M., & Ames, B. N. (1983). Revised methods for the Salmonella mutagenicity test. *Mutation Research, 113*, 173–215.

Marzulli, F. N., & Maibach, H. L. (1996). *Dermatotoxicity* (5th ed.). Philadelphia: Taylor and Francis.

Matoltsy, A. G., Downes, A. M., & Sweeney, T. M. (1968). Studies of the epidermal water barrier. II. Investigation of the chemical nature of the water barrier. *The Journal of Investigative Dermatology, 50*(1), 19–26.

Maurer, T., Thomann, P., Weirich, E. G., & Hess, R. (1975). The optimization test in the Guinea pig. *Agents and Actions., 5*, 174–179.

MHW. (2000). *Guidelines for basic biological tests of medical materials and devices,* notification 99, MHW (Japan).

Miller, A. E., Jr., & Levis, W. R. (1973). Studies on the contact sensitization of man with simple chemicals, I. specific lymphocyte transformation in response to dinitrochlorobenzene sensitization. *The Journal of Investigative Dermatology, 61*, 261–269.

Milner, J. E. (1970). In vitro lymphocyte responses in contact hypersensitivity. *The Journal of Investigative Dermatology, 55*, 34–38.

Milner, J. E. (1971). In vitro lymphocyte responses in contact hypersensitivity II. *The Journal of Investigative Dermatology, 56*, 349–352.

Milner, J. E. (1983). In vitro tests for delayed skin hypersensitivity: lymphokine production in allergic contact dermatisis. In F. N. Marzulli & H. D. Maibach (Eds.), *Dermatotoxicology* (pp. 185–192). New York: Hemisphere Publishing.

Monash, S., & Blank, H. (1958). Location and reformation of the epithelial barrier to water vapor. *A.M.A. Archives of Dermatology, 78*(6), 710–714.

Montagna, W. (1961). *The structure and function of the skin* (2nd ed.). New York: Academic Press.

Morimoto, K., Sato, M., & Koizumi, A. (1983). Proliferative kinetics of human lymphocytes in culture measure by autoradiography and sister chromatid differential staining. *Epithelial Cell Research, 145*, 249–356.

MRC. (1981). *Guidelines for work with chemical carcinogens in Medical Research Council establishments*. London: Medical Research Council.

Muller, P., Raabe, G., Horold, J., & Juretzek, U. (1988). Action of chronic peracetic acid (Wofasteril) administration on the rabbit oral mucosa, vaginal mucosa and skin. *Experimental Pathology, 34*, 223–228.

NAS Publication 1138. (1977). *Principles and procedures for evaluating the toxicity of household substances.* Prepared for the Consumer Product Safety Commission, National Academy of Sciences, Washington, D.C., pp. 36–39.

Natarajan, A. T., & Obe, G. (1982). Mutagenicity testing with cultured mammalian cells: Cytogenetic assays. In J. A. Heddle (Ed.), *Mutagenicity, new horizons in genetic toxicology* (pp. 172–213). New York: Academic Press.

Nelson, D. L., & Cox, M. M. (Eds.). (2016). *Principles of biochemistry* (7th ed.). New York: W. H. Freeman and Company.

Neudecker, T., & Henschler, D. (1985). Allyl isothiocyanate is mutagenic in Salmonella typhimurium. *Mutation Research, 30*, 143–148.

Newton, M. F., & Lilly, L. J. (1986). Tissue specific clastogenic effects of chromium and selenium salts in vivo. *Mutation Research, 169*, 61–69.

NIH. (1985). *Guidelines for blood-materials interactions*. Report of the National Heart, Lung, and Blood Institute-working group, (NIH publication no. 85–2185), September, 1985.

Nixon, G. A., Buehler, E. V., & Newman, E. A. (1972). Preliminary safety assessment of disodium etidronate as an additive to experimental oral hygiene products. *Toxicology and Applied Pharmacology, 22*, 661.

OECD. (1983). OECD guidelines for the testing of chemicals. No. 475. *Genetic toxicology: in vivo mammalian bone marrow cytogenetic test—chromosomal analysis*. Adopted 4 April 1984.

Pagano, D. A., & Zeiger, E. (1985). The stability of mutagenic chemicals tested in solution. *Environmental Mutagenesis, 7*, 293–302.

Pennisi, I. (1996). Teetering on the brink of danger. *Science, 271*, 1665–1667.

Piegorsh, W. (1989). Quantification of toxic response and the development of the median effective dose (ED50) - a historical perspective. *Toxicology and Industrial Health, 5*, 55–62.

Pizzoferrato, S., Savarino, L., Stea, S., & Tarabusi, C. (1988). Result of histological grading on 100 cases of hip prosthesis failure. *Biomaterials, 9*, 314–318.

Preston, R. J., San Sebastian, J. R., & McFee, A. F. (1987). The in vitro human lymphocyte assay for assessing the clastogenicity of chemical agents. *Mutation Research*, *189*, 175–183.

Preston, R. J., Au, W., Bender, M., et al. (1981). Mammalian in vivo and in vitro cytogenetic assays. *Mutation Research, 87*, 143–188.

Pueyo, C., & Lopez-Barea, J. (1979). The L-arabinose-resistance test with Salmonella typimurium strain SV3 selects forward mutations at several ara genes. *Mutation Research, 64*, 249–258.

Rahn, B. A., Geret, V., Capaul, C., Lardi, M., & Solothurnmann, B. (1982). Morphometric evaluation of tissue reaction to implants using low cost digitizing techniques. In A. J. C. Lee, T. Albrektsson, & P. I. Branemark (Eds.), *Clinical application of biomaterials* (pp. 112–131). New York: Wiley.

Rhodes, C. (2000). Principles of Testing for acute toxic effects. In B. Ballantyne, T. Marrs, & T. Szversen (Eds.), *General and applied toxicology* (pp. 33–54). London: Macmillan References, Ltd.

Richardson, C., Williams, D. A., Allen, J. A., Amphlett, G., Changer, D. O., & Phillips, B. (1989). Analysis of data from in vitro cytogenetic assays. In D. J. Kirkland (Ed.), *UKEMS subcommittee on guidelines for mutagenicity testing. Report. Part III. Statistical Evaluation of Mutagenicity Test Data* (pp. 141–154). Cambridge: Cambridge University Press.

Rosenbluth, S. A., Weddington, G. R., Guess, W. L., & Autian, J. (1965). Tissue culture method for screening toxicity of plastic materials to be used in medical practice. *Journal of Pharmaceutical Sciences, 54*, 156–159.

Scott, D., Danford, N., Dean, B., Kirkland, D., & Richardson, C. (1983). In vitro chromosome aberration assays. In B. J. Dean (Ed.), *UKEMS subcommittee on guidelines for mutagenicity testing, report part I, basic test battery* (pp. 63–86). Swansea: UKEMS.

Scott, D., Dean, B. J., Danford, N. D., & Kirkland, D. J. (1990). Metaphase chromosome aberration assays in vitro. In D. J. Kirkland (Ed.), *UKEMS subcommittee on guidelines for mutagenicity testing, report. Part I: Revised basic mutagenicity tests, UKEMS recommended procedures* (pp. 63–86). Cambridge: Cambridge University Press.

Sina, J. F., Ward, G. J., Laszeh, M. A., & Gautheron, P. D. (1992). Assessment of cytotoxic assays as predicators of ocular irritation of pharmaceuticals. *Fundamental and Applied Toxicology, 18*, 515–521.

Tates, A. D., Neuteboom, I., Hofker, M., & den Engelese, L. (1980). A micronucleus technique for detecting clastogenic effects of mutagens/carcinogens (DEN,DMN) in hepatocytes of rat liver in vivo. *Mutation Research, 74*, 11–20.

Thompson, E. D. (1986). Comparison of in vivo and in vitro cytogenetic assay results. *Environmental Mutagenesis, 8*, 753–767.

Thorne, P. S., Hillebrand, J. A., Lewis, G. R., & Karol, M. H. (1986). Contact sensitivity by diisocyanates: Potencies and cross-reactivities. *Toxicology and Applied Pharmacology, 87*, 155–165.

Thulin, H., & Zacharian, H. (1972). The leukocyte migration test in chromium hypersensitivity. *The Journal of Investigative Dermatology, 58*, 55–58.

Thust, R., Mendel, J., Schwarz, H., & Warzoki, R. (1980). Nitrosated urea pesticide metabolites and other nitrosamides. Activity in clastogenicity and SCE assays, and aberration kinetics in Chinese hamster V79-E cells. *Mutation Research, 79*, 239–248.

Topham, J., Albanese, R., Bootman, J., Scott, D., & Tweats, D. (1983). In vivo cytogenetic assays. In B. Dea (Ed.), *Report of UKEMS sub-committee on guidelines for mutagenicity testing. Part I* (pp. 119–141).

Traina, V. M. (1983). The role of toxicology in drug research and development. *Medicinal Research Reviews, 3*, 45–72.

TSCA. (1979). Toxic Substance Control Act.

Tsuchimoto, T., & Matter, B. E. (1979). In vivo cytogenetic screening methods for mutagens with special reference to the micronucleus test. *Archives of Toxicology, 42*, 239–248.

Turner, J. E., Lawrence, W. H., & Autian, J. (1973). Subacute toxicity testing of biomaterials using histopathologic evaluation of rabbit muscle tissue. *Journal of Biomedical Materials Research, 7*, 39–58.

USP. (2007). *Biological tests-plastics, the United States pharmacopoeia* (3rd ed., pp. 1235–1238). Rockville: United States Pharmacopoeial Convention.

USP. (2013). *United States pharmacopoeia 36th/National Formulary 31st,* United States Pharmacopoeia, Philadelphia.

Vos, J. G. (1977). Immune suppression as related to toxicology. *CRC Critical Reviews in Toxicology, 5*, 67.

Weary, M., & Baker, B. (1977). Utilization of the limulus amebocyte lysate test for pyrogen testing of large volume parenterals, administration sets and medical devices. *Bulletin of the Parenteral Drug Association, 31*, 1127–1133.

Williams, P. D., Masters, B. G., Evans, L. D., Laska, D. A., & Hattendorf, G. H. (1987). An in vitro model for assessing muscle irritation due to parenteral antibiotics. *Fundamental and Applied Toxicology, 9*, 10–17.

Wilsnack, R. E. (1976). Quantitative cell culture biocompatibility testing of medical devices and correlation to animal tests. *Biomaterials, Medical Devices, and Artificial Organs, 4*, 235–261.

Wilsnack, R. E., Meyer, F. J., & Smith, J. G. (1973). Human cell culture toxicity testing of medical devices and correlation to animal tests. *Biomaterials, Medical Devices, and Artificial Organs, 1*, 543–562.

Wolff, S. (1977a). Lesions that lead to SCEs are different from those that lead to chromosome aberrations. *Mutation Research, 46*, 164.

Wolff, S. (1977b). *In vitro inhibition of mono-oxygenase dependent reactions by organic solvents*, international conference on industrial and environmental Xenobiotics, Prague, Czechoslovakia.

Xu, L., Whong, W. Z., & Ong, T. M. (1984). Validation of the Salmonella (SV50)/L-arabinose-resistant forward mutation assay with 26 compounds. *Mutation Research, 130*, 79–86.

Yahagi, T., Degawa, M., Seino, Y., et al. (1975). Mutagenicity of carcinogen azo dyes and their derivatives. *Cancer Letters, 1*, 91–96.

Yahagi, T., Nagao, M., Seino, Y., Matsushima, T., Sugimura, T., & Okada, M. (1977). Mutagenicities of N-nitrosamines in Salmonella. *Mutation Research, 48*, 120–130.

Yanofsky, C. (1971). Mutagenesis studies with Escherichia coli mutants with known amino acid and base-pair changes. In A. Hollaender (Ed.), *Chemical mutagens, principles and methods for their detection* (Vol. 1, pp. 283–287). New York: Plenum Press.

Yokoyama, M., Nakahashi, T., Nishimura, T., Maeda, M., Inoue, S., Kataoka, K., & Sakurai, Y. (1986). Adhesion behavior of rat lymphocytes to poly (ether)-poly (amino acid) block and graft copolymers. *Journal of Biomedical Materials Research, 20*, 867–878.

Zaslavsky, B. Y., Ossipov, N. N., & Rogozhin, S. V. (1978). Action of surface-active substances on biological membranes. Ill. Comparison of hemolytic activity of ionic and nonionic surfactants. *Bhiochima et Biophysica Acta, 510*(1), 15–159.

Chapter 4
Considerations for Leachables and Extractables Testing

Amy L. Mihalchik-Burhans and Erica N. Rogers

Abstract This chapter provides a summary on the regulation, experimental design, and analytical techniques commonly utilized in extractables and leachables (E&L) testing approaches for medical devices. Extractables are indicated as chemical compounds generated under aggressive, exaggerated experimental conditions, while leachables are chemical compounds expected to be present in a medical device under clinical use conditions. Key regulatory guidances have been established to define the requirements for E&L assessment; the most important guidances commonly used in the USA for medical device testing are summarized here. As the purpose and testing design for E&L assessment of medical devices differ from container closure systems for drug products, the framework of E&L testing including characteristics and usage of various solvents, optimization of extraction conditions, and implementation of various analytical techniques to analyze the extractable profile for a device under a "worst-case" patient exposure scenario is outlined in this chapter. Furthermore, the chapter is written from the perspective of toxicologists who routinely assess the safety of medical devices and therefore is a suitable resource for beginning risk assessors to grasp a basic understanding of the E&L testing process and appropriate points of contact with analytical laboratory staff.

Keywords Extractables · Leachables · Controlled extraction study · ISO 10993:18 · ISO 10993:12

4.1 Introduction

Amy L. Mihalchik-Burhans

This chapter, written from the perspective of toxicologists who often assess extractables and leachables (E&L) data for pharmaceuticals and medical devices, is intended to provide a brief overview of experimental design and techniques commonly employed

A. L. Mihalchik-Burhans, PhD (✉) · E. N. Rogers, PhD
Gad Consulting Services, Raleigh, NC, USA
e-mail: amy.mihalchik@gadconsulting.com; erica.rogers@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_4

in E&L testing strategies for those with general interest in the topic and summarize recently updated regulatory materials. *Extractables* are chemical species forcibly extracted from a material using a selected solvent under specified experimental conditions (i.e., time, temperature, and pressure), while *leachables* are chemical species expected to be present in a product under clinical use conditions (Feilden 2011; Jenke 2012). For these substances, a general summary of the most common US regulatory guidelines and standards used is discussed in this chapter. The testing schema for medical devices differs from container closure systems for drug products in general design and intent; medical devices are usually assessed via a *controlled extraction study*, which uses multiple solvents of differing polarities, optimized extraction conditions, and various analytical techniques to determine the extractables profile for a device, which should represent a "worst case" patient exposure scenario without resulting in destruction or material degradation of the device. Following an initial assessment by a toxicologist, compounds belonging to the cohort of concern (e.g., aflatoxin-like compounds, N-nitroso compounds, azo compounds, polyhalogenated-dibenzodioxins, polyhalogenated-dibenzofurans, and polyhalogenated-biphenyls, strained heteronuclear rings, heavy metals, alpha-nitro furyl compounds, hydrazines/triazines/azides/azoxy compounds, polycyclic amines, steroids, and organophosphorus compounds) and compounds with TE values exceeding the potential daily exposure, in addition to compounds lacking adequate toxicity data for assessment and present above the TTC, should be assessed further in an additional targeted leachables study under conditions more reflective of clinical exposure (Feilden 2011; ISO 2019).

For toxicologists and risk assessors, detailed experimental protocols and study design utilized in E&L testing are often out of our purview, but here, we highlight time points in which E&L testing may be required as well as the roles of toxicologists and analytical chemists in tandem throughout general product development. In general, E&L assessment may be necessary or suggested:

- As part of a 510(k) premarket notification dependent upon degree of difference between the subject and predicate devices.
- Prior to submitting an investigational device exemption (IDE) as part of the premarket approval process (PMA).
- Prior to a de novo device submission (for novel devices only).
- Following a significant change in material(s) of construction used in a device.
- In order to better understand potential local toxicities observed in biocompatibility testing.
- To assess raw materials of construction prior to use as a control strategy for known leachables in a device.

The following figures also provide a general overview of the overlapping roles toxicologists and analytical laboratory staff may have during the testing process. Figures 4.1 and 4.2 describe the activities toxicologists and analytical staff, respectively, may be responsible for or should inquire about during the initial extractables evaluation.

In the event that additional testing is required, Figs. 4.3 and 4.4 describe the same respective roles for toxicologists and analytical laboratory staff for targeted leachables assessments.

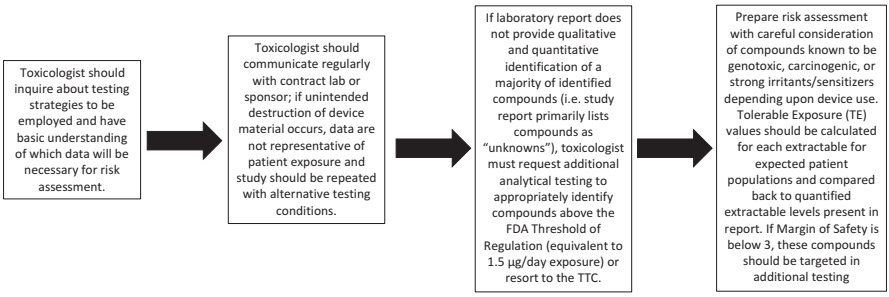


Fig. 4.1 Role of the toxicologist in extractables testing

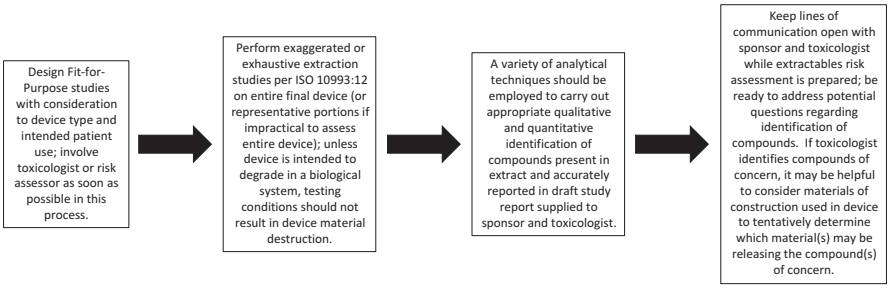


Fig. 4.2 Role of the analytical laboratory in extractables testing

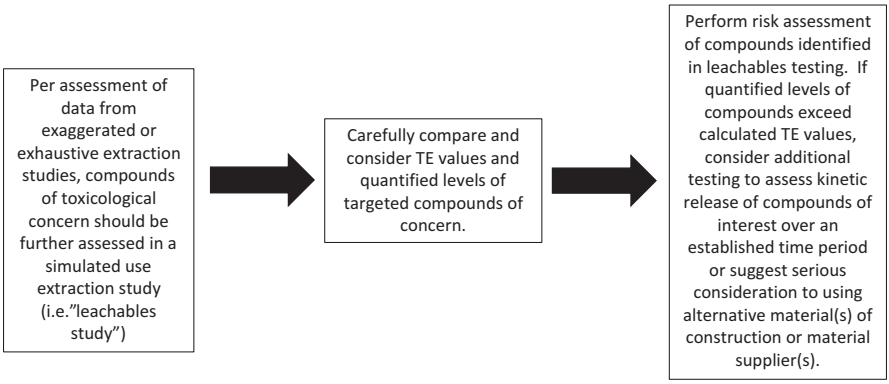


Fig. 4.3 Role of the toxicologist in targeted leachables testing

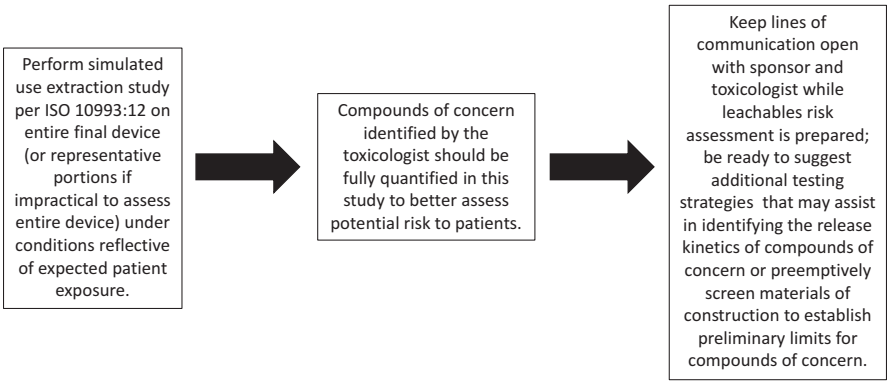


Fig. 4.4 Role of the analytical laboratory in targeted leachables testing

4.2 US Regulatory Basis for E&L Testing in Medical Devices

Erica N. Rogers

4.2.1 Definitions of an Extractable and Leachable

During the process validation period of a medical device, extractables and leachables must be identified and addressed. *Extractables* are compounds that are released from the contact surface of a medical device or material under exaggerated conditions which may include elevated temperature, extended contact time, or aggressive solvent system. Generally, the exaggerated conditions of extraction conditions are conducted relative to those which a material is normally used. Although the extraction procedures are performed at reasonably severe conditions than normally found in a biopharmaceutical process, extractions should not be carried out in such a manner to cause degradation of a medical device to a point of not being mechanically functional. *Leachables* are compounds that are released from a medical device or material under clinically relevant use. Typically, clinical use of the medical device or material should mimic normal conditions which are expected for patient exposure. Overall, while the goal of an extractable study is to identify as many compounds as possible that have the potential to be released as leachables, a leachable study identifies chemicals which may be released during clinical applications. Therefore, the amount of identified leachables should be smaller than the quantity of identified extractables. As such, leachables are a subset of extractables, and not all leachables may be found as extractables as illustrated in Fig. 4.5.

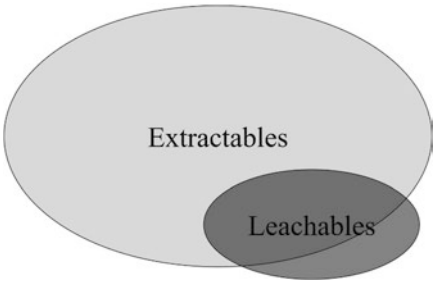


Fig. 4.5 Relationship between extractables and leachables

4.2.2 Current, Common Regulatory Guidance and Standards Exercised in the USA

Presently, several regulatory standard and guidance documents have been established to define the requirements of E&L assessment. For simplicity, this chapter will only focus on four of the most common ISO guidelines that are used quite often in the USA in reference to medical devices. Under the current International Organization for Standardization (ISO) 10,993 and FDA regulatory guidelines and expectations, a critical function of safety assessment for medical devices with internal body ("systemic") contact for potentially at least 29 days or more includes the recognition and measurement of chemical compounds which may be transmitted out from a medical device to an individual as well as performance of a risk analysis and evaluation of the medical device of interest (FDA 2016). The guidelines that will be discussed in this chapter include ISO 10993 parts 1, 12, and 18 and ISO 14971.

4.2.2.1 ISO 10993-1: Biological Evaluation of Medical Devices: Evaluation and Testing in the Risk Management Process

ISO 10993-1 provides guidance on which biological safety tests that should be performed in the ISO 10993 series for a specific application. This guidance requires that the biological safety tests selected and performed on a medical device should be documented, informed decisions that evaluate the advantages and disadvantages as well as relevance (among other things) of current toxicological and biological safety data on product and component materials, breakdown products, and metabolites of medical device materials. Although initial identification of material chemical constituents and consideration of chemical characterization emphasized to occur prior to any biological testing, this objective is not always the case. Overall, a systematic approach to biological evaluation of a medical device as part of the risk management process should begin with collection of material identification information and evaluation of a chemical characterization of material identified in the medical device. In theory, all material information on the construction of the medical device is provided from the manufacture(s). However, suppliers rarely share all material detail on the composition and production of a medical device.

In 2016, FDA issued a finalized, updated document which provided further clarification and updated information on the utilization of ISO 10993-1, "Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process" in support of present applications to the FDA. This finalized guidance replaces the Office of Device Evaluation (ODE) Blue Book Memorandum #G95–1 (1995), entitled "Use of International Standard ISO-10993, 'Biological Evaluation of Medical Devices – Part 1: Evaluation and Testing." In this guidance, the FDA discussed the selection of appropriate endpoints for biological evaluation of a medical device which includes consideration of chemical characteristics of the device materials as well as the nature, degree, frequency, and duration of intended exposure of the device to the body (FDA 2016). Additionally, this guidance document considers new topics including the use of risk-based approaches in the determination of whether biocompatibility testing is necessary, chemical assessment recommendations, and suggestions for biocompatibility preparation for medical device testing with regard to submicron or nanotechnology components and for medical devices made from in situ polymerizing and/or absorbable materials. Some medical devices are composed of materials that have been well characterized chemically and physically in published data and/or possess a lengthy record of safe use in legally US-marketed medical devices. In this case, it may be necessary to perform some, if not all, of the biocompatibility endpoints as discussed in this FDA guidance document. For instance, if a manufacture is able to demonstrate through documentation that the use of a particular material in a legally marketed predicate device or a legally marketed device with comparable tissue exposure as well as provide a convincing explanation why manufacturing is not expected to adversely impact biocompatibility, additional testing may not be needed to address some or all biocompatibility endpoints recommended for consideration.

When extractables and leachables have existing toxicological or biocompatibility data relevant to intended dose, and for which route and frequency of exposure indicate adequate safety margins exists, an acceptable exposure level shall be determined in accordance with *ISO 10993-17 Biological evaluation of medical devices – Establishment of allowable limits for leachable substances*, and the need for further biocompatible testing can be either excluded or reduced. For devices with known leachable chemical mixtures, potential synergies and recommended implementation of multiple ISO 10993 standards shall be considered (ISO 2018a).

4.2.2.2 ISO 10993-12: Biological Evaluation of Medical Devices – Sample Preparation and Reference Materials

ISO 10993-12 assumes that the quantity of extractable during the extraction period which is influenced by time, temperature, surface-area-to-volume ratio regarding extraction solvent, and extraction medium in the final product forms for a test article. At higher temperature or other aggressive extraction conditions, the identification of extractables should be carefully considered. Extraction time and temperatures, as outlined in this guidance, should be high enough to maximize the amount of

extractable possible without resulting in degradation or deformation of the test article. Surface-area-to-volume extraction ratios include the thickness of such materials as film, tubing wall, natural elastomers, and molded items. Additionally, the surface-area-to-volume extraction ratio includes the evaluation of porous surfaces provided that these materials simulate the conditions during clinical use or result in a measurable hazard. In relation to surface-area-to-volume ratio, measurements are dependent on the shape and thickness, among other factors. The standard surface area can be used to determine the volume of extraction medium required. This area may include a combination of areas for all sides of the test article with the exception of irregular-shaped devices which are determined based on mass-to-volume ratio. For most devices, extraction medium should be selected and prepared based on conditions to mimic and exaggerate clinical usage for determination of leachable and extractable substances, respectively. A more detailed discussion of sample preparation with regard to the assessment of extractables and leachables is further discussed in Sect. 4.3 of this chapter (ISO 2011).

4.2.2.3 ISO 10993-18: Biological Evaluation of Medical Devices – Chemical Characterization of Medical Device Materials Within a Risk Management Process

The focus of ISO 10993-18 is to provide an overall outline for the identification of biological hazards as well as estimated quantitation and qualitative regulation of biological risks from chemical constituents. The requirements specified in this guidance document are intended in a stepwise manner to conduct chemical characterization on a medical device through (1) identification of material construction; (2) characterization of material construction through recognition and estimation of chemical constituents; (3) characterization of chemical substances present in the manufacturing of medical device; (4) approximation of potential chemicals during or following the construction of a medical device; and (5) estimation of exposure levels to leachables released in a medical device, that is, chemical constituents released from a device under clinical conditions (ISO 2018b).

4.2.2.4 ISO 14971: Medical Devices – Application of Risk Management to Medical Devices

ISO-14971 provides a fundamental guidance on a product's intended use, determination of potential hazards, and consideration of biological and chemical risks as well as biocompatibility of a medical device. The primary focus of this guidance document is the identification, evaluation, regulation, and management of chemical risks associated with a test article which is defined as a medical device, a component or material, or an extract or portion that undergoes biological or chemical testing or evaluation. In accordance, ISO-14971 provides information on the preparation of a test article for assessment of chemical risks or chemical characterization.

Evaluation of chemical risks should consider the exposure of airway, tissues, and/or environment to foreign materials such as residues, contaminates, additives or processing aids, degradants, and acids. Results obtained from chemical characterization testing shall be evaluated by a toxicologist or trained individual in order to decide the overall biocompatibility of a medical device. Evaluation of chemical characterization testing should be conducted on extractables and/or leachables which may potentially be released from a test article. Overall, ISO 14971 requires that a manufacture should either decide whether a test article is safe for exposure based on criteria defined in the risk management plan or if the test article is considered unsafe the manufacture must gather and review existing published data and literature in determination of the benefits outweighing the overall chemical risk (ISO 2010).

4.3 Fit for Purpose Experimental Design in E&L Testing

Amy L. Mihalchik-Burhans

4.3.1 Identification of Study Needs Based Upon Device Type

Although E&L testing strategies for pharmaceutical products and medical devices share some basic similarities, the intent and progression of testing may vary between these products as generally described below in Table 4.1. Of course, program design may reflect specific drug classes (i.e., orally inhaled and nasal drug products, parenteral and ophthalmic drug products, etc.) and intended duration of patient use, which is not fully considered here. However, these subtle differences in testing strategy and intent are essential for toxicologists and risk assessors to understand when considering and interpreting E&L data.

General study design to identify extractables from finished medical devices or representative portions of devices (which are expected to be of similar surface area and proportion to the finished medical device, undergo similar processing and include any coating materials) is typically device specific, dependent upon expected route and duration of patient contact, as well as physicochemical characteristics of materials of construction used in the device (ISO 2011). Generally, three study condition paradigms are considered for testing including exhaustive, exaggerated, and simulated-use extraction conditions. Exhaustive extraction conditions are recommended for testing of permanent implanted medical devices in order to ascertain the maximum quantity of extractables (and therefore, potential leachables) from device materials throughout the expected period of clinical use (or up to a lifetime) without resulting in chemical changes to device materials or extracted compounds. Exhaustive extraction study design requires sequential extractions of the device in

Table 4.1 Comparison of E&L testing strategies for pharmaceuticals and devices

Product	Intent of study program	Potential origin of E&L compounds	Appropriate guidance	General study design and progression
Pharmaceutical product	Derive qualitative and quantitative data on E&L compounds derived from container closure or packaging system over a specified time period	Primary container closure, secondary packaging, adhesives, and labeling inks	USP <1663>, USP <1664>, USP <661>, FDA *Container Closure Systems for Packaging Human Drugs and Biologics*	Extractables testing followed by leachables correlation study to identify and qualify leachates for safety in a drug product
Medical device	Derive qualitative and quantitative data regarding compounds extracted from finished device or representative portions of device dependent upon expected duration of patient use	Materials of construction with expected direct or indirect surface contact with patient (i.e., implanted devices, externally communicating devices such as IV tubing)	ISO 10993 series (especially Part 12 and 18), ISO 14971, and FDA *Use of International Standard ISO 10993-1, "Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process"*	Dependent upon expected patient use and contact, exaggerated, exhaustive, or simulated-use extraction study which may be followed up with additional targeted leachables testing to assess release characteristics of compounds

appropriate solvents replaced after each extraction period until the level of an extracted substance is less than 10% of the level identified in the initial extraction *or* the total peak area, total organic carbon, or nonvolatile residue level is less than 10% of that observed in the initial extraction. Exaggerated extraction studies may be justified for devices used for <24 h or for prolonged use (>24 h to 30 days) and are designed to result in an extractables profile which generates an increased number or level of extractable compounds than expected during clinical use without resulting in device material degradation. Per ISO 10993:12, the following conditions are suggested for exaggerated extractables testing:

$$(37 \pm 1)^{\circ}\text{C for } (72 \pm 2)\text{ h.}$$
$$(50 \pm 2)^{\circ}\text{C for } (72 \pm 2)\text{ h.}$$
$$(70 \pm 2)^{\circ}\text{C for } (24 \pm 2)\text{ h.}$$
$$(120 \pm 2)^{\circ}\text{C for } (1 \pm 0.1)\text{ h.}$$

However, conditions should be carefully selected based upon the materials of construction used in the device (i.e., an appropriate solvent and temperature should be selected to avoid destruction of device materials which is not expected to reflect

Table 4.2 Testing conditions to meet exaggerated testing conditions

Testing conditions	Temperature exceeds clinical use scenario (commonly referred to as an accelerated extraction)
	Testing duration exceeds clinical use scenario
	Test solvent(s) used which exceed extraction power of solution (i.e., drug, IV fluids, etc.) which makes clinical contact between device and patient
	Using a surface area/solvent volume ratio exceeding clinical exposure
	Use of exhaustive testing protocol for a single-use device

[a]Derived from ISO/DIS 10993–18

standard clinical use). It is also important to consider that release kinetics of all extractables may not be known and that alterations in temperature can significantly impact cross-linking or polymerization of a polymer or result in degradation products that are irrelevant to clinical use (Feinberg et al. 2012; Gad-McDonald et al. 2016; ISO 2018b). Exaggerated extractables testing should meet at least one or more of the following conditions as described in ISO/DIS 10993:18 in order to identify an extractables profile reflective and protective of a worst-case leachables profile from a device (Table 4.2).

As testing conditions (which should be as aggressive as possible without resulting in device degradation) are expected to reflect a worst-case patient exposure scenario to expected E&L compounds, additional E&L testing is not required if toxicological risk assessment of all identified compounds may be qualified as safe at levels identified in testing. If E&L compounds cannot be safely justified, additional E&L testing in the form of simulated-use testing or kinetics testing may be required.

Simulated-use testing may be used for two purposes: (1) identify extractables in devices with <24 h clinical use in some instances and (2) reflect clinical use of a prolonged or permanent use device in a targeted leachables study following identification of an extractables profile using more aggressive testing conditions. The testing output is expected to reflect a realistic leachables profile during clinical use and should provide quantitative data reflective of the daily patient exposure level (i.e., presented in μg extractable/day) of an extractable leaching from a device throughout clinical exposure.

4.3.2 Test Article Considerations and Solvent Selection in E&L Testing

A variety of factors can significantly impact the outcome of E&L testing. Time, temperature, surface- area-to-volume ratio, extraction solvent, phase equilibrium of materials of construction, solubility of extractables in the selected solvent, and rate of transfer of the extractables from the bulk matrix of the material to the surface may all impact testing results (Albert 2012). Clearly, test article preparation may also significantly impact test results. Although test articles assessed in E&L testing should be reflective of either the finished device or as

representative portions of the device per ISO 10993:12, and if at all possible, tested in triplicate, it is not always feasible to do so due to test article size or configuration within the device. If a device must be cut or ground for testing purposes, this should be accounted for in the safety evaluation and in comparing test replicates, especially if mechanical or thermal test article preparation techniques are not uniform among samples. Test conditions (i.e., duration of extraction) may be altered if surface area of the test article is greatly increased by processing. If possible, prepared and unprepared test article samples should be compared in a separate study to identify if additional or new monomers or oligomers are released from the prepared sample, especially for polymer materials of construction (Feinberg et al. 2012; Stults and Creasey 2012). Although other surface-area-to-volume ratios may be justified, the following surface-area-to-solvent or mass-to-solvent ratios are recommended for testing purposes per ISO 10993:12 (Table 4.3):

Alternatively, in order to reach the recommended ratio, multiple devices or test articles may be assessed within the same experiment, and replicate or microlevel extraction techniques may be recommended (Stults and Creasey 2012).

Assuming the composition of the test article is known, solvent selection should be based upon expected or known compatibility; conversely, test article-incompatible solvents should be avoided, as should solvents which result in extreme swelling of the material which could result in artificial inflation of extractables levels observed. Overall, and based upon regulatory recommendations, solvent selection is predicated upon several considerations (ISO 2018b; Stults and Creasey 2012):

1. A polar, nonpolar, and in some cases, semipolar solvent should be used in testing to generate the most reflective extractables profile.
2. Solvent choice may impact which analytical methods may subsequently be used or require additional processing steps which may result in the loss or alternation of E&L compounds.
3. Solvent(s) must effectively extract compounds from device or device material without resulting in degradation or destruction of test article.

Table 4.3 Standard surface areas and extract liquid volumes[a]

Sample thickness (mm)	Extraction ratio (± 10%)	Examples of forms of materials
<0.5	6 cm^2/mL	Film, sheet, tubing wall
0.5–1.0	3 cm^2/mL	Tubing wall, slab, small molded items
>1.0	3 cm^2/mL	Larger molded items
>1.0	1.25 cm^2/mL	Elastomeric closures
Irregularly shaped solid devices	0.2 g/mL	Powder, pellets, foam, nonabsorbent molded items
Irregularly shaped porous devices (low-density materials)	0.1 g/mL	Membranes, textiles

[a]Derived from ISO 10993:12

4. Solvent(s) should also not result in swelling or deformation of the test article, which may artificially increase concentration of extractables, result in identification of compounds that would otherwise not be reflective of potential patient exposure, or damage the test article resulting in particulate matter formation.
5. Solvent(s) selected should be compatible with the study conditions selected (especially temperature and extraction method selected).

Potential solvents suggested for use in ISO 10993:12 and 10,993:18 based upon polarity include:

Polar: water, physiological saline, phosphate-buffered saline, and cell culture media without serum.
Semipolar: dimethyl sulfoxide, acetonitrile, methanol, acetone, ethanol (polarity may vary if in aqueous solution), tetrahydrofuran, *n*-propyl alcohol, *i*-propyl alcohol, and dichloromethane.
Nonpolar: vegetable oil, toluene, cyclohexane, heptane, and *n*-hexane.

Additional solvents (or solvent "variations," e.g., solvents with differing ethanol/water ratios) may be permitted in simulated-use testing, with especial consideration to the intended clinical use of a device. For instance, it may be appropriate to utilize synthetic blood, sweat, sebum, or other contact-specific vehicle as an extraction solvent for testing assuming that the solvent has an extraction power equivalent to the bodily solution which mediates device/patient contact.

Solvent selection for generating extracts for biological assessment may not reflect solvents used in generating an extractables or leachables profile. Solvents intended for the correlation of biological and chemical assay results are discussed more fully in Annex D of ISO/DIS 10993:18 but are not considered here.

4.3.3 Common Extraction Techniques

Dependent upon the device materials and study intent, a variety of extraction techniques may be employed to generate extracts for further analytical testing. "Traditional" extraction methods as described in ISO/DIS 10993:18 include Soxhlet extraction, boiling under reflux, shaking, and sonication, although newer techniques such as microwave-assisted extraction, pressurized fluid extraction, and supercritical fluid extraction may also be used. Selection of an extraction method is predicated upon the practical limitations of each method (i.e., "traditional" methods may take an extensive amount of time, solvent, and resources) balanced with the intended use of the medical device.

4.3.3.1 Soxhlet Extraction

Soxhlet extraction is one of the most commonly utilized techniques in E&L testing. The Soxhlet extractor equipment allows for maintenance of a favorable concentration gradient due to periodic replenishing of the extraction solvent, which is dis-

tilled in a reflux condenser. Extraction is also usually carried out at an intermediate temperature (as opposed to the solvent boiling point).

4.3.3.2 Reflux Extraction

Reflux extraction carried out by placing the test article in a round-bottomed flask in solvent attached to a condenser (may be a Graham, Allihn, or Friedrich condenser) to recycle the solvent vapor formed at the boiling point of the solvent for a specified amount of time. This technique is poorly suitable for thermally labile materials of construction and creates an unfavorable concentration gradient of extractables in the solvent compared to device as time proceeds.

4.3.3.3 Maceration

Maceration is the controlled soaking of a test article in a sealed vessel at a controlled temperature (usually at the boiling point for water or ~10 °C below the boiling point for organic solvents) and humidity in an appropriate solvent. Agitation in the form of stirring may also be included.

4.3.3.4 Shaking Extraction

Shaking extraction is simple in that it only requires the sample to be placed in solvent and for mechanical energy in the form of shaking to be applied. Elevated temperatures may be used with this technique.

4.3.3.5 Sonication

Sonication is one of the most basic techniques available and can be carried out with basic laboratory equipment. A sample is placed in solvent and sonicated in an open or closed container for a specified period of time. However, there is limited control of the temperature, intensity, and repeatability of the specific ultrasonic energy applied to the sample and efficiency of the extraction when comparing sonicators.

4.3.3.6 Sealed Container Extraction

In a sealed container extraction, a sample placed with solvent in a sealed container is heated under reduced pressure usually at a high temperature (i.e., 121 ± 2 °C) for a specified (and usually short) amount of time.

4.3.3.7 Pressurized Solvent Extraction

These approaches, which are termed accelerated solvent extraction, pressurized fluid extraction, and pressurized liquid extraction, may be used as an alternative to Soxhlet or sonication approaches. Pressurized solvent extraction allows for the use of more extreme temperatures and pressures and less solvent while improving matrix penetration and extraction of compounds in less time.

4.3.3.8 Headspace, Thermal Desorption, and Dynamic Headspace

Headspace allows for extraction of a volatile and semi-volatile compounds from a material without the use of solvent. A sample is placed in a sealed vial and heated to form a gaseous headspace which is extracted and processed using gas chromatography. Thermal desorption may be used to analyze trace levels of impurities in a gaseous sample, while dynamic headspace which combines static headspace and thermal desorption techniques concentrates volatile compounds for further assessment.

4.3.3.9 Microwave-Assisted Extraction

Microwave-assisted extraction, which may occur under open or pressurized conditions, utilizes microwave energy to heat the solvent as part of the extraction process. However, this technique requires significant set up and specialized equipment. Due to dissipation of the microwave energy throughout the entire solvent volume resulting in a solvent temperature higher than under atmospheric conditions, material degradation may be more likely.

4.3.3.10 Supercritical Fluid Extraction

Supercritical fluid extraction uses extreme temperature and pressure such that the mobile phase becomes a supercritical fluid which is collected in a small volume of organic solvent. While this technique allows for a high degree of selectivity, optimization of parameters may require extensive time and resources to do so (Feilden 2011; Feinberg et al. 2012).

4.3.4 Common Analytical Methods Used to Assess Extracts

Common analytical methods used to identify and, in most cases, quantify compounds present in device extracts are described in greater detail in Sect. 4.4 of this chapter. Considering that an extract from a device may contain a wide variety of

chemical constituents, it is important for the analytical testing strategy to include. The following table briefly describes each general method and compound classes identified by these methods. Of course, additional and/or specialized analytical methods may also be employed at the discretion of the analytical chemist conducting the assessment (Albert 2012; Gad-McDonald et al. 2016; ISO 2018b) (Table 4.4).

4.4 Common Analytical Methods Used in E&L Testing

Erica N. Rogers

The are several analytical methods that can be used in E&L testing. For simplicity, this book will only discuss five of the most common analytical techniques utilized in E&L testing. Because many of these analytical techniques involve mass spectrometry and/or chromatography, this section will be divided into "analytical techniques that involve mass spectrometry and/or chromatography" and "other analytical methods."

4.4.1 Definition Mass Spectrometry and/or Chromatography

Because many of the different analytical techniques that involve mass spectrometry and/or chromatography essentially involve similar background information, general descriptions of chromatography, mass spectrometry, and ionization techniques are discussed in this subsection.

4.4.1.1 Chromatography

In general, chromatography is an analytical technique used to separate mixtures of chemical substances into individual components, so that the individual components can be thoroughly analyzed. Separation of chemicals occurs when the sample mixture is introduced (injected) into a mobile phase. The mobile phase is different depending on the phase of matter (i.e., liquid or gas form) utilized. In a liquid chromatography (LC) process, the mobile phase is a solvent. During gas chromatography (GC), the mobile phase is an inert or unreactive gas.

The mobile phase transports the sample mixture through the stationary phase and carries the components of the sample mixture with it. The stationary phases are either a solid or a liquid supported on a solid. Usually, the stationary phase uses a chemical that can selectively attract components in a sample mixture. The stationary phase is contained in a tube termed as a column which is typically made up of glass or stainless steel of various dimensions. The relative affinity of the different chemicals in the mixture for the stationary phase controls the separation of substances as they travel the length of the column.

Table 4.4 Summary of common analytical methods used to identify E&L compounds

Analytical method	General purpose of method	General identified compound class
Gas chromatography with headspace sampling (HS-GC)	Most common analytical techniques assist in identification and quantification of volatile organic compounds (VOCs) and semi-volatile organic compounds (SVOCs)	Organic compounds – VOC
Gas chromatography-mass spectroscopy (GC/MS)		Organic compounds – SVOC
Liquid chromatography with mass spectrometry or ultraviolet detectors (LC/MS or LC/UV)	Most common analytical techniques assist in identification and quantification of nonvolatile organic compounds (NVOCs)	Organic compounds – NVOC
Inductively coupled plasma atomic emission spectroscopy (ICP-AES); inductively coupled plasma mass spectroscopy (ICP-MS)	Identification of elemental compounds (although unable to determine specific elemental form)	Elemental compounds
Ion chromatography	Assists in identification and quantification of extracted inorganic anions and low molecular weight organic acids	Anions and cations
Gravimetric and Fourier-transform infrared spectroscopy (FTIR)	Estimate quantity of extractables present in sample and ensure major extractable compound(s) have been identified by other targeted methods without specific compound identification	Nonvolatile residue (NVR), total organic carbon (TOC), ash

The mixture of compounds in the mobile phase interacts with the stationary phase at a different rate. Compounds of the mobile phase that interact with the stationary phase the fastest will elute from the column first; those that interact the slowest exit the column last. Changing properties of the mobile phase and stationary phase directly leads to the separation of different mixtures of chemicals at different retention times. In addition, the separation process can be altered by changing the temperature of the stationary phase or the pressure of the mobile phase.

4.4.1.2 Mass Spectrometry

Mass spectrometry is an analytical, separation technique that measures the mass-to-charge (m/z) ratio of ions. This tool involves that production and subsequent separation and identification of charged species. The basic components of a mass spectrometer involve the ion source, the mass analyzer, the detector, and the data and vacuum systems. The ion source is where the components of the sample are introduced to the mass spectrometer, subsequently ionized, and send the ions to the mass analyzer. The mass analyzer applies an electric and a magnetic field to sort the ions by their masses, while the detector measures and amplifies the ion current to calculate the abundances of each mass-resolved ion.

In a standard mass spectrometry procedure, a sample is ionized which may cause it to either break into charged fragments or become charged without fragmenting. Generation of ions then is separated according to their mass-to-charge ratio, and ions are detected by an electron multiplier. The results following the identification of charged species are typically presented as a mass spectrum which plots the ion signal as a function of the mass-to-charge (*m/z*) ratio.

4.4.1.3 Ionization Techniques

Depending on the information desired from the mass spectrometry analysis, different ionization techniques yield different, desired results. For simplicity, this section will discuss some of the most common techniques during extractable and leachable assessments; these common ionization techniques include hard ionization, soft ionization, and inductively coupled plasma.

Hard ionization, sometimes referred to as high fragmentation ionization, is a technique that uses highly energetic electrons to produce ions. Because of the high quantities of residual energy in the molecule large degrees of fragmentation occur which rupture bonds and subsequently result in the removal of excess energy thereby restoring stability to the resulting ion. The most common example of hard ionization is electron impact ionization (EI) which is typically coupled with gas chromatography (GC). *Soft ionization* is a technique which imparts little residual energy in the subject molecule which does not fragment macromolecules into smaller charged particles but rather alters the macromolecules into ionized small droplets. Common examples of soft ionization techniques include electrospray ionization (ESI) and atmospheric pressure chemical ionization (APCI) which are usually combined with liquid chromatography (LC). *Inductively coupled plasma* (ICP) uses a plasma torch generated by electromagnetic induction to result in the ionization of a sample. Because the effective temperature of the plasma is about 10,000 °C and samples are broken down to ions of their constituent elements, ICP is typically used for analysis of trace elements.

4.4.2 Analytical Techniques that Involve Mass Spectrometry and/or Chromatography

Four of the most common analytical techniques which involve mass spectrometry and/or chromatography are discussed in this subsection. These analytical techniques include liquid chromatography-mass spectrometry (LCMS), liquid chromatography/ultraviolet mass spectrometry (LC/UV MS), gas chromatography-mass spectrometry (GCMS), and inductively coupled plasma mass spectrometry (ICP-MS).

4.4.2.1 Liquid Chromatography-Mass Spectrometry (LCMS)

LCMS is commonly used to identify nonvolatile or non-thermostable polymer additives such as antioxidants, fillers, plasticizers, polymerization or hydrogenation catalysts, anti-slip agents, and other polymer additives during the extractable and leachable assessments. When compared to GC/MS, LC/MS is best suited for the analysis of this extensive range of large, polar, and thermally unstable organic compounds. The method of LCMS utilizes two separation techniques: the LC process, which separates mixtures with multiple components, and the MS process, which separated the compounds according to mass and detects and identifies the structural integrity of the individual compounds found in the liquid mixture. Ultimately, the components separated in the liquid phase of a mixture must be converted to charged ion species via ionization. The overarching theory of the LCMS technique is identification of target compounds is based on both the retention time of the analytes in the chromatogram and the mass spectrum of the eluting compound at this specific time. Prior to LCMS analysis, an aliquot of the extract for extractable and leachable assessment is typically obtained using dichloromethane (DCM), which is denser than water, and concentrated to a level ten times the undiluted solution. Subsequently, internal calibration is performed during the analysis of the extract to determine the concentration of the polymer additive.

Typical LCMS analysis can be divided into four main stages: separation of the liquid mixture, conversion of the separated sample into ions, sorting of the ions according to their mass-to-charge ratio, and detection and quantitation of the ions.

Separation of the Liquid Mixture Separation of the liquid mixture takes place during the LC process. During this first stage of the LCMS method, physical separation of the components in a liquid mixture sample is dispersed between the stationary and mobile phases. Typically, the mobile phase is a mixture of a polar, soluble solvent such as water, methanol, isopropanol, and acetic acid. By contrast, the stationary phase is a nonpolar solid phase that is coated onto a support material and packed into a column. Currently, there are several types of stationary columns used during the LC process. Examples of LC methods include hydrophilic interaction liquid chromatography (HILIC), normal phase chromatography, and ion-exchange chromatography. Examples of commonly used columns for the LC method include silica, amino, and mixed mode (such as alkyl diol, alkyl carboxyl, and aromatic cyano) for HILIC; silica, amino, and diol for normal-phase chromatography; and ionic groups such as sulfonate and tetraalkylammonium for ion-exchange chromatography.

Conversion of the Separated Sample into Ions In order for a MS to detect and identify the individual compounds, the components must be in the gas phase and must be charged. Therefore, the neutral separated compounds during the LC process are converted into ions. Several common ionization methods can achieve conversion of the separated sample into ions including electrospray ionization (ESI), atmospheric pressure chemical ionization (APCI), and atmospheric pressure photo ionization (APPI). Of these common ionization methods, the tech-

nique, ESI, is the most widely used LCMS technique to date; this soft ionization technique applies little residual energy which does not cause extensive fragmentation and preserves the structure integrity of the molecule being ionized. ESI is a popular ionization technique utilized during the ionization process due to its ability to assess a wide range of compounds which include large molecules (such as intact proteins and oligonucleotides) and thermally, liable compounds. The APCI technique is another soft ionization technique which employs a highly charged probe to generate ions from neutral compounds; this method is suited for low, nonpolar compounds such as steroids, aromatic structures, and some lipids. APPI is an ionization technique which generates ions by exposing an aerosol of droplets to photoirradiation. APPI is ideal for highly nonpolar compounds that are not easily ionized by ESI or ACPI; these nonpolar compounds include lipids, steroids, and polycyclic aromatic hydrocarbons (PAHs). Production of ions differed based on the ionization technique used. For the ESI process, ionization occurs in the condensed phase, whereas ionization of molecules for APCI and APPI occurs in the gas phase.

Sorting of the Ions According to Their Mass-to-Charge Ratio A MS is used to sort and isolate the ions created on separated molecules according to the *m/z* ratio. Some of the most popular analyzers include quadrupole and time-of-flight (TOF) instruments. Quadrupole analyzers overall employ electrostatic potentials to the components of the MS which filter out ions according to their *m/z* ratio. Single- and triple-quadrupole instruments are two commonly used quadrupole analyzers. The single-quadrupole instrument uses a quadrupole mass filter to determine the *m/z* ratio of ions produced in a sample and can either scan a range of masses from screening analyses or select specific ions of particular *m/z* ratios for targeted analyses. The triple-quadrupole instrument, sometimes referred to as a tandem mass spectrometer, is the most widely used quadrupole instrument. The triple-quadrupole LCMS instrument uses three mass filtering quadrupoles in tandem to isolate, fragment, and analyze the resulting fragment ions. Overall, the reproducibility and inexpensive cost of quadrupole mass analyzers are appealing to scientists; however, these instruments produce low resolution and may contain a mass bias because the peak height vs. mass response must be frequently tuned. TOF instruments separate ions of different *m/z* by their time of travel between the ion source and detector, through a filed-free region after acceleration by a constant voltage in the source. As a result, ions in the TOF instrument will have different velocities depending on their mass. While advantages of using TOF instruments include high ion transmission, a high mass range in excess of 2×10^4 *m/z*, the sensitivity of this instrument may be affected by the scan speed.

Detection and Quantitation of the Ions Once the ions are generated and separated, they are detected and transformed into a usable, quantifiable signal. MS detection on nonvolatile and non-thermostable compounds is tailored for these target compounds because of its ability to provide additional molecular and structural information of these substances, its highly sensitive ability, and enhanced capability of compound identification. Electron multiplier, dynode, photodiode,

and multichannel plate (MCP) are examples of detectors used to count the ions emerging from the mass analyzer. Overall, the detector is essential in the MS instrument which produces a signal based on the generation of secondary electrons, which are further amplified, or induction of a current from moving charges. Ion detectors are divided into two classes: point detectors and array detectors. While ions in a point detector are detected on a sequential basis in a not spatially resolved manner, array detectors detect ions on a spatial manner, and all ions arrive simultaneously.

4.4.2.2 Gas Chromatography-Mass Spectrometry (GCMS)

The overarching principle for use for GMS techniques is the identification of target compounds based on the retention time of the analytes in the chromatogram and the mass spectrum of the eluting compound at this specific retention time. This technique also has the capability of scanning for wide variety of unknown analytes. As a result, the obtained spectra from unknown analytes can be compared with the electronic mass spectrum libraries. This will aid in the qualitative and semi-qualitative identification of more than 190,0000 nontarget compounds. GCMS is usually applied for the identification of volatile organic components (VOCs) and semi-volatile organic components and (SVOCs).

For VOCs, GCMS has the capability of determining such compounds as monomers, residual solvents, volatile degradants, and other substances present in sample extracts in a precise and rapid manner through a purge-and-trap sampling or headspace techniques. Typically, these components are characterized by boiling points which range between 50–100 and 240–260 °C. The volatility of VOCs when compared to SVOCs is much higher because these compounds have lower boiling points.

SVOCs are thermostable, potential organic migration products that may not be volatile enough to detect using the purge-to-trap or headspace GCMS techniques but volatile enough for detection using other GCMS methods. SVOCs include, but not limited to, hydrocarbons, aldehydes, ethers, esters, phenols, organic acids, ketones, amines, amides, nitroaromatics, polychlorinated biphenyls (PCBs, also known as Aroclors), polycyclic aromatic hydrocarbons (PAHs), phthalate esters, nitrosamines, haloethers, and trihalomethanes. Commonly, SVOCs are found in lubricants, plasticizers, antioxidants, polymer degradants, and solvents with boiling points in the range of 240–260 to 380–400 °C. SVOCs refer to compounds that possess Henry's law constant (H) in the range of 10^{-5} to 3×10^{-7} atm m^3/mol and possess higher boiling greater than water with correspondingly low vapor pressure between 10^{-14} and 10^{-4} atm.

Resulting analytical results for VOCs and SVOCs may include the following:

- Identified compounds (ICs): confirmed, identified compounds based on MS and retention time of a standard compound.

- Most probable compounds (MPCs): identified compounds in which unknown substances are matched at least 80% to known chemicals in the electronic mass spectral libraries. These unknown chemicals are not present in a standard.
- Tentatively identified compounds (TIC): identified compounds in which unknown substances are matched less than 80% to known chemicals in the electronic mass spectral libraries.
- Unknown compounds: compounds that cannot be matched to any degree with known chemicals in the electronic mass spectral libraries. Mass spectrum for these substances is provider at the sponsor's request.

During the GC process, separation of different components of a mixture is carried out between a liquid stationary phase and a gas mobile phase. Helium, hydrogen, and nitrogen are the most common carrier gases used as the mobile phase of GC (Hites 1997; Kupiec 2004). While each carrier gas has its benefits and systems which it is best suited for, helium is the typical gas used with GCMS systems in the USA (Bartram and Froehlich 2010; Kupiec 2004). Helium is commonly used in the USA, because of its inert chemical properties, its ability to provide good separations, and its overall recognition as a safe carrier gas. In other countries, helium is not as readily available and/or is considered an expensive carrier gas. Therefore, hydrogen is used as a substitute carrier gas which is considered to allow shorter run times and increased throughput of the laboratory, longer column lives and less generation of phase breakdown ("column bleeding"), and execution at lower temperatures for the separation process, and this carrier gas is generally recognized as a green, non-environmental gas (Kupiec 2004). Although nitrogen is also identified as an inexpensive and readily available carrier gas to use during the GC process, the run times for the GC process using this gas are considerably longer when compared to hydrogen and helium gas (Hinshaw 2001).

The difference in the chemical properties between different components of the mixture and the relative affinity for the stationary phase of the column will encourage separation of the molecules as they are transport along the length of the column during the GC process. Molecules retained by the column and then eluted at various retention times are introduced into a mass spectrometer downstream via a jet separator or direct connection. The mass spectrometer serves to capture, ionize, accelerate, deflect, identify, and quantitate ionized molecules separately. The hard ionization method, EI, is the most common ionization source combined with GC during mass analysis (Alves et al. 2013; Vékey 2001). EI was used due to its extensive fragmentation, subpicomole to picomole sensitivity, and structural information obtained from fragmentation pattern. While EI contains a number of advantages, this hard ionization technique possesses some limitations which include the fragmentation of organic compounds at or below 600 daltons, the rearrangement process may complicate the spectra, extensive fragmentation limits value of molecular weight determination and recognition, and fragmentation is usually accomplished on nonpolar organic compounds.

4.4.2.3 Inductively Coupled Plasma-Mass Spectrometry (ICP-MS)

The ICP-MS technique is best suitable for the detection and quantitation of many metals and nonmetals in the periodic table. The range of metal content is typically observed in polymer additives including fillers (e.g., talc), pigments, and catalyst residues. Extraction solutions are usually acid digestions in order to reduce the interference by organic matter and to allow conversion of metals associated with particulates to form that are measurable by ICP spectroscopy. Currently, the presence of 22 standard metals can be determined by ICP analyses following acid digestions. The detection limit for each metal is dependent on the amount of material used in the digestion method. Typically, detection limits for metals range between 1 and 10 μg/L.

In brief, the process of ICP-MS can be divided into four stages: sample introduction, ICP torch, interface, and MS.

Sample Introduction There are multiple methods a sample can be introduced during the ICP-MS technique. Sample introduction is dependent on the phase of matter for the sample as illustrated in Fig. 4.6. While there are several ways to introduce a sample during the ICP-MS process, the fundamental objective for each method is to convert the sample into a fine droplet aerosol suitable for ionization in plasma discharge (Thomas 2001). In order to achieve fine aerosol generation, the sample is pumped into a nebulizer which is subsequently converted into a fine aerosol by pneumatic action of gas flow shattering the sample into varied size droplets. Because the plasma discharge cannot effectively dissociate large droplets, the spray chamber removes nebulizer-generated droplets larger than 10 μm in diameter (Thomas 2001, Gaines 2005).

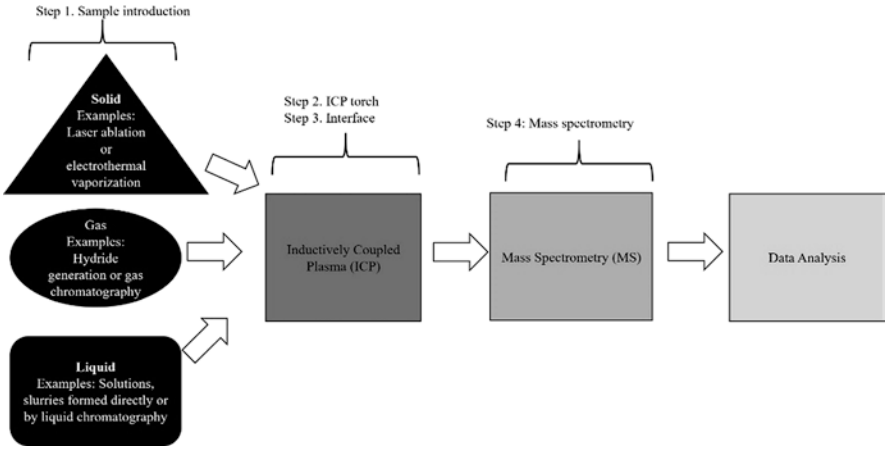


Fig. 4.6 Common steps used during the ICP-MS method and types of examples used during the sample introduction

ICP Torch Once the fine droplets in a gaseous or aerosol form emerge from the spray chamber, they are transferred into the sample injector of the ICP torch. The tiny droplets enter the high- temperature plasma, where atomization occurs, or the initial dried of the gas to a solid and then heated to a gas. The atoms will continue to be transported through the plasma thereby absorbing energy until they release an electron and undergoes ionization. The newly formed ions are then transported out of the torch and enter the interface.

Interface The interface is the point at which the sample from the ICP portion of the instrument comes into contact with the MS portion of the instrument. The interface involves a step pressure reduction process which serves to transmit ions in an effective and consistent manner with electrical integrity from the plasma (at an atmospheric pressure of 760 Torr) to the MS region (at an atmospheric pressure of about 10^{-6} Torr). Typically, the interface is composed of two or three metallic, electrically neutral cones. Cones are often constructed from nickel or platinum. While nickel cones are more cost-effective, platinum cones withstand deterioration from the presence of some acids and are more durable during ICP-MS applications. In order to reduce the effects of the high-temperature plasma, the interface chamber consists of a water-cooled apparatus. After passing through the cones, the generated stream of positively charged ions is directed into the main vacuum chamber by a series of electrostatic lens, known as ion optics. In the ion optical region, the created ion beam is electrostatically focused toward the MS.

MS A number of MS devices are used, but the most common types are quadrupole, magnetic sector, time of flight (TOF), and collision/reaction cell technology. The main aim of the MS procedure in the ICP-MS technique only allows the transmission and analysis of analyte ions of a specific mass-to-charge ratio through the detector.

4.4.3 Other Analytical Methods

There are other types of analytical methods for this analysis of extractables and leachables that do not involve mass spectrometry and/or chromatography. One of the most common analytical method, Fourier Transform Infrared (FTIR) spectroscopy, is discussed in this subsection.

4.4.3.1 Fourier-Transform Infrared (FTIR) Spectroscopy on Nonvolatile Residue (NVR)

FTIR spectroscopy is another analytical technique that can be used in the detection of nonvolatile extractables and leachables including several organic compounds, polymers, adhesives, lubricants, coolants, gases, inorganics, and minerals. FTIR spectroscopy detects the presence of nonvolatile residue in an extraction solution

and provides general information with regard to the release of components using nanograms of material. FTIR spectroscopy uses an infrared spectrum to determine the structural integrity of compounds included in a mixture. FTIR spectra illustrate absorption bands with characteristic frequencies associated with different functional groups within a molecule. Some of the examples of functional groups with different characteristic frequencies include carbonyls (C=O), alkyls (C-H), and multiple bond carbon atoms (C=C and C ≡ C). An infrared spectrometer identified light absorbed between about 400 cm^{-1} and 4000 cm^{-1}. Shifts in the frequency of a molecule in the infrared range can occur based on intramolecular and non-bonded intermolecular interactions. Overall, the use of a FTIR can provide information regarding the chemical composition of nonvolatile residues. In some cases, thin films of residue may also be identified using FTIR spectroscopy.

4.4.3.2 Liquid Chromatography-Ultraviolet (LCUV)

Detectors are useful in increasing the sensitivity of the LCMS instrument. For example, liquid chromatography/ultraviolet mass spectrometry (LC/UV MS) instruments contain a diode-array detector (DAD) which obtain data on selected ultraviolet (UV) and visible wavelengths and spectra. Normally, the use of a DAD in a LC/UV instrument aids in the screening process to identify unknown peaks and/or determination of peak purity. This screening method is best utilized with the LC technique, reversed-phase chromatography, with an assortment (gradient) range of solvent strength as detailed in the Product Quality Research Institute (PQRI) document, *Safety Thresholds and Best Practices for Extractables and Leachables in Orally Inhaled and Nasal Drug Products*. Initially, individual extractables will be identified based on retention time and characteristic ions in the mass spectra. Subsequently, the LC chromatogram obtained using a UV detector will allow determination of nontarget peaks.

References

Albert, D. E. (2012). Material and chemical characterization for the biological evaluation of medical device biocompatibility. In J.-P. Boutrand (Ed.), *Biocompatibility and performance of medical devices* (pp. 65–94). Elselvier.

Alves, S., Rathahao-Paris, E., & Tabet, J. C. (2013). Chapter five – potential of Fourier transform mass spectrometry for high-throughput metabolomics analysis. In D. Rolin (Ed.), *Advances in botanical research* (Vol. 67, pp. 219–302). Academic Press.

Bartram, R. J., & Froehlich, P. (2010). Considerations on switching from helium to hydrogen. *LCGS North America, 28*(10), 890–900. Retrieved from http://www.chromatographyonline.com/considerations-switching-helium-hydrogen.

FDA, U. S. (2016). *Use of International Standard ISO 10993-1: "Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process"*. U.S. FDA Retrieved from https://www.fda.gov/media/85865/download

Feilden, A. (2011). *Update on undertaking extractable and leachable testing.* iSmithers.

Feinberg, T. N., Norwood, D. L., Granger, A. T., & Jenke, D. (2012). Extractables: The controlled extraction study. In D. J. Ball, D. L. Norwood, C. L. M. Stults, & L. M. Nagao (Eds.), *Leachables and extractables handbook* (pp. 289–330). Hoboken, NJ: Wiley.

Gad-McDonald, S., Albert, D. E., & Gad, S. C. (2016). Leachables and extractables from medical devices. In S. C. Gad & S. Gad-McDonald (Eds.), *Biomaterials, medical devices, and combination products: Biocompatibility testing and safety assessment* (pp. 419–468). Boca Raton, FL: CRC Press.

Hinshaw, J. V. (2001). Selecting carrier gases and conditions. *LC-GC North America, 19*(10), 1056.

Hites, R. A. (1997). Gas chromatography mass spectrometry. In *Handbook of instrumental techniques for analytical chemistry* (pp. 609–626). American Chemical Society.

ISO. (2010). *Medical devices – Application of risk management to medical devices*. American National Standard.

ISO. (2011). *Biological Evaluation of medical devices – Part 12: Sample preparation and reference materials*. American National Standard.

ISO. (2018a). *Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process*. International Organization for Standardization.

ISO. (2018b). *Draft International Standard: Biological evaluation of medical devices – Part 18: Chemical characterization of medical device materials within a risk management process.* International Organization for Standardization.

ISO. (2019). *Biological evaluation of medical devices – Application of the threshold of toxicological concern (TTC) for assessing biocompatibility of medical device constituents*. International Organization for Standardization.

Jenke, D. (2012). A general overview of the suitability for intended use requirements for materials used in pharmaceutical systems. In D. L. N. D. J. Ball, C. L. M. Stults, & L. M. Nagao (Eds.), *Leachables and extractables handbook* (pp. 21–36). Hoboken, NJ: Wiley.

Kupiec, T. (2004). Quality-control analytical methods: Gas chromatography. *International Journal of Pharmaceutical Compounding, 8*(4), 305–309.

Stults, C. L. M., & Creasey, J. M. (2012). Development, optimization, and validation of methods for routine testing. In D. J. Ball, D. L. Norwood, C. L. M. Stults, & L. M. Nagao (Eds.), *Leachables and extractables handbook* (pp. 449–506). Hoboken, NJ: Wiley.

Thomas, R. (2001). A beginner's guide to ICP-MS. *Spectroscopy, 16*(5), 56.

Vékey, K. (2001). Mass spectrometry and mass-selective detection in chromatography. *Journal of Chromatography A, 921*(2), 227–236. https://doi.org/10.1016/S0021-9673(01)00825-1.

Chapter 5
Where the Data Is and What Is It?

Sherrie Smith

Abstract Once testing has been done on a device, the next step is assessment of the identified compounds. This chapter discusses strategies and resources to consult in order to find the data necessary for the said assessment. This chapter will start with the online sources used to help identify compounds without proper names, then sources to find pertinent data. Sources include online and print, both free and paid.

Keywords Extractables · Leachables · CAS registry number · ChemSpider · PubChem · Permissible daily exposure (PDE) · Minimal risk level (MRL) · Hazardous substances data bank (HSDB) · TOXNET · Registry of toxic effects of chemical substances (RTECS) · Leadscope toxicity database · Patty's toxicology · Encyclopedia of toxicology

Once a device has been tested and extractables (or potentially leachables) of component material have been identified and quantified, the next step is to determine the potential risk of each such chemical entity. In order to accomplish this task, either testing or a literature search should be conducted. The purpose of this chapter is to discuss strategies and sources for the search. Before starting any search, it is important to identify the correct name, any synonyms, and Chemical Abstracts Service (CAS) Registry Number (if available) for each chemical. These pieces of information will be the key to locating all the information needed. Some categories that are important in any chemical risk assessment are lethality, mutagenicity, carcinogenicity, and reproductive toxicity. Other categories such as local tissue tolerance and potential skin sensitization may be important depending on the patient contact and duration of such contact with the medical device to patients.

In the case that the only identification provided for a chemical is the structure, there are a few sites available to help identify the name and CAS Number, if available. ChemSpider is a chemical structure database providing access to over 67 million structures. This site allows for structure searches to be conducted by drawing

S. Smith, MS (✉)
Gad Consulting Services, Raleigh, NC, USA
e-mail: sherrie.smith@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_5

structures on the web site, uploading a saved structure file (.mol, .sdf, .cdx), or uploading an image file (.png, .jpg, .gif). A structure search will provide key information such as the chemical name, synonyms, and CAS Number(s) (ChemSpider 2015). A second site that is helpful for this purpose is PubChem. PubChem is an open chemistry database provided by the US National Institutes of Health (NIH). This site also allows for structure searches to be conducted by drawing structures on the web site or uploading a saved structure file (PubChem 2019). In addition to identification information, PubChem can also provide information useful to the chemical assessment. This information can include physical and chemical properties, occupational and environmental health data, safety data, and toxicity data (PubChem 2019).

Once the correct name(s) and CAS Number(s) are identified (or lack of CAS Number confirmed), there are numerous online and print sources available when assessing chemicals. The first place that should be consulted are the ICH guidelines, particularly M7, Q3C, and Q3D. ICH M7 [Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk] lists some particularly worrying compounds and the actions that should be taken if these compounds are present (ICH 2017). ICH Q3C (Impurities: Guideline for Residual Solvents) has some common residual solvents seen in the pharmaceutical industry and calculates permissible daily exposures (PDEs) for them (ICH 2018). ICH Q3D (Guideline for Elemental Impurities) is similar to Q3C but for elemental impurities and includes PDEs for oral, inhalation, and parenteral exposure (ICH 2019). Any compound appearing in any of these guidelines should be assessed in a manner in line with the guidelines.

The next source that should be consulted are any assessments or reviews by regulatory agencies. The Agency for Toxic Substances and Disease Registry (ATSDR) is a federal public health agency working as part of the US Department of Health and Human Services. ATSDR provides toxicological profiles and assessments for environmental health threats. The assessments provide minimal risk levels (MRLs) for acute, intermediate, and chronic duration for exposures via the oral and inhalation routes. The MRL provides an estimate of the daily exposure to a substance that is unlikely to cause a significant risk of adverse noncancer health effects (ATSDR 2019). The EPA provides toxicological reviews and summaries for chemicals or chemical groups that may present a health hazard. These assessments may provide an oral reference dose (RfD) and/or inhalation reference concentration (RfC) for these hazardous chemicals. The RfD and RfC are estimates of a daily exposure that is unlikely to cause a significant risk of adverse noncancer effects during a lifetime (EPA 2017). The European Union (EU) has agencies that may also be of assistance when assessing chemicals, such as the European Food Safety Authority (EFSA) and the European Chemicals Agency (ECHA). EFSA provides guidance and opinions on chemicals that may pose a risk to any part of the food chain. The guidance may consist of acceptable daily intake (ADI) levels or tolerable daily intake (TDI) levels. An ADI is an estimate of the amount of a chemical to food or drinking water that is unlikely to cause a significant risk of adverse effects over a lifetime. This value applies to chemicals intentionally used at some point in the food chain (such as a food additive or veterinary medicine). TDI has the same definition but applies to chemicals that are not intentionally added (such as contaminants) (EFSA 2019).

ECHA regulates chemicals on the EU market and aims to make sure chemicals are being used safely. They provide opinions and guidance on chemicals that may need further assessment. ECHA also provides a vast database of chemicals that have been registered under REACH (Registration, Evaluation, Authorisation and Restriction of Chemicals). The database provides a great array of toxicological information, including toxicokinetic data, genotoxicity, irritation, and sensitization data. Data on carcinogenicity, acute, repeated dose, and reproductive toxicity are also provided for multiple routes when available. In addition, the database provides physical and chemical properties and ecotoxicological information (ECHA 2019).

There are also several intergovernmental agencies that provide assessments. The Organisation for Economic Co-operation and Development (OECD) provides hazard and exposure assessments for industrial chemicals. OECD also provides the eChemPortal, which is a database providing physical and chemical properties, environmental fate and behavior, ecotoxicity, and toxicity data for chemicals. The World Health Organization (WHO) provides several different types of reports that may be of assistance when assessing chemicals. These assessments can provide physical and chemical properties, toxicological data, evaluations from other international bodies, and conclusions on how to protect human health. The WHO also partly administers the Joint FAO/WHO Expert Committee on Food Additives (JECFA), which provides risk assessments and evaluations on food additives, processing aids, flavoring ingredients, contaminants, natural toxins, and veterinary drug residues that may be present in animal products. JECFA has evaluated over 2500 chemicals (JECFA 2018).

If there are no regulatory or interagency reports available, there are several other free sources of toxicological data available online. A great place to start is the US National Library of Medicine (NLM) TOXNET (toxicology data network) database. TOXNET is a collection of databases that contain information on hazardous substances, toxicology data, and environmental health information (TOXNET 2019). These databases can be searched by subject or CAS Registry Number. The most commonly used databases are:

- The Hazardous Substances Data Bank (HSDB) a toxicology database focusing on providing human exposure, industrial hygiene, safe handling procedures, and environmental fate for potentially hazardous chemicals. Data are derived from a variety of sources including government documents and special reports.
- TOXLINE a database providing citations for articles discussing the biochemical, pharmacological, physiological, and toxicological effects of drugs and other chemicals. Citations date from the 1840s up to the present, with new citations added weekly.
- ChemIDplus a dictionary of over 400,000 chemicals that provides links to other resources including other NLM databases and federal and international web sites.
- Although archived, the Chemical Carcinogenesis Research Information System (CCRIS) can also be helpful. It was developed by the National Cancer Institute and provides carcinogenicity, mutagenicity, tumor promotion, and tumor inhibition test results for over 8000 chemicals. Data were obtained from the years 1985–2011.

- Other databases include the Drugs and Lactation Database (LactMed), Developmental and Reproductive Toxicology Database (DART), TOXMAP, Toxics Release Inventory (TRI), Comparative Toxicogenomics Database (CTD), Household Products Database, Haz-Map, Integrated Risk Information System (IRIS), International Toxicity Estimates for Risk (ITER), ALTBIB, the archived Carcinogenic Potency Database (CPDB), and Genetic Toxicology Data Bank (GENE-TOX).

Outside of the TOXNET databases, the US NLM also maintains MEDLINE and PubMed. MEDLINE is a bibliographic database containing over 25 million references to life science journal articles concentrating on biomedicine. PubMed is a database that holds over 29 million citations from MEDLINE, life science journals, and online books. Fields covered include the life sciences, behavioral sciences, chemical sciences, and bioengineering. Links to full-text articles that are located in PubMed Central or on publisher web sites are also provided for many citations. PubMed also provides links to related articles and additional relevant web sites and resources. PubMed Central is a free archive of full-text journal articles provided by participating publishers as well as author manuscripts submitted in compliance with the NIH Public Access Policy. These resources can be searched by subject, CAS Registry Number, author, journal name, date, etc (PubMed 2019; Medline 2019).

Another free source is the COSMOS database, developed by the European Union from information obtained during the COSMOS project. The COSMOS project ran from 2011 to 2015 and was an effort to determine the human safety of cosmetic ingredients without using animal testing. During the project, several free tools and workflow were developed to achieve the goal, one of which was the COSMOS database. The database contains more than 40,000 unique structures searchable by name, CAS Registry Number, or structure search. The database contains toxicological data, such as repeat dose and genotoxicity for more than 1600 compounds (Cosmos 2019).

Although it is not free, the Registry of Toxic Effects of Chemical Substances (RTECS) database has a vast amount of toxicological data available if one has the proper identification of the chemical entity. The RTECS database was originally built and maintained by the National Institute of Occupational Safety and Health (NIOSH) from 1971 through 2001. It is now maintained and updated quarterly through BIOVIA, using the data selection criteria and rules established by NIOSH. RTECS is a toxicological database containing information for chemical entities in commerce or commercial use in the USA including prescription and non-prescription drugs, food additives, solvents, chemical wastes, pesticides, and diluents. The database lists available data for acute toxicity, tumorigenicity, skin and eye irritation, mutagenicity, reproductive effects, and repeat-dose effects. As of 2011, aquatic and in vitro toxicology data are also available (RTECS 2019). The database is available through purchase of a subscription through BIOVIA or other authorized resellers. There are also options available for single searches for a fee.

The Cosmetic Ingredient Review (CIR) was established by the Personal Care Products Council and is supported by the US FDA. The CIR reviews and assesses the safety of ingredients commonly used in cosmetic products. These reviews are then published in peer-reviewed science journals (Find Ingredient Reviews and Documents 2016).

There are several well-known scientific journals that publish in the field of toxicology and may contain pertinent information. *Annual Review of Pharmacology and Toxicology* has been in publication since 1961 and publishes articles dealing with significant developments on a wide variety of related topics. Topics include chemical agents, drug development science, and body systems such as the gastrointestinal tract, cardiovascular system, and immune system. *Critical Reviews in Toxicology* has been in publication since 1971 and publishes comprehensive reviews of research findings in toxicology related to the mechanism of action, responses, and assessment of health risks due to exposure to various substances. Reviews are provided on many types of substances and may include pharmaceutical agents, pesticides, or ingredients found in consumer products. *Toxicological Sciences* has been in publication since 1981 and is the official journal of the Society of Toxicology. This journal aims to publish significant contributions to toxicology research and provide expert insight in the form of contemporary reviews and editorials relating to important toxicological topics. *Toxicology* has been in publication since 1973 and is associated with the German Toxicology Society. Articles focus on understanding the mechanisms of toxicity associated with exposure to toxicants as it relates to human health. *Food and Chemical Toxicology* has been in publication since 1982 and publishes articles and reviews focusing on toxic effects of natural or synthetic food, drugs, or chemicals to animals and humans. *Toxicology Letters* has been in publication since 1977 and publishes novel articles related to determining mechanisms of toxicity and reviews on various areas of toxicology. *The International Journal of Toxicology* has been in publication since 1982 and is associated with the American College of Toxicology. Articles cover a range of topics important to toxicologists including contemporary issues, safety assessments, risk assessment, new approaches to testing, mechanisms of toxicity, and biomarkers. *Archives of Toxicology* has been in publication since 1930 and publishes articles on the latest advances in toxicology, focusing on research related to mechanisms of toxicity in humans and animals.

Additionally, a great resource that compiles data from several of the above listed sources is the Leadscope Toxicity Database. This database contains over 180,000 chemical structures and can be searched by several metrics including CAS Registry Number, structure, or name. When available, the database provides toxicological data including acute, repeat dose, carcinogenicity, genotoxicity, reproductive/developmental, and irritation studies. There are over 400,000 study results in the database, and searches can be conducted by the type of toxicological study, duration of study, or route of exposure. Data comes from the National Toxicology Program chronic database, Carcinogenic Potency Database, FDA PAFA database, and published journal articles (Leadscope 2019). In addition, Leadscope is a value-added reseller of the RTECS database. The RTECS database is enhanced with normalization of dose regimen, additional end point calculations, and mapping to allow for easier searching. In order to use the Leadscope Toxicity Database, you would need to purchase one of the Leadscope applications. The applications come with a host of other features that are useful, including the ability to add your own structures and data to strengthen your personal database.

Finally, there is a great deal of information present in books that never makes it to electronic sources. These include reference volumes such as *Patty's Toxicology*,

Encyclopedia of Toxicology, *Sittig's Handbook of Toxic and Hazardous Chemicals and Carcinogens*, and *Sax's Dangerous Properties of Industrial Materials* (Pohanish 2011; Bingham and Cohrssen 2012; Wexler 2014). If the potential leachable (or extractable) happens to be a commonly used ingredient in pharmaceuticals or food, sources such as the *Handbook of Pharmaceutical Additives*, *Handbook of Pharmaceutical Excipients*, *and Handbook of Food Additives* may have relevant data (Sheskey 2017; Ash 2007; Ash 2008).

References

Ash, M., & Ash, I. (Eds.). (2007). *Handbook of pharmaceutical additives* (3rd ed.). Synapse Information Resources, Inc, New York.

Ash, M., & Ash, I. (Eds.). (2008). *Handbook of food additives* (3rd ed.). New York: Synapse Information Resources, Inc.

ATSDR Toxicological Profiles. (2019). ATSDR. https://www.atsdr.cdc.gov/toxprofiledocs/index.html. Accessed 20 June 2019.

Bingham, E., & Cohrssen, B. (Eds.). (2012). *Patty's toxicology* (6th ed.). New York: John Wiley & Sons.

ChemSpider. (2015). *Royal society of chemistry*. Cambridge, UK. http://www.chemspider.com/. Accessed 20 June 2019.

Cosmos, D. B. (2019). *Cosmos*. https://cosmosdb.eu/cosmosdb.v2/dbsearch/full. Accessed 20 June 2019.

ECHA Search for Chemicals. (2019). *ECHA*. https://echa.europa.eu/. Accessed 20 June 2019.

eChemPortal: The Global Portal to Information on Chemical Substances. (2019). *OECD*. https://www.echemportal.org/echemportal/index.action. Accessed 20 June 2019.

EFSA Scientific Outputs at a Glance. (2019). *EFSA*. http://www.efsa.europa.eu/en/publications. Accessed 20 June 2019.

EPA Risk Assessment for Other Effects. (2017). *EPA*. https://www.epa.gov/fera/risk-assessment-other-effects. Accessed 20 June 2019.

Evaluations of the Joint FAO/WHO Expert Committee on Food Additives. (2018). *WHO*. http://apps.who.int/food-additives-contaminants-jecfa-database/search.aspx Accessed 20 June 2019.

Find Ingredient Reviews and Documents. (2016). *Cosmetic ingredient review*. Washington, DC. https://www.cir-safety.org/ingredients. Accessed 20 June 2019.

ICH. (2018). *Q3C(R7) impurities: Guideline for residual solvents*. Geneva, Switzerland. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q3C/Q3C-R7_Document_Guideline_2018_1015.pdf. Accessed 20 June 2019.

ICH. (2019). *Q3D(R1) guideline for elemental impurities*. Geneva, Switzerland. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q3D/Q3D-R1EWG_Document_Step4_Guideline_2019_0322.pdf. Accessed 20 June 2019.

ICH. (2017). *M7(R1) assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk*. Geneva, Switzerland. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_R1_Addendum_Step_4_2017_0331.pdf. Accessed 20 June 2019.

Leadscope Toxicity Database. (2019). Leadscope, Inc, Columbus, OH. http://www.leadscope.com/product_info.php?products_id=78. Accessed 20 June 2019.

Medline: Description of the Database. (2019). *US NLM*. Rockville, MD. https://www.nlm.nih.gov/bsd/medline.html Accessed 20 June 2019.

Pohanish, R. P. (Ed.). (2011). *Sittig's handbook of toxic and hazardous chemicals and carcinogens* (6th ed.). Oxford, UK: Elsevier.

PubChem. (2019). *US NLM NCBI*. Rockville, MD. https://pubchem.ncbi.nlm.nih.gov/. Accessed 20 June 2019.
PubMed. (2019). *US NLM NCBI*. Rockville, MD. https://www.ncbi.nlm.nih.gov/pubmed/. Accessed 20 June 2019.
RTECS. (2019). *BIOVIA*. San Diego, CA. https://www.3dsbiovia.com/products/collaborative-science/databases/bioactivity-databases/rtecs.html Accessed 20 June 2019.
Sheskey, P. J., Cook, W. G., & Cable, C. G. (Eds.). (2017). *Handbook of pharmaceutical excipients* (8th ed.). Philadelphia: Pharmaceutical Press.
TOXNET. (2019). *US NLM*. Rockville, MD. https://toxnet.nlm.nih.gov/. Accessed 20 June 2019.
Wexler, P. (Ed.). (2014). *Encyclopedia of toxicology* (3rd ed.). San Diego, CA: Academic Press.

Chapter 6
Bridging Issues of Route

Amy L. Mihalchik-Burhans and Dexter W. Sullivan Jr.

Abstract Current regulatory guidelines call for the identification and quantification of chemical compounds which may migrate out of a medical device, combination product, or a drug container closure system, and into an individual. Compounds that migrate from these devices often have little available toxicity data resulting in a major challenge in performing an adequate assessment of biological or toxicological risks and thus potentially compromising patient safety. These limited toxicity data may exist by alternative routes. Therefore, route-to-route extrapolation of the available data may be considered for use in these instances. Here we discuss the factors to consider when performing route-to-route extrapolations for L&E compounds beginning with historic and regulatory approaches. This will be followed by a review of pharmacokinetic and pharmacodynamic considerations regarding the applicability of route-to-route extrapolation. Finally, we will discuss route-specific considerations and the derivation of safe exposure limits.

Keywords Route-to-route · Extrapolation · Leachables · Extractables · L&E · Risk assessment · TTC

6.1 Route-to-Route Extrapolation and Relevance to Assessment of Medical Devices

Current ISO 10993 and FDA regulatory guidelines call for the identification and quantification of chemical compounds which may migrate out of a medical device, combination product, or a drug container closure system, and into an individual (FDA 2016; ISO 2018). Often times, compounds that migrate from these devices have little available toxicity data resulting in a major challenge in performing an adequate assessment of biological or toxicological risks from leachables &

A. L. Mihalchik-Burhans, PhD (✉) · D. W. Sullivan Jr., MS, DABT
Gad Consulting Services, Raleigh, NC, USA
e-mail: amy.mihalchik@gadconsulting.com; dexter.sullivan@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_6

extractable (L&E) compounds that can potentially compromise patient safety. These limited toxicity data may exist by alternative routes. Therefore, route-to-route extrapolation of the available data may be considered for use in these instances where adequate route-specific data for a compound are not available.

Here we discuss the factors to consider when performing route-to-route extrapolations for L&E compounds beginning with historic and regulatory approaches. This includes discussions on EPA's use of safety factors as well as approaches considered by ECHA, occupational exposure groups, cosmetic ingredient review panels, and ISO and ICH guidelines consideration of route-to-route extrapolation. This will be followed by a review of pharmacokinetic and pharmacodynamic considerations regarding the applicability of route-to-route extrapolation. Finally, we will discuss route-specific considerations and the derivation of safe exposure limits.

6.2 Considerations Regarding Applicability of Route-to-Route Extrapolation

Although there is a significant need to develop methods of route-to-route extrapolation that achieve regulatory acceptance, standardized approaches have yet to be created especially in the field of pharmaceutical and medical device risk assessment. The majority of proposed approaches have primarily been developed in the European Union (EU), in the US Environmental Protection Agency (EPA), and by independent academic or industry research groups for the purposes of assessing risk associated with environmental, occupational, and consumer exposures to chemicals via dermal and inhalation routes extrapolated from oral data in various test species. Few groups have developed methods to carry out route-to-route extrapolation from oral (or dermal or inhalation) exposure data to parenteral exposures, which compose a significant portion of medical device use (e.g., externally communicating medical devices and implanted medical devices), an area for much-needed research (Broschard et al. 2016; ICH 2014; ISO 2018). Regardless, the basic tenants historically set forth for route-to-route extrapolation (typically with oral data as the point of departure) should apply to medical devices and provide the basis for further exploration of this topic. Most route-to-route extrapolation methods discussed in the literature are predicated upon the use of readily available oral data as the point of departure.

Below, key factors required to consider regarding route-to-route extrapolation are described in greater detail.

6.2.1 Critical Target Tissue of Toxicity Following Exposure

When possible, the critical toxic effect of interest should be independent of route. Route-to-route extrapolation is generally only performed on compounds with known systemic toxicity (toxicities noted at distant site from exposure route) and should not be considered if local or portal of entry effects, such as corrosion, moderate to severe irritation, or sensitization, is expected. Regarding local effects,

toxicity is often associated with the concentration of a substance present at the site of administration as opposed to the total systemic dose received (ECETOC 2003; ECHA 2012; Geraets et al. 2014; IGHRC 2006; Pepelko and Withey 1985; Schroder et al. 2016; Sharratt 1988).

6.2.2 *Metabolism of the Compound*

When considering route-to-route extrapolation, tissue-specific metabolism should be carefully considered, especially when considering non-oral routes of exposure. While hepatic and extrahepatic biotransformation enzymes often share similarities, specific enzyme distribution, induction, and forms may vary among tissues. Upon oral administration, compounds may be metabolized by stomach acid, mucosal cells of the gastrointestinal tract, or intestinal microorganisms prior to reaching the liver. Dermal metabolism, which is fairly limited for a majority of chemicals (aside from esters), is typically quite low compared with the liver, suggesting that significant biotransformation is typically unlikely. Similarly, the lungs are expected to also have limited metabolic function compared to the liver; additionally, considerations such as particle size, deposition, and compound solubility may significantly impact compound metabolism as related to distribution of metabolic enzymes throughout the respiratory tract. Furthermore, species-specific effects should be considered whenever possible, considering that the ratio of activating/deactivating metabolic enzymes within pulmonary tissue may exhibit a 100-fold difference between rats and humans (Bond 1993; IGHRC 2006). In cases where a compound is expected to undergo significant metabolism by one route, route-to-route extrapolation may significantly underestimate potential toxicity associated with exposure via the route of interest and should be approached with caution (IGHRC 2006).

6.2.3 *Toxicokinetic Characteristics*

When performing route-to-route extrapolation, it is suggested that compound absorption be independent of potential local effects and that absorption efficiency data of the starting route and intended route of exposure be available to calculate a modifying factor (Geraets et al. 2014). As it is usually challenging to accurately quantify differences in other TK characteristics including distribution, metabolism, and excretion between routes and test species, absorption is the key characteristic to examine extrapolation is also often more reliable for compounds with relatively long half-life values as toxicity may be more closely related to stable systemic exposure as opposed to dose-related exposure factors (e.g., a bolus dose may result in high exposure concentration which may not be reflected by other routes of exposure) (IGHRC 2006; Sharratt 1988). In the absence of experimental data, the use of physiologically based pharmacokinetic models to more accurately predict target tissue dose exposure concentrations at sites of interest may also be considered (ECHA 2012; IPCS 2010).

6.2.4 First-Pass Effects

Following oral administration, a compound may undergo extensive first-pass metabolism in the liver which may significantly alter its bioavailability and toxicity. For instance, a compound rapidly metabolized in the liver following oral exposure may result in limited systemic exposure, suggesting that exposure by an alternative route (such as parenteral) may elicit increased toxicity. Toxicity may also vary by exposure route for compounds requiring metabolic activation, as tissue-specific metabolism of a compound may not result in the production of the same metabolites at the same rate (Geraets et al. 2014; IGHRC 2006; Sharratt 1988).

6.2.5 Physicochemical Properties of the Compound in Regard to Toxicokinetics

Physicochemical properties may impact the absorption, distribution, metabolism, and excretion of a compound in the body, especially by varying routes of exposure (Schroder et al. 2016). Characteristics that may impact absorption and systemic exposure dose include dissociation size, molecular size, molecular weight, partition coefficient, pKa, reactivity, solubility, and volatility (Gerrity et al. 1990).

6.2.6 Intended Route/Duration of Patient Exposure

When considering route-to-route extrapolation as part of a risk assessment for a compound of interest present in or leached from a medical device, it is essential to consider all available data as related to the intended patient exposure route and duration when possible.

Per ISO 10993-17, medical devices may be broadly categorized by two factors: expected body contact and duration of contact. Expected patient use of a device may impact potential routes and durations of exposure to L&E compounds, which may migrate from the device or device packaging. Medical device categories based upon body contact and exposure routes of interest are summarized in Table 6.1. Routes of interest for the purposes of risk assessment per expected patient contact are also included; if available, data by these routes should be used in derivation of a safe exposure level or extrapolation to the most appropriate exposure route should be considered (ANSI 2008).

Three primary categories of exposure duration are described in 10993-17 as shown in Table 6.2. When considering available animal or human data for risk

Table 6.1 Medical device categorization and exposure routes of interest[a]

Medical device category	Application site(s)	Exposure route(s) of interest	Data of interest for risk assessment purposes if available	Examples
Surface contacting	Skin, intact mucosal membranes, or breached body surfaces	Dermal, ocular, vaginal, rectal	Dermal, subcutaneous, ocular, buccal, vaginal/rectal	Electrodes, bandages, wound dressings, contact lenses, urinary catheters, colonoscopes, etc.
Externally communicating (indirect	Indirect contact with blood; indirect contact with tissue, bone, or dentin (teeth); contact with circulating blood	Parenteral	Intravenous, intraperitoneal, intramuscular	Blood administration sets, laparoscopes, arthroscopes, intravascular catheters, dialysis tubing, and accessories
Implant	Direct contact with the bone or tissues, direct contact with circulating blood and cardiovascular system	Parenteral	Intravenous, intraperitoneal, intramuscular, implant	Orthopedic implants, bone cements, pacemakers, breast implants, artificial heart valves

[a]Table contents adapted from 10993-17

Table 6.2 Contact duration of medical devices[a]

Contact duration category	Definition
Limited exposure	Cumulative sum of single, multiple, or repeated exposures up to 24 h
Prolonged exposure	Cumulative sum of single, multiple, or repeated exposures over 24 h but not in excess of 30 days
Long-term exposure	Cumulative sum of single, multiple, or repeated exposures exceeds 30 days

[a]Table contents adapted from 10993-17

assessment purposes, the experimental exposure duration should be considered in regard to expected patient exposure. The use of chronic exposure data (generally from a 2-year study in rodents or duration of half a lifetime in other test species) when calculating a safe exposure limit is the most appropriate and conservative approach when considering devices with long-term patient use (ECHA 2012). If only subacute or subchronic data are available when assessing a device with expected long-term use, an additional modifying factor should be applied to account for this added uncertainty. In cases where patient exposure is expected to be limited or prolonged, subacute and subchronic data may be appropriate without application of a modifying factor. LD_{50} values are not recommended as a point of departure in calculating safe exposure limits for a compound due to inherent uncertainty associated with this endpoint, failure to identify earlier critical toxic effects, and extreme uncertainty regarding potential long-term effects associated with compound exposure.

6.2.7 Data Quality and Availability

As with any standard risk assessment, data quality and availability may significantly impact uncertainty associated with a calculated safe exposure level (IGHRC 2006; Sharratt 1988). When carrying out route-to-route extrapolation, it is ideal if high-quality data (i.e., studies performed to GLP or OECD guidelines with appropriate sample size in animals of both sexes) pertaining to systemic toxicity, reproductive and developmental toxicity, carcinogenicity, and genotoxicity are readily available for a compound or can be extrapolated from surrogate compounds expected to have similar toxicities, toxicokinetic parameters, and structure/chemical space. It is helpful if potential mechanism(s) of toxicity have been elucidated for a compound to determine if local or portal of entry effects may occur or if a mechanism of toxicity may be conserved in a variety of bodily tissues. For the purposes of route-to-route extrapolation, lack of quality data may increase uncertainty or restrict feasibility of using this approach (Fig. 6.1).

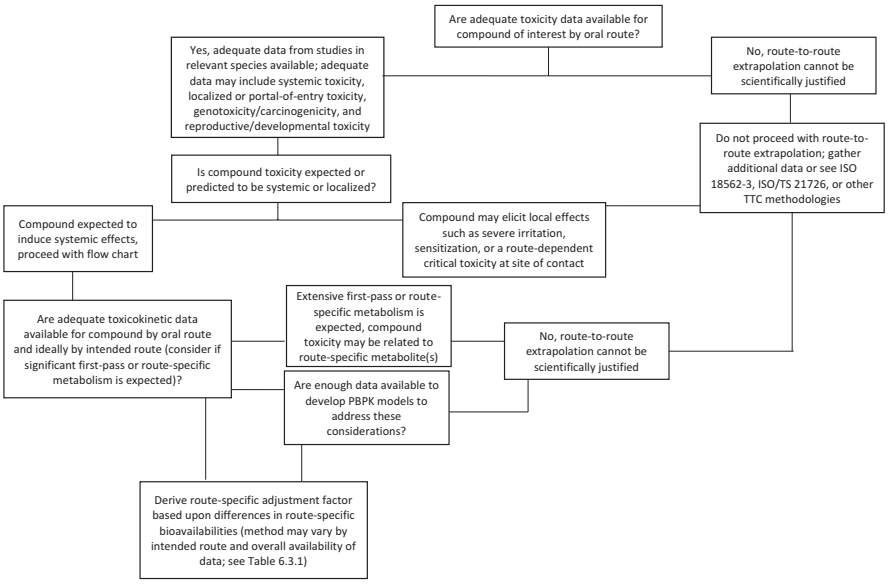


Fig. 6.1 Decision tree for establishing suitability of route-to-route extrapolation

6.3 Regulatory Approaches to Route-to-Route Extrapolation

6.3.1 US Environmental Protection Agency (US EPA)

US EPA published "Methods for Derivation of Inhalation Reference Concentrations and Application of Inhalation Dosimetry" in 1994 to provide general guidance on derivation of reference concentration (RfC) values for inhalation exposure. It was suggested that route-to-route extrapolation could only be used in circumstances where the following criteria were established:

- Local respiratory effects and first-pass metabolic effects (via oral or inhalation metabolism) were not expected to occur.
- Compound is not expected to elicit route-specific toxicities.
- Dosimetry comparison between routes cannot be established.
- Respiratory tract data are unavailable in oral (or starting route) study.
- Short-term inhalation, dermal irritation, in vitro assessments, or physicochemical characteristics of the compound suggest compound may result in local or portal of entry effects within respiratory tract but are insufficient to form basis of RfC derivation.

If criteria are met, the development of a physiologically based pharmacokinetic (PBPK) model to determine the pharmacokinetic parameters of the compound for routes of interest should be completed in order to model the effective dose required to elicit the critical toxic effect of concern via the starting route and intended route

for extrapolation. This guidance did not discuss the use of default extrapolation factors or a default method for route-to-route extrapolation (EPA 1994).

6.3.2 *Interdepartmental Group on Health Risks and Chemicals (IGHRC)*

IGHRC, which is a working group with members from numerous scientific agencies in the United Kingdom, drafted guidelines on the use of route-to-route extrapolation for the purposes of assessing human health risk to chemicals present in the environment and in "normal use." Potential default extrapolation values were described for the purposes of route-to-route extrapolation, as well as methods for extrapolation to dermal and inhalation exposure in cases where bioavailability data are available for either the starting route or intended route (IGHRC 2006).

6.3.3 *ECHA*

ECHA published "Guidance on information requirements and chemical safety assessment: Chapter R.8: Characterization of dose [concentration]-response for human health" as part of a guidance series in order to meet REACH regulation standards. The guidance suggested that route-to-route extrapolation should only be considered if the compound of interest is not expected to elicit local effects or undergo first-pass metabolism. Although it is suggested that compound- and route-specific absorption values be used to derive a modifying factor, default factors for oral-to-inhalation, inhalation-to-oral, and oral-to-dermal extrapolations were provided. A preferred general method of extrapolation was proposed, which accounted for allometric scaling and route-specific absorption (see Table 6.3) (ECHA 2012).

6.3.4 *European Medicines Agency*

The EMA "Guideline on setting health based exposure limits for use in risk identification in the manufacture of different medicinal products in shared facilities" published in 2014 established general guidelines for route-to-route extrapolation to establish safe exposure limits to pharmaceutical contaminants during manufacturing processes for drugs that may be administered via different routes. In cases where compound bioavailability is expected to be significantly different between routes, a correction factor assuming 100% bioavailability of a contaminant compared to the bioavailability of starting route may be calculated. If data are unavailable and systemic bioavailability to the contaminant is expected to be lower than that of the intended route, a correction factor is not required (see Table 6.3) (EMA 2014).

Table 6.3 Summary of suggested default extrapolation factors for compounds amenable to route-to-route extrapolation

Agency, guidance, or group	Applicability	Suggested adjustment factor(s) per route (generally to be multiplied by POD unless otherwise noted)		Reference
Interdepartmental group on health risks from chemicals (IGHRC)	Environmental chemicals, occupational exposures, pesticides	**Additional data**	**Dermal**	IGHRC (2006)
		Oral and dermal bioavailability	$\frac{\%\text{oral bioavailability}}{\%\text{dermal bioavailability}}$	
		Oral bioavailability data and data suggesting compound readily crosses the skin	Assume equivalent bioavailability	
		Oral bioavailability data and experimental data indicating low dermal absorption	$\frac{\%\text{oral absorption}}{10}$	
		Dermal bioavailability data and data suggesting good oral bioavailability	$\frac{50}{\%\text{dermal bioavailability}}$	
		Dermal bioavailability data and data suggesting poor oral bioavailability	$\frac{10}{\%\text{dermal bioavailability}}$	
		No bioavailability data for either route; assumed low dermal bioavailability based upon physicochemical factors	$\frac{50}{10}$	
		No data	Assume equivalent bioavailability	
		Additional data	**Inhalation**	
		Oral and inhalation bioavailability	$\frac{\%\text{oral absorption}}{\%\text{respirable absorption}}$	

(continued)

Table 6.3 (continued)

Agency, guidance, or group	Applicability	Suggested adjustment factor(s) per route (generally to be multiplied by POD unless otherwise noted)		Reference
		Oral bioavailability data and unknown inhalation toxicity	$\frac{\%\text{oral absorption}}{100}$	
		Acute oral and inhalation toxicity data	Correction factor based upon comparison of relative potency	
		Inhalation bioavailability data for compound known to be toxic or very toxic in acute oral toxicity study	$\frac{50}{\%\text{inhalation bioavailability}}$	
		No inhalation bioavailability data for compound known to be toxic or very toxic in acute oral toxicity study	$\frac{50}{100}$	
		Inhalation bioavailability data for compound known to harmful or not classified in acute oral toxicity study	$\frac{10}{\%\text{inhalation bioavailability}}$	
		No inhalation bioavailability data for compound known to harmful or not classified in acute oral toxicity study	$\frac{10}{100}$	

EMA	Cross-contaminants in drug manufacturing facilities	**Additional data**	**Unspecified**	EMA (2014)
		Species-specific bioavailability data (i.e., human data for human drug product, animal data for veterinary drug product) with assumption that compound may exhibit enhanced bioavailability by other route	$\frac{\%\text{absorption starting route}}{100\%\text{absorption by other route}}$	
		Data unavailable; systemic exposure to contaminant expected to be less than that by drug administration route	None; assumed that exposure by other route will be less than that of drug administration route	
ICH Q3D	Elemental impurities in pharmaceuticals (but often applied to medical devices as well)	**Additional data**	**Inhalation or parenteral**	ICH (2014)
		If route-specific data were absent or insufficient, modifying factors based upon oral bioavailability were employed	Oral bioavailability <1%, oral bioavailability data unavailable, or occupational exposure limit unavailable: divide by modifying factor of 100	
			Oral bioavailability ≥1% and <50%: divide by modifying factor of 10	
			Oral bioavailability ≥50% and <90%: divide by modifying factor of 2	
			Oral bioavailability ≥90%: divide by a modifying factor of 1	

(continued)

Table 6.3 (continued)

Agency, guidance, or group	Applicability	Suggested adjustment factor(s) per route (generally to be multiplied by POD unless otherwise noted)		Reference
ECHA	Provides guidance for calculation of derived no-effect levels under REACH; includes guidance regarding bioavailability, extrapolation, and allometric scaling for data among different routes and species	**Additional data**	**Inhalation or dermal**	ECHA (2012)
		To convert from an oral (or dermal) N(L)OAEL to inhalation N(L)OAEC, should consider species-specific default respiratory volume and route- and species-specific absorption values	$\text{Inhalation N(L)OAEC} = \text{oral N(L)OAEL} \times \frac{1}{\text{sRV}} \times \frac{\text{absorption}_{\text{oral-rat}}}{\text{absorption}_{\text{inhalation-human}}}$ where: N(L)OAE(C/L) = no or lowest observed adverse effect concentration/level sRV = species default daily respiratory volume [a]**may need to correct for differences in experimental and human exposure conditions (i.e., account for exposure duration observed in animal study compared to human exposure)**
		To convert from inhalation N(L)OAEC to oral (or dermal) N(L)OAEL, should consider species-specific default respiratory volume and route- and species-specific absorption values	$\text{Oral N(L)OAEL} = \text{inhalation N(L)OAEC} \times \text{sRV} \times \frac{\text{absorption}_{\text{inhalation-rat}}}{\text{absorption}_{\text{oral-human}}} \times \text{allometric scaling factor}$ where: N(L)OAE(C/L) = no or lowest observed adverse effect concentration/level sRV = species default daily respiratory volume [a]**may need to correct for differences in experimental and human exposure conditions (i.e., account for exposure duration observed in animal study compared to human exposure)**

		To convert from oral N(L)OAEL to dermal N(L)OAEL, should consider route- and species-specific absorption values	$\text{Dermal N(L)OAEL} = \text{oral N(L)OAEC} \times \frac{\text{absorption}_{\text{oral-rat}}}{\text{absorption}_{\text{dermal-human}}} \times \text{allometric scaling factor}$ where: N(L)OAE(C/L) = no or lowest observed adverse effect concentration/level [a]**may need to correct for differences in experimental and human exposure conditions (i.e., account for exposure duration observed in animal study compared to human exposure)**
		To convert from dermal N(L)OAEL to oral N(L)OAEL, should consider route- and species-specific absorption values	$\text{Oral N(L)OAEL} = \text{dermal N(L)OAEC} \times \frac{\text{absorption}_{\text{dermall-rat}}}{\text{absorption}_{\text{oral-human}}} \times \text{allometric scaling factor}$ where: N(L)OAE(C/L) = no or lowest observed adverse effect concentration/level [a]**may need to correct for differences in experimental and human exposure conditions (i.e., account for exposure duration observed in animal study compared to human exposure)**
		Additional data	**Inhalation**
		If bioavailability data are unavailable, worst-case assumptions apply assuming limited absorption by the starting route (e.g., oral) and 100% exposure via intended route (e.g., inhalation)	$\frac{50\%\text{bioavailability by starting route}}{100\%\text{bioavailability by intended route}}$
		Additional data	**Dermal**
		Assumed that dermal absorption will not exceed oral absorption	Default factor should not be introduced to calculation (as systemic exposure is expected to be less than that for oral exposure)

[a]In all cases described in this table, it is expected that all conditions allowing for route-to-route extrapolation were met

6.3.5 ICH Harmonized Guideline Q3D

Within the ICH "Guideline for Elemental Impurities" ICH Q3D, parenteral PDE values for elemental compounds lacking sufficient route-specific toxicity data were derived by applying additional modifying factors to the available point of departure. Modifying factors (MFs) ranging from 1 to 100 were applied based upon oral bioavailability as follows:

Oral bioavailability <1%: divide by MF of 100
Oral bioavailability <50%: divide by MF of 10
Oral bioavailability between 50% and 90%: divide by MF of 2
Oral bioavailability >90%: divide by MF of 1

For elemental compounds lacking adequate toxicity data by the inhalation route for derivation of a PDE, a MF of 100 was applied to the calculated oral PDE. The use of 100 as a MF was rooted in an analysis of a subset of compounds with oral and inhalation data in the EPA HEAST dataset, which summarized provisional risk assessment data and compound-specific oral and inhalation reference dose (RfD) and concentration (RfC) values, respectively. Within the subset, safe daily human exposure doses in mg/kg body weight/day derived from RfC values were up to 100-fold lower than those from accompanying RfD values, therefore providing a basis for general application of a MF of 100 for route-to-route extrapolation of data from oral to inhalation exposures (Ball et al. 2007; ICH 2014).

ICH Q3D also provides recommendations regarding derivation of route-specific PDEs for alternative routes (e.g., dermal, rectal, vaginal, and ocular) not described in the guidance, suggesting that the risk assessor (1) select the most appropriate Q3D PDE as a starting point, utilizing the oral PDE as the default; (2) determine if localized portal of entry effects are expected when administered by the intended route and adjust an existing PDE as necessary; (3) if local effects are absent, consider the bioavailabilities of the compound by the oral (or parenteral or inhalation) route and new proposed route for deriving a compound-specific correction factor; and (4) consider product quality attributes if a higher PDE is derived compared to an established PDE (Fig. 6.1; Table 6.3) (ICH 2014).

6.4 Route-Specific Considerations

Risk assessments for compounds which may migrate from a medical device are typically challenging. These compounds typically have little to no safety data. The limited safety data that may be available are almost never in the correct route of exposure. Typically, most available toxicity data are reported by the oral route, whereas exposure to medical devices occurs by multiple routes not limited to oral, inhalation, dermal/transdermal, parenteral, rectal, vaginal, ocular, and intrathecal.

While some routes of exposure have barriers that limit exposure to compounds, others provide little to no protection against chemical exposure. For example, and as previously mentioned, the oral rout of exposure has the first-pass effect. Dermal exposure is often limited as the skin provides an excellent layer of protection, thus limiting the amount and types of compounds that enter the body. On the other hand, parenteral exposure offers little to no such protection. As a result, the same compound can have much different bioavailability based on the route of exposure. These differences in bioavailability are key when performing route-to-route extrapolations for compounds. For the most part, these route-to-route extrapolations for medical device components are performed using oral data, going from oral route of exposure to inhalation or intraperitoneal.

When using oral data to derive safe exposure limits by the parenteral or inhalation route, modifying factors (MFs) are typically used as previously discussed here (see Sect. 6.2.5) and described in ICH Q3D. MFs typically range from 1 to 100 based on the bioavailability of the compound by the oral route of exposure. For parental exposure, as the oral bioavailability increases, the MF decreases. In instances where the oral bioavailability is greater than 90%, the MF is 1. When the oral bioavailability is unknown, a worst-case scenario of oral bioavailability <1% should be assumed and a MF of 100 used. Similarly, for inhalation exposures based on oral data, a MF of 100 is generally used.

Similar methodologies may be used for specialized routes of exposure including rectal, vaginal, ocular, intravenous, and intrathecal. However, in addition to differences in bioavailability, localized effects should also be considered.

6.5 Derivation of Safe Exposure Limits

A variety of methods exist regarding derivation of safe exposure limits to compounds that may be present in, on, or leach from medical devices or combination products. Possible methods and regulatory guidance on the topic are summarized below.

6.5.1 Derivation of a Permissible Daily Exposure (PDE) Value

The PDE calculation is intended to provide a safe lifetime daily exposure limit to a compound based upon assessment of pertinent toxicity data to identify the critical effect (the most sensitive indicator of an adverse effect at the lowest identified dose) as a point of departure for applying additional safety factors as described in ICH Harmonized Guideline ICH Q3C, “Impurities: Guideline for Residual Solvents Q3C(R6)” and elsewhere. Regarding medical devices, it is more common to derive a tolerable intake (TI) and subsequent tolerable exposure (TE) value based upon standardized patient body weights in kg, but PDEs may be calculated when assess-

ing combination products, container closure systems or packaging materials, and manufacturing components. Standard safety factors (each up to a value of 10) used in calculating a PDE are described in further detail below (ICH 2016):

$$\text{PDE} = \frac{\text{POD} \times \text{Body Weight}\left(\text{kg}\right)}{\text{F}_1 \times \text{F}_2 \times \text{F}_3 \times \text{F}_4 \times \text{F}_5}$$

where:

PDE = permissible daily exposure

POD = point of departure; preferably a no-observed-adverse-effect-level (NOAEL) or no-observed-effect-level (NOEL); in the absence of NOAEL or NOEL values, a lowest-observed-adverse-effect-level (LOAEL) value may be used. LD_{50} values are not recommended in deriving a PDE as these are crude markers of toxicity which fail to identify the earliest critical effect essential in risk assessment.

F1 = interspecies extrapolation (accounts for comparative surface area to body weight ratios for various test species and humans)

F2 = intraspecies extrapolation (accounts for variability in humans)

F3 = duration of exposure (accounts for duration of acute or subchronic studies)

F4 = severe toxicity (accounts for non-genotoxic carcinogenicity, neurotoxicity, or teratogenicity)

F5 = accounts for data quality and availability (i.e., extrapolation from a LOAEL or TD_{LO} to NOAEL)

Regarding medical devices, the risk assessor when considering route-to-route extrapolation or especially sensitive patient populations, an additional safety factor, F6, may also be included. F6, if possible, should be predicated upon experimental data such as bioavailability data for routes of interest to derive an appropriate adjustment factor. If bioavailability data are unavailable for the intended route of exposure, absorption of 100% may be used as a default for comparison (especially in considering parenteral exposures).

$$\text{F6} = \frac{\text{Absorption via intended route}}{\text{Absorption via oral route}\left(\text{or route of POD}\right)}$$

For example, when considering a compound (such as a leachable or extractable substance) derived from an implanted device resulting in parenteral exposure, F6 may be calculated as follows assuming the risk assessor has absorption data for the compound via the route of the POD value and assumes 100% absorption by the intended route of exposure in the absence of experimental ADME data for this route.

$$\text{F6} = \frac{100\%}{80\%} = 1.25$$

If a range of absorption values are available regarding the route of the POD, the most conservative (i.e., the lowest) absorption value should be utilized in calculating F6 in order to arrive at the most conservative adjustment factor. For instance, for a compound with adequate toxicity data and extremely low oral absorption (>1%), F6 would be 100 assuming 100% bioavailability by the intended route of exposure. However, this approach should be used with caution and expert judgment, especially in considering whether the POD may be an accurate representation of potential toxicity associated with the intended route of exposure (Broschard et al. 2016).

A similar approach to route-to-route extrapolation has also been presented by the European Medicines Agency (EMA) in a 2014 guidance document, "Guidelines on setting health based exposure limits for use in risk identification in the manufacture of different medicinal products in shared facilities." The guidance describes calculation of a PDE using precepts from ICH Q3C for the purposes of establishing cleaning validation limits in pharmaceutical manufacturing facilities and provides additional recommendations regarding route-to-route extrapolation. In instances where bioavailabilities between the POD and intended routes exhibit clear differences, a correction factor (preferably based upon data from same species) assuming 100% bioavailability via the intended route may be derived as follows:

$$\text{Correction Factor} = \frac{\%\ \text{bioavailability POD}}{100\%\ \text{bioavailability intended route}}$$

This factor may then be multiplied by the calculated PDE to derive a safe exposure limit accounting for route (EMA 2014).

In the event that bioavailability data are entirely unavailable for the compound of interest, standardized default values for route-specific extrapolations have not been established (Broschard et al. 2016). Furthermore, in the absence of compound-specific data capable of meeting criteria described in Sect. 6.3, the use of established generic safety thresholds as described in ICH M7, the Cramer classification scheme (which is based upon TTC concept), ISO 18562 (general TTC values for compounds in medical devices resulting in gas pathway exposure), or ISO/TS 21276 (general TTC values for medical device constituents) may be more appropriate (ICH 2017; ISO 2017b, 2019).

6.5.2 Derivation of Tolerable Intake (TI) and Tolerable Exposure (TE) Values

The International Standard ISO 10993 guidance series regarding biological evaluation of medical devices provides the international framework for biological assessment of sterile and non-sterile medical devices that may come in indirect or direct

contact with patients for and by a variety of purposes and routes, respectively (FDA 2016). ISO 10993-1, "Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process," and ISO 10993-17, "Biological evaluation of medical devices – Part 17: Methods of the establishment of allowable limits for leachable substances," suggest that patient contact duration with a medical device and route of exposure play a key role in determining (1) if an overall device, material of construction, or leachable compound may result in patient toxicity and (2) if safe exposure doses to leachable compounds may be calculated based upon available toxicity data [route, duration of study, applicability of study dose(s), and overall data quality]. While ISO 10993-1 describes the overarching premise for evaluating safety of a device based upon a number of biocompatibility and toxicity assays, ISO 10993-17 provides guidance on the derivation of TI and patient population-specific TE values for leachable/extractable substances identified in a medical device or its packaging materials (ANSI 2008; ISO 2018).

Although specific methods of route-to-route extrapolation are not defined in ISO 10993-1 and ISO 10993-17, it is strongly encouraged that risk assessors compare and consider route and duration of exposure to a device in actual patient populations with available literature data when calculating a safe exposure level to a leachable/extractable compound. Route-to-route extrapolation approaches should be employed with expert judgment, careful consideration of portal of entry or localized effects, potential use of physiologically based pharmacokinetic modeling of systemic- and tissue-specific exposure, and relevance of data (e.g., dose, species, study duration) to expected patient exposure and population(s). Upon consideration of these factors, ISO 10993-17 suggests that uncertainty regarding route extrapolation be included in uncertainty factor 3 (UF3) of the TI calculation, which is provided below. A TI provides a value, in mg/kg body weight, that should be a safe limit for daily exposure to compounds present in or on a device.

$$\mathrm{TI} = \frac{\mathrm{POD}}{\mathrm{MF}}$$

$$\mathrm{MF} = \mathrm{UF1} \times \mathrm{UF2} \times \mathrm{UF3}$$

where:

TI = tolerable intake

POD = point of departure; preferably a NOAEL or LOAEL if possible

MF = modifying factor which is the multiplied product of uncertainty factors 1, 2, and 3 (UF1, UF2, and UF3); intended to account for uncertainty associated with estimating a safe dose while considering potential toxicities associated with a compound of interest; total factor up to 1000 is usually protective but may vary if only poor or inappropriate (e.g., study duration or route of exposure) data are available

UF1 = accounts for interhuman variability; default of 10 usually

UF2 = accounts for extrapolation of data from other species to human; default of 10 usually but may be altered based upon expected differences in test species and human

UF3 = accounts for quality and relevance of available data, ranges from 1 to 100 based upon applicability of POD based upon study duration and route of exposure, identification of a NOAEL, breadth of data available on compound, rate of exposure, and overall confidence in available data, may exceed 100 if necessary if only acute data are available for deriving safe level of compound in device with permanent patient contact

Once a TI is calculated for a compound, patient population-specific TE values may be calculated based upon standard patient population body weights (kg) and a utilization factor (UTF) based upon expected concomitant exposure to a leachable from other medical devices and days of actual use of a device compared to duration of an exposure category [e.g., expected device use of 7 days for device determined to have prolonged (up to 30 days) contact with patient].

The TE calculation is as follows and described in greater detail below:

$$\mathrm{TE} = \mathrm{TI} \times \text{body weight}\left(\mathrm{kg}\right) \times \mathrm{UTF}$$

$$\mathrm{UTF} = \mathrm{CEF} \times \mathrm{PEF}$$

where:

TE = tolerable exposure (in mg/day)

TI = tolerable intake (in mg/kg)

Body weight = up to 70 kg for adults (other regulatory agencies may use standard body weight of 60 kg or 50 kg), 10 kg for pediatric patients, 3.5 kg for infants, and 0.5 kg for neonates

UTF = utilization factor

CEF = concomitant exposure factor, default of 0.2, may also be calculated using additional equations present in ISO 10993-17 if leachable is expected to be present in >5% of medical devices or increased to 1 if <5% of devices solid in calendar year or <5 devices containing the leachable are used in any single medical procedure

PEF = proportional exposure factor, default of 1, may be calculated to result in higher value by dividing number of days in exposure category by number of days of device use if less than exposure category

Despite mention of route-to-route extrapolation as necessary for an accurate assessment of toxicity associated with leachable compounds, specific methods or recommended resources for further reading are not described in ISO 10993 guidances. Therefore, extrapolation is based upon expert judgment and methods otherwise discussed throughout the literature without regulatory guidance or consensus.

6.5.3 Adjustment of Cramer Classification Values

The Cramer classification scheme (decision tree) is a predictive toxicology method used to assess chemicals in order to generate a toxicological threshold concern (TTC) estimation. It is typically intended to estimate a TTC for chronic exposure via the oral route but may be a useful tool to gauge the potential risk associated with a compound. The decision tree uses chemical structures, estimates of total human intake, recognized pathways for metabolic deactivation and activation, toxicity data, and the presence of a substance as a component of traditional foods or as an endogenous metabolite to evaluate the compound of interest. The Cramer classification recommended threshold (Cramer TTC) is intended to be protective for daily oral exposure to compounds and was designed using oral NOAEL values from chronic, subchronic, reproductive, and developmental toxicity studies carried out in rodents and rabbits. The lowest 5% of NOELs for a group were then modified with a safety factor of 100 to ensure a margin of safety. Acute toxicity of the compounds was not taken into consideration when these levels were determined, nor were other routes of administration (Cramer et al. 1978; Munro et al. 1996). The Cramer TTC is still useful in estimating overall potential risk from a compound and can be used as a basis for determining the safe level relevant for alternative routes of exposure (Broschard et al. 2016; ICH 2014).

Substances are classified and threshold levels recommended as follows (Broschard et al. 2016; Health Canada 2016; Kroes et al. 2004; WHO/EFSA 2016):

- Class I: substances with simple structures and for which efficient modes of metabolism exist, suggesting a low order of oral toxicity. Recommended threshold is 1800 μg/day for oral products.
- Class II: substances which possess structures that are less innocuous than class I substances, but do not contain structural features suggestive of toxicity like those substances in class III. Recommended threshold is 540 μg/day.
- Class III: substances with chemical structures that permit no strong initial presumption of safety or may even suggest significant toxicity or have reactive functional groups. Recommended threshold is 90 μg/day.
- Compounds with anti-acetylcholinesterase activity including organophosphates or carbamates. Recommended threshold is 18 μg/day.
- FDA threshold of regulation for genotoxic substances. Recommended threshold is 1.5 μg/day. Compounds identified at levels below 1.5 μg are generally not assessed, and 1.5 μg should be protective against effects of genotoxic carcinogens.

Although a consensus or default safety factor value has not been established by regulatory authorities for route-to-route extrapolation, it has generally been suggested that compound-specific factors may be derived based upon bioavailability of a compound by the oral and intended exposure route, although caveats may apply (see Sect. 6.3). For instance, assuming conservative oral bioavailability of ~1% and potential parenteral bioavailability of 100%, a safety factor of 10 should be adequate

to address extrapolation as described by Broschard and others (Broschard et al. 2016; ICH 2014). Furthermore, numerous groups have researched and published possible methods regarding extrapolation of the TTC method primarily to inhalation and dermal exposures resulting from occupational, environmental, and consumer products (e.g., household products, cosmetics, and personal care products) which are not discussed in detail here and have not been fully validated by regulatory agencies (Blackburn et al. 2005; Carthew et al. 2009; Escher et al. 2010; Hoersch et al. 2018; Kroes et al. 2007; Rennen et al. 2004; Safford 2008; Williams et al. 2016; Yang et al. 2017).

6.5.4 Consideration of ISO 18562-3

ISO 18562-3 specifically provides TTC values for volatile organic compounds (VOCs) lacking appropriate data for calculation of TI and TE values that may be conveyed to patients through the gas pathway throughout the use of respiratory devices including devices, component parts, and accessories.

The TTC values listed in Table 6.4 above are not intended to be protective against local effects or assume biocompatibility of a compound. The TTC values are also presented in μg/day assuming a 70 kg adult. ISO 18562-1 indicates that patient population-specific TTC values be calculated for patient populations of lower body weight to derive a safe limit. Body weights used in adjustment include 0.5 kg for neonates, 3.5 kg for infants, and 10 kg for pediatric patients. Sex-specific body weight adjustments are not discussed in this guidance, although ISO 10993-17 discusses the use of a reduced standardized body weight for adult females (ANSI 2008; ISO 2017a, b).

6.5.5 Consideration of ISO/TS 21726

ISO/TS 21276, *Biological evaluation of medical devices – Application of the threshold of toxicological concern (TTC) for assessing biocompatibility of medical device constituents*, was published in early 2019. This standard describes the derivation

Table 6.4 TTC limits for VOCs identified in gas pathway from respiratory medical devices

Device exposure category	Duration of patient exposure	TTC (μg/day)		
Limited	≤24 h	360	–	–
Prolonged	>24 h and <30 d	360 (in first 24 h exposure)	120 (for subsequent 29 days)	–
Permanent	≥30 d	360 (in first 24 h exposure)	120 (for subsequent 29 days)	40, beyond 30 days

Table 6.5 TTC limits for medical devices[a]

Medical device contact category	Limited (<24 h)	Prolonged (24 h to 30 d)	Long term (>30 d)[b]		
Duration of body contact	≤1 month		>1 month to 12 months	>1 year to 10 years	>10 years to lifetime
Daily intake of any one constituent (μg/day)	120		20	10	1.5[c]

[a]TTC does not apply to compounds that are expected to be highly toxic ("cohort of concern" compounds) including aflatoxin-like compounds, N-nitroso compounds, azo compounds, polyhalogenated-dibenzodioxins, polyhalogenated-dibenzofurans, and polyhalogenated-biphenyls, strained heteronuclear rings, heavy metals, α-nitro furyl compounds, hydrazines, triazines, azides, or azoxy compounds, polycyclic amines, steroids, and organophosphorus compounds, as well as high molecular weight polymers, particles, ceramics, proteins, and radioactive substances
[b]Long-term includes devices generally defined as permanently contacting in ISO 10993-1
[c]Based upon 10^{-5} cancer risk and 60 kg adult body weight

and applicability of TTC values specifically for compounds present in, or, or released from a medical device that lack adequate data to derive TI and TE values. The TTC values described in ISO/TS 21726 are intended to be protective against effects elicited by exposure to genotoxic and non-genotoxic carcinogenic, systemic, and reproductive toxicants, but do not encompass endpoints otherwise assessed in standard ISO 10993 testing including cytotoxicity, irritation and sensitization, hemocompatibility, pyrogenicity, or local tissue effects at the site of bodily contact. The values were based upon TTC limits described in ICH M7 for control of DNA reactive impurities in pharmaceutical products regardless of patient population or route of administration ordinarily without the need for further route-related adjustment (ICH 2017; ISO 2019). However, it is important to note that FDA CDRH has not recognized this standard yet for use in regulatory submissions. Table 6.5 below presents these TTC values as defined in ISO/TS 21726:

6.6 Conclusions

While route-to-route extrapolation has been explored for decades in the field of regulatory toxicology, formal guidance from regulatory agencies specifically in reference to chemical exposures resulting from medical devices is lacking. The US EPA, EU agencies and programs (ECHA, EU SCOEL, COSMOS, EMA), and numerous academic and regulatory scientists have proposed a variety of methodologies and considerations regarding appropriate application of route-to-route extrapolation. However, a consensus standard regarding these approaches is unavailable, which significantly reduces compliance with standardized (and comparable) route-specific extrapolation methods among regulatory scientists. Considering that regulatory scientists are often required to assess compounds with incomplete or inadequate data to fulfill all proposed criteria for route-to-route extrapolation, more research

regarding derivation and implementation of route-specific "default" values for regular use in risk assessment is required. Alternatively, use, and perhaps extrapolation, of Threshold of Toxicological Concern (TTC) values as presented in ICH M7, ISO 21726, or ISO 18562 guidance may assist in deriving safe exposure values for compounds with inadequate data by all potential routes of exposure.

References

ANSI. (2008). *Biological evaluation of medical devices - Part 17: Methods for the establishment of allowable limits for leachable substances*. Association for the Advancement of Medical Instrumentation (AAMI).

Ball, D., Blanchard, J., Jacobson-Kram, D., McClellan, R. O., McGovern, T., Norwood, D. L., et al. (2007). Development of safety qualification thresholds and their use in orally inhaled and nasal drug product evaluation. *Toxicological Sciences, 97*(2), 226–236. https://doi.org/10.1093/toxsci/kfm058.

Blackburn, K., Stickney, J. A., Carlson-Lynch, H. L., McGinnis, P. M., Chappell, L., & Felter, S. P. (2005). Application of the threshold of toxicological concern approach to ingredients in personal and household care products. *Regulatory Toxicology and Pharmacology, 43*(3), 249–259. https://doi.org/10.1016/j.yrtph.2005.08.007.

Bond, J. A. (1993). Metabolism and elimination of inhaled drugs and airborne chemicals from the lungs. *Pharmacology & Toxicology, 72*(Suppl 3), 36–47.

Broschard, T. H., Glowienke, S., Bruen, U. S., Nagao, L. M., Teasdale, A., Stults, C. L., et al. (2016). Assessing safety of extractables from materials and leachables in pharmaceuticals and biologics - current challenges and approaches. *Regulatory Toxicology and Pharmacology, 81*, 201–211. https://doi.org/10.1016/j.yrtph.2016.08.011.

Carthew, P., Clapp, C., & Gutsell, S. (2009). Exposure based waiving: The application of the toxicological threshold of concern (TTC) to inhalation exposure for aerosol ingredients in consumer products. *Food and Chemical Toxicology, 47*(6), 1287–1295. https://doi.org/10.1016/j.fct.2009.02.024.

Cramer, G. M., Ford, R. A., & Hall, R. L. (1978). Estimation of toxic hazard--a decision tree approach. *Food and Cosmetics Toxicology, 16*(3), 255–276.

ECETOC. (2003). *Derivation of assessment factors for human health risk assessment*. Brussels: European Centre for Ecotoxicology and Toxicology of Chemicals.

ECHA. (2012). Guidance on information requirements and chemical safety assessment. In *Chapter R.8: Characterisation of dose [concentration]-response for human health*. European Chemicals Agency.

EMA. (2014). *Guideline on setting health based exposure limits for use in risk identification in the manufacture of different medicinal products in shared facilities*. European Medicines Agency.

EPA. (1994). *Methods for derivation of inhalation reference concentrations and application of inhalation dosimetry*. U.S. Environmental Protection Agency.

Escher, S. E., Tluczkiewicz, I., Batke, M., Bitsch, A., Melber, C., Kroese, E. D., et al. (2010). Evaluation of inhalation TTC values with the database RepDose. *Regulatory Toxicology and Pharmacology, 58*(2), 259–274. https://doi.org/10.1016/j.yrtph.2010.06.009.

FDA. (2016). *Use of International Standard ISO 10993-1, "Biological evaluation of medical devices - Part 1: Evaluation and testing within a risk management process"*. U.S. Food and Drug Administration Retrieved from https://www.fda.gov/media/85865/download

Geraets, L., Bessems, J. G., Zeilmaker, M. J., & Bos, P. M. (2014). Human risk assessment of dermal and inhalation exposures to chemicals assessed by route-to-route extrapolation: The necessity of kinetic data. *Regulatory Toxicology and Pharmacology, 70*(1), 54–64. https://doi.org/10.1016/j.yrtph.2014.05.024.

Gerrity, T. R., Henry, C. J., Andersen, M., Bond, J., Bronaugh, R., Bruckner, J., et al. (1990). *Principles of route-to-route extrapolation for risk assessment*.

Health Canada. (2016). *Threshold of Toxicological Concern (TTC)-based Approach for Certain Substances* Government of Canada Retrieved from https://www.ec.gc.ca/ese-ees/326E3E17-730A-4878-BC25-D07303A4DC13/HC%20TTC%20SciAD%20EN%202017-03-23.pdf

Hoersch, J., Hoffmann-Doerr, S., & Keller, D. (2018). Derivation of an inhalation TTC for the workplace based on DNEL values reported under REACH. *Toxicology Letters, 290*, 110–115. https://doi.org/10.1016/j.toxlet.2018.03.030.

ICH. (2014). *Guideline for Elemental Impurities*. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH).

ICH. (2016). *Impurities: Guideline for residual solvents*. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH).

ICH. (2017). *Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk*. International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH).

IGHRC. (2006). *Guidelines on route-to-route extrapolation of toxicity data when assessing health risks of chemicals*. Bedfordshire: Institute of Environment and Health.

IPCS. (2010). *Characterization and application of physiologically based pharmacokinetic models in risk assessment.* Retrieved from http://www.inchem.org/documents/harmproj/harmproj/harmproj9.pdf

ISO. (2017a). *Biocompatibility evaluation of breathing gas pathways in healthcare applications - Part 1: Evaluation and testing within a risk management process*. International Organization for Standardization.

ISO. (2017b). *Biocompatibility evaluation of breathing gas pathways in healthcare applications - Part 3: Tests for emissions of volatile organic compounds (VOCs)*. International Organization for Standardization.

ISO. (2018). *Biological evaluation of medical devices - Part 1: Evaluation and testing within a risk management process*. International Organization for Standardization.

ISO. (2019). *Biological evaluation of medical devices - Application of the threshold of toxicological concern (TTC) for assessing biocompatibility of medical device constituents*. International Organization for Standardization.

Kroes, R., Renwick, A. G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., et al. (2004). Structure-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. *Food and Chemical Toxicology, 42*(1), 65–83.

Kroes, R., Renwick, A. G., Feron, V., Galli, C. L., Gibney, M., Greim, H., et al. (2007). Application of the threshold of toxicological concern (TTC) to the safety evaluation of cosmetic ingredients. *Food and Chemical Toxicology, 45*(12), 2533–2562. https://doi.org/10.1016/j.fct.2007.06.021.

Munro, I. C., Ford, R. A., Kennepohl, E., & Sprenger, J. G. (1996). Correlation of structural class with no-observed-effect levels: A proposal for establishing a threshold of concern. *Food and Chemical Toxicology, 34*(9), 829–867.

Pepelko, W. E., & Withey, J. R. (1985). Methods for route-to-route extrapolation of dose. *Toxicology and Industrial Health, 1*(4), 153–170. https://doi.org/10.1177/074823378500100410.

Rennen, M. A., Bouwman, T., Wilschut, A., Bessems, J. G., & Heer, C. D. (2004). Oral-to-inhalation route extrapolation in occupational health risk assessment: A critical assessment. *Regulatory Toxicology and Pharmacology, 39*(1), 5–11.

Safford, R. J. (2008). The dermal sensitisation threshold- a TTC approach for allergic contact dermatitis. *Regulatory Toxicology and Pharmacology, 51*(2), 195–200. https://doi.org/10.1016/j.yrtph.2008.02.010.

Schroder, K., Escher, S. E., Hoffmann-Dorr, S., Kuhne, R., Simetska, N., & Mangelsdorf, I. (2016). Evaluation of route-to-route extrapolation factors based on assessment of repeated dose toxicity studies compiled in the database RepDose((R)). *Toxicology Letters, 261*, 32–40. https://doi.org/10.1016/j.toxlet.2016.08.013.

Sharratt, M. (1988). Assessing risks from data on other exposure routes. Possibilities and limitations of using toxicity data derived from one exposure route in assessing risks from other exposure routes. *Regulatory Toxicology and Pharmacology, 8*(4), 399–407.

WHO/EFSA. (2016). *Review of the Threshold of Toxicological Concern (TTC) approach and development of new TTC decision tree*. European Food Safety Authority Retrieved from https://efsa.onlinelibrary.wiley.com/doi/epdf/10.2903/sp.efsa.2016.EN-1006

Williams, F. M., Rothe, H., Barrett, G., Chiodini, A., Whyte, J., Cronin, M. T., et al. (2016). Assessing the safety of cosmetic chemicals: Consideration of a flux decision tree to predict dermally delivered systemic dose for comparison with oral TTC (Threshold of Toxicological Concern). *Regulatory Toxicology and Pharmacology, 76*, 174–186. https://doi.org/10.1016/j.yrtph.2016.01.005.

Yang, C., Barlow, S. M., Muldoon Jacobs, K. L., Vitcheva, V., Boobis, A. R., Felter, S. P., et al. (2017). Thresholds of toxicological concern for cosmetics-related substances: New database, thresholds, and enrichment of chemical space. *Food and Chemical Toxicology, 109*(Pt 1), 170–193. https://doi.org/10.1016/j.fct.2017.08.043.

Chapter 7
Risk Assessments for Medical Devices

Erica N. Rogers

Abstract This chapter focuses on providing a brief but thorough summary of risk assessment in relation to medical devices. This chapter is divided into four key sections: (i) Overview of Device Regulation, (ii) Classification of Medical Devices, (iii) Medical Devices and Risk Assessment, and (iv) Case Studies. In the Overview of Device Regulation section, information is provided on the definition of a medical device as it applies to the US Food and Drug Administration (FDA). For the Classification of Medical Devices section, a summary of the key classes of medical devices, as given by the FDA, is discussed. The Medical Devices and Risk Assessment section includes information on set standards as detailed by the International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH), the key steps of human health risk assessment, examples of sources for data retrieval and analysis, and calculation of a tolerable exposure limit with some implementation of uncertainty and/or safety factors as well as determination of whether a device is safe to use. Lastly, this chapter provides two fictious case studies which assist readers in understanding how to apply the steps of human health risk assessment to determine an appropriate TE limit.

Keywords Medical device · Risk assessment · Tolerable exposure (TE) · Uncertainty factor (UF)

7.1 Overview of Device Regulation

Medical devices have been recognized as inert constructs composed of a single biomaterial or a combination of biomaterials, diagnostic devices, microelectronics, computer components, and software (Gad and Schuh 2018). Medical devices are primarily constructed from such materials as polymers, metals, textiles,

E. N. Rogers, PhD (✉)
Gad Consulting Services, Raleigh, NC, USA
e-mail: erica.rogers@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_7

nanomaterials, ceramics, and animal-derived materials or tissues (Gad and Schuh 2018). Overall, medical devices have been accepted as an apparatus that supports a limited range of interaction in the human body with limitations mainly to be due to direct contact of the medical device with the host tissue. Majority of medical devices are present as solid entities. However, some medical devices are constructed as injected liquids and injectable suspensions or particles or liquids. Additionally, medical devices can be grouped based on their rate of resorption in the human body.

In 1976, the FDA enacted the Medical Device Amendments of 1976 (1976 Amendments) (Pub. L. 94-295). In the Medical Device Amendments of 1976, the FDA established standards for some devices and required premarket clearance and approval for others. Devices classified as posing minimal risk to patients during use were excluded from both these standards and premarket clearance. According to Section 201(h) of the Food Drug and Cosmetic (FD&C) Act, a medical device was defined as an instrument, apparatus, implement, machine, contrivance, implant, in vitro reagent, or other similar related article, including a component, part, or accessory, which is:

1. Recognized in the official National Formulary, or the United States Pharmacopeia, or any supplement to them
2. Intended for use in the diagnosis of disease or other conditions, or in the sure mitigation, treatment, or prevention of disease, in man or other animals
3. Intended to affect the structure or any function of the body of man or other animals, and which does not achieve its primary intended purposes through chemical action within or on the body of man or other animals and which is not dependent upon being metabolized for the achievement of its primary intended purposes (CDRH 1992)

7.2 Classification of Medical Devices

In addition to implementing an exact definition for the term "medical device," the Medical Device Amendments of 1976 created a classification system to categorize medical devices based on their required level of control to ensure the safety and efficacy on the device. In this classification system, medical devices are assigned to one of three regulatory classes which depends on the intended use of the device and the indications for use in the device of interest. Additionally, as the classification for a device increases, the risk to the patient and therefore regulatory control of the device also is enhanced (i.e., Class I includes device with the lowest risk, whereas Class III devices include devices with the greatest amount of risk). The three regulatory classes and their respective requirements are described as follows:

1. *Class I: General Controls*
   Defined as medical devices that possess low to moderate risk to the patient and/or user. Devices in this class have the least amount of regulatory control as they potentially present minimal harm to the patient and/or user. Generally, devices in

this class constitute the broad enforcement, and regulatory requirements as established by Congress should be sufficient (if achieved) to assure effective usage and safety to the patient and/or user. Class I devices typically are straightforward in design and manufacture detail and contain a history of safe usage. Within the Class I category, devices on this group can be subdivided according to exemption controls. Today, approximately 47% of currently used medical devices are designated as Class I devices (USFDA 2017). About 95% of Class I devices are exempt from the regulatory process (USFDA 2017). As defined in 21 Code of Federal Regulation (CFR) Parts 862–892, a Class I device labeled with exemptions does not require compliance with the good manufacturing practices regulation, a premarket notification application, and FDA clearance prior to its promotion in the USA. If a Class I device is determined as not exempt, a 510(k) is require for marketing. Examples of common Class I devices include tongue depressors, an arm sling, and manual surgical instruments (such as clip applier, manual dermabrasion brush, scrub brush, ligature carrier, and chisel).

2. *Class II: General Controls and Special Control*
   Class II medical devices are identified as devices where General Controls are insufficient to provide reasonable assurance of safety and effectiveness. These types of devices are suggested to exert a moderate to high risk to the patient and/or user. To assure the safety and effectiveness of these devices are adequately met, existing methods, standards, and guidance documents have been established. Besides being compliant with General Control measures, it is mandatory that Class II devices comply with Special Controls requirements which include special labeling fields; patient registries and guidelines; adherence to necessary performance standards; post-market surveillance; and FDA medical device specific guidance(s). Like Class I devices, Class II devices contain the subcategories with or without exemption requirements. Similar to Class I exempt devices, Class II devices deemed as exempt do not have to complete a premarket notification application as well as do not require FDA clearance before marketing of the device in the USA. Class II medical devices that are not exempt require FDA-submitted premarket notifications and a complete FDA review of a 510(k) clearance to market submission. To date, about 43% of medical devices are considered as Class II devices (USFDA 2017). Examples of Class II devices include powered wheelchairs, X-ray systems, gas analyzers, pumps, and surgical drapes.
3. *Class III: General Controls and Premarket Approval*
   Class III medical devices have the most stringent regulatory controls as sufficient information is not readily available to assure safety and effectiveness through the application of General Controls and Special Controls. These devices contain the most risk to the patient and/or user and are generally characterized to support or sustain human life, are of substantial importance in preventing impairment of human health, or present potential unreasonable risk illness or injury. Notably, while many implant devices are classified as Class III devices, many have been categorized as in the Class II group. Typically, a Class III device that is not

considerably equivalent to a legally marketed predicate device must be subjected to a premarket approval (PMA) application prior to it being commercially distributed in interstate commerce; exceptions to this requirement can be mandated if the device is down-classified by the FDA or the device is cleared through de novo review process. Currently, about 10% of medical devices regulated by the FDA are classified as Class III devices (USFDA 2017). Examples of Class II devices that require PMA include heart valves, silicone gel-filled breast implants, and implanted cerebella stimulators.

7.3 Medical Devices and Risk Assessment

Although manufacturing companies constructing medical devices strive to achieve the most effectiveness with the least amount of health risk possible, complete absence of potential danger through use of a device is not realistic. In actuality, use of all medical devices over an extended period of time may result in potential health risk to the patient and/or user. In general, a "risk" is defined as the possibility of loss, harm, or other adverse effect to occur as a result of a given situation. A health risk is defined as the possibility that an adverse event will harm or otherwise affect an individual's health. Overall, the risk of an adverse event can be quantified as the combination of the likelihood (probability) and magnitude (severity) of a negative outcome to occur:

$$\text{Risk} = \text{Severity} \times \text{Probability}$$

Although this equation is useful due to its simplicity, health risk is not one-dimensional when in relation to a patient's and/or user's operation of medical devices and potential human health outcomes. To address this problem, many have considered the process known as "health-based risk assessment" which provides an alternative approach for evaluating health risk perspectives (Gad and Gad-McDonald 2015; ISO14971-1 2019; Stark 1997).

7.3.1 Standards and Guidances

A detailed paradigm for the medical device risk management process has been described in the American National Standard/Association for the Advancement of Medical Instrumentation (ANS/AAMI) standards under the International Organization for Standardization (ISO) Standard, *ISO 14971: Medical devices – Application of risk management to medical devices* as shown in Fig. 7.1 (ISO14971-1 2019).

In *ISO 14971* standard, information is provided detailing that a manufacture must establish, document, and maintain accurate identification of hazards through-

out the life cycle and ongoing process of medical devices. During the ongoing process of identifying hazards associated with a medical device, the following elements must be included: risk analysis, risk evaluation, risk control, and production and postproduction information. A schematic illustration of the risk management process, which is a truncated version of event display in Fig. 7.1, is shown in Fig. 7.2 (Ecobichon 1997; ISO14971-1 2019).

For the purposes of this chapter, we will only focus on the risk analysis and risk evaluation which are the essential events for risk assessment as indicated by *ISO 14971*.

7.3.2 Elements of Human Health Risk Assessment

Health risk assessment can be defined as the characterization of the potential adverse effects of human exposure to a particular agent or substance. The aim of health risk assessment in reference to a specific agent or substance is to either establish a safe exposure level or determine the likelihood of its harm during exposure (White et al. 1999). As generalized by the ISO standard, *ISO 14971: Medical devices – Application of risk management to medical devices*, a human health risk assessment with regard to medical devices is a process which encompasses risk analysis and risk evaluation. Thus, human health risk assessment involves systematic use of available information to identify hazards to estimate a risk (i.e., risk analysis) and utilize judgment, according to hazard identification, to determine whether an acceptable risk has been achieved in a given context based on current, available scientific information (i.e., risk evaluation) (ISO14971-1 2019).

Human health risk assessment can be divided into four steps:

- *Step 1: Hazard Identification*
  Hazard identification involves the evaluation of whether exposure to an agent or substance can result in an increased incidence of an adverse health outcome such as birth defects, cancer, and systemic toxic effects. Hazard identification of agent(s) in a device can be qualitatively assessed by conducting some or all of the biological tests as outlined in the medical device category of the *ISO 10993-1* guidance (Table 7.1) (ISO10993-1 2018).

  In general, qualitative extracts of the material or device are prepared according to ISO 10993-12 guidance when short-term, less expensive tests will be performed (Table 7.1). The most common tests during hazard identification include cytotoxicity, sensitization, irritation or intracutaneous reactivity, systemic (acute) toxicity, genotoxicity, short-term implantation analysis, and hemocompatibility evaluation. Chronic testing and/or carcinogenic studies are usually not performed during the hazard identification step as they are time-laboring and more expensive to conduct. Alternatively, analysis of available scientific data may be considered during hazard identification in place of biological testing.

- *Step 2: Dose-Response Assessment*

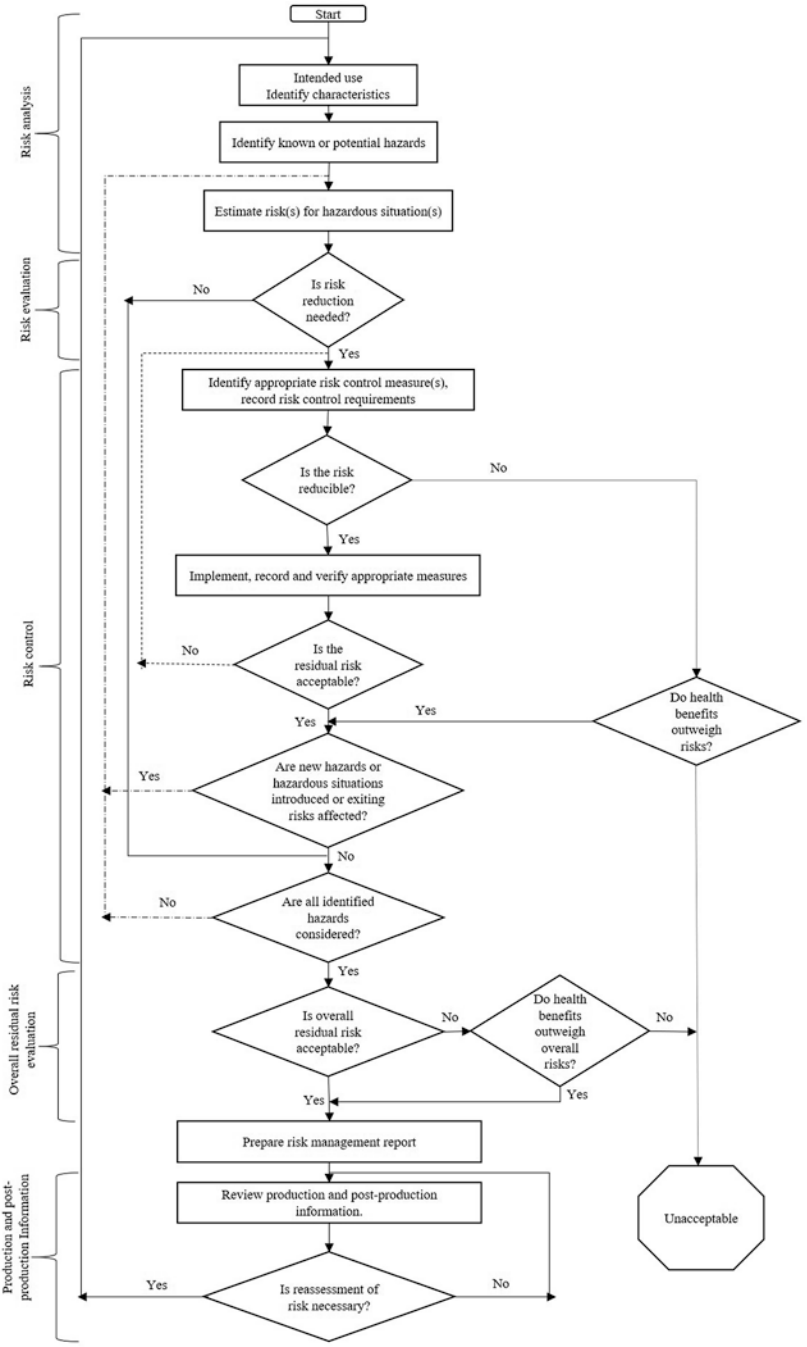


Fig. 7.1 Overview of risk management process for medical devices as described by ISO 14971:2007/(R)2010 (ISO14971-1 2010)

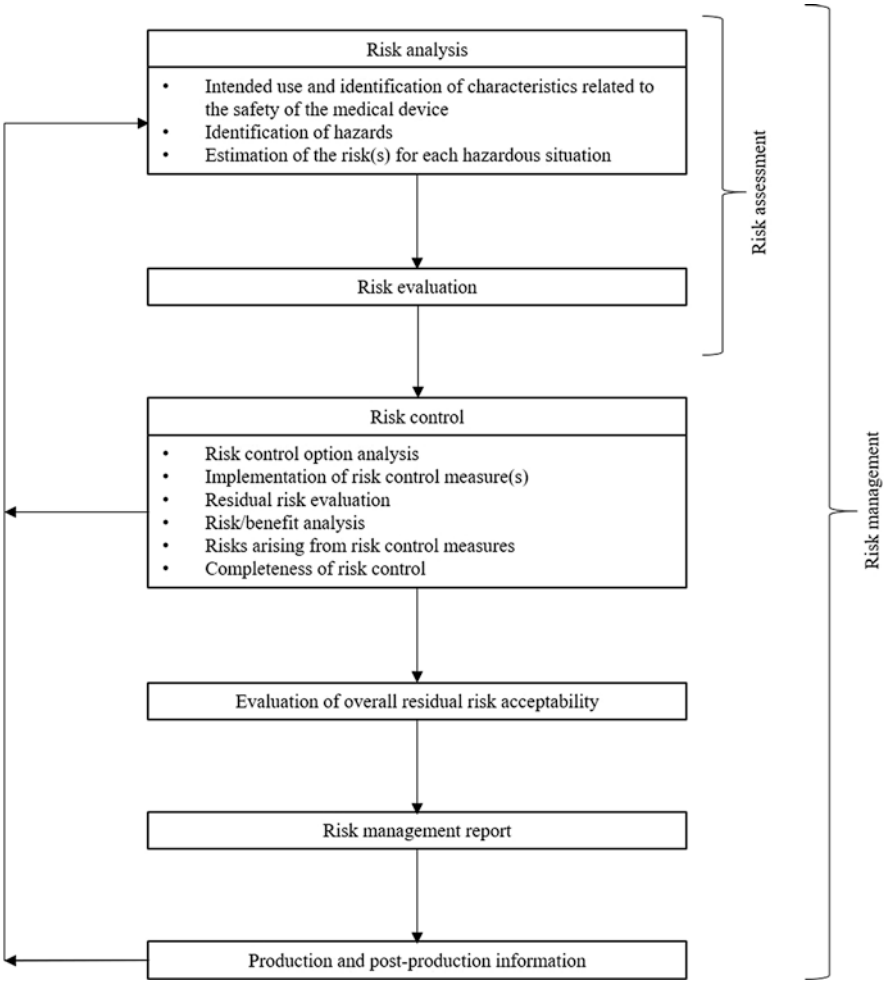


Fig. 7.2 A schematic illustration of the risk management process as described by ISO 14971:2007/(R)2010 (ISO14971-1 2010)

Dose-response assessment is the relationship between the highest exposure dose that does not result in an adverse effect. In theory, all chemicals are toxic at a particular dose; toxicity of a given chemical is dependent on the exposure dose. That is, the toxic effect as a result of exposure to a chemical is reliant on such factors as availability of the substance, form of contact to the substance, and duration of contact to the substance. Compilation and subsequent analysis of adequate animal and human scientific data are essential for determining the dose-response relationship. Overall, the dose-response relationship determines the threshold in which an adverse effect is produced. Many calculated values considered during analysis of the dose-response relationship are provided as a

Table 7.1 ISO 10993 – Biological evaluation of medical devices

Part	Title	Most recent publication year
10993-1	Guidance on selection of tests	2018
10993-2	Animal welfare requirements	2006
10993-3	Tests for genotoxicity, carcinogenicity, and reproductive toxicity	2014
10993-4	Selection of tests for interaction with blood	2017
10993-5	Tests for cytotoxicity: in vitro methods	2009
10993-6	Tests for local effects after implantation	2016
10993-7	Ethylene oxide sterilization residuals	2008
10993-8	Guidance for reference materials	No longer operative
10993-9	Degradation of materials related to biological testing	2009
10993-10	Tests for irritation and sensitization	2010
10993-11	Tests for systemic toxicity	2017
10993-12	Sample preparation and reference materials	2012
10993-13	Identification and quantification of degradation products from polymers	2010
10993-14	Identification and quantification of degradation products from ceramics	2001
10993-15	Identification and quantification of degradation products from coated and uncoated metals and alloys	2000
10993-16	Toxicokinetic study design for degradation products and leachables	2017
10993-17	Method for establishment of allowable limits for leachable substances using health-based risk assessment	2002
10993-18	Chemical characterization of materials	2005
10993-19	Physicochemical, mechanical, and morphological characterization	2006
10993-20	Principles and methods for immunotoxicology testing of medical devices	2006
10993-22	Guidance on nanomaterials	2017
10993-23	Tests of irritation	Under development
10993-33	Guidance on tests to evaluate genotoxicity – supplement to ISO 10993-3	2015

numerical estimate, including the no observed adverse effect level (NOAEL), lowest observed adverse effect level (LOAEL), and no adverse effect level (NOEL). Generally, the most conservative dose, which is the highest dose that will not cause an adverse effect, is selected for the dose-response assessment; this is designated as the tolerable exposure (TE) limit. A thorough overview of data sources is included in the "Sources of Data" section of this chapter.

- *Step 3: Exposure Assessment*
  The exposure assessment is the quantification of the frequency or duration of human exposure to the agent or substance of interest. Typically, the exposure assessment is calculated by determining the number of devices to which a patient

and/or user will potentially be exposed to in a sequential period or use or over a lifetime. Knowledge on how the device will be used is essential for estimating what the experienced or expected human exposure to an agent or substance will likely be. The exposure assessment can be divided into three main groups: (i) determination of the concentration of chemicals(s) which a patient and/or user will be exposed to; (ii) a thorough grasp of anticipated patterns of usage for the medical device; and (iii) the amount of chemical per human bodyweight (commonly expressed in mg/kg-bw) can be used for comparisons with available toxicological data.

- *Step 4: Risk Characterization*
  Risk characterization is a cumulation of the all three previous steps to estimate the likelihood of an adverse effect to occur following exposure to the medical device of interest. In short, risk characterization is a direct comparison between the TE limit selected (i.e., Step 2 – Dose-Response Assessment) and the estimated exposure level (i.e., Step 3 – Exposure Assessment). As a general rule, when the TE limit is greater than the estimated exposure level, no adverse toxic effects are expected for the agent or substance of interest.

  As described in Step 2 for Dose-Response Assessment, the estimated exposure level is determined based on available toxicological data which include animal studies, published literature, other relevant sources, and in some cases human data. Extrapolation of selected data to the human exposure to a medical device of interest may occur based on such areas as inter- and intraspecies differences, route-to-route differences, and reliability of available (or selected) data. When extrapolation of data is needed, uncertainty factors may be implemented. (A detailed explanation of uncertainty factors is contained in the "Uncertainty Factor" section of this chapter.)

7.4 Sources of Data

Data for calculating the TE limit of exposure to a material can come from multiple resources. The data sources included in this section are not intended to represent the major sources for analysis of available toxicological data but rather some examples of resources that can aid during the dose-response assessment process. The types of resources an individual can use during the dose-response assessment procedures can be divided into topics which include:

- Toxicity Databases
- Biomedical Search Engines
- Published Toxicological Guidance
- Industrial and Environmental Regulations Guides and Other Regulatory Sources

7.4.1 Toxicity Databases

- *Leadscope Toxicity Database* is a portal which is composed of more than 180,000 chemical structures with over 400,000 toxicity study results. Their database takes a novel, interactive approach to browsing and interpreting chemical and biological screening data in an organized fashion through the use of such detail as reference source, dose duration, dose amount, and study results (Leadscope 2019; Roberts et al. 2000).
- *TOXNET (Toxicology Data Network)* is a group of databases which is centered around chemicals and drugs, diseases and environment, environmental health, occupational safety and health, poisoning, risk assessment and regulations, and toxicology. TOXNET can be divided into two main subsections: chemical nomenclature and structure and toxicology data and toxicology literature. The toxicology data and toxicology literature subsections are further divided into subdivisions which are related to data collection and literature sources. While the toxicology data and toxicology literature subsections are most pertinent during the dose-response assessment procedure, all subsections of may provide valuable information for risk assessment (TOXNET 2019; Wexler 2001).

7.4.2 Biomedical Search Engines

- *PubMed* is a free search engine which is contained under the support of the National Center for Biotechnology Information (NCBI). PubMed is composed of more than 29 million citations for biomedical literature from the MEDLINE database of references and abstracts, life journals, and online books. Many of the citations included provide direct links to full-text information from a variety of publisher web sites such as PubMed Central. Many filter options (including publication dates, abstract and summary text, and article types) are available to conduct a more focused data search (Geer et al. 2010; Lu 2011).

7.4.3 Published Toxicological Guidance

- *Sittig's Handbook of Toxic and Hazardous Chemicals and Carcinogens* is a well-known, reliable, and well-accepted reference on more than 2000 of the most common hazardous chemicals. This handbook provides an enormous array of critical information with regard to the use, transportation, and regulation of substances of occupational and environmental concern. The supplied data is extensive yet displayed in a uniform format which allows locating relevant information with ease. Currently, this handbook is in its 7th edition, under the editorial management of Richard Pohanish, and was published in 2017 (Pohanish 2017).

- *Carcinogenically Active Chemicals: A Reference Guide* is a reference manual with a focus on identification of carcinogens, evaluation of implanted materials with carcinogenic potential, and a thorough analysis of some of the most common chemicals termed as "suspected" or "questionable" carcinogens. Presently, only one edition exists for this handbook which was published in 1991 under the editorial direction of Richard J. Lewis.
- *Patty's Toxicology* uses a logical six-volume grouping system to sort comprehensive toxicological data for industrial compounds. Provided data for each compound is established in an easy-to-read format which include available information as CAS number, physical and chemical properties, exposure limits, and biological tolerance values for occupational exposure. Types of industrial compounds discussed in this handbook include aliphatic hydrocarbons, glycols, organic peroxides, and halogenated one-carbon compounds. Further, this book provides information of novel, innovative topics such as nanotechnology, human health effects of nonionizing electromagnetic fields, and occupational chemical carcinogenesis. This handbook is in its 6th edition which was published in 2012 under the editorial guidance of Eula Bingham and Barbara Cohrssen.

7.4.4 Industrial and Environmental Regulations Guides and Other Regulatory Sources

- *Threshold Limit Values (TLVs)* are guidelines (not standards) prepared by the American Conference of Governmental Industrial Hygienists, Inc. (ACGIH) to mainly assist industrial hygienists in formulating decisions with regard to safe exposure values for chemicals in the workplace. The TLVs are time-weighted average (TWA) concentrations of airborne substances that assume workers are repeatedly exposed to a substance during their entire working lifetime which is estimated to occur at 7–8 hour/day, 5 days/week. Generally, ACGIH publishes TLVs on a yearly basis and focus on chemical air concentration.
- *Biological Exposure Indices (BEIs)* are another type of guidelines prepared by the ACGIH. BEIs represent determinant levels as an index of an individual's uptake of a chemical which is considered biological monitoring. Unlike TLVs, which implement air monitoring to determine the potential inhalation exposure of an individual group, the biological monitoring parameters may differ for a variety of reasons. The uptake measures for individuals within a workgroup can differ based on physiological makeup and health status (e.g., body build, diet, and metabolism), occupational exposure (e.g., work-rate intensity, exposure duration, and temperature), and non-occupational exposure factors (e.g., exposure to community/home air pollutants, personal hygiene, and alcohol intake). Similar to TLVs, BEIs apply to approximately an 8 hour workday, 5 days/week and are published by the ACGIH on an annual basis.

- *Workplace Environmental Exposure Levels (WEELs)* are recommended by the American Industrial Hygiene Association (AIHA) and managed by the 501(c)(3) nonprofit organization Toxicology Excellence for Risk Assessment (TERA) under the Occupational Alliance for Risk Science (OARS) initiative. WEELs are airborne concentration limits that provide protection guidance for most workers against the development and onset of adverse health effects induced by occupational chemical exposure. All WEELs are expressed as either TWA concentrations (i.e., the average concentration a worker is exposed to a particular chemical for approximately 7–8 hour workday) or ceiling values (i.e., occupational exposure levels which indicate the airborne concentration of a substance that should not be exceeded in a worker's breathing zone); different time periods are specified based on the properties of the chemical of interest.
- *Short-Term Exposure Limits (STELs)* are the maximum concentration levels of substances which workers may be continually exposed to for a period up to 15 minutes, provided that no more than four excursions per day are permitted, and with at least 60 minutes between each excursion period. The STEL value as defined by the ACGIH is a short-term exposure period to a hazardous chemical in which workers are not expected to suffer from irritation, chronic or irreversible tissue damage, or narcosis of sufficient degree to increase a worker's likelihood of accidental injury, impairment of self-rescue, or a reduction in work efficiency.
- *International Council on Harmonisation of Technical Requirements for Registration of Pharmaceuticals for Human Use (ICH)* is a joint initiative between regulatory authorities and key participants in the pharmaceutical industry to discuss scientific and technical aspects of pharmaceutical product development, guidelines, and registration. The ICH topics are separated into four key categories: quality guidelines, safety guidelines, efficacy guidelines, and multidisciplinary guidelines. While all guidelines in this initiative are extremely informative, the ICH Q3 and ICH M7 guidelines provide much detail in relation to safety aspects of impurities based on essential properties such as characterization as an element, an organic volatile substance, or a mutagenic compound.

7.5 Establishment of the Tolerable Intake (TI) for Chemicals in Medical Devices

The TI is an estimation of the average daily intake of a chemical over a specified time period based on the body mass of an individual (in terms of milligrams per kilogram bodyweight per day) and is considered as a safe level. As stated in the 10993-17 guidance, a review of toxicological data is necessary to determine the conservative and appropriate NOAEL or another relevant toxicological endpoint. This toxicological endpoint is then adjusted for duration of patient exposure, route of exposure, cancer or noncancer endpoints, and irritation potential using UFs (and in some cases SFs).

7.5.1 Calculating the TI

The TI is calculated from the most conservative endpoint (i.e., NOAEL, LOAEL, etc.) from the most relevant route of administration/exposure. The calculated TI must take into account the degree of severity of the chemical of interest and the uncertainty as characterized during the risk assessment. Majority of the time, a modifying factor (MF) is used when calculating the TI to account for route-to-route extrapolation, interspecies differences, and the relevancy and quality of available data for the identified chemical in a medical device.

The formula for calculating TI level, in milligrams per kilogram bodyweight per day, using the MF approach is

$$\mathrm{TI} = \frac{\text{Conservative endpoint (NOAEL, LOAEL, etc.)}}{\mathrm{MF}}$$

where the MF is calculated using the UFs as described in the "Uncertainty Factor" section of this chapter. The MF formula is

$$\mathrm{MF} = \mathrm{UF}_1 \times \mathrm{UF}_2 \times \mathrm{UF}_3$$

7.6 Establishment of the Tolerable Exposure (TE) for Chemicals in Medical Devices

Once the TI has been determined for an identified chemical in a medical device, it is necessary to adjust the TI to determine the exposure level that would be tolerable to an individual. The TE takes into account populations being exposed, body mass of exposed population, intended usage pattern of the device, and the potential exposure of the identified chemical from multiple devices. Together, the TE is calculated based on the estimated TI, body mass of intended patient population using the medical device of interest, and the utilization factor (UTF). The formula for TE is as follows:

$$\mathrm{TE} = \mathrm{TI} \times m_{\mathrm{B}} \times \mathrm{UTF}$$

where

TE is the tolerable exposure level which takes into account the populations being exposed, body mass of exposed population, intended usage pattern of the device, and the potential exposure of the identified chemical from multiple devices

TI is the tolerable intake after modification based upon the device evaluation. Typically, this value is expressed in milligrams per kilogram bodyweight per day (mg/kg bw/day)

m_{B} is the bodyweight specific to the intended patient population

UTF is the utilization factor which takes into account the frequency of the use of the device and the use conjunction with other medical devices that may be expected to contain the same identified chemical as the substance present in the medical device of interest

7.6.1 Determining the Bodyweight of the Intended Patient Population

The intended patient population groups are usually grouped into three categories: adults, children, and neonates. The body masses of 3.5 kg for neonates (<1 year) and 10 kg for children (usually up to 12 years of age) are used in TE calculations. For adults, the bodyweight for TE calculations vary between 50 and 70 kg based on the government or other regulation standard being used. For example, in the USA the standard bodyweight of an adult is accepted at 60 kg; however, in the ISO 10993-17 standards, the accepted bodyweight for an adult is set at 70 kg. Though bodyweight standards have been set for neonates, children, and adults, the bodyweight can vary and may be derived based on special considerations including devices specifically intended for use with uniquely sensitive groups such as adult women.

7.6.2 Calculation of the Utilization Factor (UTF) from Intended Use Pattern

The TI and bodyweight are adjusted using a UTF which takes into account the anticipated use patterns of medical devices. The UTF is a numerical value which factors in the utilization of the device in terms of frequency of use (i.e., the proportional exposure factor [PEF]) and utilization in combination with other medical devices (i.e., the concomitant factor [CEF]). These factors are multiplied together to obtain the UTF as given in the following equation:

$$\mathrm{UTF} = \mathrm{CEF} \times \mathrm{PEF}$$

- *Concomitant exposure factor (CEF)* – this factor takes into account the extent of exposure of the chemical of interest arising from use of multiple devices. The value for CEF ranges normally between 0.2 and 1.0, where the CEF is 0.2 if the UTF is unknown and 1.0 if either fewer devices can release the chemical of interest (i.e., <5% of the devices sold in a calendar year) or less than five devices in any single medical procedure are used. If many medical devices can release the chemical of interest, the CEF can be calculated based on formulas described in the *ISO 10993-17* guidance. For simplicity, this book will not provide detail on

calculating the CEF for multiple devices that can release the same chemical of interest.

- *Proportional exposure factor (PEF)* – this factor takes into consideration the frequency of use of a medical device. The PEF value is calculated as the number of days in the exposure category (n_{exp}) during which the actual number of days of device use (n_{use}) as shown as

$$\text{PEF} = \frac{n_{exp}}{n_{use}}$$

If the number of days of device use is not specified, a reasonable upper limit is used. By default, the PEF is set at 1, which is the most conservative value for this factor.

7.7 Extrapolation Factors

7.7.1 Uncertainty Factor

The UF is a concept which is integrated into health risk assessment for animal-to-human extrapolation when determining acceptable levels for humans based on animal studies. When calculating the TI, the UF is typically utilized. As a common practice, at least three UFs are applied which accounts for human variability, interspecies extrapolation, and the quality and relevance of experimental data. Each of the three UFs used are:

UF_1 (human (interindividual) variability) – this UF value ranges between 1 and 10. As a default, the UF for human (interindividual) variability is normally set at 10. This default value accounts for variations that were not considered if an occupational limit, such as WEEL, was selected. If by chance an occupation limit is selected the human variation is considered as intermediate, and an intermediate UF should be applied. In cases where an animal study was chosen as the point of departure, the default value of 10 was used because animal studies do not account for human variations.

UF_2 (interspecies differences) – this UF value ranges between 1 and 12 and accounts for animal-to-human extrapolation. Some of the most common UF_2 values for interspecies extrapolation are listed in Table 7.2 based on an adaptation of values provided in ICH Q3D Step 4 guidance.

UF_3 (quality and relevance of experimental data) – the UF_3 value ranges between 1 and 100. Of the three UF values, UF_3 varies the most. A number of components influence this variability including the point of departure selected (e.g., TLV, NOAEL, LOAEL, LD_{50}) and available data (e.g., carcinogenicity, genotoxicity, acute toxicity, and repeat dose toxicity studies) for the chemical of interest. Although scientists try to be as similar as possible in their selection of a UF_3 value, variations may arise because of scientific bias.

Table 7.2 UF_1 for extrapolation between species

UF_1 value	Extrapolation from species to human
1	Extrapolation from rats to humans
12	Extrapolation from mice to humans
2	Extrapolation from dogs to humans
2.3	Extrapolation from rabbits to humans
3	Extrapolation from monkey to humans
10	Extrapolation from other animals to humans

7.7.2 Safety Factor (SF)

In addition to UFs, other SFs may be used to account for any other remaining uncertainties, and it clearly allows for the incorporation of scientific judgment based on the overall quality and relevance of available studies to human health risk assessment. Typically, the SF value varies between <1 and 10. Some of the most commonly cited reasons for use of a SF includes remaining uncertainties in relation to exposure scenarios, severity of selected point of departure, uncertainties with regard to bioavailability measure for route of administration, and route-to-route extrapolations.

7.8 Margin of Safety (MOS)

The MOS in principle is a comparison of the TE with the estimated maximum daily human exposure level. To calculate the MOS with regard to a chemical in a medical device, the TE (after correction for UFs and/or MFs) is divided by the maximum potential daily exposure to the chemical identified from medical device. Instead of the maximum potential daily exposure level, the total amount of the chemical in a device can also be used as a (conservative) basis for the calculation with the assumption that all the available chemical is released in a single day.

$$\text{MOS} = \frac{\text{TE}}{\text{Maximum Potential Daily Exposure Level in a Medical Device}}$$

As a general rule, when the TE is greater than the estimated human exposure level or dose, then no adverse human effects are predicted. Additionally, the greater the estimated MOS value, the fewer potential adverse effects are probable for a chemical of interest. A MOS value >100 is most desirable as there is minimum prediction for toxicological effects or death to occur for the chemical of interest; a MOS value >1 is the minimum value which is accepted to pose a human health risk.

7.9 Case Studies

In order to provide a better sense of how to perform a risk assessment on medical devices, two fictious case studies will be discussed.

7.9.1 Metered-Dose Inhaler

A company developed a metered-dose inhaler that delivers a specific amount of medication to alleviate respiratory problems such as asthma and chronic obstructive pulmonary disease (COPD). The company has requested a risk assessment on bis(2-ethylhexyl) phthalate (DEHP) which is suggested to be released from the device into the delivered drug product for adults and children. The worst-case scenario level of DEHP in drug product was estimated at 0.015 mg/day.

Hazard Identification According to the Toxicology Data Network/Hazardous Substance Database (TOXNET/HSDB), the US Environmental Protection Agency Integrated Risk Information System (USEPA IRIS) and the International Agency for Research on Cancer (IARC) have concluded DEHP as a possible human carcinogen (HSDB 2015). Further, the National Institute for Occupational Safety and Health (NIOSH) considered DEHP as a potential occupational carcinogen. The determination of DEHP as a potential human carcinogen was based on a dose-dependent response of liver tumors in rodents fed DEHP (Kluwe 1982). Additionally, dietary exposure to DEHP resulted in benign testicular tumors (Leydig cell tumors) (Voss et al. 2005) and benign pancreatic tumors (acinar cell and islet cell adenoma) in male rats (David 2000; Rao et al. 1990). Inhalation of DEHP in various animal studies reported reproductive effects to male and female rodents (HSDB 2015).

Dose-Response Assessment Female rats which inhaled 25 mg/m^3 DEHP for 6 hours/day for 5 continuous days/week from postnatal days (PND) 22 to 41 and to PND 84 showed advancement of the age for vaginal opening (which is related to indicative of puberty) and the age of the first estrous cycle (Ma et al. 2006). Inhalation of DEHP by male rats at 5 or 25 mg/m^3 for 6 hr/day for 4 and 8 weeks significantly increased the concentration of plasma testosterone and weight of seminal vesicles (Kurahashi et al. 2005). The ACGIH recommends the TLV-TWA level for DEHP at 5 mg/m^3 (NTP 2016).

Because the TLV-TWA level for DEHP at 5 mg/m^3 is the most conservative concentration based on inhalation studies, this will be used to determine the TI for this compound. Since the value is an occupational exposure level that is intended to occur during an approximate 8 hr limit during a 5-day work week, the TLV-TWA limit will be adjusted for continuous exposure (8 hr/24 hr and 5 days/7 days). Before determining the TE for DEHP, the TLV-TWA limit (which is the effect level) must be converted to mg/kg using the FDA standards for adult respiratory volume at

20 m^3/day and the default human bodyweight of 60 kg. Thus, the effect level was determined as

$$\text{Effect Level} = \text{TLV} - \text{TWA value} = \frac{5\dfrac{\text{mg}}{\text{m}^3} \times 20\dfrac{\text{m}^3}{\text{day}} \times \dfrac{8\,\text{hr}}{24\,\text{hr}} \times \dfrac{5\,\text{days}}{7\,\text{days}}}{60\,\text{kg}} = 0.4\frac{\text{mg}}{\text{kg}-\text{day}}$$

Because inherent uncertainties in estimating the potential effects of DEHP on humans when using the metered-dose inhaler are possible, UF was used. The UF and the MF were:

UF_1 – Takes into account variations between humans (1–10); here 10 to be conservative and human variability was not accounted for in this occupational exposure limit.

UF_2 – Extrapolates from data derived from species other than humans (generally 1–10, can be larger if differences are toxicologically significant); here 1 is used because this is a human study and therefore extrapolation is not needed.

UF_3 – Accounts for quality and relevance of experimental data (1–100); here 10 was used because although the USEPA and IARC have suggested this as a potential human carcinogen, a TLV-TWA limit has been established in humans based on the regulatory agency ACGIH.

The MF was therefore determined as

$$\text{Modifying factor}\left(\text{MF}\right) = \text{UF}_1 \times \text{UF}_2 \times \text{UF}_3 = 10 \times 1 \times 10 = 100$$

The TI for DEHP based on the TLV-TWA was calculated as

$$\text{TI} = \frac{\text{Conservative endpoint}}{\text{MF}} = \frac{0.4\dfrac{\text{mg}}{\text{kg}-\text{day}}}{100} = 0.004\frac{\text{mg}}{\text{kg}-\text{day}}$$

Because it is assumed that less than five devices in any single medical procedure will be used (CEF = 1), and the number of days in which the medical device will be used is uncertain (the PEF = 1), the UTF was set at

$$\text{UTF} = \text{CEF} \times \text{PEF} = 1 \times 1 = 1$$

Finally, to calculate the TE it is assumed that the medical device is intended for adults and children with a bodyweight of 60 kg and 10 kg, respectively (m_B). The TE value for DEHP was determined as

$$\text{TE}\left(\text{adults}\right) = \text{TI} \times m_B \times \text{UTF} = 0.004\frac{\text{mg}}{\text{kg}-\text{day}} \times 60\,\text{kg} \times 1 = 0.24\frac{\text{mg}}{\text{day}}$$

$$\mathrm{TE}\left(\text{children}\right) = \mathrm{TI} \times m_{\mathrm{B}} \times \mathrm{UTF} = 0.004 \frac{\text{mg}}{\text{kg} - \text{day}} \times 10\,\text{kg} \times 1 = 0.04 \frac{\text{mg}}{\text{day}}$$

Exposure Assessment The maximum daily dose (MDD) of DEHP present in the metered-dose inhaler is estimated at 0.015 mg/day.

Risk Characterization Comparing the MDD of DEHP (at 0.015 mg/day) to the TE limit of DEHP (at 0.24 mg/day for adults and 0.04 mg/day for children), the margin of safety (MOS) values were determined as

$$\mathrm{MOS}\left(\text{adults}\right) = \frac{0.24 \frac{\text{mg}}{\text{day}}}{0.015 \frac{\text{mg}}{\text{day}}} = 16$$

$$\mathrm{MOS}\left(\text{children}\right) = \frac{0.04 \frac{\text{mg}}{\text{day}}}{0.015 \frac{\text{mg}}{\text{day}}} = 2.6$$

Because the MOS values for both adults and children were above 1, exposure to DEHP in the metered-dose inhaler is considered safe.

7.9.2 Implant Blood Vessel Support Device

A company has created an external scaffold that is intended for adult patients with kidney failure in need of dialysis requiring vascular access through the arteriovenous (AV) fistula. This implanted blood vessel external support device is placed over the fistula, thereby reducing the tension in the vein. The metal, nickel (Ni), was detected in the support device at 2.71 μg/g when evaluated by inductively coupled plasma optical emission spectrometry (ICP-OES). The extraction procedure was performed in purified water at 70 °C for 24 hours on two test articles which together weighed 3.1 g. The amount of Ni in the device was therefore determined as

$$\text{Amount of nickel per tested device} = \frac{2.71\,\mu g\,\text{of nickel}}{\text{g test article}} \times \frac{3.1\,\text{g test article}}{2\,\text{devices}}$$

$$\text{Amount of nickel per tested device} = 4.2 \frac{\mu g}{\text{device}} = 0.0042 \frac{\text{mg}}{\text{device}}$$

Hazard Identification Nickel particulate (e.g., elemental and subsulfide) has been associated with nasal and lung cancer after workplace inhalation exposures. The International Agency for Research on Cancer (IARC) classifies nickel and nickel compounds as having sufficient evidence of cancer in humans (Group 1); however, IARC notes that the evaluation applies to the group in general and not necessarily to all compounds in the group, in particular, zero valence metals. Serious reproductive or developmental effects related to nickel exposure were not readily observed in rodent studies upon acute and chronic exposure to various nickel compounds. However, female workers exposed to 0.08–0.196 mg/m^3 nickel sulfate on the job experienced more spontaneous abortions and gave birth to more infants with structural malformations compared to non-exposed workers, but a causative relationship was not established. Chromosomal aberrations have been noted in lymphocytes in occupationally exposed individuals (ATSDR 2005).

Dose-Response Assessment The EMA has published a guidance for allowable levels of heavy metals in chronically administered pharmaceuticals (EMA 2008). For nickel (a Class 1C metal in the guidance), the acceptable limit on daily exposure is 30 μg/day by the parenteral route of administration (0.6 mg Ni/kg/day in a 50 kg person and based on a bioavailability of 10%). The *ICH Q3D(R1) Elemental Impurities* has estimated a parenteral permissible daily exposure (PDE) limit for nickel at 22 μg/day (ICHQ3D(R1) 2019).

The parenteral dose for Ni of 22 μg/day (0.022 mg/day) under ICH Q3D guidance was used as the TE level for this element. This limit was chosen because it was the most conservative dose for the most relevant route of administration. Because this is a well-established dose set by a regulatory agency, no further uncertainty or safety factors need to be implemented.

$$\mathrm{TE} = 0.022\,\frac{\mathrm{mg}}{\mathrm{day}}$$

Exposure Assessment The maximum daily dose (MDD) of nickel present in the implant blood vessel support device is estimated at 0.0042 mg/day.

Risk Characterization Comparing the MDD of Ni (at 0.0042 mg/day) to the TE limit of Ni (at 0.022 mg/day for adults), the margin of safety (MOS) value was determined as

$$\mathrm{MOS} = \frac{0.022\,\frac{\mathrm{mg}}{\mathrm{day}}}{0.0042\,\frac{\mathrm{mg}}{\mathrm{day}}} = 5$$

The TE for Ni is approximately fivefold higher for adults when compared to the amount detected in the implant blood vessel support device during an extraction procedure using ICP-OES. Therefore, adults exposed to Ni in this support device should not result in adverse toxic effects.

References

ATSDR. (2005). *Toxicological profile for nickel*. Agency for Toxic Substances and Disease Registry, Atlanta.

CDRH. (1992). *Regulatroy Requirements for Medical Devices: A Workshop Manual*. (FDA 92-4165). Center for Device and Radiological Health: HHS Publication.

David, R. M. (2000). Exposure to phthalate esters. *Environmental Health Perspectives, 108*(10), A440–A440.

Ecobichon, D. J. (1997). *The basis of toxicity testing*. CRC press, Boca Raton, FL.

EMA. (2008). *Guideline on the specification limits for residues of metal catalysts or metal reagents*. European Medicines Agency, London.

Gad, S. C., & Gad-McDonald, S. (2015). *Biomaterials, medical devices, and combination products: Biocompatibility testing and safety assessment*. CRC Press, Boca Raton, FL.

Gad, S. C., & Schuh, J. C. L. (2018). Toxicologic pathology forum opinion paper: Considerations for Toxicologic pathologists evaluating the safety of biomaterials and finished medical devices. *Toxicologic Pathology, 46*(4), 366–371. https://doi.org/10.1177/0192623318768719.

Geer, L. Y., Marchler-Bauer, A., Geer, R. C., Han, L., He, J., He, S., et al. (2010). The NCBI BioSystems database. *Nucleic Acids Research, 38*(Database issue), D492–D496. https://doi.org/10.1093/nar/gkp858.

HSDB. (2015). Hazardous substances data bank (Online Database). Available from National Institutes of Health, Health & Human Services TOXNET Retrieved June 29, 2019, from U.S. National Library of Medicine.

ICHQ3D(R1). (2019). *Guideline for Elemental Impurities*. ICH Harmonised Guideline.

ISO10993-1. (2018). *Biological evaluation of medical devices -- Part 1: Evaluation and testing within a risk management process*. Arlington: Association for the Advancement of Medical Instrumentation.

ISO14971-1. (2019). *Medical devices—Application of risk management to medical devices*. Arlington: Association for the Advancement of Medical Instrumentation.

Kluwe, W. M. (1982). Overview of phthalate ester pharmacokinetics in mammalian species. *Environmental Health Perspectives, 45*, 3–9.

Kurahashi, N., Kondo, T., Omura, M., Umemura, T., Ma, M., & Kishi, R. (2005). The effects of subacute inhalation of di (2-ethylhexyl) phthalate (DEHP) on the testes of prepubertal Wistar rats. *Journal of Occupational Health, 47*(5), 437–444. https://doi.org/10.1539/joh.47.437.

Leadscope. (2019). Leadscope Enterprise, Inc.

Lu, Z. (2011). PubMed and beyond: a survey of web tools for searching biomedical literature. *Database, 2011*.

Ma, M., Kondo, T., Ban, S., Umemura, T., Kurahashi, N., Takeda, M., & Kishi, R. (2006). Exposure of prepubertal female rats to inhaled di(2-ethylhexyl)phthalate affects the onset of puberty and postpubertal reproductive functions. *Toxicological Sciences, 93*(1), 164–171. https://doi.org/10.1093/toxsci/kfl036.

NTP. (2016). *Report on carcinogens* (14th ed.). National Toxicology Program, Research Triangle Park.

Pohanish, R. P. (2017). *Sittig's handbook of toxic and hazardous chemicals and carcinogens*. William Andrew, Kidlington, Oxford, United Kingdom; Cambridge, MA, United States.

Rao, M., Yeldandi, A. V., & Subbarao, V. (1990). Quantitative analysis of hepatocellular lesions induced by di (2-ethylhexyl) phthalate in F-344 rats. *Journal of Toxicology and Environmental Health, Part A Current Issues, 30*(2), 85–89.

Roberts, G., Myatt, G. J., Johnson, W. P., Cross, K. P., & Blower, P. E. (2000). LeadScope: Software for exploring large sets of screening data. *Journal of Chemical Information and Computer Sciences, 40*(6), 1302–1314.

Stark, N. (1997). *Case studies: using the world literature to reduce biocompatability testing.* Paper presented at the Proceedings of the Medical Design and Manufacturing East 97 Conference and Exposition.

TOXNET. (2019). US National Library of Medicine/Toxicology Data Network.

USFDA. (2017). Learn if a medical device has been cleared by FDA for marketing.

Voss, C., Zerban, H., Bannasch, P., & Berger, M. R. (2005). Lifelong exposure to di-(2-ethylhexyl)-phthalate induces tumors in liver and testes of Sprague–Dawley rats. *Toxicology, 206*(3), 359–371.

Wexler, P. (2001). TOXNET: An evolving web resource for toxicology and environmental health information. *Toxicology, 157*(1–2), 3–10.

White, R., Cammack, J., Eyre, R., & Wilson, D. (1999). Medical device risk assessment paradigm: Use and case history. *International Journal of Toxicology, 18*(4), 269–274.

Chapter 8
An Introduction to (Q)SAR with Respect to Regulatory Submissions

Jerry L. Bettis Jr.

Abstract (Q)SAR ((quantitative or qualitative) structure-activity relationships) methodologies are used to predict physical and biological properties of small molecules. These methods are used to support pharmaceutical research and regulatory submissions. Primarily, (Q)SARs are used to predict the activity of untested chemicals based on structurally related compounds with known activity. The term (Q)SAR is often used to refer to predictive models, especially computer-based models; however, in reality (Q)SAR encompasses a wide variety of computerized (i.e., in silico) and non-computerized tools and approaches. As a tool, (Q)SAR is accepted internationally for predicting mutagenicity; however, its applicability for predicting additional endpoints (e.g., skin sensitization or hepatotoxicity) is still an active debate topic, particularly debates about the acceptability of the (Q)SAR models for additional endpoints and how they are either explained or interpreted. After discussing the basics of (Q)SAR, we relate (Q)SAR methodologies to inexpensive and practical applications.

Keywords (Q)SAR · ICH M7 · Mutagenicity · Structural alerts · DEREK · Leadscope · Cramer classifications

8.1 Introduction

Quantitative structure-activity relationships (QSARs) are techniques that attempt to identify correlations between chemical structure and associated activity (e.g., biological), primarily to predict the activity of untested chemicals based on structurally related compounds with known activity. The parentheses around "Q" in (Q)SAR indicates that the term refers to both qualitative predictive tools (i.e., structure-activity relationships (SARs)) and quantitative predictive methods (i.e., quantitative structure-activity relationships (QSARs)). In practice, the term (Q)SAR is often

J. L. Bettis Jr., PhD (✉)
Gad Consulting Services, Raleigh, NC, USA
e-mail: jerry.bettis@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_8

used to refer to predictive models, especially computer-based models; however, in reality (Q)SAR is encompassing a wide variety of computerized (i.e., in silico) and non-computerized tools and approaches.

Using a ligand-based approach, (Q)SARs attempt to quantify the relationship between an aspect of chemical structure (i.e., a molecular fragment) and an activity or property (e.g., mutagenicity) imparted by that structure. The molecular structure of molecules is often described by descriptors (e.g., electrophilicity, hydrogen bonding, molecular fragments) or physical-chemical properties (e.g., LogP) which are then used to develop a mathematical correlation between a group of structures and a defined activity or endpoint (e.g., carcinogenicity, sensitization, HERG channel inhibition, etc.). The mathematical correlations usually take the form of statistical algorithms developed through a variety of techniques (e.g., univariate regression, multiple linear regression, partial least squares analysis).

Although the term (Q)SAR is often used to refer to predictive models, especially computer-based models, it should be noted that (Q)SAR is inclusive of a wide variety of tools and approaches such as analogs, chemical categories, and computer-based or non-computer-based SAR/QSAR models.

8.2 Key Principle of (Q)SAR

Crum-Brown and Fraser published Equation 1.1 in 1868, which is considered to be the first formulation of a QSAR: the "physiological activity" (Φ) was expressed as a function of the chemical structure C (Crum-Brown and Fraser 1868):

$$\Phi = f(C) \tag{8.1}$$

However, given time and experience, (Q)SAR's most general mathematical form is:

$$\Phi = f(\text{physicochemical properties and / or structural properties}) \tag{8.2}$$

Based on Eqs. 8.1 and 8.2, three key components may be inferred that are required to develop a viable QSAR model:

- Some measure of the activity (in this case toxicity) for a group of chemicals in a biological or environmental system – toxicological endpoint
- A description of the physicochemical properties and/or structure for this group of chemicals (i.e., molecular descriptors)
- A form of statistical relationship to link activity and descriptors

These three key components are pivotal in toxicological risk assessment of mutagenic compounds using (Q)SAR.

Expert models generally follow a process of identifying active and inactive chemicals based on the presence or absence of specific structural features. For example, expert systems (e.g., DEREK) use "rules" (i.e., decision logic) to categorize the potential activity of untested chemicals based on expert knowledge gathered

from the analysis of experiments data. Statistically based (Q)SAR models (e.g., Leadscope Model Applier), which rely on a statistical association between structure and activity, may be developed objectively with little mechanism of action expertise and are useful for detecting structural features/molecular descriptors predictive of toxicity (EPA 2012).

8.3 Molecular Descriptors Used in (Q)SAR

8.3.1 Physicochemical Properties

Physicochemical molecular descriptors may be defined as a numerical representation of chemical information encoded within a molecular structure via mathematical procedures (Jhanwarb et al. 2011). Physicochemical molecular descriptors are used extensively as predictors for toxicity in statistical-based models.

The three major types of physicochemical molecular descriptors commonly employed are (1) hydrophobic (partition coefficient, distribution coefficient, solubility parameter, etc.), (2) electronic (Hammett constant, ionization constant, etc.), and (3) steric (Taft's steric parameter, molar volume, van der Waals, etc.) (Jawarkar and Game 2018).

Physicochemical properties tend to describe fundamental molecular effects (e.g., hydrophobicity). As a general rule, these properties are considered related to a molecule's mechanism of action and, therefore, are less susceptible to spurious correlation. A spurious correlation is a mathematical relationship in which two or more events or variables are associated but not causally related.

Although determining the physicochemical properties of a pure chemical substance is a relatively simple and straightforward task, the computation of physicochemical properties is relatively cheap and advantageous when investigating untested molecular entities with little to no experimental data.

Molecular orbital theory is an interesting candidate for use as a molecular descriptor. The theory has proven applicable and appropriate for predicting metabolism, persistence, or biochemical reactivity (Braga and Andrade 2012). The rapid increase in the speed of computers has enabled the rapid calculation of numerous atomic and molecular orbital descriptors, such as charges, dipole moment, energy levels, etc.

Among the physicochemical factors that modulate and may hinder the potential biological activity of the chemicals with SAs are (1) molecules with very high molecular weights and size because they have little chance of being absorbed in significant amounts; (2) physical state, which influences the capability of the compounds to reach critical targets; (3) solubility in consideration of highly hydrophilic compounds being poorly absorbed and are readily excreted; and (4) the chemical reactivity of highly reactive compounds which may not be carcinogenic because they spontaneously hydrolyze or polymerize or react with noncritical cellular constituents before interacting with critical cellular targets (Benigni and Bossa 2006).

To become prolific in the field of molecular descriptors requires a diverse skillset because it is inherently interdisciplinary and encompasses many theories and

disciplines. For example, those successful in the field of molecular descriptions must possess an in-depth knowledge of algebra, graph theory, information theory, computational chemistry, physical chemistry, quantum chemistry, and organic chemistry. But those skills are insufficient to fully develop a viable (Q)SAR model. Sophisticated software and good programming skills are also required.

8.3.2 *Structural Alerts*

Structural alerts (SAs), otherwise known as "expert rules," are substructural features (i.e., fragments of a molecular structure) that are associated with a particularly adverse outcome (i.e., toxicity). SAs are widely accepted in chemical toxicology and regulatory decision support as a simple and transparent means to flag potential chemical hazards or group compounds into categories. Concerning (Q)SAR methodologies, SAs are the primary predictors of toxicity in expert/rule-based models (Alves et al. 2016).

SAs for predicting genotoxicity were introduced in 1985 by John Ashby (Fig. 8.1) (1). Ashby depicted the SAs in the form of a hypothetical chemical, namely, poly-carcinogen which contains most of the known SAs. Subsequently, Benigni et al. (2008) summarized SAs as follows: "The Structural Alerts are

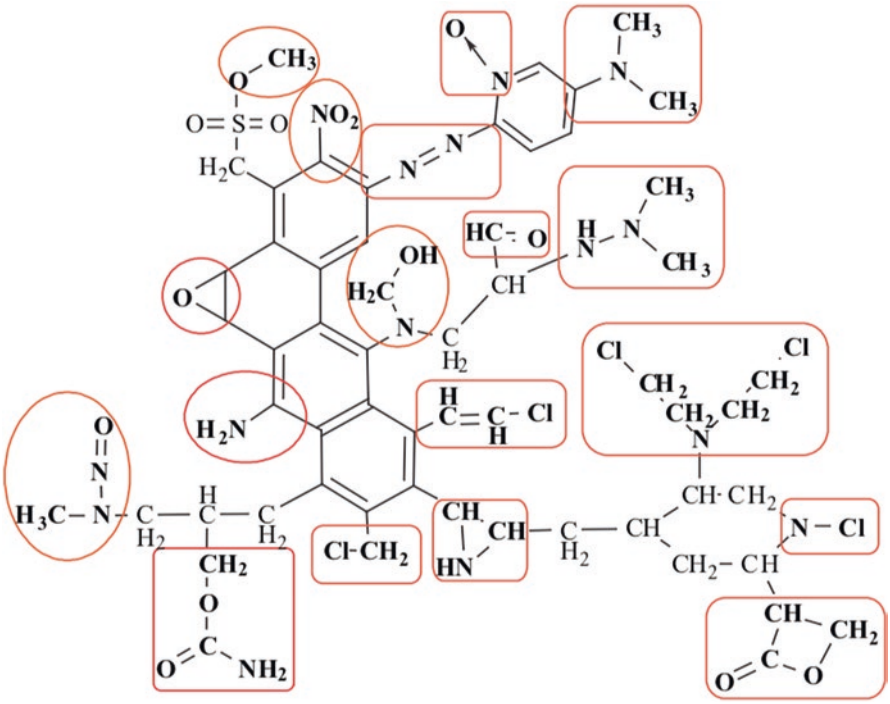


Fig. 8.1 The figure displays the Ashby's poly-carcinogen SAs

Table 8.1 The SAs represented in Fig. 8.1

Ashby's structural alerts	
Alkyl esters of either phosphonic or sulfonic acids	β-Haloethyl mustards; N-chloroamines
Aromatic nitro groups	Propiolactones and propiosultones
Aromatic azo groups (because of the possible reduction to aromatic amines)	Aromatic and aliphatic aziridinyl derivatives
Aromatic ring N-oxides	Aromatic and aliphatic substituted primary alkyl halides
Aromatic mono- and dialkylamino groups	Derivatives of urethane (carbamates)
Alkyl hydrazines	Alkyl N-nitrosoamines
Alkyl aldehydes	Aromatic amines (including their N-hydroxy derivatives and the derived esters)
N-methylol derivatives; monolakenes	Aliphatic and aromatic epoxides

molecular substructures or reactive groups that are related to the carcinogenic and mutagenic properties of the chemicals, and represent a sort of 'codification' of a long series of studies aimed at highlighting the mechanisms of action of the mutagenic and carcinogenic chemicals." Thus, SAs for genotoxicity aid in the categorization of potential carcinogens and/or mutagens and also provide insight into the underlying mechanisms of genotoxicity (Table 8.1).

The key point of this section is that the research in the field of modeling the structural properties of mutagens and carcinogens is highly interdisciplinary work. In particular, mechanistic research based on experimental systems together with human ingenuity in the interpretation of the results has provided the essential basis for the identification of the SAs that characterize the mutagens and carcinogens.

8.4 Endpoints Encoded in (Q)SARs

In general, an endpoint is defined as the recorded observation coming from a biological effect (e.g., NOAEL, NOAEC, LD50, or LC50) determined from an in vitro or an in vivo assay (JRC 2016). A large number of endpoints (e.g., lethality, carcinogenicity, immunological responses, organ effects, developmental and reproductive effects) are used in regulatory assessments of chemicals. State-of-the-art (Q)SAR methodologies develop endpoint-specific models for evaluating individual toxic endpoints (NAFTA 2011). Endpoint-specific (Q)SAR models are defined by molecular descriptors. Numerous biological activities have been successfully modeled using this approach. In general, endpoints may be divided into two major groups, namely, human health endpoints and environmental toxicity endpoints (Piir et al. 2018). If it isn't evident already, we are focusing solely on human health endpoints.

Two very important human health endpoints are chemical carcinogenicity and mutagenicity. Carcinogens can be genotoxic, which interact directly with DNA and are thought to work by inducing mutations, or epigenetic, which act through mechanisms that do not involve direct DNA damage; however, no unifying theory exists

for their mode of action. However, mutagens provoke heritable changes to the genetic material. The modeling of chemical carcinogenicity and mutagenicity is a very important goal in toxicology because of the huge impact they have on the quality of life and because of the enormous investment in time, money, and animal lives needed to test chemicals adequately (Plošnik et al. 2016). Currently, the only (Q)SAR endpoints with enough historical data to be recognized by the ICH are carcinogenicity and mutagenicity.

The chemical metabolism and biotransformation of chemicals within biological organisms are also endpoints receiving a lot of attention in the field of drug development. The driving force behind this attention is due to the pharmaceutical industry shifting toward more intelligent and time-efficient screenings to avoid the failure of candidate drug substances. Thus, pharmacokinetic parameters (e.g., absorption, distribution, metabolism, and excretion) are becoming prioritized toxicological endpoints. The pharmacokinetic properties must be optimized such that the drug will be readily absorbed, transported to the appropriate site, and eliminated from the body promptly. Remember this, approximately 40% of drug candidates fail preclinical development tests due to unacceptable pharmacokinetics (Caldwell 2000).

Undeniably, (Q)SAR models that reliably predict toxicity endpoints of untested compounds are invaluable tools; however, one of the major problems in (Q)SAR modeling is the availability of high-quality experimental data for building the models. If the untested compound contains structural feature not included in the experiment data, then the model may not correctly predict the compound's toxicity due to gaps in the experiment data used to create the model. Additionally, the input data must be both accurate and precise to develop a meaningful model.

Remember, a (Q)SAR model is as valid as the experimental data that led to its development. Spoken plainly, if the experimental data used to build the model is small and of low quality (i.e., bad data), then the model itself may also be of low quality; and thus, the predicted toxicity of an endpoint may be unreliable (i.e., bad model).

8.5 Regulatory Submission Requirements for Compounds Evaluated in Silico

The ICH M7 (2017) guideline entitled "assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk" states that expert rule-based and statistical-based models ((Q)SAR models) can be used to predict the outcome of a bacterial mutagenicity assay to support hazard assessment. In certain cases, these predictions can avoid having to test impurities or degradants for bacterial mutagenicity.

Toxicologists using state-of-the-art (Q)SAR methodologies receive predictions and supporting information from several (Q)SAR models relevant to the particular compound and the particular toxicological endpoint. The reason for employing several QSAR models within a software program is that each model uses a different

correlation algorithm; together, these algorithms can increase and cross-check the reliability of the prediction. Significantly, it indicates difficulties in prediction, and consequently lower levels of reliability are revealed to the investigator by a lack of consistency across the values from each model (Benfenati et al. 2013).

8.6 Widely Accepted Expert-Based and Statistical-Based (Q)SAR Models

Given that the only in silico endpoints recognized by the ICH are carcinogenicity and carcinogenicity, it comes as no surprise that the ICH M7 (2017) guideline stipulates that two complementary (Q)SAR prediction methodologies must be applied to untested compounds to satisfy regulatory submission requirements (e.g., US FDA, EMA, and Health Canada). The first methodology should be expert rule-based (e.g., DEREK), and the second methodology should be statistical-based (e.g., Leadscope Model Applier (LSMA)). Both DEREK and Leadscope Model Applier employ the general validation principles set forth by the Organisation for Economic Co-operation and Development (OECD) (LSMA 2016; OECD 2007).

8.6.1 Statistical Models

Statistical methods are the mathematical foundation for the development of QSAR models. The application of multivariate analysis, data description, classification, and regression modeling is combined with the ultimate goal of interpretation and prediction of untested compounds (Cherkasov et al. 2014).

Leadscope Model Applier (LSMA) is a statistical (Q)SAR methodology built using the Leadscope Predictive Data Miner software and training datasets compiled at the US Food and Drug Administration (FDA) by the Division of Applied Regulatory Science (DARS). The LSMA statistical models are implemented with molecular descriptors that include structural features and seven physicochemical properties. The structural features include a set selected from Leadscope's 27,000 pre-defined structural features, predictive scaffolds (i.e., toxicophores with larger structural features that show association or lack of association to the toxicity endpoint), and SAs identified from the literature or through an analysis for larger databases. The seven calculated physicochemical properties used are parent molecular weight, aLogP, polar surface area, hydrogen bond acceptors, hydrogen bond donors, number of rotational bonds, and Lipinski score (rule violation). The models are built using molecular descriptors (i.e., substructural features and properties) also described as x- or independent variables. The models encode the relationship between these descriptors and the toxicity endpoint, such as the results of the bacterial mutagenesis assay (i.e., y-variable or response variable). The modeling technique used to generate these models is referred to as partial logistic regression (LSMA 2016).

When a prediction is made on an untested compound, the same structural features and properties in the model are calculated for the test compounds (i.e., dataset). These descriptors are then used with the models to calculate a probability of a positive result, as long as the untested compound is within the applicability domain of the model (i.e., the majority of substructural features of the untested compound are known to the model).

8.6.2 Expert Models

Expert systems are composed of structural rules derived from specific toxicological mechanisms or plausible modes of action of chemical agents in combination with pattern recognition routines to identify substructures associated with specific toxic effects (Sanderson and Earnshaw 1991).

DEREK is an expert/rule-based (Q)SAR methodology that indicates whether a specific toxic response may occur; it does not provide a quantitative (e.g., statistical) estimate of the prediction. DEREK has several rule bases, consisting of descriptions of molecular substructures (SAs) that have been associated with toxic endpoints (e.g., mutagenicity, carcinogenicity, or skin irritation). Since substructures can exist in a variety of molecular contexts, the rules are not compound-specific, but rather serve as broad generalizations concerning the chemical structure (e.g., alkylating agent, acid, or halogen-containing molecule). The development of the rules is a continuous process that is monitored by the DEREK Users Group (Cariello et al. 2002).

Generally, when a structure is evaluated by DEREK, it is standardized and then compared to the certified Lhasa knowledge base, and a toxicity prediction is generated. The prediction includes an overall conclusion about the likelihood of toxicity of a structure and detailed reasoning information for the likelihood. The prediction is generated by applying expert rules in toxicology to the data returned from the knowledge base (Cariello et al. 2002). The structure standardization in DEREK uses a set of transform rules including, but not limited to, aromaticity perception, transforming pentavalent nitrogens, and removing specific stereochemistry. The standardization aims to interpret structures more accurately, in order to optimize predictions (Lhasa 2019).

8.6.3 Cramer's Classification

The Cramer decision tree is one the best-known approaches used to estimate the thresholds of toxicological concern (TTCs) for an untested compound based on its chemical structure (Cramer et al. 1978; Munro et al. 1996; Kroes et al. 2004). Initially, the TTC approach was developed to protect humans against carcinogenicity and systemic toxicity. The Cramer decision tree (aka Cramer's rules) is intended to help rank and/or exclude compounds from formal testing based on their predicted toxicities.

Table 8.2 Descriptions of Cramer classes

Cramer class	Description
Class I	Substances with simple chemical structures and for which efficient modes of metabolism exist, suggesting a low order of oral toxicity
Class II	Substances which possess structures that are less innocuous than Class I substances, but do not contain structural features suggestive of toxicity like those substances in Class III
Class III	Substances with chemical structures that permit no strong initial presumption of safety or may even suggest significant toxicity or have reactive functional groups

Computerized implementations of the Cramer decision tree are not representative of traditional (Q)SAR systems because they do not predict the presence or absence of potential bioactivities for specific endpoints.

Cramer decision tree categorizes a compound's potential toxicity based on the presence or absence of structural alerts/expert rules (Cramer et al. 1978; Curios-IT 2009). Cramer decision tree consists of 33 "yes" or "no" questions (i.e., rules) (Cramer et al. 1978). The answers to these questions lead to other questions until the decision tree categorizes the compound into one of the three distinct classes (see Table 8.2).

Cramer first tested the decision tree against 81 chemicals (e.g., pharmaceuticals, excipients, and pesticides,) with oral no-observed-effect level (NOEL) reported in the literature (Cramer et al. 1978). Although NOELs between the three classes overlapped, in general, NOELs of Class I compounds were higher than those of Class II, and NOELs of Class II compounds were higher than those of Class III compounds. Most importantly, the decision tree did not underestimate the toxicity of the compounds tested when compared with oral toxicity data obtained from chronic studies (Lapenna and Worth 2011).

Subsequently, the Cramer decision tree was modified by Munro et al. (1996) to estimate human TTCs for additional toxicity endpoints. The dataset evaluated by Munro et al. contained more than 613 organic chemicals associated with 2941 NOEL values. The non-cancer NOEL values were obtained from sub-chronic, chronic, reproductive, and developmental toxicity studies conducted in rodents and rabbits. Munro et al. classified the dataset compounds based on the Cramer scheme and estimated human TTC limits by taking the lower fifth percentile value of the distribution of NOELs for each three Cramer classes, multiplying by 60 to convert the values expressed as mg/kg/day into mg/day, and then dividing by a factor of 100 to ensure a margin of safety. Using these methods, Munro et al. proposed oral TTC limits for Class I, II, and III compounds, respectively (Table 8.3).

The TTC levels proposed by Munro et al. are widely used in the food safety industry. Additionally, three independent non-food committees (SCCP, SCHER, and SCHENIHR) evaluated the potential applications of the TTC and concluded that the TTC approach is scientifically acceptable for human health risk assessment.

Thus, to promote a consistent application of the Cramer decision tree as defined by Cramer et al. (1978) with TTC limits defined by Munro et al. (1996), the JRC commissioned the development of Toxtree, an open-source software, for use by professionals. In principle, Toxtree's implementation of the Cramer decision tree may be applied to organic molecules, organic salts, organometallics, oligomers, and polymers.

Table 8.3 TTC values of Cramer classes

Cramer class	TTC values (μg/day)
Class I	1800
Class II	540
Class III	90

8.6.4 (Q)SAR Domains of Applicability

The importance of considering the domain of applicability of the (Q)SAR tool has already been mentioned concerning the applicability of a (Q)SAR tool to an untested compound. Evaluating the relationship of untested compound to the domain of applicability of the (Q)SAR tool essentially involves determining whether the compound lies within the domain of applicability or outside of it. As implied, predictions for untested compounds outside of the domain of applicability of a (Q)SAR tool are not necessarily inaccurate but are generally considered less reliable than predictions for compounds falling with the domain of applicability (NAFTA 2011).

The domain of applicability may be defined in different ways (e.g., descriptor, structural fragment, mechanistic, and metabolic domains). To determine if an untested compound is within the domain of a molecular descriptor-based (Q)SAR model, the untested compound's molecular descriptor values are compared to the range of values for the chemicals in the training set (i.e., experimental data) encoded in the model. Expert/ruled-based domain analyses typically ensure that the untested compound doesn't contain substructural features that are not present in the training set of the model. For the mechanism of action or metabolic domain, the key question is whether the untested compound is likely to act via the same mode/mechanism of action and/or be metabolized in the same manner as other chemicals for which the (Q)SAR model is applicable (NAFTA 2011).

OECD has noted that because domains of applicability, which are programmer-defined, vary from methodology to methodology, a prediction for an untested compound that is within the domain of applicability of a (Q)SAR model based on structural and physicochemical parameters may still not be reliable if it has a unique mechanism of action not covered by the mechanistic domain(s) of applicability of the (Q)SAR model (OECD 2007).

The domain of applicability and the reliability of a model's prediction may also depend on the age and/or training set (i.e., empirical data encoded in the model). For example, old global type (Q)SAR model may generate a negative prediction for an untested compound because its training set is populated with a limited number of chemicals that contain the key substructural features present in the untested compound and that all tested negative in historical empirical studies. However, the training set of a newer model will likely be validated against modern in vitro or in vivo studies. This new training set will be segregated into groups according to mechanism of action. If the new training set contains a larger number of compounds, many of which have positive empirical test results, from the same chemical class as the

untested compound, then the model may generate a positive prediction that may be considered more reliable even though the untested compound falls within the domains of applicability of both models.

In short, for the purpose of regulatory submission, we recommend using of the most up-to-date versions of models and training sets. This point may be particularly important when combining information from multiple prediction methodologies.

8.7 How to Interpret (Q)SAR Prediction Results

8.7.1 DEREK

All the rules in DEREK are based either on hypotheses relating to mechanisms of action of a chemical class or on observed empirical relationships (Sanderson and Earnshaw 1991). Information used in the development of rules includes published data and suggestions from toxicological experts in industry, regulatory bodies, and academia.

The toxicity predictions are the result of two processes:

1. The model attempts to identify SAs in the knowledge base that are also present in the untested compound.
2. The reasoning engine then assesses the likelihood of a structure being toxic. There are nine levels of confidence: certain, probable, plausible, equivocal, doubted, improbable, impossible, open, and contradicted (Table 8.4).

Genotoxicity alerts in DEREK include alerts for mutagenicity (in bacteria and mammals) and alerts for chromosome damage based on the in vitro chromosomal aberration assay and including effects that do not involve direct DNA damage (Lhasa 2019).

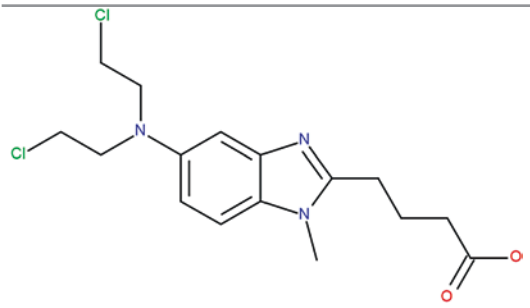
To demonstrate the capabilities of DEREK, the mutagenic potential of bendamustine (CAS N°. 3543-75-7) was evaluated and predicted to be PLAUSIBLE for mutagenicity in vitro (Table 8.5) due primarily to the presence of two (2) SAs, namely, an alkylating agent and nitrogen mustard.

The alkylating agent triggering the SA is alkyl chloride (Fig. 8.2a). Alkylating agents are reactive compounds that replace hydrogen with an alkyl group. DNA, RNA, and cellular proteins contain many groups that are susceptible to these alterations. The carbon attached to the halide in an alkyl halide group is slightly positive, making it easily attracted to and reactive with DNA. Alkylation of DNA can cause abnormal nucleotide sequences, breaking of DNA strands, and DNA cross-linking, resulting in interruption of DNA replication. Consequently, many compounds containing these groups are mutagenic in the Ames test, especially in *Salmonella* strains TA100 and TA1535. In general, bromo and iodo alkyl groups are equally mutagenic and more mutagenic than chloro alkyl groups, with mutagenicity decreasing the longer the alkyl chain. Nitrogen mustards (Fig. 8.2b) are alkylating agents that can

Table 8.4 Terminology used in DEREK reports

Terminology	Description
Certain	There is proof that the proposition is true
Probable	There is at least one strong argument that the proposition is true, and there are no arguments against it
Plausible	The weight of evidence supports the proposition
Equivocal	There is an equal weight for and against the proposition
Doubted	The weight of evidence opposes the proposition
Improbable	There is at least one strong argument that the proposition is false, and there are no arguments that it is true
Impossible	There is proof that the proposition is false
Open	There is no evidence that supports or opposes the proposition
Contradicted	There is proof that the proposition is both true and false

Table 8.5 Summary of DEREK (Q)SAR results for bendamustine (CAS N°. 3543-75-7)

Structures	DEREK prediction
 Bendamustine CAS N°. 3543-75-7	***Mutagenicity* in vitro *is PLAUSIBLE in:*** Bacterium *S. typhimurium* *E. coli*

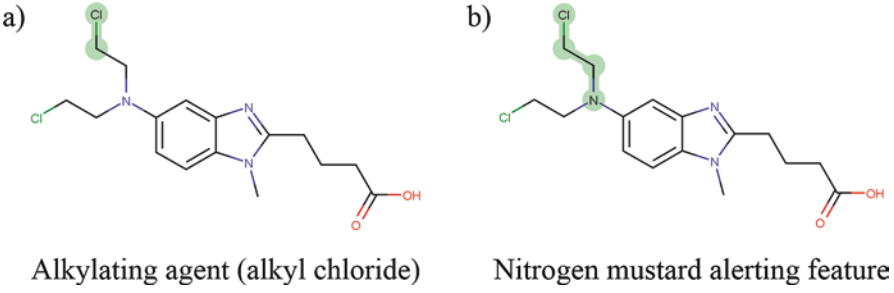


Fig. 8.2 The potentially mutagenic features triggering SAs in bendamustine. Although not highlighted, one (1) more instance of each SA is present in the molecule. (**a**) Alkylating agent (alkyl chloride). (**b**) Nitrogen mustard alerting feature

induce interstrand cross-links, and DNA adducts especially at the N7 guanine position. Nitrogen mustards are genotoxic in most assays including Ames, in vitro chromosome, and in vivo micronucleus.

8.7.2 Leadscope Model Applier (LSMA)

The statistical QSAR model encoded in LSMA calculates a probability for mutagenic activity. Mutagenicity prediction values ≤0.4 and ≥0.6 correspond to NEGATIVE and POSITIVE mutagenicity predictions, respectively. Mutagenicity prediction values between 0.4 and 0.6 correspond to intermediate mutagenicity predictions. These are the cutoffs presently used by the US FDA in evaluating M7 QSAR submissions using Leadscope. Alternatively, an out-of-domain call is generated if the query compound was not in the model's domain of applicability (LSMA (2016).

The full statistical (Q)SAR report from Leadscope Model Applier includes a picture of the chemical structure(s), a description of the models that were used, a table of the results, an explanation of how the results were calculated (including a presentation of the structural feature responsible for the positive or negative prediction), and a listing of the structural analogs from the training set.

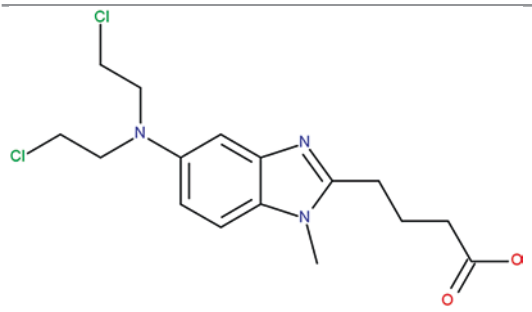
To illustrate some of the features described, the potential mutagenicity of bendamustine was also evaluated using the Leadscope Model Applier Version 2.2.2. Bendamustine generated a POSITIVE consensus call for mutagenicity. Both the *E. coli* 102 A-T Mut. and *Salmonella* Mut. models generated a high mutagenic prediction probabilities (>0.8) (Table 8.6).

Again, as with DEREK, the presence of an alkylating agent and nitrogen mustard in both compounds triggered structural alerts for mutagenicity (Fig. 8.3). Neither model covered the methylated imidazole ring well. The API has feature count for both models that was ≥15. The model feature count aids in determining whether or not the training set (i.e., experimental data) contains a significant number of features used in the prediction model. More confidence is placed in predictions when the model feature count completely covers the entire chemical space of the untested compound. Both models identified analog compounds for bendamustine. The majority of analog compounds were correctly predicted to be positive, which adds to the confidence of the predictions generated by Leadscope.

8.7.3 Toxtree

To demonstrate the capabilities of Toxtree's implementation of the Cramer decision tree, the mutagenic potential of bendamustine was again evaluated. The following outputs were derived in Toxtree v 3.1.0.1851, indicating that the compound is within Class III and should therefore be limited below 90 μg/day. However, based on the

Table 8.6 Summary of Leadscope Model Applier (Q)SAR results for bendamustine (CAS N°. 3543-75-7)

Structures	Leadscope result[a]
 Bendamustine CAS N°. 3543-75-7	Consensus: **Positive** for mutagenicity ***E. coli-Sal.* 102 A-T Mut.** Positive, prediction probabilities = 0.816 ***Salmonella* Mut.** Positive, prediction probabilities = 0.864

[a]Leadscope prediction probabilities less than 0.4 are considered negative predictions, prediction probabilities greater than 0.6 are considered positive, and prediction probabilities between 0.4 and 0.6 are considered indeterminate for mutagenicity

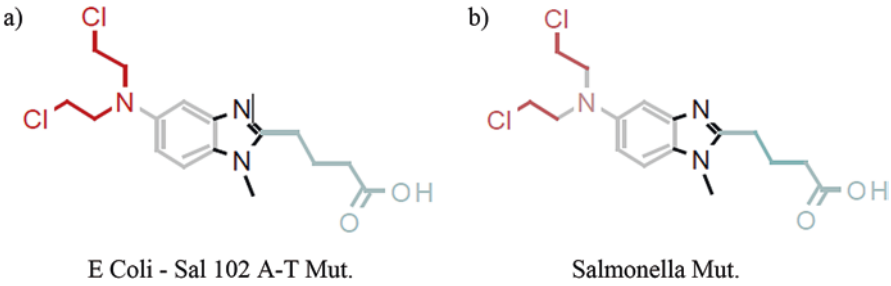


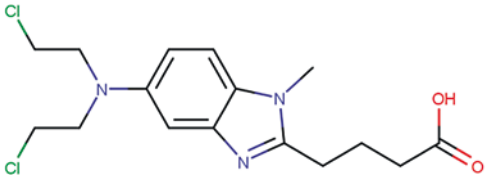
Fig. 8.3 Leadscope's statistical model coverage of bendamustine in the (**a**) *E. coli-Sal* 102 A-T Mut. and (**b**) *Salmonella* Mut. models. Red (blue) bonds and atoms represent positive (negative) SAs that contribute to the overall mutagenicity prediction. Black bonds and atoms represent areas not covered by the model (i.e., not within the model's domain of applicability), and gray bonds make neither negative nor positive contributions

DEREK and Leadscope Model Applier results, it may be more prudent to limit bendamustine to the TTC of 1.5 μg/day for potentially mutagenic compounds (Table 8.7).

8.7.4 Disscussion

The criteria for interpreting predictions that have been developed by the originator of the (Q)SAR tool and the rationale for them should also be considered when evaluating the reliability of predictions. Statistical-based QSAR models often generate probabilities (i.e., 0–1.0) for dichotomous (e.g., positive/negative) endpoints,

Table 8.7 Toxtree's Cramer classification results for bendamustine (CAS N°. 3543-75-7)

Compound	Toxtree predictions
 Bendamustine (CAS N°. 3543-75-7)	*Cramer rules* Q1 – Normal constituent of the body: No Q2 – Contains functional groups associated with enhanced toxicity: No Q3 – Contains elements other than C, H, O, N, divalent S: Yes (i.e., Cl) Q4 – Elements not listed in Q3 occurs only as a Na, K, Ca, Mg, N salt, sulfamate, sulfonate, sulfate, hydrochloride, etc.: No Result: Class III with an oral TTC 90 μg/day

and the model developers recommend specific criteria for interpreting the predicted probabilities (e.g., LSMA criteria: prediction probabilities less than 0.4 are considered negative predictions, prediction probabilities greater than 0.6 are considered positive, and prediction probabilities between 0.4 and 0.6 are considered indeterminate for mutagenicity). Such criteria are usually developed and monitored based on internal and/or external validation testing to optimize the predictive performance of the model.

8.8 Conclusion

This chapter described expert-based and statistical-based (Q)SAR methodologies designed to support the ICH M7 guideline for drug impurities. The rules and molecular descriptors encoded in an expert-based (e.g., DEREK) and statistical-based (e.g., Leadscope Model Applier) methodologies, respectively, are based on well-defined experimental data (e.g., Ames test). Using two orthogonal (Q)SAR models, based on different sets of experimental data and different molecular descriptors, to investigate the toxic potential of an untested compound may potentially increase or decrease the confidence in predictions generated by either methodology. Both DEREK and Leadscope are recommended for use in regulatory submissions because both platforms display good validation results that adhere with the ICH M7 guidance and OECD validation principles.

The structural alerts encoded in the (Q)SAR models described in this chapter are based on well-defined mutagenicity structural alerts from the literature which have been further refined to include additional activating/deactivating factors as well as active subclasses (which represent increased concern). We also described how predictions are generated by the alert system and the various factors influencing the final prediction. Although the (Q)SAR models described here can't extrapolate the potential toxicity of substructural features unknown to the model's database, their good validation results and adherence with the ICH M7 guidance allow these models to be used in the regulatory assessment of impurities with confidence.

References

Alves, V., Muratov, E., Capuzzi, S., Politi, R., Low, Y., Braga, R., Zakharov, A. V., Sedykh, A., Mokshyna, E., Farag, S., Andrade, C., Kuz'min, V., Fourches, D., & Tropsha, A. (2016). Alarms about structural alerts. *Green Chemistry, 18*(16), 4348–4360.

Ashby, J. (1985). Fundamental structural alerts to potential carcinogenicity or noncarcinogenicity. *Environmental Mutagenesis, 7*(6), 919–921.

Benfenati, E., Pardoe, S., Martin, T., Gonella Diaza, R., Lombardo, A., Manganaro, A., & Gissi, A. (2013). Using toxicological evidence from QSAR models in practice. *ALTEX, 30*(1), 19–40. https://doi.org/10.14573/altex.2013.1.019.

Benigni, R., & Bossa, C. (2006). Structural alerts of mutagens and carcinogens. *Current Computer-Aided Drug Design, 2*, 1–19. 169–176.

Benigni, R., Bossa, C., Jeliazkova, N., Netzeva, T., Worth, A. (2008). The Benigni/Bossa rulebase for mutagenicity and carcinogenicity-a module of Toxtree 2(008) Available at https://eurl-ecvam.jrc.ec.europa.eu/laboratories-research/predictive_toxicology/doc/EUR_23241_EN.pdf

Braga, R. C., & Andrade, C. H. (2012). (2012) QSAR and QM/MM approaches applied to drug metabolism prediction. *Mini Reviews in Medicinal Chemistry, 12*(6), 573–582.

Caldwell, G. W. (2000). Compound optimization in early- and late-phase drug discovery: Acceptable pharmacokinetic properties utilizing combined physicochemical, in vitro and in vivo screens. *Current Opinion in Drug Discovery & Development, 3*(1), 30–41.

Cariello, N. F., Wilson, J. D., Britt, B. H., Wedd, D. J., Burlinson, B., & Gombar, V. (2002). Comparison of the computer programs DEREK and TOPKAT to predict bacterial mutagenicity. Deductive estimate of risk from existing knowledge. Toxicity prediction by Komputer assisted technology. *Mutagenesis, 17*(4), 321–329.

Cherkasov, et al. (2014). QSAR modeling: Where have you been? Where are you going to? *Journal of Medicinal Chemistry, 57*(12), 4977–5010.

Cramer, G. M., Ford, R. A., & Hall, R. A. (1978). Estimation of toxic hazard—a decision tree approach. *Food and Chemical Toxicology, 16*, 255–276.

Crum-Brown, & Fraser. (1868). On the connection between chemical constitution and physiological action; with special reference to the physiological action of the salts of the ammonium bases derived from strychnia, brucia, thebaia, codeia, morphia, and nicotia. *Journal of Anatomy and Physiology, 2*(2), 224–242.

Curios-IT. (2009). *Cramer rules with extensions*. Leiden: Curios-IT.

EPA. (2012). (Quantitative) Structure Activity Relationship [(Q)SAR] guidance document. Retrteived online 2019 from https://archive.epa.gov/pesticides/news/web/pdf/qsar-guidance.pdf

ICH M7. (2017). *Assessment and control of DNA reactive (mutagenic) impurities in pharmaceuticals to limit potential carcinogenic risk*.

Jawarkar, S. G., & Game, M. D. (2018). Quantitative structure activity relationship: A tool for new drug design. *World Journal of Pharmaceutical Sciences*. ISSN (Online): 2321–3086.

Jhanwarb, et al. (2011). QSAR - Hansch analysis and related approaches in drug design. *Pharmacology, 1*, 306–344.

JRC.(2016).JRCTechnicalreports;NANoREGharmonisedterminologyforenvironmentalhealthandsafetyassessmentofnanomaterials.Retrievedonlinefromhttps://ec.europa.eu/jrc/en/publication/nanoreg-harmonised-terminology-environmental-health-and-safety-assessment-nanomaterials

Kroes, R., Renwick, A. G., Cheeseman, M., Kleiner, J., Mangelsdorf, I., Piersma, A., et al. (2004). Structure-based thresholds of toxicological concern (TTC): Guidance for application to substances present at low levels in the diet. *Food and Chemical Toxicology, 42*(1), 65–83.

Lapenna S & Worth A (2011), Analysis of the Cramer classification scheme for oral systemic toxicity - implications for its implementation in Toxtree. EUR 24898 EN. Publications Office of the European Union, Luxembourg.

Lhasa (2019) ICH M7 Assessment using Derek Nexus. Retrieved online from https://www.lhasa-limited.org/products/ICH-M7-assessment-using-derek-nexus.htm

LSMA. (2016). Leadscope model applier and the ICH M7 impurities guidelines frequently asked questions. Retrieved online from https://www.leadscope.com/faq/LSMA-ICHM7-FAQs-June2016.pdf.

Munro, I. C., Ford, R. A., Kennepohl, E., & Sprenger, J. G. (1996). Correlation of structural class with no-observed-effect levels: A proposal for establishing a threshold of concern. *Food and Chemical Toxicology, 34*(9), 829–867.

NAFTA. (2011). Technical Working Group on Pesticides (TWG). (Quantitative) Structure Activity Relationship [(Q)SAR] guidance document. Retrieved online from http://www.oecd.org/chemicalsafety/testing/49963576.pdf.

OECD. (2007). Guidance document on the validation of (Quantitative)Structure-Activity Relationships [(Q)Sar] models. Retrieved online 2019 from http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=env/jm/mono%282007%292&doclanguage=en

Piir, G., Kahn, I., García-Sosa, A. T., Sild, S., Ahte, P., & Maran, U. (2018). Best practices for QSAR model reporting: Physical and chemical properties, ecotoxicity, environmental fate, human health, and toxicokinetics endpoints. *Environmental Health Perspectives, 126*(12), 126001.

Plošnik, A., Vračko, M., & Dolenc, M. S. (2016). Mutagenic and carcinogenic structural alerts and their mechanisms of action. *Arhiv za Higijenu Rada i Toksikologiju, 67*(3), 169–182.

Sanderson, D. M., & Earnshaw, C. G. (1991). Computer prediction of possible toxic action from chemical structure. *Human & Experimental Toxicology, 10*, 261–273.

Chapter 9
Pathology and Histopathology Evaluations of Biomaterials and Medical Devices

JoAnn C. L. Schuh

Abstract This chapter will focus on pathology- and histopathology-based study interactions that will optimize the study design, tissue collection and preparation, and evaluation, interpretation, and documentation of biologic responses to biomaterials and finished medical devices. Much of provided information is also applicable to pathology and histopathology evaluations of combination products (device and pharmaceutical or biologics) and regenerative medicine products that include engineered or polymer scaffolds. The reader should be familiar with and consult the most recent ISO and country-specific regulatory standards and reviews to ensure regulatory compliance with the pathology components of any study.

Keywords 2016 · Biofilms · Biopsy samples · Carcinogenicity studies · Causality · Determination of complete biodegradation · FDA · Foreign body response (FBR) · Histopathology evaluations · Imaging techniques · Immunohistochemistry · Immunotoxicity · Infections · International Harmonization of Nomenclature and Diagnostic (INHAND) Criteria · Local and systemic immune reactions · Local implantation sites · Macroscopic anatomical pathology · Modifiers of tissue responses · Nomenclature · International Organization for Standardization · Quantitative (objective) morphometry · Resin embedment · Risk assessment plan · Standardization for Exchange of Nonclinical Data (SEND) · Tissue lists

9.1 Introduction

Whether intentionally working with medical devices or not, in their career pathologists have usually encountered tissue responses to biomaterials or medical devices such as ear tags, catheters, internal fixation devices for fractures, and suture materials. The response to biomaterials and medical devices

J. C. L. Schuh, CL, DVM, PhD, DACVP (✉)
JCL Schuh, PLLC, Bainbridge Island, WA, USA
e-mail: schuhj@jclschuh.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_9

is also familiar to pathologists as the classic foreign body response (FBR), which has similar features whether induced by surgical implantation of biomaterials or a finished medical device or by accidental tissue introduction of plant material, fragments of metal, glass or porcupine quills, hair fragments released into the dermis by surgery or inflammation, and migrating parasites or other exogenous microorganisms. Although no tissue response is often considered the ultimate goal, any foreign material will induce a FBR and a minimal or controlled tissue response while retaining the tissue and device function is a more realistic goal for determination of a biologically compatible biomaterial or medical device.

9.1.1 Regulatory Standards and General References

Toxicologists and toxicologic pathologists are usually familiar with Good Laboratory Practice (GLP) guidance documents, which also apply to medical devices, but the standards from the International Organization for Standardization (ISO; https/www.ISO.org/) are the primary documents that govern biomaterial and finished medical device testing globally. Readers should consult other chapters in this book and ISO and country-specific regulatory standards for additional guidance on expected minimal testing, as well as recent publications on definitions, types of biomaterials and devices, development programs types, and a tabulation of regulatory information and online resources (Funk et al. 2018; Gad and Gad-McDonald 2016; Gad and Schuh 2018; Schuh and Funk 2019).

9.1.2 Risk Assessment Plan

The testing plan for each biomaterial and finished medical device should be on a case-by-case basis using a risk-based approach rather than based on the lack of demonstrable toxicity in a list of standard tests (International Organization for Standardization 2009, 2017b; U.S. Food & Drug Administration 2016). This also applies to selection of the histopathology evaluations (International Organization for Standardization 2016). A thorough characterization of the physiochemical properties of the biomaterial, implant size relative to body size, intended location and expected duration of exposure in the body, degradation characteristics, leachables and extractables, and an informational review of similar biomaterials and available nonclinical data help to assess known risk and need for additional nonclinical testing, including histopathology evaluations. The available regulatory standards and guidelines present lists of tests to consider in the development program to assess risk. Unique biomaterials and medical devices will often require specialized in vivo and histopathology evaluations that reflect biomaterial and device- or organ-specific characteristics.

9.1.3 Biocompatibility and In Vivo Biologic Response Determinations

Biocompatibility, safety, and efficacy are terms that are often used inconsistently and interchangeably for characterization of responses to biomaterials and finished medical devices (Funk et al. 2018). A PubMed (https://www.ncbi.nlm.nih.gov/pubmed/) search for "biocompatibility" identified 26,819 papers, "biocompatibility implant" identified 2884 papers, and "biocompatibility medical device" identified 8776 papers (last accessed 1 Sept 2018). Scanning the content of these papers indicates that biocompatibility is a topical word for engineering design of biomaterials, equated with inertness or lack of a tissue/material interface response of implanted biomaterials; refers to biodegradation patterns after in vitro, ex vivo, and in vivo testing; and may be used to describe integration and function of a finished medical device (including confounding variables such as surgical implantation). Regulatory agencies such as the US Food and Drug Administration (FDA) Center for Devices and Radiological Health (CDRH) often evaluate results of in vitro and in vivo biocompatibility of biomaterials separately from the safety and efficacy or the biologic response evaluation of the finished medical devices tested in animals (Funk et al. 2018; U.S. Food & Drug Administration 2015, 2016). The FDA use of 10993-1:2016 guide Attachment G (U.S. Food & Drug Administration 2016) and other proposed definitions (Williams 2008; Williams 2014) define biocompatibility by context and performance, but a widely accepted and specific usage pattern for the term remains elusive. In reality, the final determination for compatibility of a biomaterial or finished medical device should encompass the totality of the response of the biomaterial or finished device to the biologic test system and the response of the biologic system to the biomaterial or device (Sect. 9.2.3.1). This broader view is similar to pharmacokinetic (what the body does to the drug) and pharmacodynamic (what the drug does to the body) determinations for drugs, respectively. Single studies can report the specific biodegradation patterns and ex vivo or in vivo tissue responses, but the term biocompatibility should be used sparingly and only if modified by the characteristics of the study design and criteria used to determine compatibility (Sect. 9.5.6).

9.2 General Pathology Considerations for Study Designs

9.2.1 ISO and GLP Compliance

The FDA has a stated preference that in vitro and in vivo local biocompatibility and safety and efficacy studies in animals should be conducted in accordance with GLP principles in 21 CRF Part 58 (U.S. Food & Drug Administration 2015, 2016). The Organisation for Economic Co-operation and Development allows for the application of GLP principles by individual jurisdictions (The Organisation for Economic Co-operation and Development 2015). Other global regulations often only focus on

ISO standards, but the study director, toxicologist, and pathologist should be trained under GLP principles, and the CRO selected to conduct the study should be GLP compliant to ensure proper conduct of studies for global registration (Funk et al. 2018). An overview of study types, selection of studies, and general study design of implantation, safety, and efficacy studies are provided in these ISO, FDA, and GLP regulatory documents.

9.2.2 Standardized Nomenclature and Controlled Terminology

9.2.2.1 Anatomy and Histology

Macroscopic anatomical nomenclature and histological nomenclature have been standardized for animals in the Nomina Anatomica Veterinaria and Nomina Histologica, respectively (http://www.wava-amav.org/). While the standardized nomenclature focuses on domestic animals, these standardized terms also apply to laboratory animals (Popesko and Getty 1971; Popesko et al. 1990, 1992). Use of standardized anatomical and histologic terminology enhances accuracy and consistency of tissue identification and collection, and the use of standardized terms is encouraged for any toxicology, safety, or efficacy studies.

Histopathology terminology for toxicologic pathology is being standardized through the International Harmonization of Nomenclature and Diagnostic (INHAND) Criteria (https://www.toxpath.org/inhand.asp#pubg). For organ systems and species criteria that have not been finalized, the older Standardized System of Nomenclature and Diagnostic Criteria guides can be used for rodents, or individual publications may need to be consulted until the INHAND series is finalized. INHAND pathology terminology is fundamental to generation of Standard for Exchange of Nonclinical Data (SEND) datasets which do not yet apply to studies on biomaterials and medical devices (Sect. 9.2.3.2).

Despite attempts to harmonize anatomic, histology, and histopathology terminology, recent publications and textbooks may not follow these nomenclature standards. Terminology used in older publications, particularly in publications in bioengineering and toxicology journals and textbooks, make it difficult to identify which tissues were collected, particularly when tissues such as lymph nodes have multiple locations that are modified by anatomical location, such as medial or lateral. Consistent use of standardized terms for such tissues is particularly important to properly document that the correct lymph nodes draining the implantation site were collected.

9.2.2.2 Best Practices and Position Papers

The Society of Toxicologic Pathology (STP) has published numerous best practices and position papers for histopathology conduct of studies. Currently there are no STP best practices or position papers (https://www.toxpath.org/best-practices.asp)

specifically for biomaterials, medical devices, and combination, regenerative, or 3D printed products. However, multiple best practices and position papers on data interpretation, study design, study reports, peer review, and organ-specific guidelines directed toward pharmaceuticals and biologics can be useful and applied to evaluation of biomaterials, medical devices, and related products.

9.2.2.3 International Harmonization of Nomenclature and Diagnostic (INHAND) Criteria

INHAND has been completed for multiple organ systems in rodents, and work continues on remaining rodent organs and for terminology in large animals (rabbits, minipig, dogs, and nonhuman primates). The STP, along with related international societies (British, European, and Japanese), has also published these documents in *Toxicologic Pathology* (http://journals.sagepub.com/home/tpx) and the *Journal of Toxicologic Pathology* (https://www.jstage.jst.go.jp/browse/tox). While these standardized terms can be applied to histopathology evaluations of biomaterials and medical devices, INHAND does not address all biocompatibility findings used for semiquantitative scoring such as skeletal muscle fatty infiltration (Roberts et al. 2013) nor details of some tissue/device interface reactions. INHAND guides are also not proposed for some species/strains of animals frequently used for medical device testing (e.g., ruminants, full-size swine, hound dogs, chinchillas, and guinea pigs). Published veterinary diagnostic information for farm animals, pet animals, and caged pets should be researched to identify the incidence and occurrence of sex- and age-related background and spontaneous findings, diseases, and induced pathology in these species/strains.

9.2.3 FDA Considerations for ISO Testing and Animal Studies for Medical Devices

9.2.3.1 FDA-Specific Guidance Documents

In 2016, the FDA issued a guidance on Use of International Standard ISO 10993-1, Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process (U.S. Food & Drug Administration 2016). This guidance clarifies and modifies the ISO 10993-1 standard testing matrix and also adds additional details for in vivo implantation biocompatibility testing and for medical devices.

The FDA has also issued a draft guidance specific to the design, evaluation, and reporting of animal studies used to test medical devices (U.S. Food & Drug Administration 2015). Important points from this document are that the animal model selected should be scientifically justified, efficacy is secondary to determination of safety, and in vivo assessments should be bidirectional. This would include

device effect on the biologic system such as injury, inflammation, resorption, healing, and integration and include animal effects on device structural and functional integrity, surgical success, imaging characteristics, and system compatibility with ancillary devices. A full tissue evaluation is encouraged to assess local, downstream, and systemic effects. The study should be conducted according to GLP, and justification should be provided if the test site is not a GLP facility.

9.2.3.2 Standard for Exchange of Nonclinical Data (SEND)

SEND was implemented in Dec 2017 as a requirement by the FDA for submission of standardized electronic data along with nonclinical report submissions to the FDA. Specific component datasets, including pathology data, are now required for regulatory submissions to the Center for Drug Evaluation and Research and the Center for Biologics Evaluation and Research. As of Sept 2018, there was no published plan to incorporate SEND requirements into biocompatibility and medical device study submissions to the FDA CDRH.

9.3 Workflow Involving the Pathologist

Pathologists can provide broad input on the study design and help to optimize all tissue-related interactions, but study plans frequently fail to assign the pathologist to the early planning or study team (Maul et al. 2011). Unfortunately, the necropsy and histopathology interpretation may be provided by two different pathologists; the in vivo, necropsy, and tissue trimming services may be provided by separate contract research organizations (CRO) or an academic or other test facility distant to one another; and the histopathologist may be at another site or in another country from the CRO or other test facilities. Integrating the interactions of one or more CRO, test facilities, and one or more study pathologists with the sponsor, study director, clinical veterinarian, or surgeon and study toxicologists needs to be incorporated into the study plan.

For any studies, the necropsy pathologist and histopathologist should be consulted prior to study initiation. The pathologist can provide input on specific study design components including: 1) the need for pilot or proof-of-concept studies, 2) selection of appropriate species/strains or animal models of disease, 3) the use of surrogate rather than clinical devices, 4) sacrifice intervals and selection of endpoints, 5) tissue samples and tissue sampling, 6) histology methods and special stains (histochemistry and immunohistochemistry), 7) digital photographic documentation in vivo and ex vivo, 8) light and electron microscopic evaluation, 9) need for blind, peer, or expert reviews, and 10) to advise on supplemental tissue imaging or analytical techniques (Funk et al. 2018; Nikula and Funk 2016).

9.3.1 Limitations of Test Animals and Animal Models of Disease

Pathologists can contribute to selection of appropriate test species and strains and suitable animal models of human disease through their familiarity with comparative physiology, anatomy, and pathology and broad training in mechanisms of spontaneous and induced diseases (Funk et al. 2018). The complexity of the biologic response does not allow for in vitro or ex vivo models to substitute for evaluating the ultimate safety, efficacy, and fate of medical devices. Selection of the wrong species, strain, or animal model can also be a costly mistake. Examples of advantages and disadvantages of animal models that can be problematic for medical devices or the tissue assessment by the study pathologist are included in Table 9.1. Biomaterial and medical device placement should be discussed with the study pathologist to determine if device size and location may cause confounding tissue responses, affect tissue sampling, or increase procedural and by-design device failure.

9.3.2 Pilot and Proof-of-Concept Studies

In some development programs, the pivotal studies should be postponed while a non-GLP pilot or proof-of-concept study is conducted to evaluate the species selection and feasibility of any surgical procedures (U.S. Food & Drug Administration 2015). The pilot study may also be required to optimize the necropsy, tissue processing, and histology procedures (Sect. 9.4) and to evaluate the robustness of proposed biologic response scores (Funk et al. 2018) as discussed further in Sect. 9.5.2.

9.3.3 Pathology Observers at Implantation or Surgery

If possible, the pathologist responsible for the necropsy and the histopathology should also be present at the surgery or insertion of the biomaterial or device. This allows the pathologist to observe possible confounding features introduced by the surgical methods, to observe orientation and placement of the biomaterial or device that may affect the tissue sampling, and to optimize plans for photographic documentation, tissue collection, trimming, and processing procedures. The surgeon can also be a valuable observer at the necropsy to assist in separating surgical from medical device changes and in retrieving the biomaterial or medical device (Funk et al. 2018; Nikula and Funk 2016).

Table 9.1 Examples of the advantages and disadvantages of animal models that may affect the suitability, biologic responses, and histopathology evaluation of biomaterials and medical devices[a]

						Disadvantages		
Model criteria	Species or strain[b]	Clinical indications	Animal availability	CRO needs custom husbandry	Advantages	Anatomy	Pathology databases[c]	Additional comments
Surgically accessible site	Chinchilla	Otic (ear) surgery and implants	Low	Yes	Easy surgical access to external ear and auditory bulla	Not well documented Horizontal spatial orientation	No	Accepted surgical/efficacy model but insufficient histopathology data to recommend for safety studies
Physiologic similarity to humans	Nonhuman primate (NHP) Pigs (mini or full size)	Multiple	Moderate	Yes (NHP) Sometimes (pigs)	Anatomical and physiologic similarity, but biological differences exist for NHP and pigs relative to human	Transient biped spatial orientation (NHP) Typical *Macaca* spp. too small for many finished medical devices Horizontal spatial orientation (pig)	Yes	Use of chimpanzee or large apes not ethically acceptable Some strains of pigs susceptible to anesthetic-induced death and malignant hypothermia
Proof-of-concept or main surgical model	Ruminant	Cardiovascular grafts	High	Sometimes	Ease of access to saphenous vein for engraftment to femoral artery on medial side of inner hind leg	Horizontal spatial orientation Difficult to do end-to-end anastomosis as in human surgery	Yes	Graft may be compressed and thrombosed when animal lays in a ventral recumbency with legs tucked under the body

Large animal species/strain for clinically relevant surgically placed device	Adult ruminant versus calf, kid, or lamb Hound versus beagle dog	Cardiovascular	Moderate (dog) High (ruminants)	Yes (ruminants) Sometimes (hound dog)	Surgical implantation of large finished medical devices	Horizontal spatial orientation Narrow chest cavity of dog may require device placement modifications	No (hound) Yes (farm animals)	Goats browse upright walking on their hind legs but are still quadrupeds Bone and cartilage (Os cordis) in dog and ruminant hearts[d]
Clinically relevant for nonsurgical device	Beagle or hound dog	Intraurethral device or catheter	High (beagle dog) Moderate (hound dog)	No	Hound dog accepts large devices; scalable (pediatric) devices are suitable for beagle dog	Horizontal spatial orientation	Yes	Sharp curve at urethral transition in males may require surgical insertion of this typically nonsurgical device[e]
Clinically relevant pig wound model	Minipig Standard-size pig Duroc strain	Integument	High (minipigs) Moderate (standard pig) Low (Duroc)[f]	No (minipigs) Yes (standard pig and Duroc pig)	Tissues physiologically similar to humans, but biological differences exist for pigs relative to human Full-size pig accepts large devices	Horizontal spatial orientation	Yes, but specifics limited for some strains (Duroc)	Maximum feasible wound size (±10% body surface area) is usually much smaller than human patient population with skin damage or loss

(continued)

Table 9.1 (continued)

						Disadvantages		
Model criteria	Species or strain[b]	Clinical indications	Animal availability	CRO needs custom husbandry	Advantages	Anatomy	Pathology databases[c]	Additional comments
Large number of animals for test groups	Rodents Guinea pigs Rabbits	Multiple	High Low (hairless strains and pigmented rabbits)	Sometimes (hairless strains, guinea pigs, and rabbits)	More economical than large animals Accepts small and scalable devices (e.g., bone replacement) May require development and use of a smaller surrogate device	Horizontal spatial orientation Limited tissue, body mass, and surgical accessibility	Yes (rodents) Limited (rabbits) Sparse (guinea pigs)	Uniform response in inbred strains but diverse response in outbred strains may better reflect risk assessment for humans Guinea pigs susceptible to anesthetic-induced death Rabbits susceptible to spine fractures during handling

[a]Stress induced by surgery, frequent handling, and test procedures (Everds et al. 2013) and opportunistic infections in animals not raised behind filter barriers complicate the husbandry and interpretation of tissue responses during any experimental procedure

[b]Unusual species may have a specific research, organ, or surgical use or may be larger than or a different strain of nonrodent than that typically used in toxicology studies

[c]No database for clinical pathology and incidence and occurrence of spontaneous and induced gross pathology and histopathology of tumor and nontumor changes

[d]Os cordis (heterotopic ossification) is a cardiac skeleton in the base of the heart or atrium which supports large vasculature of the aortic ring, but can also be present in ventricular walls and papillary muscle. This change is reported as a normal variation or incidental histopathology change in multiple species including dogs (Douglass et al. 2003; Sato et al. 2012a, b) and possibly as metaplasia secondary to chemical damage to the heart (www.ema.europa.eu/ema/pages/includes/document/open_document.jsp?webContentId=WC500013083)

[e]Firm intraurethral devices may not be able to negotiate the sharp curve as the urethra transitions from the penile to the pelvic part in male animals (Evans and De Lahunta 2013). This may require a surgical incision and insertion of the device above the curve to avoid trauma trying to insert stiff devices past this transition point

[f]Female Duroc pigs are used to model hypertrophic scar formation, and in addition to erratic and low availability, they are difficult to manage

9.3.4 Submission of Study Materials to the Pathologists

At a minimum, the necropsy and the histopathologist should have access to a copy of the physiochemical characteristics of the biomaterials or configuration of the finished device, manufacturing processes, extractables and leachables, the study protocol with any amendments, animal source/vendor health and background data, and surgical procedures in life clinical observations and body weights, life specialized analyses (e.g., bone morphometry, echocardiograms, cytokine analysis, etc.), clinical pathology, gross pathology, and organ weight data. With access to complete datasets, both the pathologist responsible for the necropsy and the pathologist responsible for the histopathology can provide macroscopic and microscopic correlates with in life observations or imaging and microscopic changes, clinical pathology versus microscopic changes, and gross finding and organ weight changes with microscopic data. Neglecting or intentionally withholding physiochemical and study in vivo data frequently leads to suboptimal histopathology preparations or incomplete interpretation of the pathology data.

9.4 Pathology Evaluations

9.4.1 Necropsy

A recent special issue of *Toxicologic Pathology* provides both general and device-specific information for gross pathology and histology techniques for biomaterials and medical devices (Rouselle and Paulin 2019).

9.4.1.1 Responsibility

A common practice for implantation and medical device studies is to have the technical staff trained as prosectors conduct all necropsy procedures or to use the study director, clinical veterinarian, or surgeon to supervise this phase of the study. Ideally, an experienced toxicologic or veterinary pathologist should supervise the gross pathology and be on site to help enter the necropsy and organ weight data, to guide the prosectors on identification of unusual or unexpected gross findings, to make the determination on collecting additional tissues or samples outside of the protocol-required list, and to direct photographic documentation. A pathologist is also the most likely individual to meet GLP standards for necropsy responsibilities for a necropsy, except for simple studies such as intramuscular implantation. Less desirable is to have a pathologist familiar with the study design available on-call and to have them direct or assist in the tissue collection based on remote gross photographs. Insufficient or inappropriate recording of necropsy data and necropsy data that conflicts with the histopathology data may have a negative effect on the risk assessment and may require studies to be unnecessarily repeated.

9.4.1.2 Scheduled Necropsy

A complete necropsy should include macroscopic evaluation of general body appearance, body cavities, and all organ systems and their visible tissue components. Even if systemic toxicity of the biomaterial and medical device is known not to occur, the necropsy should not be restricted only to local implant sites and examination of protocol-required tissues.

9.4.1.3 Unscheduled Deaths

In the case of sudden deaths or euthanasia of a moribund animal, qualified CRO or test facility staff members need to be available outside of normal operating hours. The necropsy should be conducted the same day the death is identified, and in no case should the examination be delayed by refrigerating the carcass overnight or over a weekend. Autolysis begins in most tissues within 5–15 min of death, and delaying the necropsy may preclude macroscopic and microscopic examinations. To assist the pathologist in making a determination of proximate or ultimate cause of death and to correlate to morbidity observations, the necropsy should be optimized to collect tissues in a timely manner and to collect a comprehensive list of local and systemic tissues even if there are no visible gross findings. For planned euthanasia of a moribund animal, any feasible protocol-required assessments, including veterinary clinical evaluation, clinical pathology (hematology, chemistry, and coagulation), imaging, and specialized organ-specific physiological evaluations (e.g., ECG or echocardiography), should be undertaken prior to the necropsy. Animals euthanized should be subjected to gross pathology, sampling for microbial evaluation of implant infections, organ weights, collection of the entire implantation site and adjacent tissues, and collection of tissues (ISO 10993-11:2017 Annex E) including all protocol-required tissues and gross abnormalities and all treatment sites. Similar procedures and sampling, except for clinical pathology and organ weights, apply to animals found dead.

9.4.1.4 Organ Weights

Organ weights along with relative weights (relative to body weight and brain weight) can be supporting indicators of deleterious local and systemic effects. The ISO 10993-11:2017 (International Organization for Standardization 2017b) standard and Society of Toxicologic Pathology (STP) best practice paper (Sellers et al. 2007) provide guidelines for organ weight measurements and interpretation.

9.4.1.5 Recording Necropsy Findings

Any macroscopic findings but particularly abnormalities in and around the site of implantation or insertion of the medical device and of any protocol-required tissues should be recorded. If the implant is unstable or not visible without invasive cutting

(e.g., injected bone substitutes), gross observations of the implant may have to be postponed until the tissues are trimmed or skipped entirely. In this situation, the records should note why no gross observations are recorded. For implants in tissues where the entire organ and device need to be collected together, care should be taken in manipulating the fresh or fixed tissue to observations as the tissue/device interface should not be disturbed. Electronic data capture using standardized nomenclature glossaries is preferred for GLP studies but may be too restrictive for recording observations for all or part of some implantation and medical device studies. If needed, electronic data records can be supplemented with handwritten records that capture the use of nonstandard tools and methods, study-specific needs such as diagrams of implant locations and detailed tissue collection maps, and photographic documentation.

9.4.1.6 Digital Gross Photography (Macrophotography)

Regulatory requests are increasing for gross and microscopic photographic documentation of implants at intermittent stages and at the terminal phase of studies. This may include digital photographs to document the surgery, interim in vivo photographs for visible devices, device/tissue interface, and general normal or distorted organ/tissue architecture in the implantation area. Photographs may need to be taken immediately after surgical implantation, at multiple time points during in-life procedures, at necropsy, and at tissue trimming or explantation of the device. These digital photographs can be used to follow the biologic response to and fate of the medical materials (Sects. 9.7 and 9.8) and correlate clinical observations with necropsy and histopathology findings. Macrophotographs themselves may also be a specific endpoint for vascular devices (Rousselle and Wicks 2008) and burn models (Greenwood and Dearman 2012). Photographs should be high-resolution color images with appropriate identifiers (date, study and animal number, tissue, and laterality) and a numeric scale. For comparison within and between groups, images should be captured consistently from the same tissue regions and at the same magnification. Digital macrophotographs for regulatory submission should comply with GLP standards and should be presented with no or minimal digital post-processing or manipulation (Tuomari et al. 2007).

9.4.2 Tissue Collection

9.4.2.1 Study Endpoints

The time intervals, number, and type of endpoints need to be carefully considered to ensure that each endpoint sample contributes to the objectives of the study. Tissue collection for histopathology most often occurs at study termination, but interim necropsies may also be desirable. Interim tissue samples may also be collected by biopsy of externally accessible implants in areas such as the skin or mucosa. Interim

samples help to monitor the progression of the tissue responses and the fate of implanted materials (Sects. 9.7 and 9.8). Bioengineering sampling needs may also compete with the needs for histopathology. The timing, preservation method, sample size and orientation, need to retain the device in situ, or device explantation may create competition for both interim and study termination samples. If necessary to ensure proper sampling, additional animals or satellite animals for a specific sample may need to be added to the groups. These considerations will also need to meet the requirements and ethical considerations of the institutional animal care and use committee.

9.4.2.2 Tissue Handling

To avoid artifacts, prosectors should be trained to avoid compressing tissues with forceps and to not wipe or rub tissue surfaces by hand, with instruments, or with gauze or paper products. Blood, excreta, and other contaminants can be reduced or removed by carefully rinsing tissues in a small quantity of buffered saline. All tissues, particularly gross abnormalities and the treatment sites, should be cleaned of any normal extraneous tissues such as muscles or unneeded membranes. Care should be taken to clean gross lesions and treatment sites sparingly and to retain implanted devices and inflammatory reactions with adjacent tissues to preserve the tissue/device interface.

9.4.2.3 Marking Implant Sites and Tissues

Implant sites and tissue/device interface margins, such as skin wound margins, can be permanently marked on animals during surgery using indelible tattoo ink that will be visible at the gross necropsy and tissue trimming and will be microscopically detectable. Similar to marking tissues prior to or during tissue collection, the location of an implant site can be permanently marked during surgery with a piece of inert US Pharmacopeia (USP) control material. This material is often visible during tissue collection and trimming, and the material acts as a marker during the microscopic examination (Elmore et al. 2017). This marker can serve as an accurate indicator of the original implantation site tissue/device interface and may be useful in determining the fate of biodegraded biomaterials that may disappear during chronic studies.

Rather than free floating in a jar of formalin, small tissues and serial sections or multiple regional samples taken from larger tissues can be collected into labeled tissue cassettes used for trimming, processing, and embedding to easily retrieve and retain the sequence or sub-location of the tissue samples. At necropsy, implant sites can also be marked with indelible dye before fixation to facilitate accurate recovery at trimming and sectioning. Tissue orientation or laterality can also be marked using small selective cuts into the tissues, using indelible dyes, or placing the tissues directly into marked cassettes. Flat or long tissues with elastic components that are prone to shrinkage, contracture, or curling (skin, blood vessels, nerves, muscle, and

gastrointestinal tract) can be immobilized prior to fixation in a more relaxed state by lightly attaching the tissues by wrapping, stapling, or inserting the ends into a V-notch of a solid surface such as cardboard or perforated plastic sheets. These extraneous materials need to be pretested to ensure that they will not disintegrate in the fixative and not affect the subject tissues and that the tissues will fix properly.

9.4.2.4 Tissue Lists

Tissues to be collected at necropsy will be listed in the study protocol or plan. As per ISO 10993-11:2017 Annex E (International Organization for Standardization 2017b) and the FDA guidance (U.S. Food & Drug Administration 2015), the test material or device should be evaluated for local, downstream, and systemic exposure. This may involve collecting, preserving, processing, and histopathologically evaluating a full complement of tissues (ISO 10993-11:2017 Annex E). In the absence of systemic exposure, all tissues should be collected, but a more restricted histopathology examination of representative tissues from major organs may be appropriate (Tier I tissues in ISO 10993-11:2017 Table F.1.). Collection of an abbreviated tissue list for histopathology examination should always be based on the study and device requirements and knowledge acquired through prior studies. Not all suggested Tier I tissues may be appropriate, and tissues not listed may provide important information on the local, upstream, and systemic tissue responses. It should be noted that (1) ISO 10993-11:2017 Table F.1. includes the liver but not major gastrointestinal tract components (stomach, small and large intestine) as Tier 1 tissues, (2) excretory is a function and not a recognized organ system and many excretory tissues are not included in the list,(3) thymus rather than bone marrow (representing primary lymphoid tissue) may be preferable or an important adjunct tissue for evaluation of systemic and lymphohematopoietic effects, (4) lymph nodes draining the implantation site (and distant comparator lymph nodes) rather than just spleen (representing secondary lymphoid tissue) may be needed as a Tier I tissue used to interpret local inflammatory events, (5) a consistent location should be chosen for the muscle sample, and (6) a bone marrow smear is better than a tissue section for evaluation of hematopoietic activity. A histopathologist can provide advice on the major organs and a scientifically justified abbreviated list of representative tissues to collect from major organs, as well as additional tissues within the major organs that would be relevant to the individual study and medical device.

9.4.2.5 Sample Selection from Test Sites

The number and location of samples collected for fixation and for histopathology evaluation of the medical device site are dependent on the implant size, type, location and characteristics, number of test sites, the size of the implant relative to the size of the animal, degradation patterns, the integration characteristics and reparative processes, and need for efficacy evaluations. Additional samples may

also need to be collected for bioengineering analyses. Where possible, a consistent number of tissue samples should be collected from each test subject and site, and the biomaterial or medical device should be retained with adjacent tissues to allow histopathology evaluation of the tissue/device interface. Individual tissue samples should also be consistent in size, orientation, location, and laterality. Tissues larger than 0.5 cm thick and dimensionally larger than 5–6 cm usually need to be cut into multiple smaller representative pieces to allow for penetration of tissue fixatives.

Along with the full or abbreviated tissue list, the number of test site samples included in the study protocol should be determined collaboratively by the sponsor, CRO, study director, and study toxicologist and pathologist. Tissue collection plans may also need to be modified and the protocol amended after evaluation of interim data collected by noninvasive imaging, biopsies, or necropsy. Interim imaging data may also aid the pathologist in selection of specific areas of the test site that need to be correlated with abnormalities identified by imaging (Alves et al. 2019; Rousselle and Wicks 2008). During the necropsy, the supervising pathologist should collect any gross lesions in any tissues and, with discretion and additional documentation, collect additional tissues not predetermined as protocol required. Unlike rodents, it is not feasible to retain all residual body parts from a large animal necropsy, and selection of samples during the necropsy is critical. Therefore, duplicate or extra tissues should be preserved in all studies to allow for future testing or for confirmation of unusual or unexpected histopathology findings. Any tissues collected during the study should be retained until the study report is finalized and any GLP retention policies have expired.

9.4.2.6 Sampling Large and Complex Implants

Although the tissue/device interface should be preferentially preserved, large or complex implants may need to be separated from the tissues to ensure and optimize tissue fixation, processing, and sectioning. Retrieval of the device for bioengineering analyses may also be necessary. In these situations, multiple small samples of the tissue should be carefully cut and removed from the tissue/device interface or other critical areas prior to removing the entire device from the remaining tissue. Other tissue samples are then collected for fixation from the residual tissue after removal of the device. Physical tissue extraction by maceration or enzymatic methods (stable polymers) may allow biomaterials to be identified macroscopically (Rousselle and Wicks 2008), but these techniques should not be applied to tissues intended for histopathology.

9.4.2.7 Identification and Collection of Regional Lymph Nodes

Gross identification of lymph nodes and differentiation from the fat pads that they are often enclosed in can be enhanced by using acetone-, alcohol-, hydrochloric acid-, or acetic acid-based revealing agents and fixatives during the necropsy or at

tissue trimming (Horne et al. 2014). In ISO 10993-6:2016 (International Organization for Standardization, 2016), lymph nodes draining the site of implantation are included as an endpoint for biocompatibility studies, and evaluation of regional lymph nodes is an important part of evaluation of safety and efficacy of finished medical devices (Wancket 2019). Veterinary anatomical textbooks (Budras et al. 2007; Evans and De Lahunta 2013; Grossman and Getty 1975a, b; Popesko and Getty 1971; Popesko et al. 1990, 1992) and some specific publications (Soto-Miranda et al. 2013; Suami et al. 2013, Ito and Suami 2015; Suami and Scaglioni 2017) are useful to identify patterns of regional lymph nodes drainage from implantation sites. Multiple lymph nodes in some locations and overlapping or interconnected drainage regions complicate the choice of the primary draining lymph node and mapping dye studies (Harrell et al. 2008), and lymphoscintigraphy or lymphangiography (Zhang et al. 2011) may be required. A concern is that the presence of inflammation or concurrent disease may change the lymphatic drainage pattern and that unexpected reactivity in distant lymph nodes (Kesler et al. 2013) may be an indicator of systemic effects of the medical device or may be a false negative related to a confounding tissue finding such as infections.

9.4.2.8 Special Sampling Procedures

Biopsy Samples

In addition to samples collected at necropsy, accessible implants may be intermittently biopsied throughout the duration of the study to monitor the device/tissue interface as part of a time course study. This technique is most often applied to skin samples, but transcutaneous or surgical re-opening of deeper implantation sites for biopsy collection is also possible. For biodegradable polymers, biopsies can be used to monitor and determine the point of complete dissolution of the matrix and resolution of any inflammatory response (Sect. 9.8.2.1). This intermittent tissue sampling is particularly useful for chronic studies (1 or more years in duration) where complete biodegradation of the biomaterial and resolution to a functionally normal tissue is an endpoint. Biopsy samples may be able to demonstrate early biodegradation and return to normal function and thus justify early termination of the study. For permanent implants, biopsy may have a role in evaluating tissue responses and infections and monitoring unexpected degradation of the biomaterials. However, biopsy samples require the risk of additional anesthesia and surgery, and the sampling introduces a confounding acute inflammatory and reparative process on top of a healing or healed implantation site. Biopsy sample sites should be carefully documented to avoid accidentally collecting these areas at study termination. The major limitation for biopsy samples is that the resulting tissue sample is small (1–20 mm diameter) and of short depth (usually no more than 15 mm). Multiple samples may be required to adequately evaluate an implantation site for residual biomaterials. As with samples collected at necropsy, biopsy samples may be evaluated by qualitative, semiquantitative, and quantitative histopathology evaluations and can be stained for immunohistochemistry evaluations or used for molecular and chemical analyses.

Fresh Frozen Tissues

Some analytical procedures are not successful when performed on formalin-fixed tissues, but flash frozen tissues stored at greater than −60 °C can be held until the analyses are conducted. If antigenic epitopes are not stable in formalin, immunohistochemistry may have to be conducted on tissues that are placed in a protective embedding medium (optimal cutting temperature [OCT] compound) prior to flash freezing in liquid nitrogen. Laser capture microdissection (Sect. 9.6.1) and molecular analyses (Sect. 9.6.3) may be affected by the use of OCT, and alternative media may be required (Datta et al. 2015). Usually, tissues no larger than $1cm^3$ should be flash frozen to ensure complete freezing. For molecular analyses, tissues are usually directly frozen or macerated in saline (without special media) and are held at greater than -60C until analyzed.

Hemocompatibility

In vivo hemocompatibility of devices that contact blood (International Organization for Standardization 2017a) is generally incorporated into safety studies rather than done as a free-standing study (Wolf and Andwraon 2012). Tissues collected and evaluated for in vivo hemocompatibility should include collection of vasculature upstream of the insertion site as control tissue and at the insertion site to assess immediate local tolerance and multiple samples downstream of the insertion site (starting 1–2 cm or further downstream) to assess potential thrombosis and cytotoxicity. Selected distant end organs that readily display thromboemboli (particularly the lung, but also the brain, liver, spleen, and kidney) should also be evaluated.

Immunotoxicity

The need for immunotoxicity assessment of biomaterials and medical devices may be detected early in the development program by one or more changes in hematologic parameters, lymphoid tissue weights, microscopic changes of primary (thymus and bone marrow) and secondary (spleen, lymph nodes, and mucosa-associated lymphoid tissues), and susceptibility to opportunistic infections. Histopathology screening of lymphoid tissue, particularly lymph nodes draining the implantation site compared to distant external and internal lymph nodes, is useful to fully characterize the inflammatory response and for detecting immunological abnormalities. The typical lymph nodes collected in GLP toxicology studies are often not the lymph nodes draining the implantation sites (Sect. 9.4.2.7). Simple qualitative and semiquantitative scoring of lymph node changes may not be informative, and detailed reporting of individual lymphoid compartments using enhanced histopathology principles may be needed for lymph nodes (Elmore 2006b), as well as the spleen (Elmore 2006c), thymus (Elmore 2006d), and mucosa-associated lymphoid tissues (Elmore 2006a). Bone marrow smear and hematologic evaluations are preferable to bone marrow tissue sections for characterization of bone marrow changes (Tomlinson et al. 2013).

Carcinogenicity Studies

Carcinogenicity testing of devices should be considered for all permanent devices and when certain leachables, extractables, and manufacturing residuals are identified, unless other data (genotoxicity, structure activity relationships, and supportive literature) is sufficient to justify omission with a regulatory waiver (International Organization for Standardization 2014; International Organization for Standardization 2017c). Carcinogenicity studies are generally 2-year studies in two rodent species with the possibility of a 6-month study in genetically modified mice (International Council for Harmonisation 1995; International Organization for Standardization 2014). Biomedical and medical device inquiries may only require the use of a single rodent species, with the rat most often selected. The transgenic CByB6F1-Tg(HRAS)2Jic oncogenic model (rasH2) is the most commonly used transgenic model. Transformation-related protein 53 deletion ($Trp53^{\pm}$ or $p53^{\pm}$) model should not be used due to the high incidence of implanted microchip transponder-induced sarcomas which indicates an enhanced carcinogenicity response to foreign material (Blanchard et al. 1999; Schuh 2015). The Tg.AC model, often used for dermal testing, had a high incidence of false positives, and these animals are no longer commercially maintained. Rarely are carcinogenicity studies in large animals considered economically or practically feasible, and tumor databases, similar to those in rodents, are lacking for large species. As carcinogenicity assays focus on tumor development, assessment of clinical pathology and organ weights and special tissue evaluations are generally not necessary. The maximum implantable dose (MID; often the finished medical product or a multiple of expected clinical exposure in mg/kg) may be the only treatment group. An additional dose group at one-half the MID (MID_{50}) may be considered to identify a minimal effect or no effect dose. Due to foreign body-induced carcinogenesis, control groups implanted with USP control material is a questionable practice. Large, smooth, flat, and nonporous reference materials are most often associated with tumor induction. The gross pathology and histopathology examination in a carcinogenicity study should include a complete gross examination of the animal, and tissues collected will generally include a screen of all tissues (ISO 10993-11:2017 Annex E) and all tumors, particularly tumors around the implant in untreated control and the MID groups. In an MID_{50} group, collection only of gross lesions and tumors may be appropriate. Histopathology evaluations should conform to reporting standards used for pharmaceuticals and biologics. The International Council for Harmonisation (ICH) safety guidelines S1A-S1C (http://www.ich.org/products/guidelines/safety/article/safety-guidelines.html) should be consulted for additional information on recommended carcinogenicity study parameters. Tumor evaluations and reporting should include a peer review and may need a pathology working group review (Sect. 9.5.4). Tumorigenicity of biomaterials and medical devices is discussed further in Sect. 9.7.2.3.

9.4.3 Tissue Fixation

General principles used for collection and fixation of tissues should also be rigorously applied to optimize the quality of tissues to be used for histopathology and other related investigations of biomaterials and medical devices. Whenever possible, the device should be fixed together with the adjacent tissue to maintain the tissue/device interface for evaluation. If this is not possible, small sections of the interface should be carefully collected for fixation prior to explanting the device. When the medical device is encompassed by a large piece of tissue, the animals will need to be anesthetized and the entire animal or the implant region fixed by intravascular perfusion with 4% buffered paraformaldehyde, followed by submersion fixation, usually in 10% neutral buffered (Gage et al. 2012).

Tissues are collected immediately after death into fixatives to prevent autolysis (begins within 5–15 min for most tissues) and to allow clean cuts to be made in the hardened tissues at trimming. Most fixatives do not penetrate rapidly or deeply into whole organs or large tissues and multiple cuts should be made into large tissues and organs and bones should be cut open or representative sections, no larger than 5 mm thick, and 2–3 cm of the tissue should be taken as a sub-sample. Whole eyes and bones from rodents and rabbits usually do not have to be pre-cut prior to fixation but may require longer than 48 hr of fixation. An exception for bones is that the skull and spinal cord should always be opened and the nervous tissues removed for direct fixation. Filling the lumen of hollow tissues (lung, intestine, and blood vessels) with fixative will also allow larger pieces of tissue to be successfully fixed. Blood and excreta reduce the effectiveness of fixatives so any surface contaminants should be rinsed off of tissues prior to placing in a fixative.

The most common fixative solution for tissues is 10% neutral buffered formalin, but other buffered fixatives such as paraformaldehyde (perfusion), glutaraldehyde (electron microscopy), and for eyes, modified Davidson's (Latendresse et al. 2002) may be appropriate. Alcohol (70% or greater) is a preservative rather than a fixative, but it may be a component of special fixatives including those for immunohistochemistry. Tissues are normally fixed at room temperature for 24–48 hr, but less or greater time may be optimal for small or larger tissues, respectively. A rotation table helps to keep the solution moving gently between the tissues and helps to prevent the tissues from compacting at the bottom of chemically resistant collection jars. Changing out the formalin half way through the optimal fixation time also helps to improve the quality of fixation. Sufficient fixative should be used to provide a 1:20 ratio of tissue/fixative, but never less than a 1:10 ratio. Conical centrifugation tubes seldom provide sufficient volume for fixation, and tissues will invariably be trapped in the conical end and may autolyze. Tissues should be trimmed for histologic processing as soon as fixation is complete. Excess fixation time is particularly adverse for formalin-fixed tissues used for immunohistochemistry (Sect. 9.4.5.4), and sectioning artifacts (e.g., fracturing) are often present in tissues held long term in fixatives.

9.4.4 *Tissue Trimming*

Illustrated guides for tissue sampling and trimming have been published for tissues of rodents (Kittel et al. 2004; Morawietz et al. 2004; Ruehl-Fehlert et al. 2003) and pigs (Albl et al. 2016). Depending on the list of study required tissues, not all tissues that are collected and fixed will need to be trimmed and histologically processed. Once fixation is complete, the protocol-required tissues need to be trimmed down in size to ensure that they will fit histologic cassettes used for tissue processing (Knoblaugh and Randolph-Habecker 2017). Unless the device was separated from the tissue, tissues from the test site should be trimmed to retain the tissue/device interface. Tissues should be trimmed using sharp and disposable razor or necropsy blades appropriate to the size of the tissue. Cutting blades should be changed frequently and the cutting surface regularly cleaned to minimize artifacts and contamination. All tissues should be consistently trimmed from the same part and area of the fixed tissues and in the same orientation to assist in making direct comparisons between sites and animals. Consistent sectioning is particularly important for skin or tissues that tend to curl during fixation. These tissues should be fixed as flat as possible (Sect. 9.4.2.3) to ensure that the tissues are not trimmed tangentially which changes the proportions of the different tissue compartments and presents structures such as hair in cross-section rather than longitudinally. Tissues are placed in labeled cassettes for subsequent processing, embedding, and sectioning by the histology laboratory.

Identification of the tissue/device interface may be difficult unless the tissues were marked during implantation (Sect. 9.4.2.3). For intraosseous metal implants and eye implants (Short 2008), the implant interface may not be readily visible. Marking or imaging with positional data (Roberts et al. 2013) provides the location of implantation which is critical to reducing or preventing the need for blind trimming. If the appropriate tissue is not present on the slides, the pathologist will need to request recuts, deeper sections, or that additional tissue be retrieved from the fixative. Any diagrams of the trimming locations and orientation should be archived with the study and made available to the histopathologist. After all tissues are trimmed, the remaining tissue should be preserved for the duration of the study or until GLP retention policies are met.

9.4.5 *Histology*

9.4.5.1 Pre-study Preparations

The study director, toxicologist, pathologist, and histotechnologist will benefit from being able to see and also feel the native biomaterials and the composition of a finished medical device. Photographic examples of common biomaterials and medical devices are included in numerous individual biomaterial-specific publications and in reviews (Funk et al. 2018; Goad and Goad 2013; Hassler et al. 2011; Long 2008;

Ratner et al. 2013), or the sponsor should be asked to provide a sample. The pathologist and histotechnology laboratory should always be offered a sample of any novel biomaterials or materials with novel processes (e.g., biospinning) so that the hardness and suitability of standard histologic processing, embedding, and sectioning processes can be tested and to determine if the biomaterial is visible in hematoxylin and eosin (H&E) tissue sections. Polymers or soft materials can be stabilized in these preparations by creatively inserting or wrapping the material in skeletal muscle or other suitable tissue purchased from a grocery store. Embedding the biomaterial within a block of solidified agar may not be useful as both the agar and biomaterial may disappear during histologic processing and will not provide the needed contrast. Using the proposed fixation and trimming methods, the biomaterial within a tissue can also be prepared by the histopathology laboratory and examined by the histopathologist. Adjustments can then be made to the histology methods prior to receiving the often irreplaceable study tissues.

9.4.5.2 Tissue Embedding and Sectioning

An algorithm for selecting embedment, cutting procedures, and stains for biomaterials has been published (Alves et al. 2019).

Formalin-Fixed and Paraffin-Embedded Tissues

Trimmed tissues are processed by dehydration through graded alcohol and then embedded in a block of paraffin wax (Callis et al. 2002). Thin sections (5–6 μm) of the tissue blocks are cut onto a glass microscope slide using a microtome, deparaffinized, and stained, and the tissue is covered with a glass coverslip. Ultrathin sections (2–3um) or larger sections (De Jong et al. 2005) can be cut to increase cellular detail for microscopic examination, or thick sections (5–15um) of unstained and stained sections can be cut for laser capture microdissection. Serial (adjacent) or step sections may be necessary to capture the best histologic sections for the pathologist to evaluate. Extra-large paraffin blocks can be used to embed a small whole mouse, whole tissues (eyes), and large implants in toto. These tissue sections are cut from the blocks with a specialized sledge (slab) microtome onto extra-large microscopic slides. Flash frozen tissues for immunohistochemistry are sectioned in a low temperature cryotome (Sect. 9.4.5.4) and thin sections placed on specially charged slides to enhance section adherence for staining. All nonstandard methods used in preparation of the slides should be documented for the report.

Resin Embedment

Hard tissues, such as bone and teeth, and metal or other hard devices cannot be directly trimmed or paraffin embedded and sectioned, and resin embedment is preferred (Long 2008). Bone and teeth are usually fixed and then softened by decalci-

fied (formic acid or EDTA) for paraffin embedment (Sanjai et al. 2012). Hard tissues and devices, some firm devices, and critical tissue/device interfaces should be cut down with a saw and small sections embedded in plastic resin. Thin sections are cut using a diamond sawing or grinding of the resin block with a polishing machine (Caropreso et al. 2000; Long 2008; Malik et al. 1998; Ramot et al. 2016; Rousselle et al. 2019).

Troubleshooting Histology Preparation of Implants in Tissues

Firm biomaterials embedded in paraffin may tear or be ejected during microtomy and may have to be re-embedded or a new sample prepared. Firm biomaterials (e.g., contact lenses) may distort when embedded in certain resins. Novel biomaterials should be pretested for the suitability of the histologic reagents and procedures. Polymers seldom stain with the standard histochemical dyes, and if they do, they consist of lightly stained amorphous material. Empty spaces will be left as a placeholder representing a negative image of the polymer. Some polymers are dissolved by solvents during processing or, like polylactic acid and poly(lactic-co-glycolic) acid, dissolve during staining. Processing may also damage the polymer, change its appearance in tissue sections, or create artifactual fragments. Implants that are successfully sectioned may not adhere properly to the glass microscope slides and may be lost before or during staining. Positively charged slides, used for immunohistochemistry, may help to stably attach the thin section to the slide.

9.4.5.3 Tissue Stains

Stains are used on tissues to provide cellular differentiation and contrast for microscopic evaluation. Standard H&E staining and special histochemical stains are applicable to formalin-fixed paraffin-embedded tissues. These histochemical stains are also applicable to staining resin-embedded tissues, although modified stains may be required (Alves et al. 2019; Long 2008; Malik et al. 1998; Rousselle et al. 2019). Metals, some textiles, fibers, glass, and ceramics may be readily visible or have edge contrast without absorbing histochemical stains. Other than polyaryletheretherketone (PEEK), most polymers will not stain or absorb very little stain, and polymers may be dissolved by the tissue processing and staining procedures. Occasionally polymers may stain with connective tissue stains or special stains in combination with polarized light (dermal fillers stained with Alcian blue and dental/bone material stained with von Kossa stain). Some polymers autofluoresce, and this fluorescence can also be quenched with Sudan black (Jaafar et al. 2011). More difficult is the visualization and identification of degrading fragments or particles of damaged biomaterials. Permanent metal and polymer particles may be birefringent under polarized light and polyethylene debris stains with oil red-O stain (Bauer 1996; Hansen et al. 2002; Krenn et al. 2014; Morawietz et al. 2006; Schmalzried et al. 1993; Shea et al. 1996). For non-staining biomaterials, special histochemical

stains are useful to highlight and characterize the reaction in the tissues and identify device-associated infections. Special stains such as a modified gram stain are useful to identify and confirmation bacterial infections (Becerra et al. 2016). Contrasting stains may also outline the edge of polymeric materials. Commonly used histochemical stains include Masson's trichrome, Masson-Goldner trichrome, and toluidine or Alcian blue (connective tissue and bone), Safranin-O (cartilage), Movat's pentachrome (blood vessels), periodic acid-Schiff (carbohydrates), and von Kossa (mineral). These special stains will differentially stain specific tissue compartments that may not be otherwise visibly differentiated, highlight fibrotic reactions and encapsulation, and may help to differentiate intact and fragments of biomaterials from tissue reactions (e.g., Prussian blue stain for hemosiderin tissue reactions and von Kossa for mineral changes from bone implants) (Alves et al. 2019; Krenn et al. 2014; Morawietz et al. 2006; Rentsch et al. 2014; Rodriguez et al. 2014; Schwartz et al. 2004).

9.4.5.4 Immunohistochemistry

Identification and localization of cells, subcellular and molecular features, and tissue response characteristics can be highlighted in tissues using labeled antibodies against tissue antigens. If antigenic epitopes are retained or can be unmasked, formalin-fixed tissues can be a suitable investigative matrix. Labile antigens may require use of flash frozen tissues (Sect. 9.4.2.8). As with histochemical stains, immunohistochemistry is most applicable to characterization of tissue changes and cellular responses rather than identification of biomaterials (De Jong et al. 2005; Diller et al. 2015; dos Santos et al. 2016; Funk et al. 2018; Krenn et al. 2014; Morawietz et al. 2006; Rentsch et al. 2014), but immunohistochemistry will also monitor biologically derived medical materials (Brown et al. 2015; DeLustro et al. 1986).

9.5 Histopathology Evaluations

A recent special issue of *Toxicologic Pathology* provides both general and device-specific information for histopathology of, and special techniques for biomaterials and medical devices (Rouselle and Paulin 2019). Histopathology consists of qualitative (subjective), semiquantitative (subjective), and quantitative (objective) evaluation or microscope slides and tissue images, of local, systemic, and confounding changes. There is a tendency to assume that the semiquantitative scoring schemes in Annex E of the ISO 10993-6:2016 (International Organization for Standardization 2016) are the only microscopic evaluations needed for local biocompatibility implantation and finished medical device studies. While convenient, the minimalistic to no qualitative content that is provided by these scoring schemes does not properly evaluate the entirety of the biologic responses to biomaterials and the

safety and efficacy of finished medical devices. Histopathology evaluations should not be restricted to semiquantitative scoring based on ISO 10993-6:2016 templates, and peer and expert reviews should be used to confirm the accuracy of the reported pathology findings. Masked (blinded) study evaluation should be used appropriately and only when warranted. Reporting and interpretation of the pathology findings should include correlation of the gross and histopathology findings with organ weight alterations, clinical observations, clinical pathology observations, and any related imaging or analytical tests.

9.5.1 ISO 10993-6:2016 Limitations in Qualitative and Quantitative Histopathology

Historically, histopathology evaluations of biomaterials and medical devices have not been as rigorous as that applied to pharmaceuticals and biologics. In these other therapeutic categories and even combination drug/device or biologics/device evaluations, toxicologic pathology evaluations start with recording morphological diagnoses, followed by or in parallel scoring using a template with well-defined numerical limits. For safety (toxicology) studies of biomaterials, it is not unusual to see full histopathology evaluation and reporting of all other tissues, but surprisingly, the tissue changes in the implanted test site may only be evaluated by a scoring scheme based on ISO 10993-6 Annex E templates.

The use of simple scoring schemes (ISO 10993-6 Annex E) as the sole record of microscopic changes for biomaterials and medical devices is seldom adequate and should not be relied upon for characterization of the tissue response. Tissue response scores for implanted biomaterials using the ISO suggested scoring templates are incomplete unless footnotes and other detailed descriptive notes are included to fully capture the breadth and depth of tissue changes. Reporting the qualitative features with interpretation enhances the reaction score (formerly, irritancy score) of the test versus control material and provides a full histopathology accounting and risk assessment. Detailed reporting of the qualitative morphologic diagnoses is particularly useful to support the interpretation of the biomaterial reaction scores when the physiochemical characteristics of the USP polymer control material differ from that of the test item as the test item may score better or worse than the control material (Elmore et al. 2017; Schuh 2008). Templated responses such as fatty infiltration are difficult to apply in subcutaneous implantation and in deep wounds. Subcutaneous fat is anatomically normal, and adipocytes may be captured and readily incorporated into the edge of implant and reparative sites of full-thickness skin wounds and subcutaneous implants. Scoring templates also do not directly capture bystander reactions in tissues such as reactions to surgically place sutures or staples (which are themselves medical devices), inflammation, and degeneration of nerves, blood vessels, and fat in subcutaneous or intramuscular implantation site. Wound dressings placed in full-thickness skin wounds often have two separate foreign body responses:

one associated with the degradation of the wound dressing material and the second associated with hair fragments (hair granulomas) inadvertently deposited in the wound bed during wound creation or study procedures. To properly characterize and interpret the tissue response, the FBR to the hair should be reported separately from the FBR to the medical device. Examination of local draining lymph nodes is also useful to properly interpreting the local inflammatory response in tissues (Sect. 9.4.2.7), and this examination should be interpreted relative to the tissue and not using an independent scoring scheme. Focusing on response scores also does not usually document problems with the tissue sections (tangential or incomplete sections), the differences between heterophils (rabbits) and granulocytes (most animals), changes known to be spontaneous findings for the tissue, patterns of inflammatory changes (local cellular clustering) and small reparative details (collagen fiber orientation, cellularity and density, and breakage or loss of vascular elastic lamina), and shape and size of degraded material that may be an important but subtle feature of the biologic response. Scores may also not incorporate tissue changes that may be biomaterial-induced metaplasia or spontaneous/background changes in that tissue leading to false-positive reporting of a potentially unacceptable tissue change (Table 9.1; metaplastic bone formation versus os cordis in the heart). The incidence and severity of concurrent spontaneous or background changes may change after surgical implantation. Complex devices with multiple biomaterials usually have different reactions to the different biomaterials, and combining the responses as a single score is seldom informative. Also, unusual findings such as tertiary lymphoid tissue (Sect. 9.7.3.1) may appear only months to years after study initiation, and if interim samples (Sect. 9.4.2.1) have been evaluated, the original scoring scheme may have to be abandoned or amended part way through the study, and the microscopic evaluations repeated to incorporate the new finding. Tertiary lymphoid tissue is often preceded by subtle accumulations and organization of lymphocytes that would not be identified early by typical scoring templates.

Characterization of tissue responses includes all reactions, and ignoring inconvenient features of the response because they are not part of typical scoring templates leads to underreporting and potentially to mis-categorizing compatibility or risk. A simple scoring scheme is only justified to rank or screen the ability of the tissue or body to tolerate one or more biomaterials or multiple procedures or production systems. To properly evaluate the toxicity, safety, and efficacy of pivotal studies, a qualitative histopathology evaluation with morphologic descriptions should be applied to all tissues including the test site. The scoring scheme should supplement, rather than supplant the qualitative (subjective) evaluation.

9.5.2 Primary Histopathology Evaluations

Toxicologic or veterinary pathologists are well trained in evaluating a local FBR, in systemic tissue responses, and in identification of tumors. The typical local tissue response is a local FBR and sometimes the induction of tumors, but systemic

response or responses in tissues distant to the original implantation site are infrequently report. The details of local and systemic inflammatory responses are summarized in Table 9.2, and tumorigenicity is summarized in Sect. 9.7.2.3. Although familiar with the appearance of the FBR to common surgical materials such as sutures, accidentally introduced foreign bodies, microorganisms, and artifacts (McInnes 2005), the histopathologist may not be familiar with the appearance of some biomaterials and medical devices in tissues. The photomicrographic appearance and responses to short- and long-term implanted biomaterials and medical devices have been documented in numerous publications (Alves et al. 2019; Funk et al. 2018; Goad and Goad 2013; Greenwood and Dearman 2012; Jessen et al. 2018; Krenn et al. 2014; Krenn and Perino 2017; Leigh Perkins 2010; Long 2008; Morawietz et al. 2006; Pierce et al. 2009; Ramot et al. 2016; Rentsch et al. 2014; Rodriguez et al. 2014; Rousselle et al. 2019; Tellez et al. 2017). As suggested in Sect. 9.4.5.1, unique materials should be pretested in histology procedures which will provide the histopathologist with a chance to preview the biomaterial in a tissue.

9.5.2.1 Histopathology Qualitative (Subjective) Evaluations

Tissue response characterization of biomaterials and medical devices is enhanced by including subjective histopathology evaluations that properly capture details of the tissue change as a morphological diagnosis, lesion age (acute, subacute, chronic, chronic-active), severity (minimal, mild, moderate or marked, or similar modifiers), multiplicity (focal, multifocal, diffuse, locally distributed), laterality (unilateral, bilateral), and findings segregated by tissue compartments and separating out bystander or secondary lesions. Subjective grading applied to the tissue changes may focus on age, severity, and multiplicity and may not use the same criteria used for a subjective semiquantitative scoring schema (Sect. 9.5.2.2). Confounding surgical or procedural related and spontaneous or background changes should also be described separately to help support the interpretation of the tissue responses (Shackelford et al. 2002; Ward and Thoolen 2011).

Local Implantation Sites

Interlacing fibers of some synthetic fabric implants or sutures, solid dark metallic or metal-coated implants, granular ceramic and glass implants, solid dark fragments or spicules of some bone substitutes, and PEEK polymers can be identified in microscopic sections by appearance, shape, and color (Alves et al. 2019; Goad and Goad 2013; Long 2008; Moya et al. 2016; Pierce et al. 2009). Although firm biomaterial may survive tissue processing, during sectioning, they may eject from the block and deform or tear the tissue, and the biomaterial may not only be partially retained in the tissue section. Polyurethane foam already contains clear spaces when viewed in tissues and typically degrades with progressive scalloping and fracturing that leads to clear spaces with multiple geometric forms (Rodriguez et al. 2014). Autograft or

Table 9.2 Local cellular and secretory foreign body responses with corresponding histopathology findings of biomaterials and medical devices[a]

Event	Inflammatory response	Histopathology findings
Insertion, contact, or implantation	Breach of mechanical defenses of the body with physiologic and metabolic disturbances	Tissue loss or damage to tissues or organ
Hemostasis and acute inflammation and provisional matrix (seconds to 1–3 days)	*Activation*: Vasculature, adhesion molecules, chemotactic factors, and upregulation of innate immunity including damage-associated recognition pattern (DAMP) and pathogen-associated recognition pattern (PAMP) receptors *Secretions*: Fibrinogen, vasoactive peptides, coagulation components, complement factors, histamine, chemotactic factors, chemokines and cytokines, proteases, free radicals, and growth factors	*Vascular*: Edema, hemorrhage, blood clot or thrombus formation, and blood protein (albumin, fibrin, complement, fibronectin, vitronectin, γ-globulin) adsorption onto the surface of the device form a provisional matrix *Cells*: Activation of tissue resident macrophages, mast cells and platelets, influx of neutrophils/heterophils/eosinophils, lymphocytes, natural killer cells, monocytes (macrophages), and dendritic cells, and phagocytic cells try to remove the foreign body by phagocytosis
Subacute inflammation and repair (usually 3–14 days)	*Activation*: Acquired immunity (cell-mediated and antibody) with upregulation of T and B cells and antigen presentation; fibroblast proliferation *Secretions*: As for acute inflammation; secretion of fibrinolytic compounds and immunoglobulins (antibody)	*Vascular*: Resolution of edema, hemorrhage (heme pigment will persist), and blood clots; new blood vessels form (neovascularization) *Cells*: Variable numbers of granulocytes (neutrophils/heterophils and eosinophils), macrophages, and lymphocytes with accumulations around the biomaterial; occasional plasma cells; mononuclear macrophages engulfing biomaterial are often vacuolated and start to form multinucleated giant cells (MNGC); fibroblasts increase; osteoblasts and osteoclasts increase in bone *Biomaterial*: Visible biodegradable materials show variable changes dependent on biomaterial type (decreased size and mass, tinctorial changes, scalloping, irregular shapes, fragmentation, macro- and microparticles, amorphous or granular material, crystals); similar changes in non-staining biomaterials inferred from outlines of phagocytic cells surrounding the empty spaces; permanent biomaterials with no visible changes *Repair*: Early reparative processes with proliferation of granulation tissue (fibroblasts and neovascularization) that replaces the provisional matrix; differentiation of fibroblasts to myofibroblasts, increased production of collagen; an early fibrotic capsule may be organizing; new bone may form

(continued)

Table 9.2 (continued)

Event	Inflammatory response	Histopathology findings
Event	Inflammatory response	Histopathology findings
Chronic inflammation, repair, remodeling (weeks to indefinite), and complete resolution (replaced by normal tissue or fibrosis or encapsulation)	*Activation*: Acquired immunity (cell-mediated and antibody) with upregulation of T and B cells and antigen presentation continues or is diminishing *Secretions*: Downregulation of most except for those related to tissue repair and growth factors	*Cells*: MNGC predominate if biomaterial is still present; MNGC very large with ≥10 nuclei and may contain biomaterial in intracytoplasmic vacuoles; variable numbers of granulocytes, macrophages, lymphocytes, and plasma cells around the biomaterial; eventually elimination of most inflammatory cells although a few lymphocytes, plasma cells, and macrophages/MNGC may persist indefinitely; fibroblasts and myofibroblasts may be organizing around the implant; osteoblasts increased around bone *Biomaterial*: Depending on time point, degradation may be active or completed (biomaterial eliminated); encapsulation may inhibit further degradation of trapped residual biomaterial; permanent biomaterials may show some degradation and particles appear around tissue/device interface *Repair*: Granulation tissue remodeling; fibroblasts and myofibroblasts organizing a capsule around material that cannot be removed; fibroblasts fill tissue voids that cannot be repaired; new bone organizing and may be forming osteophytes; adjacent damaged tissues are repaired or replaced *Remodeling*: Granulation tissue and collagen remodeled to fibrosis; mature fibrotic capsule walls off residual biomaterials and inflammation; tissues that cannot be repaired are replaced by fibrosis (scarring); new bone is organized
Chronic-active inflammation, repair, and remodeling (weeks to indefinite)	*Activation*: Acquired immunity (cell-mediated and antibody) with upregulation of T and B cells and antigen presentation continues *Secretions*: Downregulation of most except for those related to tissue repair and growth factors; metal ions may be released	*Cells*: Similar to chronic inflammation except neutrophils/heterophils and eosinophils may persist along with lymphocytes, plasma cells, and macrophages/MNGC *Biomaterial*: Incompletely eliminated; particulates may be forming *Repair and remodeling*: Incomplete

allograft tissue substitutes are often difficult to separate from normal cells and connective tissue (Orenstein et al. 2012). Some polymers leave mostly empty spaces (negative images) with a granular material or a slightly translucent or discolored amorphous material that is retracted away from the wall surrounding the space, or there may be a change in light refraction at the edges of or in folds of the biomaterial

within the space. Most polymers will not be visible in tissue sections using H&E or special stains as polymers can be dissolved during the tissue and staining processes (Sect. 9.4.5.2). This frequently leaves empty spaces of varying sizes and shapes where the polymer or fragments formerly resided. These empty negative images or visible medical materials will generally be surrounded by granulocytes (neutrophils or heterophils and eosinophils), lymphocytes, tissue resident macrophages, and multinucleated giant cells (MNGC). The phagocytic cells are often large and activated and contain cytoplasmic vacuoles that are clear or finely granular, which may represent the fragments of the degrading biomaterial. In addition to empty spaces representing a negative image of biomaterials that do not survive processing, empty spaces may need to be differentiated from fat vacuoles entrapped in the tissues or from other causes of intracytoplasmic vacuolation. Fibroblasts are also capable of phagocytosis, and biomaterial particles may be identified in intracellular spaces in these cells (Woodward et al. 1985). Mononuclear cells (neutrophils/heterophils and lymphocytes) are long-lived and may persist in the tissues, particularly if any residual biomaterial or particle remains loose or is encapsulated in the tissues. Granulation tissue and eventually fibrosis will accompany the inflammatory reaction. Tissues adjacent to the implant (blood vessels, nerves, adipose tissue, muscles, cartilage, bone, and epithelium) may also show evidence of damage and reparative processes (degeneration, necrosis, abscesses, apoptosis, hemorrhage, hemosiderin, hypertrophy, hyperplasia, and metaplasia). The fate of the biologic response and biomaterial in tissues includes complete or partial elimination of biodegradable materials, stable tissue/device interface with permanent implants, complete or partial encapsulation, partial or complete termination of the inflammatory response, unresolved chronic-active inflammation with residual biomaterial, a partial or complete return to form and function, and/or scarring (Table 9.2, Figs. 9.1 and 9.2, and Sect. 9.7). Confounding inflammatory reactions include sutures and staples used to place the device, infections from contaminations during or post-surgery (Sect. 9.8.3.1), and displacement or migration of biomaterials and medical devices (Funk et al. 2018).

Lymphoid Tissues

One or more lymph nodes draining the local implant area should be compared with more distant lymph nodes, or in the case of skin or dermal implants, the peripheral lymph nodes are compared to internal lymph nodes. As drainage patterns are interconnected, care must be taken to ensure that the comparator lymph node is well out of range of the test site (e.g., for a wound dressing on the back, collect the inguinal and popliteal lymph nodes, and compare to the mesenteric rather internal iliac lymph node as the internal iliac receives drainage from the inguinal lymph node in the periphery). For large skin wounds on the back, both posterior and anterior lymph nodes should be sampled to cover the lymphatic drainage area (Sect. 9.4.2.7). If a reaction is present in the lymph nodes, the different compartments of the lymph node should be described separately, and an enhanced histopathology evaluation may be warranted (Elmore 2006b). Histopathology changes in the lymph nodes can

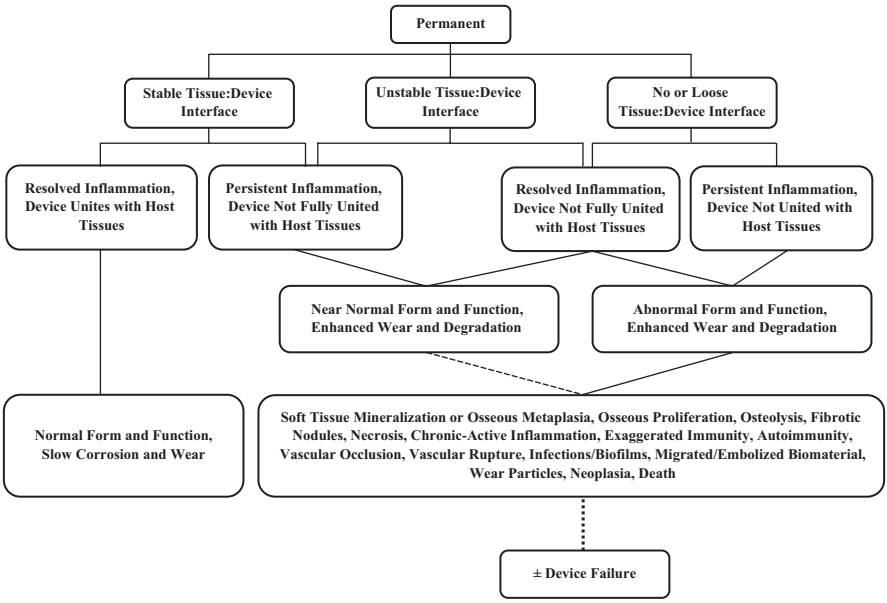


Fig. 9.1 Biologics fates and adverse sequelae to permanent biomaterials and medical devices

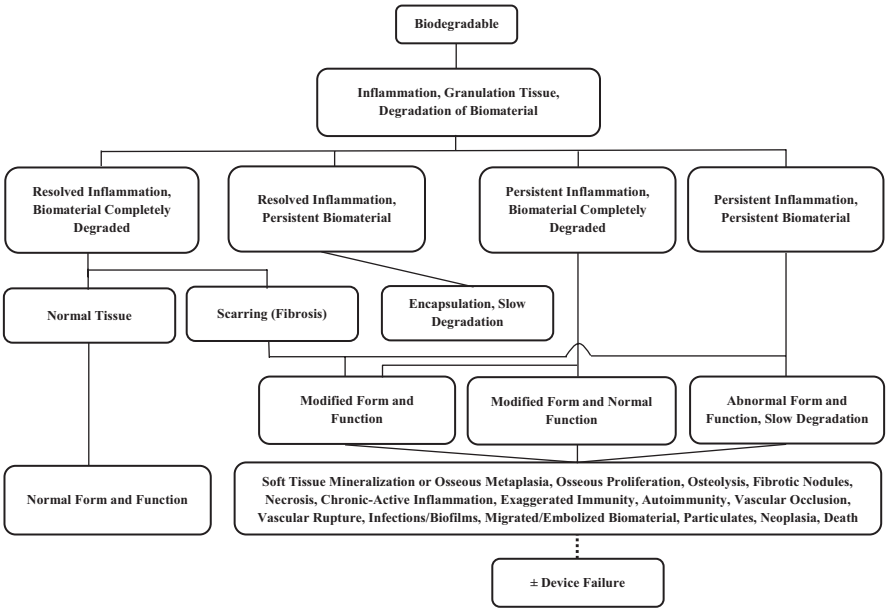


Fig. 9.2 Biologic fates and adverse sequelae to biodegradable biomaterials and medical devices

Table 9.3 Local and systemic immune-mediated and nonimmune (complement-mediated) responses to biomaterials and medical devices

Reaction	Inflammatory response	Mechanism and histopathology findings
Immediate-type hypersensitivity	IgE antibody-mediated anaphylaxis or allergy	Re-exposure to antigens simulates immunity Local (dermal or mucosal) or systemic (multiorgan), nonspecific, and inconsistent histopathology including acute inflammation involving lymphocytes, mast cells, and leukocytes
Delayed-type hypersensitivity	Cytotoxic T cell- and macrophage-mediated (antibody independent)	Memory T helper cells activate macrophages Nonspecific inflammation and necrosis in tissues
Autoimmunity	Auto-antibodies (self-antigens)	Antigen-antibody immune-complex disposition in one or more tissues (e.g., blood vessels, kidney, or skin) with local or systemic inappropriate or exaggerated immune response Inflammation, necrosis, and deposits of amorphous immune-complex material in tissues
Pseudoallergy	Complement activation-related pseudoallergy (CARPA) to nanoparticles	Nonimmune complement activation Vascular leakage identified as perivascular edema, perivascular and intravascular accumulations of leukocytes, and sometimes visible degranulation of mast cells (similar to histopathology findings for cytokine release syndrome)

be compared to changes in the primary (thymus and bone marrow) and secondary lymphoid tissues (spleen and mucosa-associated lymphoid tissues) to determine if the local reaction to the biomaterial has become a systemic reaction, if there is a confounding reaction to stress (Everds et al. 2013), or if an immunotoxicology evaluation (Table 9.3) should be considered. Fragments migrating from biodegradable biomaterials and wear particles from permanent biomaterials may be identified in regional draining and distant lymph nodes using histochemical stains, polarized light (De Jong et al. 2005; Shea et al. 1996), and autofluorescence (Sect. 9.4.5.3), but spectroscopy and molecular characterization of fragments and particles (Sect. 9.6.3) from extracted tissues is a more accurate identification method (Bauer 1996; Frydman et al. 2017). Migration versus encapsulation of microsphere or dermal fillers may depend on the size, implant location, and carriers (Lemperle et al. 2004).

Regional and Systemic Sites

Systemic reactions to biomaterials and medical devices are not well documented. However, direct systemic reactions have also not been thoroughly investigated as many nonclinical studies report data only from local implantation sites and either do not conduct or do not report systematic whole body histopathology. Regional and systemic procedure-related surgery and post-surgical problems, device failures, and infections are recognized for both nonclinical and clinical studies of medical devices.

Clinical adverse events that are reported for medical devices are indicative of the potential for systemic reactions to medical devices, even though little evidence exists for these responses in nonclinical studies. IgE-mediated and delayed-type hypersensitivity reactions, autoimmunity, and inflammatory disorders have been reported for metals, radiocontrast agents, and other biomaterials (Teo and Schalock 2016; Wawrzynski et al. 2017), but there is no evidence for these systemic responses as a sequelae to ruptured silicone breast implants (Institute of Medicine 1999; McLaughlin et al. 2007). Pseudoallergy (complement activation-related pseudoallergy [CARPA]) is a complement-mediated reaction that occurs with nanoparticles (Szebeni 2012). Systemic immune reactions can be nonspecific, and hypersensitivity is not a diagnosis made by histopathology or predicted by animal models (Table 9.3).

Wear, damage, and failure of medical devices including catheters, wires, and cardiovascular devices have resulted in vascular embolism of device components. Thrombosis with distant thromboemboli is also a sequela of intravascular devices. Collateral damage, immunotoxicity, and damage or failure of the device appear to currently be the major categories for systemic reactions reported for medical devices.

Efficacy Evaluations in Animal Models

Ideally, a complete histopathology evaluation (full tissue screen for local and systemic effects) should be conducted on animal models used to evaluate the efficacy and safety of medical devices. As many animal models are developed as research tools in species and strains of animals not commonly used for toxicology testing, the histopathology of spontaneous or background changes and databases on incidence of nonproliferative and proliferative lesions are often lacking (Table 9.1). This can make histopathology evaluations in these models challenging.

9.5.2.2 Histopathology Semiquantitative (Subjective) Scoring

When used, scoring templates should be specific to the tissue and device, and the score parameters and numerical ranges should be tested in a pilot study (Sect. 9.3.2) and adjusted for the safety and efficacy studies if needed (Shackelford et al. 2002). The parameters and numerical limits of a scoring scheme may also need to be adjusted during later time points in a long-term study, when it becomes apparent that not all evaluable parameters were identified in the short-term study, or if the device is failing. Any qualitative or numerical scoring scheme should be validated, which is a difficult procedure (Rutgers et al. 2010). Generally, what is held out as validation is that the scoring scheme was used in one or more studies that have been published or that it is adapted or is a scoring template from ISO 10993-6:2016 Annex E. Annex E templates are listed as informational, but in practice, these semiquantitative scoring sheets are frequently used without regard to applicability to device- and tissue-specific needs and that these templates are not validated. That differential weighting of tissue response other than those listed in Annex E is also

seldom addressed and justification of the weighting is need to be provided for any semiquantitative analyses. Annex E suggested scoring templates include a general template frequently used for subcutaneous and intramuscular implantation studies and a scheme for neural tissues. Appropriate implantation test sites for individual biomaterials include intra-abdominal, intrathoracic, intravascular bone (including dental implants), ocular, mucosal tissues, and central and peripheral nervous system, but individual biomaterials may not always fit into the templates. It should be noted that other ISO documents are also available for determining the biocompatibility and scoring parameters of dental, ocular, and cardiovascular devices (Schuh and Funk 2019).

General (Shackelford et al. 2002) and numerous semiquantitative or quantitative grading systems have been devised or described for a variety of site-specific implantations including bone, cartilage, cardiovascular, orthopedic, and vaginal devices (Alves et al. 2019; Batniji et al. 2002; Bergknut et al. 2013; Cook et al. 2010; De Jong et al. 2005; Funk et al. 2018; Ionita et al. 2009; Kraus et al. 2010; Masuda et al. 2005; Pierce et al. 2009; Ramot et al. 2015a; Rutgers et al. 2010; Sheth et al. 1996; Tellez et al. 2017; Wilson et al. 2009). Novel biomaterials with unique delivery (injectable and self-forming with exothermic reactions and embolic memory foams) generally resist easy adaptation to ISO 10993-6 Annex E scoring templates (Fellah et al. 2006; Rodriguez et al. 2014). Irrespective of the scoring system chosen, the pathology report should clearly provide the scoring validation or criteria (tested in a pilot study) with justification for any differential weightings that are applied.

Published scoring systems may also not be reflective of our current understanding of inflammation and usually do not assess specific regenerative changes in muscles, blood vessels, or other tissues capable of continuous cell division. Polymorphonuclear leukocytes (PMNs) are a x2 weighted inflammatory cell in the ISO 10993-6:2016 scoring template. PMNs include neutrophils or heterophils, eosinophils, and basophils, which have different functions and distributions within tissues. In general, neutrophils are the predominate PMN included in the score, without regard to the dynamics or relative proportions of inflammatory eosinophils versus neutrophils/heterophils in tissue sections. Basophils are seldom identified in tissue sections and are of lesser concern. Eosinophils are broadly involved in inflammation, their presence is not pathognomonic for hypersensitivity reactions (Kariyawasam and Robinson 2006), and they are involved in tissue and muscle repair (Heredia et al. 2013). This suggests that separation of PMN subtypes in scoring systems is scientifically justified. In rabbits, the problem is compounded by the similar appearance of heterophils and eosinophils in tissue sections.

9.5.2.3 Quantitative (Objective) Morphometry and Stereology

The quantitative (objective) evaluation provides the most rigorous measurement of the tissue responses and tissue repair to the biomaterial or medical device. Morphometry and stereology are complex procedures that should be carefully applied by an experienced and informed operator with the appropriate digital cap-

ture equipment (Boyce et al. 2010; Gundersen et al. 2013). Samples should be collected from consistent locations and with equal frequency. Quantitative sample analysis for medical devices most often includes morphometric quantification for point, dimensional, or area changes in tissue responses or the implanted material and less frequently estimations of residual biomaterial or wear particles (Bergsma et al. 1995b; De Jong et al. 2005; Diller et al. 2015; Gauthier et al. 2019; Rentsch et al. 2014; Rismanchian et al. 2012; Varela and Jolette 2018).

9.5.3 Masked (Blinded) Histopathology

Masking (blinding) the pathologist to the biomaterial, study design, and other study data should be done selectively and generally only for pivotal studies used for risk assessment and business decisions. Unnecessary and inappropriate attempts to remove observational bias from the pathology evaluation by masking the study can be costly in terms of time and money, without contributing to diagnostic accuracy (Neef et al. 2012). In order to improve the confidence of the suitability of the microscopic evaluation and interpretation, a peer or an expert review is often preferable (Crissman et al. 2004).

9.5.4 Peer and Expert Review of Histopathology

Peer review of all microscopic slides from a subset of animals, target organs/tissues, and review of the pathology report is an established mechanism to control observational bias, to increase diagnostic accuracy and interpretations, and to confirm appropriate terminology and lesion grades or scores, in evaluations of pharmaceuticals and biologics (Crissman et al. 2004; Morton et al. 2010). Peer review by a second pathologist (internally or externally to the study pathologist facility) is becoming recognized as a mechanism to also verify and improve the quality of histopathology evaluations and interpretations of tissue responses to biomaterials and medical devices. Although specific guidelines have not yet been established for peer or expert review of biomaterials, the established best practices for pharmaceuticals and biologics can be applied, as long as the review is adjusted for unique biomaterials and that finished medical device may be conducted in animal models of disease and not in the standardized species and strains of animals typical of toxicology studies. Peer and expert review procedures are particularly useful for important risk assessment and business decisions and may be required by government granting agencies. Unusual findings or additional interpretation of the study findings can be provided by a subject matter (organ/tissue or pathological process) review by an individual pathology expert for that topic. The expert can conduct their review on a single study and act as the peer review pathologist. Alternatively, they may conduct their review of a specific microscopic finding across multiple studies and provide a

supplemental interpretive pathology report that does not affect the accuracy or content of the original primary pathology report. Substantive disagreements between the primary and peer review pathologists can be handled by adding in an independent review by a subject-specific expert pathologist or by using a formal pathology working group (Morton et al. 2010).

9.5.5 *Histopathology Determination of Causality of Morbidity and Mortality*

As with any toxicology study, acute and ongoing illness and any deaths (found dead or moribund euthanasia) should be investigated. If possible, a determination should be made by the pathologist if the mortality is directly, indirectly, or not associated with the biomaterial or medical device. Surgical factors (infections and surgical failure), confounding or by-design device failure, pathophysiology at distant sites, animal model failure, and confounding or spontaneous conditions and disease need to be considered. At necropsy, more tissues or organs may have to be collected than those listed in the study protocol, particularly if an abbreviated tissue list was included in the protocol. It may also be desirable to conduct radiography, angiography, echocardiography or microcomputed tomography imaging, and scanning electron microscopy on the explanted device and photographically document the device in situ and when explanted to assist in defining the death. Microscopically, death may be attributed to systemic effects or to effects on organ systems distant to the implantation site (e.g., brain infarcts due to blood clots from cardiovascular devices). A proximate cause of death (multi-tissue inflammation with bacteria) and the ultimate cause of death (bacterial contamination of the implanted device) may need to be specified.

9.5.6 *Histopathology Reporting for Biocompatibility and Biologic Responses*

As indicated above (Sect. 9.5.2), a qualitative descriptive of findings for all tissues including the implantation or test site should be the basis of the histopathology report for safety and efficacy studies, and this detailed format should also be considered for all histopathology studies including ISO 10993-6 biocompatibility local implantation studies. The qualitative report should include individual animal reports (gross and histopathology changes with correlates between gross and histopathology findings) and at least a table of the incidence of gross and histopathology findings by group. A separate table should summarize the correlation between the gross and histopathology findings. Additional tables may include the grading of lesions. If a detailed semiquantitative scoring is applied, these results should include individual animal scores and group means and also delineate the numerical or evaluation criteria used for the scores. These data along with the clinical observations, mortality,

body weight, organ weight, clinical pathology, other data tables, images, or analyses that are evaluated by the histopathologist will be summarized and interpreted in the histopathology report. Ascribing biocompatibility to the results of a single study is generally not warranted. The results of a single histopathology report can determine the local and systemic histopathology changes and identify the ability of the tissue or body to tolerate the biomaterial or finished medical device, but only within the context of the study design. The conclusion to any study should be modified by always stating the criteria and testing limitations used to determine tolerability and the testing procedure (e.g., "the implanted device was well tolerated due to minimal chronic-active inflammation and capsule formation after 3 months of intramuscular implantation in New Zealand white rabbits" or "the wound dressing was considered safe and efficacious in a minipig model of full-thickness skin wounding, as determined by repair of the wound (fibrosis replacing the dermis and complete replacement and coverage of the epidermis), and nearly complete biodegradation of the wound dressing with few residual inflammatory cells (MNGC and lymphocytes)").

9.5.7 Digital Photomicrography

As with gross photographic documentation (Sect. 9.4.1.6), photomicrographs are useful to document the biological response and fate and to correlate microscopic changes to macroscopic observations and in vivo imaging. Photomicrographs are often a valuable adjunct document for regulatory summaries, and as such, photographs should be high-resolution color images of one or more magnifications sufficient to document the findings. For comparison within and between groups, images should be captured from consistent tissue regions and at the same magnification. A legend with the appropriate identifiers such as date, study number, animal number, group, tissue, and laterality should accompany the photograph acquisition number. Incorporation of a numeric scale within the microscope light path is preferable to manually calculating digital magnification. Photomicrographs for regulatory submission should comply with GLP standards and should be presented with no or minimal digital manipulation or post-processing (Tuomari et al. 2007).

9.6 Special Imaging Techniques and Technologies

9.6.1 Improving Histopathology Visualization of Biomaterials

There are a variety of molecular labels that can be incorporated into or administered to animals prior to necropsy to microscopically identify cells and function including tracers for localization and tracking (radiolabels and green fluorescent protein) and tracking dyes and particles (fluorescent proteins, vital dyes, labeled microspheres, and gold particles) and markers for proliferating cells (bromodeoxyuridine) and

bone growth (tetracycline). Light microscopy examination of tissue sections can be supplemented with molecular, cellular, and subcellular changes identified by using immunohistochemistry and in situ hybridization to localize cells and identify cellular content.

9.6.2 Specialized Light Microscopy Techniques

Specialized light microscopy techniques such as polarized or differential interference contrast light, autofluorescence, and laser capture microscopy can be used to help characterize tissue changes. Laser capture microdissection is an accurate and efficient microscopic dissection tool for isolating cell populations from histologic tissue sections. The cells can be further analyzed by molecular, chemical identification, and "omics" assays to characterize inflammatory and fibroblastic tissue responses (Datta et al. 2015).

9.6.3 Electron Microscopy

Ultrastructural evaluations using transmission electron microscopy (TEM) and scanning electron microscopy (SEM) can identify cellular, subcellular, and extracellular changes and characterize surface topography and subcellular localization of biomaterial particles or contaminants (Cheville and Stasko 2014). For TEM, tissues less than $5mm^3$ are typical. Slightly larger tissues, up to $1cm^3$, can be collected for scanning electron microscopy (SEM). The tissues are usually fixed in glutaraldehyde-based fixatives and will not require additional trimming prior to embedment in plastic resins (Alves et al. 2019; Rousselle and Wicks 2008). A laboratory with an ultramicrotome, knowledge of the special sectioning and staining procedures, and the TEM or SEM should be consulted for additional information on the type of tissues suitable for electron microscopy. SEM examination of graft degradation can be problematic as pitting and tearing artifacts can be induced or irregular surfaces exaggerated by water loss and metal coatings during tissue processing. The biggest limitation to using TEM and SEM for pathology examinations is that the examined tissues are small which may require examination of many samples using these time-consuming and expensive tools. SEM is frequently used by bioengineers for proof-of-concept studies, to evaluate the surface of novel biomaterials, and for failure investigations.

9.6.4 High-Resolution Imaging

Comparative high-resolution imaging modalities and their properties applicable to biomaterials have been reviewed (Alves et al. 2019; Appel et al. 2013; Funk et al. 2018; Kraus et al. 2012; Nam et al. 2015; Nyska et al. 2014; Roberts et al. 2013;

Tempel-Brami et al. 2015; Varela and Jolette 2018). High-resolution optical microscopy using confocal and atomic force microscopy and combined and related technologies are limited by use of small tissue samples. Repeated and noninvasive in situ imaging and 3D morphometry of tissues and the body can be documented by radiographs, ultrasound, echocardiography, angiography, computed tomography (CT), μCT, optical coherence tomography (OCT), magnetic resonance imaging (MRI), positron emission tomography (PET), and single-photon emission computed tomography (SPECT). Imaging techniques can be combined and applied to medical devices at necropsy or explantation, and functional images can also document device location and movement and provide repeated in situ measurements to monitor biomaterial fate in individual animals. Raman spectroscopy, attenuated total reflection/Fourier-transform infrared (ATR/FTIR) spectroscopy, and scanning acoustic microscopy imaging can also be used to analyze cellular function (Vegas et al. 2016). Multiphoton intravitreal imaging has recently been used for 3D visualization of inflammatory processes and tissue responses to implanted biomaterials (Stiers et al. 2018; Vegas et al. 2016).

Despite the value of high-resolution imaging in providing repeated and often noninvasive evaluation of biomaterial and medical device function and fate, the specialized equipment and trained operators are often not available in testing CRO, and some instruments will not accommodate large animals. It is not acceptable to temporarily move animals to another facility and then return them to the original facility, particularly for a GLP study. Ultrasound, angiography, echocardiography, μCT, and occasionally MRI and PET may be available in testing CRO. The cost, feasibility, and limitations of using any imaging technique need to be understood prior to incorporating these techniques into a study. In practice, advanced imaging techniques may only be applied during research investigations and proof-of-concept studies.

9.6.5 Microscopic Molecular and Chemical Analysis of Tissues

Matrix-assisted laser desorption/ionization imaging mass spectrometry (Walch et al. 2008) and x-ray fluorescence microscopy (Frydman et al. 2017) can be used to characterize the molecular composition within the tissue response and to identify degrading biomaterial components and particulates in tissue sections (Alves et al. 2019; Frydman et al. 2017).

9.7 Biologic Responses to Medical Materials and Finished Medical Devices

The implanted device affects the host by inducing injury, inflammation, repair, and remodeling at the same time that the biologic system affects the device by altering surgical success and altering the device structural and functional integrity through degradation and/or a stable tissue/device interface. Biologic

responses to implanted or inserted medical materials consist of the sum of the (1) biomaterial physiochemical characteristics; (2) confounding animal genetics, animal model characteristics, and confounding procedural and surgical events (summarized in Table 9.1); (3) local response of the specific tissues around the implant; and (4) response of tissues adjacent to the implant and systemic reactions of the body. Typically, most biologic responses to biomaterials and medical devices are the well-characterized local FBR. Insertion and implantation may also have unintended consequences identified in the tissue sections including nongenotoxic tumor induction. Regional tissue and systemic responses to biomaterials and medical devices are less frequently documented and are discussed in Sect. 9.7.3. Immune- and nonimmune-mediated systemic reactions are summarized Table 9.3.

9.7.1 Modifiers of Tissue Responses

9.7.1.1 Physiochemical Factors

Context of placement and desired function as well as the chemical characteristics of a medical material (Diller et al. 2015; Zhao et al. 2000) contributes to the degree and extent of the tissue responses. The ability of the tissues, organs, or body to tolerate or integrate with compatible materials is also modulated by the size, shape, and/or depot size/mass of the material (Greaves et al. 2013; Orenstein et al. 2012; Ramot et al. 2015a; Shoieb et al. 2012; Sunderman 1989; Thackaberry et al. 2017; Weyhe et al. 2015; Wood et al. 1970), body contact site (Kaminski et al. 1968; Markwardt et al. 2013; Pierce et al. 2009; Veiseh et al. 2015), surface topography (DiEgidio et al. 2014; Veleirinho et al. 2014; Wagenfuhr-Junior et al. 2012), and size of individual fibers (Sanders et al. 2000). Textured silicone gel and saline breast implants increase the risk of breast implant-associated anaplastic large cell lymphoma in women (Loch-Wilkinson et al. 2017).

9.7.1.2 Species-, Strain-, and Sex-Related Effects and Tissue Factors

Size, genetic background, sex, biological and physical differences, orientation, and load distribution are important variables when selecting animals to test biomaterials and medical devices in animal models (Funk et al. 2018). These variables plus stress-induced changes from handling and surgery (Everds et al. 2013) need to be considered in study design and tissue responses (Table 9.1). Anatomic and spatial orientation differences between dedicated bipeds (humans), terrestrial quadrupeds (nonhuman primates as transient and arboreal bipeds), and dedicated quadrupeds (most mammals) may require changes in surgical approach, implantation location, and imaging techniques. Thus, implantation of a device in a quadruped may result in unintentional movement, migration, altered load, shearing, tissue tearing,

mechanical failure, or impingement of the device that will not correlate to tissue responses in humans. Studies are often initiated when test animals are young adults and still growing, and failure of non-scalable devices occurs when the animal outgrows the device (e.g., vascular grafts and valves). Irritation and pain associated with devices may also lead animals to rub or press body parts on housing infrastructure which can lead to confounding tissue reactions and infections. For cutaneous or dermal devices, collars, wrapping, sleeves, or jackets are often used to hold or protect the device. Excess pressure by occlusive wrappings has been found to damage the liver ("corset liver") (Chandra et al. 2015) in rats and the liver (Chandra et al. 2015) and spleen (Schuh, unpublished data) of rabbits. Physiochemical, location, and animal model variables are also important to the characteristics of inflammatory responses and interpretation of microscopic findings of the tissue response (Funk et al. 2018). Marked differences in subcutaneous inflammation and biodegradation have been noted for polyurethane in rodents versus rabbits (Rigdon 1973); a copolyester tissue filler in swine versus rats (Ramot et al. 2015b); hydroxyapatite intramuscular osteoinduction in rabbits, dogs, and baboons (Ripamonti 1996); and calcium phosphate ceramic intramuscular and subcutaneous osteoinduction in rats, rabbits, dogs, pigs, and goats (Yang et al. 1996), and subtle inflammatory differences were found for polydimethylsiloxane implants in male versus female mice (Dalu et al. 2000).

9.7.2 Local Tissue Responses

9.7.2.1 Foreign Body and Tissue Responses

The ideal biomaterial would induce little to no tissue reaction, with complete resorption of biodegradable implants or biostable integration of permanent implants with host tissues, and return to normal form and function. More typical (Table 9.2) is at least a transient and sometimes persistent chronic or chronic-active inflammatory response that may be associated with or without encapsulated residual biomaterial for biodegradable materials and movement or corrosion of permanent medical devices. Often, fibrovascular proliferation (granulation tissue) followed by progressive substitution with collagen and maturing to fibrosis or osteoinduction for bone contributes to encapsulation of foreign material, replacement of missing tissues, and reparative processes (Jones 2015). In some cases, residual biomaterial may incite a granulomatous reaction which is encapsulated, forming a structure similar to a granuloma (Pagán and Ramakrishnan 2018). These responses are the classic FBR that the body uses to respond to and eliminate or harmlessly segregate exogenous materials. The FBR, including cellular and secretory mediators to biomaterials and medical devices, has been extensively investigated and reviewed (Amini et al. 2011; Anderson et al. 2008; Badylak 2015; Davies et al. 2013; Funk et al. 2018; Goad and Goad 2013; Gorbet and Sefton 2004; Hu et al. 2001; Klopfleisch and Jung 2017; Major et al. 2015; Milde et al. 2015; Pagán and Ramakrishnan 2018). Important

features of the general FBR including time course, cellular and soluble immune response, and histopathology findings are generally similar, irrespective of the foreign material. With biodegradable implants, the inflammatory response may be prolonged, and the tissue/device interface is often more limited in permanent devices. This pattern of response to biomaterials consists of overlapping hemostasis, inflammation, repair (including encapsulation), and remodeling to normal tissue or replacement by scar tissue. As indicated above, physiochemical and animal characteristics as well as concurrent disease can modify the microscopic features of the FBR (Socarrás et al. 2014). However, immunohistochemical characterization has identified differences in immunophenotypic responses to degradable (Muhamed et al. 2015) and nondegradable (Konttinen et al. 2014; Nich and Goodman 2014) implants. This suggests that immunohistochemistry assessment may identify important biologic responses that affect the accurate determination of biocompatibility and potential for device failure that is not captured by histopathology evaluation or in vitro assessments.

9.7.2.2 Unintended Tissue Sequelae Identified by Histopathology

A variety of unintended tissue sequelae are cosmetic or result in limited functional deficiencies. Other changes affect tissue or organ form and function that can contribute to morbidity, device failure and a need for removal or a revision surgery, and death (Figs. 9.1 and 9.2). These changes include tissue calcification and mineralization (Schoen et al. 1988) and osseous metaplasia in soft parenchymal tissues (Barbolt et al. 2001; Fernandez-Bueno et al. 2015) and periprosthetic osteophyte formation (Krenn and Perino 2017). Neovascularization within intimal hyperplasia of vein grafts correlates with restenosis, and thrombosis, occlusion, and rupture may affect grafted and stented blood vessels (Baklanov et al. 2003; Sheth et al. 1996; Xie et al. 2015). Fibroblastic and myofibroblastic repair remodeling may also result in contracture of wounds and fibrotic nodules that affect return to normal function (Lee and Kim 2015). With prostheses, wear particles promote persistent inflammation and skewing toward proinflammatory M1 phenotype macrophages which are associated with osteolysis and aseptic loosening of joints (Sect. 9.8.3.2).

9.7.2.3 Carcinogenicity of Medical Materials

Solid-state carcinogenesis or foreign body tumorigenicity for a variety of biomaterials and metals is well recognized. Solid-state carcinogenesis is the induction of sarcomas (benign and malignant) through nongenotoxic mechanisms related to geometry, physical form, and texture of materials. Chronic or chronic-active inflammation associated with persistence, movement, and wear particles of the biomaterials is an important mechanism of foreign body tumorigenicity (Brand et al. 1975a; Carter and Roe 1969; Kirkpatrick et al. 2000; Oppenheimer et al. 1955). Small, porous, granular, or mesh materials often have less inflammation and are less prone

to tumor induction than large, flat, impervious, and textured materials (Amini et al. 2011; Bischoff and Bryson 1964; Brand et al. 1975b). Nongenotoxic events starting with the respiratory burst of oxidation and nitric acid during acute inflammation drive the initial cytotoxic and mutagenic events. Cytokines and growth factors during chronic inflammation promote the formation of tumors (Moizhess 2008). There also appear to be species differences and a possible genetic basis for tumorigenicity of medical materials (Brand et al. 1975b; Goad and Goad 2013; Schuh 2015). Rodents with implants readily develop sarcomas compared to humans, sarcomas are frequent with telemetry or device implants in genetically modified and certain inbred mouse strains, and there is some evidence of sex predominance in inbred mice.

In toxicology studies, fibromas and fibrosarcomas are most frequently identified in rodents after subcutaneous, intramuscular, and intraperitoneal implantation of telemetry microchips and occasionally after subacute and chronic implantation studies of biomaterials and medical devices (Greaves et al. 2013; Ramot et al. 2015a; Shoieb et al. 2012; Sunderman 1989). Although some polymeric and metal materials are classified as possible carcinogens for humans, there is insufficient toxicology information to classify the carcinogenic potential of most biomaterial components and complex medical devices (McGregor et al. 2000). Epidemiological studies have identified a low incidence of tumors in humans (Altman et al. 2018; Brewster et al. 2013; Keel et al. 2001; McLaughlin et al. 2007; Moalli et al. 2014; Pinchuk 1995; Signorello et al. 2001; Visuri et al. 1996). The risk for implant-associated bone tumors around joint replacements is also low, even when manufactured with metals classified as carcinogenic (Witzleb et al. 2006). Breast implant-associated anaplastic large cell lymphoma, a peri-implant T-cell tumor-associated with textured silicone gel- and saline-filled implants, is under investigation (Loch-Wilkinson et al. 2017) and under review by regulatory agencies (https://www.fda.gov/MedicalDevices/ProductsandMedicalProcedures/ImplantsandProsthetics/BreastImplants/ucm239995.htm). There is little available data, but thus far no evidence exists for blood or distant tumors associated with permanently implanted medical devices (Mathiesen et al. 1995). Temporary use devices for radiographic imaging and use of morcellators in the presence of undiagnosed gynecologic tumors are associated with secondary and metastatic tumors, respectively (Pearce et al. 2012; Perkins et al. 2016).

9.7.3 Systemic Responses and Immunotoxicity

Systemic responses to biomaterials and medical devices are rarely reported but are also frequently not investigated. This is compounded by terminology inconsistencies used in published literature that affect the ability to effectively search the medical device literature. The term adverse event, favored for complications associated with pharmaceuticals and biologics, is used less frequently for medical devices (Farrah et al. 2016; Golder et al. 2014). Other than secondary systemic responses associated with emboli, thromboembolism, and bacterial infections, systemic immune reactions are the only reported reactions with both local and distant effects.

9.7.3.1 Local and Systemic Immune Reactions

Another unintended consequence of implanting biomaterials and medical devices is inappropriate and exaggerated immune responses. Immediate and delayed hypersensitivity reactions and autoimmunity (Table 9.3) have been associated with intact or degraded metals, polymers, elastomers, ceramics, radiocontrast reagents, biological materials, and therapeutic impregnated medical devices. Antigenic proteins or haptens and molecular mimicry (cross-reactivity between avocado and latex) may also cause hypersensitivity reactions. Animal models are poor predictors of hypersensitivity in humans, and the histopathology findings are limited and nonspecific. Immune-complex and inflammatory disorders causing tissue and vascular necrosis have been associated with metals (Athanasou 2016; Honari et al. 2008; Lie et al. 2011; Pizzoferrato et al. 1988; Ricciardi et al. 2016; Teo and Schalock 2016; Wawrzynski et al. 2017), but there is no proven association to damaged silicone breast implants (Institute of Medicine 1999; McLaughlin et al. 2007). Metal wear debris, particulates, and ions are associated with chronic local inflammatory responses that may progress to systemic metal hypersensitivity with inflammation and necrosis (Athanasou 2016; Frydman et al. 2017; Tuan et al. 2008). Aseptic joint loosening may be due to hypersensitivity inducing proliferation of osteoclasts that damage the metal and induce cytotoxicity (Wawrzynski et al. 2017). Chronic inflammation may also attract activated macrophages from distant sites to the prosthesis-induced osteolysis and prosthetic loosening (Ren et al. 2011). Pseudoallergy or complement activation-related pseudoallergy (CARPA) is a nonimmunological-mediated reaction that occurs with nanoparticles and results in hemodynamic failure (Dobrovolskaia 2015; Szebeni 2012; Szeto and Lavik 2016).

Post-inflammation, lymphocytes may also persist in tissues in a quiescent state or may organize. Tertiary lymphoid tissue (lymphoneogenesis) can form in nonlymphoid tissues during persistent infections, chronic inflammation, autoimmunity, and cancer. Lymphocytes organize into lymphoid follicle structures with germinal center formation. Tertiary lymphoid tissues have been reported in association with silicone breast and testicular implants (Abbondanzo et al. 1999), subcutaneous polymers (Schuh 2008), joint prosthetics (Ricciardi et al. 2016), and implanted metals particularly gold (Lie et al. 2011), but the functional significance of tertiary lymphoid follicles is unknown.

9.8 Biologic Fates and Adverse Tissue Responses

The potential biologic fate of biomaterials and medical devices is of interest to the regulatory agencies, sponsors, surgeons, and patients. Time course tissue samples collected by noninvasive and invasive imaging, interim biopsies, or interim necropsy can be used to monitor the tissue/device interface, integration of permanent implants, and degradation patterns of biodegradable implants. Such sampling is particularly important for early detection of device failure and to monitor the in vivo

degradation. This information may justify early termination of chronic safety studies in animals, once complete degradation is identified (Sect. 9.8.2.1). Normal host responses to biomaterials, particularly the occurrence of inflammatory macrophages and MNGC, are frequently considered an undesirable response to biomaterials. The amount and persistence of inflammation are critical to controlled and appropriate biodegradation or integration of permanent biomaterials and medical devices. Immune recognition of non-self-antigens causes the body to attack and attempt to destroy and remove any foreign material. If destruction or removal is impossible, an attempt will be made to wall off the material with a fibrous capsule (Jones 2015). If the implanted material promotes inflammation or if the biomaterial persists, then the tissue at the implant site may not return to normal form, and the tissue, organ, or body will lose normal function.

Whether the biomaterial is intended as a permanent replacement or addition to the tissues or organs (Fig. 9.1) or whether the biomaterial is intended as a temporary scaffold or replacement destined for complete biodegradation (Fig. 9.2), determines the potential biologic fates and adverse tissue reactions that contribute to normal or abnormal form and function and device failure. The biological fates and adverse reaction of the biomaterial in tissues include complete or partial elimination of biodegradable materials, stable integration into adjacent tissues, complete or partial encapsulation, partial or complete termination of the inflammatory response, unresolved chronic-active inflammation with residual biomaterial, a partial or complete return to form and function, and/or fibrotic (scar) replacement of tissues that cannot be regenerated. Minimization or modulation of excess tissue reactions is the holy grail of biomaterial and medical device implantation. Biomaterial surface manipulation and coatings with and without eluting drugs and biologics (bioactives) are used to improve attachment or promote integration into bone and tissues, to reduce bacterial adhesion, and to minimize inflammatory reactions and fibrosis (Klopfleisch and Jung 2017; Meng et al. 2016; Morais et al. 2010; Thevenot et al. 2008).

9.8.1 Stable Tissue/Device Interface (Integration) of Permanent Biomaterials and Medical Devices

In reality, permanent medical devices never become fully integrated with the host tissue, organ, or body (Fig. 9.1). Many permanent biomaterials are tolerated long term with little to no adverse changes, even though a low-grade inflammatory response, wear particles, and release of metal ions may still be present during a return to normal function. Permanent biomaterials are primarily degraded through corrosion, frictional wear of articulating components, and movement at the tissue/device interface (biotribology), but enzymes and oxidative activities from the inflammatory response are also involved (Chen and Thouas 2015; Gibon et al. 2017a, b; Shayesteh Moghaddam et al. 2016). Integration is best considered a stable

or fixed tissue/device interface. Creation of this interface may involve some inflammatory-induced surface changes to the biomaterials which help to anchor the interface. A stable tissue/device interface and anchoring into adjacent normal tissues is desirable as long as site function is retained. Adverse tissue/device interface events are illustrated by vascular damage and occlusion or restenosis secondary to placement of intravascular stents, osteophyte overgrowth of joint replacements, and tension created by fibrotic substitution in and along tissues.

Although some published scoring protocols include integration as part of the microscopic scoring (ISO 10993-6 Annex E does not), integration is an interpretative and not a valid microscopic change. An interpretation of integration based on qualitative evaluations and semiquantitative or quantitative scores should always indicate which findings were relied upon for the determination of integration.

9.8.2 Biodegradation of Biomaterials and Medical Devices

Oxidation, hydrolysis, and movement with friction are primarily responsible for degradation of biodegradable materials (Santerre et al. 2005). For biodegradable materials, integration is simply infiltration of inflammatory cells and the progression from granulation tissue to repair and remodeling of fibroblasts and fibrosis, concurrent with resorption of the biomaterials, the culmination of the FBR (Sect. 9.7.2.1 and Fig. 9.2). Differing degradation rates make it difficult to histologically compare different biomaterial timelines if samples are all collected at fixed time points. Collection times based on retention of molecular weight have been suggested as a suitable method to directly compare in vivo degradation. Using this method, poly(DTE adipate) was shown to have the mildest response (inflammation and capsule formation), followed by poly(DTE carbonate) and then poly(L-lactic acid) (Hooper et al. 1998). With biodegradable materials, strategies that focus on retention of tissue function (vascular patency, scaffold stability, and controlled and ordered inflammatory and fibrotic response) appear to be more successful than prolonging the retention of the biomaterial.

9.8.2.1 Determination of Complete Biodegradation and Return to Normal Form and Function

Proof that biodegradation is complete and that there is a return to normal tissue, organ, or body form and function is often requested by regulatory authorities as an endpoint for safety and efficacy studies, particularly for permanent implants and implants in and near vital tissues (brain, heart, liver, kidney, and lung). The final in vivo recovery point is difficult to predict in advance as it depends on the biomaterial type, location of the implant, and host response. Chronic studies to define complete degradation are often greater than 1 year and in large animals may extend beyond 2 years. Biodegradable medical materials generally degrade through hydro-

lysis and cellular oxidative activity within months to years, but with stable or slowly degrading polymers, fragmentation and loss of mass may be incomplete after years in nonclinical studies (Ramot et al. 2015a; Ramot et al. 2016; Rigdon 1973; Sun et al. 2006; Welsing et al. 2008) and after clinical explantation (Bergsma et al. 1995a; Handel 2006; Iezzi et al. 2014; Pinchuk 1995). Comparative in vitro and in vivo degradation studies may show similar or dissimilar degradation time curves, dimensional changes, and biostability, and in vitro degradation cannot predict the appearance and timelines for degradation in vivo (Adhikari et al. 2008; Ang et al. 2017; Bölgen et al. 2005; Ganta et al. 2003; Hooper et al. 1998; Ikarashi et al. 1992; Van Der Giessen et al. 1996; Wolf et al. 2014).

In histopathology studies, proper sample collection (Sect. 9.4.2) is imperative to obtaining representative samples of the implantation site, particularly during the later stages of degradation when serial sections may need to be examined to ensure biomaterial degradation is progressing or is complete. Identification of polymers is often complicated by pathologists having to use a tissue void or negative image as a marker of where biomaterials were, but were then lost during tissue preparations. These spaces change shape over time and may become difficult to differentiate from normal small tissue voids (Sect. 9.5.2.1). As degradation approaches the terminal phase, detection of small fragments of residual biomaterials is increasingly difficult as residual biomaterial macro- and microparticles may still exist within phagocytic vacuoles in the absence of other visible biomaterial in the tissues. Resolution or persistence of the inflammatory response is also variable and may persist indefinitely in small foci after general repair of tissue injury. Inflammation by neutrophils and eosinophils may resolve or be trapped within encapsulated foreign material. MNGC, resident tissue macrophages, foamy macrophages, lymphocytes, and plasma cells are long-lived, and these cells may persist in the center of and within the wall of encapsulated materials. Persistent perivascular lymphocytes are also associated with aseptic loosening of joint prostheses (Ricciardi et al. 2016). Tertiary lymphoid tissue is neogenesis of organized lymphoid follicles in tissues that usually do not contain lymphoid follicles or tissues which are not secondary lymphoid tissues. Tertiary lymphoid follicles have been occasionally reported with chronic inflammation around implants (Abbondanzo et al. 1999; Funk et al. 2018; Ricciardi et al. 2016). Encapsulated foreign bodies can also rupture leading to reactivation of a chronic-active inflammatory response to the formerly encapsulated materials (Kalimo et al. 1996). Biodegradation of materials is characterized by a change in geometry, edge features, color or tinctorial staining properties, and partial (particles, fibers, or granularity) or complete loss of the biomaterial. A reduction or absence of inflammatory cells and mature fibrosis or fibrotic encapsulation is indicative of a terminated reparative process to a foreign body in tissues. However, a few well-encapsulated residual biomaterial fragments, a small population of leukocytes, MNGC, and even tertiary lymphoid tissue within an otherwise quiescent and inactive implant site are still consistent with a completed host response. Tissues, organs, or the body may be restored to normal, modified, or abnormal form or function. Modified or abnormal form and/or function can lead to tissue sequelae that require revision or removal surgery or may result in permanent morbidity, neoplasia, or

even death. A weight-of-evidence approach to histologically determining a stable and a terminal or completed host response to a biomaterial or medical device should examine and identify features of the biomaterial resorption, inflammatory response, repair and remodeling, and other supporting features such as activity in regional draining lymph nodes (Table 9.4.).

9.8.3 Device Failures

9.8.3.1 Infections and Biofilms

Local contamination by pathogenic gram-positive and gram-negative bacteria may progress to septic joint loosening or sepsis and embolization that is difficult to treat, and death may occur if the device is not removed (Von Eiff et al. 2005). Biofilms are surface-associated phenotypically altered bacterial communities and sometimes yeast that attach to extracellular matrices and one another. Biofilm contamination is an adverse event for multiple medical devices including catheters, contact lenses, prosthetic heart valves, intrauterine devices, wounds, and joint replacements (Donlan and Costerton 2002; Phillips et al. 2016). Prevention of infections or biofilms is preferable but can be difficult. Biomaterials with incorporated antibiotics,

Table 9.4 A weight-of-evidence approach to histologically determine a stable and a terminal or completed host response

A stable or completed host response is indicated by:
Biomaterial
Little to no residual biodegradable biomaterial, unless encapsulated
Permanent biomaterials are intact
Inflammation
Few to no neutrophils/heterophils or eosinophils (PMN), MNGC, or mononuclear inflammatory cells, or
Chronic or chronic-active inflammation (PMN, MNGC, lymphocytes, and plasma cells) may be trapped in the center of completely encapsulated residual biomaterial
Tissue repair and fibrosis
Return to normal tissue morphology, or
Complete encapsulation, or
Fibrosis permanently replaces the original tissue at the implant site
Tissue changes are stable and do not progress between two sequential time points weeks to months apart
Regional draining lymph nodes
Regional draining lymph nodes may be activated (germinal centers and increased resident tissue macrophages) but do not have PMN, altered micro-architecture, or loss, or expansion of any lymphoid compartments

Abbreviations: *PMN* polymorphonuclear leukocytes (neutrophils or heterophils or eosinophils), *MNGC* multinucleated giant cells

antimicrobials (silver salts, nitrofurazone, chlorhexidine quaternary ammonium surfactants), antibacterial peptides, and anionic nanoporous hydrogels (Busscher et al. 2012; Hook et al. 2012; Monteiro et al. 2009; Veerachamy et al. 2014) have been tested to control biofilms.

9.8.3.2 Biotribology

Biotribology is the study of friction, lubrication, and wear of interacting tissue surfaces in relative motion. Permanent implants such as metals, ceramics, fabrics, and certain polymers were once considered inert, but it is now clear that these implants can degrade by corrosion, fatigue, and wear associated with friction, wear, and lack of lubrication. Macro-, micro-, and nanoparticles can be shed locally and into lymphatics (Shea et al. 1996), and biomaterials can abrade, crack and delaminate, and release metal ions. Particles can adhere and cause grooving of articulating surfaces and migrate or embolize with adverse consequences for the patients (Siddiqui et al. 2009; Stokes 2009). Wear debris is of particular concern for its association with aseptic joint loosening and peri-implant osteolysis. Particles are phagocytosed in macrophages which become activated and release cytokines that polarize the macrophages to the M1 proinflammatory phenotype. This is thought to enhance osteolysis and subsequent joint loosening (Gibon et al. 2017a, b; Konttinen et al. 2014; Nich and Goodman 2014; Ricciardi et al. 2016; Tuan et al. 2008). Friction and wear also cause cardiovascular devices to fail and contribute to hemolysis and thrombosis (Xie et al. 2015).

9.9 Conclusions

Evaluation of the pathology and histopathology of permanent and biodegradable biomaterials and medical devices is a complex topic that requires understanding and application of regulatory standards and guidances and application of available standardized nomenclature and terminology. It is imperative to include the gross pathologist, histopathologist, and histotechnology laboratory as part of the development team so that they can provide expertise to study design and appropriate tissue endpoints, collection fixation, and preparations. Histopathology evaluations should include a qualitative assessment which can be supplemented but not supplanted by semiquantitative scoring templates as per ISO 10993-6:2016. Use of a semiquantitative score in proof-of-concept studies comparing biomaterials, techniques, and processes is the one situation where only a semiquantitative score is applicable. Histopathology can also be supplemented by quantitative analyses, immunohistochemistry, special imaging techniques, and chemical analysis of tissues. Biologic responses to biomaterials and medical devices are mostly local or regional responses to the implant site and primarily consist of a foreign body response, but hypersensitivity and immune responses may be expressed as systemic abnormalities.

Implantation of foreign material may also result in a variety of unintended tissue consequences that can impact both tissue/organ form and function. The fate of the permanent and biodegradable medical materials is also important. Histopathology determination of a stable or terminated host response to biomaterials requires a weight-of-evidence approach and evaluation of biomaterial resorption, inflammatory response, repair and remodeling, and other supporting features. Finally infections, biofilms, and biotribology (friction and wear) may result in device failure that requires removal or revision surgeries.

References

Abbondanzo, S. L., Young, V. L., Wei, M. Q., & Miller, F. W. (1999). Silicone gel-filled breast and testicular implant capsules: A histologic and immunophenotypic study. *Modern Pathology, 12*, 706–713.

Adhikari, R., et al. (2008). Biodegradable injectable polyurethanes: Synthesis and evaluation for orthopaedic applications. *Biomaterials, 29*, 3762–3770.

Albl, B., et al. (2016). Tissue sampling guides for porcine biomedical models. *Toxicologic Pathology, 44*, 414–420.

Altman, D., Rogers, R. G., Yin, L., Tamussino, K., Ye, W., & Iglesia, C. B. (2018). Cancer risk after Midurethral Sling surgery using polypropylene mesh. *Obstetrics and Gynecology, 131*, 469–474. https://doi.org/10.1097/aog.0000000000002496.

Alves, A., Wanket, L., & Metz, A. (2019). Current considerations in medical device pathology. In J.-P Boutrand (Ed.), *Biocompatibility and Performance of Medical Devices* (2nd ed. pp. 489–543). Cambridge: Elsevier.

Amini, A. R., Wallace, J. S., & Nukavarapu, S. P. (2011). Short-term and long-term effects of orthopedic biodegradable implants. *Journal of Long-Term Effects of Medical Implants, 21*, 93–122.

Anderson, J. M., Rodriguez, A., & Chang, D. T. (2008). Foreign body reaction to biomaterials. *Seminars in Immunology, 20*, 86–100.

Ang, H. Y., Bulluck, H., Wong, P., Venkatraman, S. S., Huang, Y., & Foin, N. (2017). Bioresorbable stents: Current and upcoming bioresorbable technologies. *International Journal of Cardiology, 228*, 931–939.

Appel, A. A., Anastasio, M. A., Larson, J. C., & Brey, E. M. (2013). Imaging challenges in biomaterials and tissue engineering. *Biomaterials, 34*, 6615–6630. https://doi.org/10.1016/j.biomaterials.2013.05.033.

Athanasou, N. (2016). The pathobiology and pathology of aseptic implant failure. *Bone & Joint Research, 5*, 162–168.

Badylak, S. F. (2015). *Host response to biomaterials: The impact of host response on biomaterial selection* (1st ed.). San Diego: Academic Press.

Baklanov, D. V., Peters, K. G., Seidel, A. L., Taylor, D. A., & Annex, B. H. (2003). Neovascularization in intimal hyperplasia is associated with vein graft failure after coronary artery bypass surgery. *Vascular Medicine, 8*, 163–167.

Barbolt, T. A., Odin, M., Léger, M., Kangas, L., Holste, J., & Liu, S. H. (2001). Biocompatibility evaluation of dura mater substitutes in an animal model. *Neurological Research, 23*, 813–820. https://doi.org/10.1179/016164101101199405.

Batniji, R. K., Hutchison, J. L., Dahiya, R., Lam, S. L., & Williams, E. F. (2002). Tissue response to expanded polytetrafluoroethylene and silicone implants in a rabbit model. *Archives of Facial Plastic Surgery, 4*, 111–113.

Bauer, T. W. (1996). Identification of orthopaedic wear debris. *Journal of Bone and Joint Surgery, 79*, 479–483.

Becerra, S. C., Roy, D. C., Sanchez, C. J., Christy, R. J., & Burmeister, D. M. (2016). An optimized staining technique for the detection of gram positive and gram negative bacteria within tissue. *BMC Research Notes, 9*, 216. https://doi.org/10.1186/s13104-016-1902-0.

Bergknut, N., et al. (2013). Intervertebral disc disease in dogs–Part 1: A new histological grading scheme for classification of intervertebral disc degeneration in dogs. *Veterinary Journal, 195*, 156–163.

Bergsma, J., De Bruijn, W., Rozema, F., Bos, R., & Boering, G. (1995a). Late degradation tissue response to poly (L-lactide) bone plates and screws. *Biomaterials, 16*, 25–31.

Bergsma, J. E., Rozema, F., Bos, R., Boering, G., de Bruijn, W., & Pennings, A. (1995b). In vivo degradation and biocompatibility study of in vitro pre-degraded as-polymerized polylactide particles. *Biomaterials, 16*, 267–274.

Bischoff, F., & Bryson, G. (1964). Carcinogenesis through solid state surfaces. *Progress in Experimental Tumor Research, 5*, 85–133.

Blanchard, K. T., et al. (1999). Transponder induced sarcoma in the Heterozygous $p53^{+/-}$ mouse. *Toxicologic Pathology, 27*, 519–527.

Bölgen, N., Menceloğlu, Y. Z., Acatay, K., Vargel, I., & Pişkin, E. (2005). In vitro and in vivo degradation of non-woven materials made of poly (ε-caprolactone) nanofibers prepared by electrospinning under different conditions. *Journal of Biomaterials Science. Polymer Edition, 16*, 1537–1555.

Boyce, J. T., Boyce, R. W., & Gundersen, H. J. (2010). Choice of morphometric methods and consequences in the regulatory environment. *Toxicologic Pathology, 38*, 1128–1133.

Brand, K. G., Buoen, L. C., & Brand, I. (1975a). Foreign-body tumorigenesis induced by glass and smooth and rough plastic. Comparative study of preneoplastic events. *Journal of the National Cancer Institute, 55*, 319–322.

Brand, K. G., Buoen, L. C., Johnson, K. H., & Brand, I. (1975b). Etiological factors, stages, and the role of the foreign body in foreign body tumorigenesis a review. *Cancer Research, 35*, 279–286.

Brewster, D. H., Stockton, D. L., Reekie, A., Ashcroft, G. P., Howie, C. R., Porter, D. E., & Black, R. J. (2013). Risk of cancer following primary total hip replacement or primary resurfacing arthroplasty of the hip: A retrospective cohort study in Scotland. *British Journal of Cancer, 108*, 1883–1890. https://doi.org/10.1038/bjc.2013.129.

Brown, B. N., Mani, D., Nolfi, M. A. L., Liang, R., Abramowitch, S., & Moalli, P. A. (2015). Characterization of the host inflammatory response following implantation of prolapse mesh in rhesus macaque. *American Journal of Obstetrics and Gynecology, 213*, 668.e661–668.e610. https://doi.org/10.1016/j.ajog.2015.08.002.

Budras, K. D., McCarthy, P. H., Fricke, W., Richter, R., Horowitz, A., & Berg, R. (2007). *Anatomy of the dog: An illustrated text*. Hanover: Schlütersche Verlagsgesellschaft mbH &.

Busscher, H. J., et al. (2012). Biomaterial-associated infection: Locating the finish line in the race for the surface. *Science Translational Medicine, 4*, 153rv110–153rv110.

Callis, G., Sterchi, D., & National Society for H. (2002). *Animal processing manual*. Bowie: National Society for Histotechnology.

Caropreso, S., Bondioli, L., Capannolo, D., Cerroni, L., Macchiarelli, R., & Condo, S. (2000). Thin sections for hard tissue histology: A new procedure. *Journal of Microscopy, 199*, 244–247.

Carter, R. L., & Roe, F. J. (1969). Induction of sarcomas in rats by solid and fragmented polyethylene: Experimental observations and clinical implications. *British Journal of Cancer, 23*, 401–407.

Chandra, S. A., et al. (2015). Dermal toxicity studies: Factors impacting study interpretation and outcome. *Toxicologic Pathology, 43*, 474–481. https://doi.org/10.1177/0192623314548765.

Chen, Q., & Thouas, G. A. (2015). Metallic implant biomaterials. *Materials Science & Engineering R: Reports, 87*, 1–57.

Cheville, N. F., & Stasko, J. (2014). Techniques in Electron microscopy of animal tissue. *Veterinary Pathology, 51*, 28–41. https://doi.org/10.1177/0300985813505114.

Cook, J. L., Kuroki, K., Visco, D., Pelletier, J. P., Schulz, L., & Lafeber, F. P. J. G. (2010). The OARSI histopathology initiative – Recommendations for histological assessments of osteoarthritis in the dog. *Osteoarthritis and Cartilage, 18*(Supplement 3), S66–S79. https://doi.org/10.1016/j.joca.2010.04.017.

Crissman, J. W., et al. (2004). Best practices guideline: Toxicologic histopathology. *Toxicologic Pathology, 32*, 126–131.

Dalu, A., Blaydes, B. S., Lomax, L. G., & Delclos, K. B. (2000). A comparison of the inflammatory response to a polydimethylsiloxane implant in male and female Balb/c mice. *Biomaterials, 21*, 1947–1957.

Datta, S., Malhotra, L., Dickerson, R., Chaffee, S., Sen, C. K., & Roy, S. (2015). Laser capture microdissection: Big data from small samples. *Histology and Histopathology, 30*, 1255–1269. https://doi.org/10.14670/HH-11-622.

Davies, L. C., Jenkins, S. J., Allen, J. E., & Taylor, P. R. (2013). Tissue-resident macrophages. *Nature Immunology, 14*, 986–995. https://doi.org/10.1038/ni.2705.

De Jong, W. H., Bergsma, J. E., Robinson, J. E., & Bos, R. R. (2005). Tissue response to partially in vitro predegraded poly-L-lactide implants. *Biomaterials, 26*, 1781–1791.

DeLustro, F., Condell, R. A., Nguyen, M. A., & McPherson, J. M. (1986). A comparative study of the biologic and immunologic response to medical devices derived from dermal collagen. *Journal of Biomedical Materials Research, 20*, 109–120. https://doi.org/10.1002/jbm.820200110.

DiEgidio, P., Friedman, H. I., Gourdie, R. G., Riley, A. E., Yost, M. J., & Goodwin, R. L. (2014). Biomedical implant capsule formation: Lessons learned and the road ahead. *Annals of Plastic Surgery, 73*, 451–460.

Diller, R. B., Audet, R. G., & Kellar, R. S. (2015). Quantitative histopathology for evaluation of in vivo biocompatibility associated with biomedical implants. In J. S. Potts, A. D. Eberhard, & J. A. K. Wharton (Eds.), *Molecular histopathology and tissue biomarkers in drug and diagnostic development* (pp. 153–162). New York: Springer. https://doi.org/10.1007/7653_2014_37.

Dobrovolskaia, M. A. (2015). Pre-clinical immunotoxicity studies of nanotechnology-formulated drugs: Challenges, considerations and strategy. *Journal of Controlled Release : Official Journal of the Controlled Release Society, 220*, 571–583. https://doi.org/10.1016/j.jconrel.2015.08.056.

Donlan, R. M., & Costerton, J. W. (2002). Biofilms: Survival mechanisms of clinically relevant microorganisms. *Clinical Microbiology Reviews, 15*, 167–193.

dos Santos, P. L., et al. (2016). Evaluation of bone substitutes for treatment of peri-implant bone defects: Biomechanical, histological, and immunohistochemical analyses in the rabbit tibia. *Journal of Periodontal & Implant Science, 46*, 176–196.

Douglass, J. P., Berry, C. R., Thrall, D. E., Malarkey, D. E., & Spaulding, K. A. (2003). Radiographic features of aortic bulb/valve mineralization in 20 dogs. *Veterinary Radiology & Ultrasound, 44*, 20–27.

Elmore, S. A. (2006a). Enhanced histopathology of mucosa-associated lymphoid tissue. *Toxicologic Pathology, 34*, 687–696. https://doi.org/10.1080/01926230600939989.

Elmore, S. A. (2006b). Enhanced histopathology of the lymph nodes. *Toxicologic Pathology, 34*, 634–647.

Elmore, S. A. (2006c). Enhanced histopathology of the spleen. *Toxicologic Pathology, 34*, 648–655.

Elmore, S. A. (2006d). Enhanced histopathology of the thymus. *Toxicologic Pathology, 34*, 656.

Elmore, S. A., et al. (2017). Proceedings of the 2017 National Toxicology Program Satellite Symposium. *Toxicologic Pathology, 45*, 799–833. https://doi.org/10.1177/0192623317733924.

Evans, H. E., & De Lahunta, A. (2013). *Miller's anatomy of the dog* (4th ed.). St. Louis: Elsevier Health Sciences.

Everds, N. E., et al. (2013). Interpreting stress responses during routine toxicity studies: A review of the biology, impact, and assessment. *Toxicologic Pathology, 41*, 560–614. https://doi.org/10.1177/0192623312466452.

Farrah, K., Mierzwinski-Urban, M., & Cimon, K. (2016). Effectiveness of adverse effects search filters: Drugs versus medical devices. *Journal of the Medical Library Association: JMLA, 104*, 221–225. https://doi.org/10.3163/1536-5050.104.3.007.

Fellah, B. H., Weiss, P., Gauthier, O., Rouillon, T., Pilet, P., Daculsi, G., & Layrolle, P. (2006). Bone repair using a new injectable self-crosslinkable bone substitute. *Journal of Orthopaedic Research, 24*, 628–635.

Fernandez-Bueno, I., et al. (2015). Safety and biocompatibility of a new high-density polyethylene-based spherical integrated porous orbital implant: An experimental study in rabbits. *Journal of Ophthalmology, 2015*, 904096.

Frydman, G. H., et al. (2017). Local and systemic changes associated with long-term, percutaneous, static implantation of titanium alloys in rhesus macaques (Macaca mulatta). *Comparative Medicine, 67*, 165–175.

Funk, K. A., Hampshire, V. A., & Schuh, J. C. L. (2018). Nonclinical safety evaluation of medical devices. In P. S. Sahota, J. A. Popp, P. Bouchard, J. F. Hardisty, & C. Gopinath (Eds.), *Toxicologic pathology: Nonclinical safety assessment* (2nd ed.). Boca Raton: CRC Press.

Gad, S. C., & Gad-McDonald, S. (2016). *Biomaterials, medical devices, and combination products: Biocompatibility testing and safety assessment*. Boca Raton: CRC Press.

Gad, S. C., & Schuh, J. C. L. (2018). Regulatory forum opinion paper: Considerations for toxicologic pathologists evaluating the safety of biomaterials and finished medical devices. *Toxicologic Pathology, 46*, 366–371. https://doi.org/10.1177/0192623318768719.

Gage, G. J., Kipke, D. R., & Shain, W. (2012). Whole animal perfusion fixation for rodents. *Journal of Visualized Experiments : JoVE, 3564*. https://doi.org/10.3791/3564.

Ganta, S. R., et al. (2003). Vascularization and tissue infiltration of a biodegradable polyurethane matrix. *Journal of Biomedical Materials Research. Part A, 64*, 242–248.

Gauthier, B. E., Gervais, F., Hamm, G., O'Shea, D., Piton, A., & Schumacher, V. L. (2019). Toxicologic pathology forum: opinion on integrating innovative digital pathology tools in the regulatory framework. *Toxicologic Pathology, 47*, 436–443.

Gibon, E., et al. (2017a). The biological response to orthopaedic implants for joint replacement: Part I: Metals. *Journal of Biomedical Materials Research. Part B, Applied Biomaterials, 105*, 2162–2173.

Gibon, E., Córdova, L. A., Lu, L., Lin, T. H., Yao, Z., Hamadouche, M., & Goodman, S. B. (2017b). The biological response to orthopedic implants for joint replacement. II: Polyethylene, ceramics, PMMA, and the foreign body reaction. *Journal of Biomedical Materials Research. Part B, Applied Biomaterials, 105*, 1685–1691.

Goad, M., & Goad, D. (2013). Biomedical materials and devices. In W. M. Haschek, C. G. Rousseaux, M. A. Wallig, B. Bolon, R. Ochoa, & B. Mahler (Eds.), *Haschek and Rousseaux's handbook of toxicologic pathology* (Vol. 3, 3rd ed., pp. 783–806). San Diego: Elsevier (Academic Press).

Golder, S., Wright, K., & Rodgers, M. (2014). Failure or success of search strategies to identify adverse effects of medical devices: A feasibility study using a systematic review. *Systematic Reviews, 3*, 113–113. https://doi.org/10.1186/2046-4053-3-113.

Gorbet, M. B., & Sefton, M. V. (2004). Biomaterial-associated thrombosis: Roles of coagulation factors, complement, platelets and leukocytes. *Biomaterials, 25*, 5681–5703. https://doi.org/10.1016/j.biomaterials.2004.01.023.

Greaves, P., et al. (2013). Proliferative and non-proliferative lesions of the rat and mouse soft tissue, skeletal muscle and mesothelium. *Journal of Toxicologic Pathology, 26*, 1S–26S. https://doi.org/10.1293/tox.26.1S.

Greenwood, J. E., & Dearman, B. L. (2012). Split skin graft application over an integrating, biodegradable temporizing polymer matrix: Immediate and delayed. *Journal of Burn Care & Research, 33*, 7–19.

Grossman, J. D., & Getty, R. (1975a). *Sisson and Grossman's the anatomy of the domestic animals: Equine, ruminant* (Vol. I, 5th ed.). Philadelphia: Saunders.

Grossman, J. D., & Getty, R. (1975b). *Sisson and Grossman's the anatomy of the domestic animals: Porcine, carnivore, Aves* (Vol. 2, 5th ed.). Philadelphia: Saunders.

Gundersen, H. J. G., Mirabile, R., Brown, D., & Boyce, R. W. (2013). Stereological principles and sampling procedures for toxicologic pathologists. In W. M. Haschek, C. G. Rousseaux, & M. A. Wallig (Eds.), *Haschek and Rousseaux's handbook of toxicologic pathology* (Vol. 1, 3rd ed., pp. 215–286). Boston: Academic Press. https://doi.org/10.1016/B978-0-12-415759-0.00008-X.

Handel, N. (2006). Long-term safety and efficacy of polyurethane foam-covered breast implants. *Aesthetic Surgery Journal, 26*, 265–274.

Hansen, T., et al. (2002). New aspects in the histological examination of polyethylene wear particles in failed total joint replacements. *Acta Histochemica, 104*, 263–269. https://doi.org/10.1078/0065-1281-00649.

Harrell, M. I., Iritani, B. M., & Ruddell, A. (2008). Lymph node mapping in the mouse. *Journal of Immunological Methods, 332*, 170–174.

Hassler, C., Boretius, T., & Stieglitz, T. (2011). Polymers for neural implants. *Journal of Polymer Science Part B: Polymer Physics, 49*, 18–33. https://doi.org/10.1002/polb.22169.

Heredia, J. E., et al. (2013). Type 2 innate signals stimulate fibro/adipogenic progenitors to facilitate muscle regeneration. *Cell, 153*, 376–388.

Honari, G., Ellis, S. G., Wilkoff, B. L., Aronica, M. A., Svensson, L. G., & Taylor, J. S. (2008). Hypersensitivity reactions associated with endovascular devices. *Contact Dermatitis, 59*, 7–22.

Hook, A. L., et al. (2012). Combinatorial discovery of polymers resistant to bacterial attachment. *Nature Biotechnology, 30*, 868.

Hooper, K. A., Macon, N. D., & Kohn, J. (1998). Comparative histological evaluation of new tyrosine-derived polymers and poly (L-lactic acid) as a function of polymer degradation. *Journal of Biomedical Materials Research. Part A, 41*, 443–454.

Horne, J., Bateman, A. C., Carr, N. J., & Ryder, I. (2014). Lymph node revealing solutions in colorectal cancer: Should they be used routinely? *Journal of Clinical Pathology, 67*, 383–388. https://doi.org/10.1136/jclinpath-2013-202146.

Hu, W.-J., Eaton, J. W., Ugarova, T. P., & Tang, L. (2001). Molecular basis of biomaterial-mediated foreign body reactions. *Blood, 98*, 1231–1238.

Iezzi, G., et al. (2014). Peri-implant bone tissues around retrieved human implants after time periods longer than 5 years: A retrospective histologic and histomorphometric evaluation of 8 cases. *Odontology, 102*, 116–121.

Ikarashi, Y., et al. (1992). Comparative studies by cell culture and in vivo implantation test on the toxicity of natural rubber latex materials. *Journal of Biomedical Materials Research. Part A, 26*, 339–356.

Institute of Medicine. (1999). *Safety of silicone breast implants*. Washington, DC: The National Academies Press. https://doi.org/10.17226/9602.

International Council for Harmonisation. (1995). *S1A need for carcinogenicity studies of pharmaceuticals.*

International Organization for Standardization. (2009). *ISO 10993-1: 2009/Cor 1:2010. Part 1: Evaluation and testing within a risk management process*. Geneva.

International Organization for Standardization. (2014). *ISO 10993-3:2014 tests for genotoxicity, carcinogenicity and reproductive toxicity*. Geneva.

International Organization for Standardization. (2016). *ISO 10993-6:2016. Part 6: Tests for local effects after implantation*. Geneva.

International Organization for Standardization. (2017a). *ISO 10993-4: 2017. Part 4: Selection of tests for interactions with blood*. Geneva.

International Organization for Standardization. (2017b). *ISO 10993-11:2017. Part 11: Tests for systemic toxicity*. Geneva.

International Organization for Standardization. (2017c). *ISO 10993-16: 2017. Part 16: Toxicokinetic study design for degradation products and leachables*. Geneva.

Ionita, C. N., et al. (2009). The asymmetric vascular stent: Efficacy in a rabbit aneurysm model. *Stroke, 40*, 959–965. https://doi.org/10.1161/STROKEAHA.108.524124.

Ito, R., & Suami, H. (2015). Lymphatic territories (lymphosomes) in swine: an animal model for future lymphatic research. *Plastic and Reconstructive Surgery, 136*, 297–304.

Jaafar, I. H., LeBlon, C. E., Wei, M.-T., Ou-Yang, D., Coulter, J. P., & Jedlicka, S. S. (2011). Improving fluorescence imaging of biological cells on biomedical polymers. *Acta Biomaterialia, 7*, 1588–1598. https://doi.org/10.1016/j.actbio.2010.12.007.

Jessen, S. L., et al. (2018). Method for preclinical pathology evaluation and analysis of cardiovascular implantable electronic device implant sites. *Cardiovascular Pathology, 36*, 44–52.

Jones, K. (2015). Fibrotic response to biomaterials and all associated sequence of fibrosis. In S. F. Badylak (Ed.), *Host response to biomaterials* (pp. 189–237). Oxford, UK: Academic Press. https://doi.org/10.1016/B978-0-12-800196-7.00009-8.

Kalimo, K., Räsänen, L., Aho, H., Mäki, J., Mustikkamki, U. P., & Rantala, I. (1996). Persistent cutaneous pseudolymphoma after intradermal gold injection. *Journal of Cutaneous Pathology, 23*, 328–334.

Kaminski, E. J., Oglesby, R. J., Wood, N. K., & Sandrik, J. (1968). The behavior of biological materials at different sites of implantation. *Journal of Biomedical Materials Research, 2*, 81–88.

Kariyawasam, H. H., & Robinson, D. S. (2006). The eosinophil: The cell and its weapons, the cytokines, its locations. *Seminars in Respiratory and Critical Care Medicine, 27*, 117–127. https://doi.org/10.1055/s-2006-939514.

Keel, S. B., Jaffe, K. A., Nielsen, G. P., & Rosenberg, A. E. (2001). Orthopaedic implant-related sarcoma: A study of twelve cases. *Modern Pathology, 14*, 969–977.

Kesler, C. T., Liao, S., Munn, L. L., & Padera, T. P. (2013). Lymphatic vessels in health and disease. *Wiley Interdisciplinary Reviews: Systems Biology and Medicine, 3*, 111–124.

Kirkpatrick, C. J., et al. (2000). Biomaterial-induced sarcoma : A novel model to study preneoplastic change. *The American Journal of Pathology, 156*, 1455–1467.

Kittel, B., et al. (2004). Revised guides for organ sampling and trimming in rats and mice–Part 2: A joint publication of the RITA and NACAD groups. *Experimental and Toxicologic Pathology, 55*, 413–431.

Klopfleisch, R., & Jung, F. (2017). The pathology of the foreign body reaction against biomaterials. *Journal of Biomedical Materials Research. Part A, 105*, 927–940. https://doi.org/10.1002/jbm.a.35958.

Knoblaugh, S. E., & Randolph-Habecker, J. (2017). Necropsy and histology. In P. M. Treuting, S. M. Dintzis, & K. S. Montine (Eds.), *Comparative anatomy and histology: A mouse, rat, and human atlas* (2nd ed., pp. 23–51). San Diego: Elsevier.

Konttinen, Y. T., Pajarinen, J., Takakubo, Y., Gallo, J., Nich, C., Takagi, M., & Goodman, S. B. (2014). Macrophage polarization and activation in response to implant debris: Influence by "particle disease" and "ion disease". *Journal of Long-Term Effects of Medical Implants, 24*, 267–281.

Kraus, T., Fischerauer, S. F., Hänzi, A. C., Uggowitzer, P. J., Löffler, J. F., & Weinberg, A. M. (2012). Magnesium alloys for temporary implants in osteosynthesis: In vivo studies of their degradation and interaction with bone. *Acta Biomaterialia, 8*, 1230–1238.

Kraus, V. B., Huebner, J. L., DeGroot, J., & Bendele, A. (2010). The OARSI histopathology initiative – Recommendations for histological assessments of osteoarthritis in the Guinea pig. *Osteoarthritis and Cartilage, 18*(Supplement 3), S35–S52. https://doi.org/10.1016/j.joca.2010.04.015.

Krenn, V., et al. (2014). Revised histopathological consensus classification of joint implant related pathology. *Pathology, Research and Practice, 210*, 779–786. https://doi.org/10.1016/j.prp.2014.09.017.

Krenn, V., & Perino, G. (2017). *Histological diagnosis of implant-associated pathologies. Clinical management of joint arthroplasty*. Berlin: Springer.

Latendresse, J. R., Warbrittion, A. R., Jonassen, H., & Creasy, D. M. (2002). Fixation of testes and eyes using a modified Davidson's fluid: Comparison with Bouin's fluid and conventional Davidson's fluid. *Toxicologic Pathology, 30*, 524–533.

Lee, J. M., & Kim, Y. J. (2015). Foreign body granulomas after the use of dermal fillers: Pathophysiology, clinical appearance, histologic features, and treatment. *Archives of Plastic Surgery, 42*, 232–239. https://doi.org/10.5999/aps.2015.42.2.232.

Leigh Perkins, L. E. (2010). Preclinical models of restenosis and their application in the evaluation of drug-eluting stent systems. *Veterinary Pathology, 47*, 58–76. https://doi.org/10.1177/0300985809352978.

Lemperle, G., Morhenn, V. B., Pestonjamasp, V., & Gallo, R. L. (2004). Migration studies and histology of injectable microspheres of different sizes in mice. *Plastic and Reconstructive Surgery, 113*, 1380–1390.

Lie, K. I., Jaeger, G., Nordstoga, K., & Moe, L. (2011). Inflammatory response to therapeutic gold bead implantation in canine hipjoint osteoarthritis. *Veterinary Pathology, 48*, 1118–1124.

Loch-Wilkinson, A., et al. (2017). Breast implant–associated anaplastic large cell lymphoma in Australia and New Zealand: High-surface-area textured implants are associated with increased risk. *Plastic and Reconstructive Surgery, 140*, 645–654.

Long, P. H. (2008). Medical devices in orthopedic applications. *Toxicologic Pathology, 36*, 85–91. https://doi.org/10.1177/0192623307310951.

Major, M. R., Wong, V. W., Nelson, E. R., Longaker, M. T., & Gurtner, G. C. (2015). The foreign body response: At the interface of surgery and bioengineering. *Plastic and Reconstructive Surgery, 135*, 1489–1498.

Malik, N., et al. (1998). Intravascular stents: A new technique for tissue processing for histology, immunohistochemistry, and transmission electron microscopy. *Heart, 80*, 509.

Markwardt, N. T., Stokol, J., & Rennaker, R. L. (2013). Sub-meninges implantation reduces immune response to neural implants. *Journal of Neuroscience Methods, 214*, 119–125. https://doi.org/10.1016/j.jneumeth.2013.01.020.

Masuda, K., et al. (2005). A novel rabbit model of mild, reproducible disc degeneration by an anulus needle puncture: Correlation between the degree of disc injury and radiological and histological appearances of disc degeneration. *Spine, 30*, 5–14.

Mathiesen, E. B., Ahlbom, A., Bermann, G., & Lindgren, J. (1995). Total hip replacement and cancer. *A Cohort Study Bone & Joint Journal, 77*, 345–350.

Maul, T. M., et al. (2011). Pre-clinical implants of the Levitronix PediVAS(®) pediatric ventricular assist device – strategy for regulatory approval. *Cardiovascular Engineering and Technology, 2*, 263–275. https://doi.org/10.1007/s13239-011-0063-5.

McGregor, D. B., Baan, R. A., Partensky, C., Rice, J. M., & Wilbourn, J. D. (2000). Evaluation of the carcinogenic risks to humans associated with surgical implants and other foreign bodies; a report of an IARC monographs programme meeting. *European Journal of Cancer, 36*, 307–313. https://doi.org/10.1016/S0959-8049(99)00312-3.

McInnes, E. (2005). Artefacts in histopathology. *Comparative Clinical Pathology, 13*, 100–108.

McLaughlin, J. K., Lipworth, L., Murphy, D. K., & Walker, P. S. (2007). The safety of silicone gel-filled breast implants: A review of the epidemiologic evidence. *Annals of Plastic Surgery, 59*, 569–580.

Meng, H.-W., Chien, E. Y., & Chien, H.-H. (2016). Dental implant bioactive surface modifications and their effects on osseointegration: A review. *Biomarker Research, 4*, 24.

Milde, R., et al. (2015). Multinucleated Giant cells are specialized for complement-mediated phagocytosis and large target destruction. *Cell Reports, 13*, 1937–1948. https://doi.org/10.1016/j.celrep.2015.10.065.

Moalli, P., Brown, B., Reitman, M. T. F., & Nager, C. W. (2014). Polypropylene mesh: Evidence for lack of carcinogenicity. *International Urogynecology Journal, 25*, 573–576. https://doi.org/10.1007/s00192-014-2343-8.

Moizhess, T. G. (2008). Carcinogenesis induced by foreign bodies. *Biochemistry (Mosc), 73*, 763–775.

Monteiro, D. R., Gorup, L. F., Takamiya, A. S., Ruvollo-Filho, A. C., ERd, C., & Barbosa, D. B. (2009). The growing importance of materials that prevent microbial adhesion: Antimicrobial effect of medical devices containing silver. *International Journal of Antimicrobial Agents, 34*, 103–110. https://doi.org/10.1016/j.ijantimicag.2009.01.017.

Morais, J. M., Papadimitrakopoulos, F., & Burgess, D. J. (2010). Biomaterials/tissue interactions: Possible solutions to overcome foreign body response. *The AAPS Journal, 12*, 188–196.

Morawietz, G., et al. (2004). Revised guides for organ sampling and trimming in rats and mice–Part 3: A joint publication of the RITA and NACAD groups. *Experimental and Toxicologic Pathology, 55*, 433–449.

Morawietz, L., et al. (2006). Proposal for a histopathological consensus classification of the periprosthetic interface membrane. *Journal of Clinical Pathology, 59*, 591–597. https://doi.org/10.1136/jcp.2005.027458.

Morton, D., et al. (2010). Recommendations for pathology peer review. *Toxicologic Pathology, 38*, 1118–1127.

Moya, J. S., et al. (2016). Histological response of soda-lime glass-ceramic bactericidal rods implanted in the jaws of beagle dogs. *Scientific Reports, 6*, 31478. https://doi.org/10.1038/srep31478.

Muhamed, J., Revi, D., Rajan, A., Geetha, S., & Anilkumar, T. V. (2015). Biocompatibility and immunophenotypic characterization of a Porcine Cholecyst–derived Scaffold implanted in rats. *Toxicologic Pathology, 43*, 536–545. https://doi.org/10.1177/0192623314550722.

Nam, S. Y., Ricles, L. M., Suggs, L. J., & Emelianov, S. Y. (2015). Imaging strategies for tissue engineering applications. *Tissue Engineering Part B, Reviews, 21*, 88–102. https://doi.org/10.1089/ten.teb.2014.0180.

Neef, N., Nikula, K. J., Francke-Carroll, S., & Boone, L. (2012). Regulatory forum opinion piece: Blind reading of histopathology slides in general toxicology studies∗. *Toxicologic Pathology, 40*, 697–699. https://doi.org/10.1177/0192623312438737.

Nich, C., & Goodman, S. B. (2014). The role of macrophages in the biological reaction to wear debris from joint replacements. *Journal of Long-Term Effects of Medical Implants, 24*, 259–265.

Nikula, K. J., & Funk, K. (2016). Regulatory forum opinion piece: An experienced pathologist should be present at necropsy for novel medical device studies. *Toxicologic Pathology, 44*, 9–11. https://doi.org/10.1177/0192623315617035.

Nyska, A., Schiffenbauer, Y. S., Brami, C. T., Maronpot, R. R., & Ramot, Y. (2014). Histopathology of biodegradable polymers: Challenges in interpretation and the use of a novel compact MRI for biocompatibility evaluation. *Polymers for Advanced Technologies, 25*, 461–467. https://doi.org/10.1002/pat.3238.

Oppenheimer, B. S., Oppenheimer, E. T., Danishefsky, I., Stout, A. P., & Eirich, F. R. (1955). Further studies of polymers as carcinogenic agents in animals. *Cancer Research, 15*, 333–340.

Orenstein, S. B., Saberski, E. R., Kreutzer, D. L., & Novitsky, Y. W. (2012). Comparative analysis of histopathologic effects of synthetic meshes based on material, weight, and pore size in mice. *The Journal of Surgical Research, 176*, 423–429.

Pagán, A. J., & Ramakrishnan, L. (2018). The formation and function of granulomas. *Annual Review of Immunology, 36*. https://doi.org/10.1146/annurev-immunol-032712-100022.

Pearce, M. S., et al. (2012). Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: A retrospective cohort study. *The Lancet, 380*, 499–505.

Perkins, R. B., Handal-Orefice, R., Hanchate, A. D., Lin, M., & Paasche-Orlow, M. K. (2016). Risk of undetected cancer at the time of laparoscopic supracervical hysterectomy and laparoscopic myomectomy: Implications for the use of power morcellation. *Womens Health Issues, 26*, 21–26.

Phillips, P. L., Wolcott, R. D., Cowan, L. J., & Schultz, G. S. (2016). Biofilms in wounds and wound dressing. In *Wound Healing Biomaterials-Volume 2: Functional Biomaterials* (55–78). Philidelphia: Elsevier.

Pierce, L. M., Rao, A., Baumann, S. S., Glassberg, J. E., Kuehl, T. J., & Muir, T. W. (2009). Long-term histologic response to synthetic and biologic graft materials implanted in the vagina and abdomen of a rabbit model. *American Journal of Obstetrics and Gynecology, 200*, 546.e541–546.e548. https://doi.org/10.1016/j.ajog.2008.12.040.

Pinchuk, L. (1995). A review of the biostability and carcinogenicity of polyurethanes in medicine and the new generation of biostable polyurethanes. *Journal of Biomaterials Science. Polymer Edition, 6*, 225–267.

Pizzoferrato, A., Savarino, L., Stea, S., & Tarabusi, C. (1988). Results of histological grading on 100 cases of hip prosthesis failure. *Biomaterials, 9*, 314–318.

Popesko, P., & Getty, R. (1971). *Atlas of topographical anatomy of the domestic animals* (Vol. I-III vol I-III).

Popesko, P., Rajtová, V., & Horak, J. (1990). *A colour atlas of the anatomy of small laboratory animals: Rat, mouse, hamster* (Vol. 2). London: Wolfe Publishing Ltd.

Popesko, P., Rajtová, V., & Horak, J. (1992). *A colour atlas of the anatomy of small laboratory animals: Rabbit, Guinea pig* (Vol. 1). London: Wolfe Publishing Ltd.

Ramot, Y., et al. (2015a). Long-term local and systemic safety of poly (l-lactide-co-epsilon-caprolactone) after subcutaneous and intra-articular implantation in rats. *Toxicologic Pathology, 43*, 1127–1140.

Ramot, Y., Rousselle, S. D., Yellin, N., Willenz, U., Sabag, I., Avner, A., & Nyska, A. (2016). Biocompatibility and systemic safety of a novel implantable annuloplasty ring for the treatment of mitral regurgitation in a minipig model. *Toxicologic Pathology, 44*, 655–662. https://doi.org/10.1177/0192623315627217.

Ramot, Y., et al. (2015b). Interspecies differences in reaction to a biodegradable subcutaneous tissue filler: Severe inflammatory granulomatous reaction in the Sinclair minipig. *Toxicologic Pathology, 43*, 267–271. https://doi.org/10.1177/0192623314534995.

Ratner, B. D., Hoffman, A. S., Schoen, F. J., & Lemons, J. E. (2013). *Biomaterials science: An introduction to materials in medicine* (3rd ed.). Canada: Academic Press (Elsevier).

Ren, P.-G., Irani, A., Huang, Z., Ma, T., Biswal, S., & Goodman, S. B. (2011). Continuous infusion of UHMWPE particles induces increased bone macrophages and osteolysis. *Clinical Orthopaedics and Related Research, 469*, 113–122. https://doi.org/10.1007/s11999-010-1645-5.

Rentsch, C., Schneiders, W., Manthey, S., Rentsch, B., & Rammelt, S. (2014). Comprehensive histological evaluation of bone implants. *Biomatter, 4*, e27993. https://doi.org/10.4161/biom.27993.

Ricciardi, B. F., et al. (2016). Histopathological characterization of corrosion product associated adverse local tissue reaction in hip implants: A study of 285 cases. *BMC Clinical Pathology, 16*, 3. https://doi.org/10.1186/s12907-016-0025-9.

Rigdon, R. (1973). Local reaction to polyurethane—A comparative study in the mouse, rat, and rabbit. *Journal of Biomedical Materials Research, 7*, 79–93.

Ripamonti, U. (1996). Osteoinduction in porous hydroxyapatite implanted in heterotopic sites of different animal models. *Biomaterials, 17*, 31–35.

Rismanchian, M., Movahedian, B., Khalighinejad, N., Badrian, H., Mohammad Razavi, S., & Nekouie, A. (2012). Comparative evaluation of two types of immediately loaded implants using biomechanical and histomorphometric tests: An animal case study. *ISRN Dentistry, 2012*, 328945. https://doi.org/10.5402/2012/328945.

Roberts, A., et al. (2013). Integrated microscopy techniques for comprehensive pathology evaluation of an implantable left atrial pressure sensor. *Journal of Histotechnology, 36*, 17–24. https://doi.org/10.1179/2046023613Y.0000000021.

Rodriguez, J. N., et al. (2014). In vivo response to an implanted shape memory polyurethane foam in a porcine aneurysm model. *Journal of Biomedical Materials Research. Part A, 102*, 1231–1242. https://doi.org/10.1002/jbm.a.34782.

Rousselle, S., & Wicks, J. (2008). Preparation of medical devices for evaluation. *Toxicologic Pathology, 36*, 81–84.

Rousselle, S. D., Wicks, J. R., Tabb, B. C., Tellez, A., & O'Brien, M. (2019). Histology strategies for medical implants and interventional device studies. *Toxicologic Pathology, 47*, 235–249.

Rouselle, S., & Paulin, J. (2019). Medical Devices Special Issue. *Toxicologic Pathology, 47*, 201–432.

Ruehl-Fehlert, C., et al. (2003). Revised guides for organ sampling and trimming in rats and mice – Part 1. *Experimental and Toxicologic Pathology, 55*, 91–106. https://doi.org/10.1078/0940-2993-00311.

Rutgers, M., van Pelt, M. J. P., Dhert, W. J. A., Creemers, L. B., & Saris, D. B. F. (2010). Evaluation of histological scoring systems for tissue-engineered, repaired and osteoarthritic cartilage. *Osteoarthritis and Cartilage, 18*, 12–23. https://doi.org/10.1016/j.joca.2009.08.009.

Sanders, J. E., Stiles, C. E., & Hayes, C. (2000). Tissue response to single-polymer fibers of varying diameters: Evaluation of fibrous encapsulation and macrophage density. *Journal of Biomedical Materials Research, 52*, 231–237.

Sanjai, K., Kumarswamy, J., Patil, A., Papaiah, L., Jayaram, S., & Krishnan, L. (2012). Evaluation and comparison of decalcification agents on the human teeth. *Journal of Oral and Maxillofacial Pathology : JOMFP, 16*, 222–227. https://doi.org/10.4103/0973-029X.99070.

Santerre, J., Woodhouse, K., Laroche, G., & Labow, R. (2005). Understanding the biodegradation of polyurethanes: From classical implants to tissue engineering materials. *Biomaterials, 26*, 7457–7470.

Sato, J., Doi, T., Kanno, T., Wako, Y., Tsuchitani, M., & Narama, I. (2012a). Histopathology of incidental findings in Cynomolgus monkeys (Macaca Fascicularis) used in toxicity studies. *Journal of Toxicologic Pathology, 25*, 63–101. https://doi.org/10.1293/tox.25.63.

Sato, J., Doi, T., Wako, Y., Hamamura, M., Kanno, T., Tsuchitani, M., & Narama, I. (2012b). Histopathology of incidental findings in beagles used in toxicity studies. *Journal of Toxicologic Pathology, 25*, 103–134. https://doi.org/10.1293/tox.25.103.

Schmalzried, T. P., Jasty, M., Rosenberg, A., & Harris, W. H. (1993). Histologic identification of polyethylene wear debris using oil red O stain. *Journal of Applied Biomaterials, 4*, 119–125.

Schoen, F. J., Harasaki, H., Kim, K. M., Anderson, H. C., & Levy, R. J. (1988). Biomaterial-associated calcification: Pathology, mechanisms, and strategies for prevention. *Journal of Biomedical Materials Research, 22*, 11–36.

Schuh, J. C. L. (2008). Medical device regulations and testing for toxicologic pathologists. *Toxicologic Pathology, 36*, 63–69. https://doi.org/10.1177/0192623307309926.

Schuh, J. C. L. (2015). Genetically modified animal models. In S. C. Gad (Ed.), *Animal models in toxicology* (3rd ed., pp. 935–956). Boca Raton: CRC Press.

Schuh, J. C. L., & Funk, K. A. (2019). Compilation of international standards and regulatory guidance documents for evaluation of biomaterials, medical devices, 3D printed and regenerative medicine products. *Toxicologic Pathology, 47*, 344.

Schwartz, R. S., et al. (2004). Preclinical evaluation of drug-eluting stents for peripheral applications: Recommendations from an expert consensus group. *Circulation, 110*, 2498–2505.

Sellers, R. S., et al. (2007). Society of Toxicologic Pathology position paper: Organ weight recommendations for toxicology studies. *Toxicologic Pathology, 35*, 751–755.

Shackelford, C., Long, G., Wolf, J., Okerberg, C., & Herbert, R. (2002). Qualitative and quantitative analysis of nonneoplastic lesions in toxicology studies. *Toxicologic Pathology, 30*, 93–96. https://doi.org/10.1080/01926230252824761.

Shayesteh Moghaddam, N., et al. (2016). Metals for bone implants: Safety, design, and efficacy. *Biomanufacturing Reviews, 1*, 1. https://doi.org/10.1007/s40898-016-0001-2.

Shea, K. G., Bloebaum, R. D., Avent, J. M., Birk, G. T., & Samuelson, K. M. (1996). Analysis of lymph nodes for polyethylene particles in patients who have had a primary joint replacement. *Journal of Bone and Joint Surgery, 78*, 497–504.

Sheth, S., Litvack, F., Dev, V., Fishbein, M. C., Forrester, J. S., & Eigler, N. (1996). Subacute thrombosis and vascular injury resulting from slotted-tube nitinol and stainless steel stents in a rabbit carotid artery model. *Circulation, 94*, 1733–1740. https://doi.org/10.1161/01.cir.94.7.1733.

Shoieb, A., Allavena, R., Swallow, J., & Debrue, M. (2012). Peritoneal sarcomatosis associated with telemetry implants in Sprague Dawley CD rats. *Toxicologic Pathology, 40*, 113–121.

Short, B. G. (2008). Safety evaluation of ocular drug delivery formulations: Techniques and practical considerations. *Toxicologic Pathology, 36*, 49–62. https://doi.org/10.1177/0192623307310955.

Siddiqui, R. F., Abraham, J. R., & Butany, J. (2009). Bioprosthetic heart valves: Modes of failure. *Histopathology, 55*, 135–144. https://doi.org/10.1111/j.1365-2559.2008.03190.x.

Signorello, L. B., et al. (2001). Nationwide study of cancer risk among hip replacement patients in Sweden. *Journal of the National Cancer Institute, 93*, 1405–1410.

Socarrás TO, Vasconcelos, A. C., Campos, P. P., Pereira, N. B., Souza, J. P. C., & Andrade, S. P. (2014). Foreign body response to subcutaneous implants in diabetic rats. *PLoS One, 9*, e110945. https://doi.org/10.1371/journal.pone.0110945.

Soto-Miranda, M. A., Suami, H., & Chang, D. W. (2013). Mapping superficial lymphatic territories in the rabbit. *The Anatomical Record, 296*, 965–970.

Stiers, P. J., van Gastel, N., Moermans, K., Stockmans, I., & Carmeliet, G. (2018). An ectopic imaging window for Intravital imaging of engineered bone tissue. *JBMR Plus, 2*, 92–102.

Stokes, K. (2009). The biocompatibility and biostability of new cardiovascular materials and devices. In *Implantable neural prostheses 2* (pp. 1–26). New York: Springer.

Suami, H., Yamashita, S., Soto-Miranda, M. A., & Chang, D. W. (2013). Lymphatic territories (lymphosomes) in a canine: an animal model for investigation of postoperative lymphatic alterations. *PLOS ONE, 8*:e69222.

Suami, H., & Scaglioni, M. F. (2017). Lymphatic territories (lymphosomes) in the rat: an anatomical study for future lymphatic research. *Plastic and reconstructive surgery, 140*, 945–951.

Sun, H., Mei, L., Song, C., Cui, X., & Wang, P. (2006). The in vivo degradation, absorption and excretion of PCL-based implant. *Biomaterials, 27*, 1735–1740. https://doi.org/10.1016/j.biomaterials.2005.09.019.

Sunderman, F. W., Jr. (1989). Carcinogenicity of metal alloys in orthopedic prostheses: Clinical and experimental studies. *Toxicological Sciences, 13*, 205–216.

Szebeni, J. (2012). Hemocompatibility testing for nanomedicines and biologicals: Predictive assays for complement mediated infusion reactions. *European Journal of Nanomedicine, 4*, 33–53.

Szeto, G. L., & Lavik, E. B. (2016). Materials design at the interface of nanoparticles and innate immunity. *Journal of materials chemistry B, Materials for Biology and Medicine, 4*, 1610–1618. https://doi.org/10.1039/C5TB01825K.

Tellez, A., Dillon, K. N., & Rousselle, S. D. (2017). Comprehensive preclinical postmortem evaluation of valvular prosthesis. *Toxicologic Pathology, 45*, 1077–1090. https://doi.org/10.1177/0192623317740325.

Tempel-Brami, C., Schiffenbauer, Y. S., Nyska, A., Ezov, N., Spector, I., Abramovitch, R., & Maronpot, R. R. (2015). Practical applications of in vivo and ex vivo MRI in toxicologic pathology using a novel high-performance compact MRI system. *Toxicologic Pathology, 43*, 633–650.

Teo, Z. W. W., & Schalock, P. C. (2016). Hypersensitivity reactions to implanted metal devices: Facts and fictions. *Journal of Investigational Allergology & Clinical Immunology, 26*, 279–294. https://doi.org/10.18176/jiaci.0095.

Thackaberry, E. A., et al. (2017). Evaluation of the toxicity of Intravitreally injected PLGA microspheres and rods in monkeys and rabbits: Effects of depot size on inflammatory response. *Investigative Ophthalmology & Visual Science, 58*, 4274–4285.

The Organisation for Economic Co-operation and Development (2015) General questions and answers concerning OECD principles of Good Laboratory Practice (GLP) and Mutual Acceptance of Data (MAD). European Commission. https://www.oecd.org/env/ehs/testing/good-laboratory-practiceglp.htm. Accessed 30 May 2017.

Thevenot, P., Hu, W., & Tang, L. (2008). Surface chemistry influence implant biocompatibility. *Current Topics in Medicinal Chemistry, 8*, 270–280.

Tomlinson, L., et al. (2013). Best practices for veterinary toxicologic clinical pathology, with emphasis on the pharmaceutical and biotechnology industries. *Veterinary Clinical Pathology, 42*, 252–269. https://doi.org/10.1111/vcp.12059.

Tuan, R. S., Lee, F. Y.-I., Konttinen, Y., Wilkinson, J. M., & Smith, R. L. (2008). What are the local and systemic biological reactions and mediators to wear debris and what host factors determine or modulate the biological response to wear particles? *The Journal of the American Academy of Orthopaedic Surgeons, 16*, S42–S48.

Tuomari, D. L., et al. (2007). Society of toxicologic pathology position paper on pathology image data: Compliance with 21 CFR Parts 58 and 11. *Toxicologic Pathology, 35*, 450–455.

U.S. Food & Drug Administration. (2015). General considerations for animal studies for medical devices. http://www.fda.gov/downloads/MedicalDevices/DeviceRegulationandGuidance/GuidanceDocuments/UCM466358.pdf. Accessed 25 Aug 2018 2017.

U.S. Food & Drug Administration. (2016). Use of International Standard ISO 10993-1, "Biological evaluation of medical devices - Part 1: Evaluation and testing within a risk management process" U.S. Department of Health & Human Services. http://www.fda.gov/ucm/groups/fdagov-public/@fdagov-meddev-gen/documents/document/ucm348890.pdf. Accessed 23 Oct 2016.

Van Der Giessen, W. J., et al. (1996). Marked inflammatory sequelae to implantation of biodegradable and nonbiodegradable polymers in porcine coronary arteries. *Circulation, 94*, 1690–1697.

Varela, A., & Jolette, J. (2018). Bone toolbox: Biomarkers, imaging tools, biomechanics, and histomorphometry. *Toxicologic Pathology, 46*, 511–529. https://doi.org/10.1177/0192623318779565.

Veerachamy, S., Yarlagadda, T., Manivasagam, G., & Yarlagadda, P. K. (2014). Bacterial adherence and biofilm formation on medical implants: A review. *Proceedings of the Institution of Mechanical Engineers. Part H, 228*, 1083–1099.

Vegas, A. J., et al. (2016). Combinatorial hydrogel library enables identification of materials that mitigate the foreign body response in primates. *Nature Biotechnology, 34*, 345–352. https://doi.org/10.1038/nbt.3462.

Veiseh, O., et al. (2015). Size- and shape-dependent foreign body immune response to materials implanted in rodents and non-human primates. *Nature Materials, 14*, 643–651. https://doi.org/10.1038/nmat4290.

Veleirinho, B., et al. (2014). Foreign body reaction associated with PET and PET/chitosan electrospun nanofibrous abdominal meshes. *PLoS One, 9*, e95293.

Visuri, T., Pukkala, E., Paavolainen, P., Pulkkinen, P., & Riska, E. B. (1996). Cancer risk after metal on metal and polyethylene on metal total hip arthroplasty. *Clinical Orthopaedics and Related Research, 329*, S280–S289.

Von Eiff, C., Jansen, B., Kohnen, W., & Becker, K. (2005). Infections associated with medical devices. *Drugs, 65*, 179–214.

Wagenfuhr-Junior, J., Ribas Filho, J. M., Nascimento, M. M., Ribas, F. M., Wanka, M. V., & Godoi Ade, L. (2012). Histopathological reaction over prosthesis surface covered with silicone and polyurethane foam implanted in rats. *Acta Cirúrgica Brasileira, 27*, 866–873.

Walch, A., Rauser, S., Deininger, S.-O., & Höfler, H. (2008). MALDI imaging mass spectrometry for direct tissue analysis: A new frontier for molecular histology. *Histochemistry and Cell Biology, 130*, 421–434. https://doi.org/10.1007/s00418-008-0469-9.

Wancket, L. M. (2019). Regional draining lymph nodes: considerations for medical device studies. *Toxicologic Pathology, 47*, 339–343.

Ward, J. M., & Thoolen, B. (2011). Grading of lesions. *Toxicologic Pathology, 39*, 745–746. https://doi.org/10.1177/0192623311408622.

Wawrzynski, J., Gil, J. A., Goodman, A. D., & Waryasz, G. R. (2017). Hypersensitivity to orthopedic implants: A review of the literature. *Rheumatology and Therapy, 4*, 45–56.

Welsing, R. T., van Tienen, T. G., Ramrattan, N., Heijkants, R., Schouten, A. J., Veth, R. P., & Buma, P. (2008). Effect on tissue differentiation and articular cartilage degradation of a polymer meniscus implant a 2-year follow-up study in dogs. *American Journal of Sports Medicine, 36*, 1978–1989.

Weyhe, D., Cobb, W., Lecuivre, J., Alves, A., Ladet, S., Lomanto, D., & Bayon, Y. (2015). Large pore size and controlled mesh elongation are relevant predictors for mesh integration quality and low shrinkage–systematic analysis of key parameters of meshes in a novel minipig hernia model. *International Journal of Surgery, 22*, 46–53.

Williams, D. F. (2008). On the mechanisms of biocompatibility. *Biomaterials, 29*, 2941–2953.

Williams, D. F. (2014). There is no such thing as a biocompatible material. *Biomaterials, 35*, 10009–10014.

Wilson, G. J., et al. (2009). Comparison of inflammatory response after implantation of sirolimus- and paclitaxel-eluting stents in porcine coronary arteries. *Circulation, 120*, 141–149. https://doi.org/10.1161/circulationaha.107.730010.

Witzleb, W.-C., Ziegler, J., Krummenauer, F., Neumeister, V., & Guenther, K.-P. (2006). Exposure to chromium, cobalt and molybdenum from metal-on-metal total hip replacement and hip resurfacing arthroplasty. *Acta Orthopaedica, 77*, 697–705.

Wolf, M. F., & Andwraon, J. M. (2012). Practical approach to blood compatibility assessments: General considerations and standards. In J.-P. Boutrand (Ed.), *Biocompatibility and performance of medical devices* (pp. 159–206). Philadelphia: Woodhead Publishing. https://doi.org/10.1533/9780857096456.2.159.

Wolf, M. T., Vodovotz, Y., Tottey, S., Brown, B. N., & Badylak, S. F. (2014). Predicting in vivo responses to biomaterials via combined in vitro and in silico analysis. *Tissue Engineering. Part C, Methods, 21*, 148–159.

Wood, N. K., Kaminski, E. J., & Oglesby, R. J. (1970). The significance of implant shape in experimental testing of biological materials: Disc vs. rod. *Journal of Biomedical Materials Research, 4*, 1–12.

Woodward, S. C., Brewer, P., Moatamed, F., Schindler, A., & Pitt, C. (1985). The intracellular degradation of poly (ε-caprolactone). *Journal of Biomedical Materials Research, 19*, 437–444.

Xie, D., Leng, Y. X., Jing, F. J., & Huang, N. (2015). A brief review of bio-tribology in cardiovascular devices. *Biosurface and Biotribology, 1*, 249–262. https://doi.org/10.1016/j.bsbt.2015.11.002.

Yang, Z., Yuan, H., Tong, W., Zou, P., Chen, W., & Zhang, X. (1996). Osteogenesis in extraskeletally implanted porous calcium phosphate ceramics: Variability among different kinds of animals. *Biomaterials, 17*, 2131–2137.

Zhang, F., Niu, G., Lu, G., & Chen, X. (2011). Preclinical lymphatic imaging. *Molecular Imaging and Biology, 13*, 599–612.

Zhao, S., Pinholt, E. M., Madsen, J. E., & Donath, K. (2000). Histological evaluation of different biodegradable and non-biodegradable membranes implanted subcutaneously in rats. *Journal of Cranio-Maxillo-Facial Surgery, 28*, 116–122.

Chapter 10
Nanodevices

David W. Hobson

Abstract The incorporation of nanomaterials into medical devices has benefits and also new challenges for safety assessment. Understanding these challenges is a significant and important step in the process of practicing "safety by design" in the development of nano-enabled medical devices as well as for establishing methods and practices for safety testing. Toxicological issues are being identified and resolved for a growing number of emerging nano-enabled medical devices with product safety as the objective. This safety testing has identified challenges in test design, nomenclature, and global regulatory processes and harmonization. Nevertheless, it is clear that the incorporation of nanotechnology into medical device design most certainly is having and will continue to have major impacts toward advancing both our knowledge of the utility of nanotechnology in medicine and toward improving the quality of life of those with a variety of afflictions that need hope for a brighter future.

Keywords Nanotechnology · Nanomaterials · Nanoparticle · Medical device · Safety · Toxicology · Nanodevice · Nanomedicine · Nanotoxicology

10.1 Introduction

All living things on earth live and have evolved in an environment that includes daily exposure to naturally occurring microparticles and nanoparticles. These nanoparticles include dietary (lipid micelles, minerals, peptide and protein fragments, and other food particles of many types) and a myriad of non-dietary ingested and/or inhaled substances with nanoscale dimensions as well as inhaled nanoparticles of many types (smoke, dusts, ocean spray, ash, etc.). Naturally occurring nanoparticles of many types also come into contact with our skin, lungs, and

D. W. Hobson, PhD, DABT (✉)
LoneStar PharmTox LLC, Boerne, TX, USA
e-mail: dave@lonestarpharmtox.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_10

gastrointestinal system on a daily basis, and, as a result, humans as well as other animal species have developed uses for these nanomaterials as nutrients or have developed tolerances to potentially harmful nanomaterial exposures and even mechanisms for their absorption, distribution, and elimination.

With the advancement of materials science and engineering, technology exposures to nanomaterials have expanded to include an ever-increasing development and utilization of a myriad of engineered nanotechnology products in essentially every commercial product sector including the construction of medical devices. So, it follows that inclusion of these materials in medical product design requires that we investigate their beneficial effects as well as any risks that may be associated with their intended uses.

Nanomaterial safety evaluations are prudent, practical, and necessary to identify potential safety concerns that might arise from this utilization of engineered nanomaterials in medical devices of all types to ensure that adverse effects are eliminated or are as low as reasonably possible below the threshold for toxicological concern.

Because of the duality of chemical and particle exposure characteristics of engineered nanomaterials, the evaluation of safety of medical devices that incorporate nanomaterials must include methods that examine the impact of correlated qualities of these materials especially if they might substantially potentiate an adverse event. Therefore, regulatory agencies around the world are issuing guidance as to how products containing nanomaterials should be evaluated for safety prior to market approval.

Most medical device regulatory authorities worldwide recognize that the safety as well as benefit of the use of nanomaterials in medical devices is best evaluated on a "case-by-case" basis, and the guidance documents generally reflect this approach and do not consider a nanomaterial inherently harmful or beneficial without reasonable scientific evidence.

The focus of this chapter is on the fundamental definitions, nomenclature, regulations, and safety evaluation approaches for nanomaterials used in medical devices.

10.2 Nanomaterial Nomenclature

The term "nanotechnology" was coined by US Engineer Eric Drexler in the 1980s, and over the past few decades, a wide range of different nanotechnologies have been utilized in a growing number of applications in essentially every aspect of materials science including medical products that may be classified as either drugs or devices. The use of nanotechnology is already reshaping medicine and will continue to do so over the foreseeable future.

The US National Nanotechnology Initiative (NNI) defines that "Nanotechnology is science, engineering, and technology conducted at the nanoscale, which is about 1 to 100 nanometers." The NNI further defines that "Nanoscience and nanotechnology are the study and application of extremely small things and can be used across all the other science fields, such as chemistry, biology, physics, materials science, and engineering" (National Nanotechnology Initiative 2018a).

Engineered nanomaterials are products of nanoscience and nanotechnology. These products involve particles and molecular structures in the "nanoscale" which

is generally accepted as including materials having at least one dimension between 1 and 100 nanometers. This is the size region where the quantal molecular characteristics of these very small dimensions are most evident. Internationally, there is some interest in expanding this range to include all materials having at least one dimension in the nanometer range of ≥1 nm up to ≤999 nm; however, the essential and most practical biological consideration is what dimensionality for a given material does it exhibit the quantal characteristics of substances at the nanoscale.

Developing a common definitions and standards for nanomaterials has been challenging because of the need to satisfy two diverging considerations: (1) the definitions and standards should be broad enough to define materials that may warrant additional evaluation, and (2) they should not be so broad as to include those materials for which additional examination or evaluation would not be meaningful in terms of human health, reducing animal testing or protecting the environment. Definitions proposed and standards developed to date have taken a variety of approaches as they seek to support various and often specific jurisdictional mandates, and this frequently leads to proposals to incorporate contradictory inclusions and exclusions with the definition (National Nanotechnology Initiative 2018b). This creates a complex regulatory maze for producers of nanomaterials and products containing them as to when and what to report for products containing nanomaterials. Therefore a balance is necessary to ensure appropriate and practical coverage of emerging nanomaterials to most effectively protect the health and safety of humans and the environment while at the same time allowing for the potential useful applications of proven safe nanomaterials. Overcoming difficulties associated with attempting to comply with contradictory nanomaterial definitions and the different regulations where they are used can impede the development, and entry of useful and safe nanomaterials into international trade may fundamentally reduce public confidence in the adequacy of regulatory protections from exposure to nanomaterials. This is exactly why many regulatory bodies have taken the position that nanomaterials be considered for safety on a case-by-case basis and to avoid broad classification schemes that have failings for materials in such a technologically expansive landscape.

The International Organization for Standardization (ISO) has descriptively defined “nanomaterial” as a “material with any external dimension in the nanoscale or having internal structure or surface structure in the nanoscale” (ISO 2008) and “nanoparticle” as a “nano-object having all three external dimensions in the nanoscale” where nanoscale is defined as the size range from approximately 1–100 nm (ISO 2010). These technical definitions, based on size only, may be insufficient from a risk evaluation standpoint because they do not include other important elements that should be considered when determining whether a nanomaterial may need additional review and scrutiny.

Nanomaterials are neither inherently hazardous nor inherently safe (Auffan et al. 2009; Donaldson and Poland 2013). This has been broadly recognized, and it is prudent that they should not be treated as such in evaluation programs (Hamburg 2012; Holdren et al. 2011; SCENIHR 2007). This view is the evidence in the US Food and Drug Administration (FDA) current guidance for “Drug Products, Including Biological Products, that Contain Nanomaterials Guidance for Industry” of December 2017 (U.S. Department of Health and Human Services Food and Drug Administration 2017).

In the broadest view and in modern toxicological terms, particles could be considered "nano"particles if the largest dimension of the particle was in the nanorange of <999 nm and >1 nm (Boverhof et al. 2015). This is a definition that has been used in some regulatory documents (Nikalje 2015; Paradis 2012) to include essentially everything in the nanoscale. In fact it is becoming increasingly evident from the growing body of scientific evidence that some particles above 100 nm still exhibit quantum behavior in addition to their chemical characteristics such that absorption through biological membranes and diffusion in biological fluids results in exposure that is profoundly different than the same particle with even very low micrometer dimensions. This is often because, at least in part, some materials have a broad particulate range where the mean particle diameter is above 100 nm, but the particle distribution includes a substantial proportion of particles <100 nm. This is illustrated below in more detail under the topic of nanomaterial characterization.

Current regulatory guidance for nanotechnology in US FDA-regulated products states that nanotechnology "As used in this guidance, the word 'products' (or 'FDA-regulated products') is meant to include products, materials, ingredients, and other substances regulated by FDA, including drugs, biological products, medical devices, food substances (including food for animals), dietary supplements, cosmetic products, and tobacco products" (U.S. Food and Drug Administration 2014). This clearly indicates that the use of nanotechnology in medical devices is a recognized regulatory consideration for the US FDA that must be included in the submission of any device application that involves a device that incorporates nanomaterials and nanotechnology. In many cases, the nanomaterials of concern for medical device safety are nanoparticles of different compositions and molecular structures. The use of nanomaterials in a medical device should always be disclosed to the US FDA as well as other international regulatory bodies. It is never a good idea to overlook the need to report nanomaterials or to attempt nondisclosure of their presence in a medical device.

Nanoparticles are a subset of nanomaterials in general (Hobson et al. 2016). Figure 10.1 shows a practical scheme for the classification of different types of nanomaterials.

Figure 10.2 provides a biological perspective on nanomaterial dimensions in contrast to the dimensions of biological structures and biomolecules that is helpful in understanding how materials in this size range may affect normal biological processes (Hobson et al. 2016).

Current scientific studies are addressing issues such as (1) the characterization of particle-host or particle-cell interactions, (2) the mechanisms of distribution of nanoparticles in the body, especially via regions where there are surfaces conducive for entry into the body (i.e., the pulmonary system and gastrointestinal tract), (3) the cellular and systemic effects of nanoparticles to include effects on both micro- and macro-cellular and tissue biology and on biochemical process, (4) the generation of reactive oxygen species, (5) the activation of pro-inflammatory signaling, and (6) immune response mechanisms.

Since the physicochemical nature of nanoparticles with different sizes and compositions is affected by the environment in which a given particle is surrounded, it should not be difficult to understand that nanoparticles of different

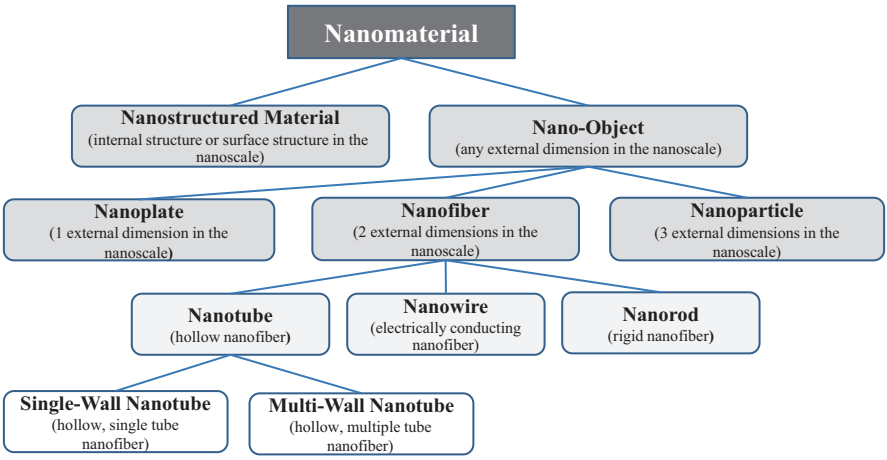


Fig. 10.1 A practical approach to the classification and description of different types of nanomaterials. (Based on Hobson et al. (2016))

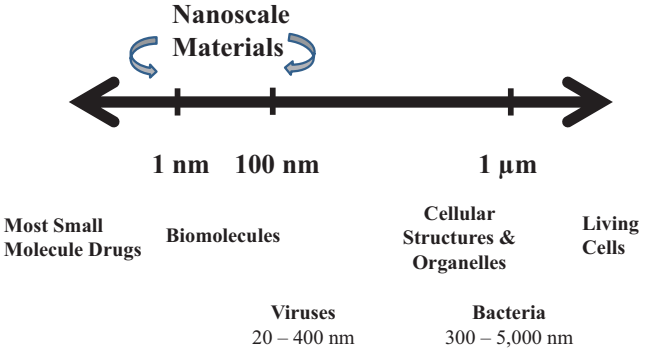


Fig. 10.2 A size-based, biological perspective on nanomaterial dimensions in contrast with the dimensions of biological structures and biomolecules. (From Hobson et al. (2016))

sizes and compositions would exhibit differential rates and degrees of absorption as they pass through the body (Obiweluozor et al. 2018; Yoshida et al. 2018). These characteristics may be helpful in determining whether or not a given nanomaterial may be innocuous or a toxicant.

The degree to which a given nanoparticle might produce toxicity depends largely on how the tissues interact with the nanoparticle and the inherent toxicity of the particle whether it may be toxicant or toxin in nature (Hobson et al. 2016).

Nanotechnology is a rapidly emerging field of great interest and promise that is heavily engaged in the development of "engineered" nanomaterials for a very wide variety of applications including use in medical devices. As new nanomaterials and devices utilizing them are developed and commercialized, hazard information also needs to be generated to reassure regulators, workers, and consumers that these materials and devices may be used safely.

10.3 Characterization of Nanomaterials

The biological properties, including tissue absorption, of nanomaterials are closely tied to the physical characteristics, including size, shape, dissolution rate, agglomeration state, and surface chemistry, to name a few. Furthermore, these properties can be altered by the medium used to suspend or disperse these water-insoluble particles. Unfortunately, some older toxicology literature often lacks much of the nanomaterial characterization information that toxicologists and regulators need to accurately assess potential hazards of emerging, engineered nanotechnologies. Toxicologists often need to know a relatively complex set of characteristics for a given particle to better investigate, evaluate, and describe its interactions with biological systems. A lack of adequate characterization also leads to different laboratories reporting discordant results on seemingly the same test of nanomaterial because subtle differences in the particle characteristics are leading to differences in the dispersion medium used that resulted in altered properties and variable toxicity of the particle. For these reasons, adequate characterization data should accompany and is now often required of all scientific publications concerned with nanomaterials biological effects.

The scientific community, regulatory agencies, environmentalists, and most industry representatives all agree that more effort is required to ensure the responsible and safe development of new nanotechnologies. Characterizing nanomaterials is a key aspect in this effort. There is no universally agreed upon minimum set of characteristics, although certain common properties are included in most recommendations. Therefore, nanomaterial characterization sometimes is more like a puzzle put together with various, potentially relevant measurements rather than having a single analytical measurement modality to rely on as with most toxicants and toxins. Therefore, the important elements of nanoparticle characterization must emphasize and illustrate with a systematic approach both on the physical and chemical characteristics of a given nanomaterial. Often, this characterization includes a comprehensive overview or summary describing the nanotechnological properties that are thought to be most significant to toxicological testing along with details as to the methods used for characterizing a nanomaterial that often includes the measurement of size, zeta potential, surface properties, and various types of imaging as appropriate for each specific nanomaterial (Boverhof and David 2010; Powers et al. 2012).

As indicated above, the need for proper characterization is essential to the advancement of reliable nanomaterial science including toxicology, pharmacology, and food science to the extent that some scientific journals now have instructions to authors regarding adequate characterization of test materials, including nanomaterials, before an article will be accepted for publication (Powers et al. 2012; International Journal of Nanotechnology and Allied Sciences 2018). Sufficient details of the nanomaterial characterization as well as the techniques employed to enable the work to be repeated with adequate fidelity to the original, reported, and scientific study must be provided before the manuscript is accepted for publication (International Journal of Nanotechnology and Allied Sciences 2018; Nature Nanotechnology 2018).

The methodology used to characterize nanomaterials is ever increasing and includes a wide array of methods and protocols that are continuously improving. The specific set of tests to be conducted is dependent on the actual nanomaterial with consideration to the intended use and conditions of exposure. Transmission electron microscopy (TEM), for example, is considered the standard for nanomaterial size characterization, but there are many other particle size determination methods including dynamic light scattering (DLS) and hyperspectral imaging (HI) that rapidly provide both particle size and size distribution data with less time in sample preparation and analysis. The methods recommended to adequately support good scientific and regulatory practice are discussed in the growing current literature on this subject and are listed in regulatory guidance documents (Boverhof and David 2010; Powers et al. 2012).

An example of some of different methods that may be used for the characterization of nanomaterials is shown in Table 10.1. Transmission electron microscopy (TEM) is often considered to be most reliable for nanoparticle characterization of size, shape, surface area, and detection in biological matrices, as well as stability evaluation in bulk and in test solutions. Scanning electron microscopy (SEM) is useful in evaluating particle size, shape, chemical composition, as well as evaluation of surface agglomeration, aggregation, and three-dimensional characteristics in test solutions as well as in biological matrices. Atomic force microscopy (AFM) may be used for particle size, agglomeration, aggregation, as well as stability

Table 10.1 Different methods used for characterization of nanomaterials

Method of analysis	A	B	C	D	E	F	G	H	I	J
Agglomeration and aggregation state		X	X			X	X			
Chemical composition and purity	X	X		X	X				X	
Crystal structure (if applicable)										X
Detection in extractable/leachable matrices	X	X		X	X		X		X	X
Detection in tissues and biological fluids	X	X		X	X		X		X	X
Evaluation of endotoxins and sterility										X
Mean particle size and size distribution	X	X	X			X	X			
Particle shape	X	X								
Protein corona formation and composition	X	X				X				X
Spectral properties and analysis									X	X
Stability in bulk and in packaging	X			X		X		X		
Stability in test media and dosing solutions	X		X			X				
Surface area	X							X		
Surface charge						X				
Surface chemistry (reactivity and hydrophobicity)										X
Surface structure and morphology		X						X		
Three-dimensional analysis										X

A = Transmission electron microscopy; B = scanning electron microscopy; C = atomic force microscopy; D = ICP-MS; E = XRF or NAA; F = particle size analysis (various); G = AFFF; H = particle surface area; I = confocal Raman; J = other standardized method (see text for more details)

characterization. X-ray fluorescence (XRF) and neutron activation analysis (NAA) are useful in the determination of chemical elemental composition of particles and nanosurfaces and can be used for nanomaterial detection in complex biological matrices including tissues. Asymmetrical flow field-flow fractionation (AFFF) separates constituents based on hydrodynamic size and is emerging as a powerful tool for obtaining high-resolution information on the size, molecular weight, composition, and stability of nanoscale particles in liquid media. Dynamic light scattering (DLS) is commonly available in many laboratories that work with nanoparticles and is often useful for particle size and size distribution evaluation. Hyperspectral imaging (HI) is a popular technique for localizing and following the movement of nanoparticles in complex matrices. Confocal Raman microscopy (CRM) is an improvement on Raman spectroscopy that provides information about depth and buried structures including nanomaterials in thin samples including biological matrices. BET (Brunauer-Emmett-Teller) analysis is considered the best technique to determine the surface area of nanoparticles. This technique is based on adsorption and desorption principle and Brunauer-Emmett-Teller (BET) theorem usually using nitrogen gas. Endotoxin and sterility determinations may be accomplished using a variety of standardized methods that are available from compendial sources such as the US Pharmacopeia (USP). Further detail for each of these methods may be easily located in current nanotechnology literature, including nanotoxicology and nanomaterials analysis literature. Some methods may be widely available, and others must be conducted in specialized facilities by experienced personnel that often do contract analytical work.

Some important practical considerations in evaluating the safety of nanomaterial use in new medical device products are the following:

- There is primary concern for "unbound" nanomaterials that may be absorbed into the body and disrupt life-essential biologic process.
- It is generally accepted that although many products may employ nanotechnology for one purpose or another, but not all products may not require extensive safety evaluation if it can be firmly established by design data and relevant testing that the nanomaterial is bound to the device and is not bioavailable.
- Even if a nanomaterial does enter the body, safety may be established if the toxicological effects of the nanomaterial are minimal to negligible and that the nanomaterial can be eliminated in a short time without any significant bioaccumulation.
- In vitro diagnostic tests employing nanotechnology generally would not require in vivo safety or biocompatibility evaluation.
- The manufacturer is solely responsible for demonstrating medical device safety clearly a lack of potential adverse effects from exposure to nanomaterials incorporated into the device design.

Central to most all of these practical considerations is the need to ensure adequate characterization of nanomaterials where exposure is possible to support state-of-the-art scientific and regulatory identity, compositional, and stability needs for inclusion as good practices (e.g., good laboratory practices or good manufacturing

practices) quality data. This characterization should, as a minimum, include the evaluation of the physical parameters of nanomaterial dimension, size and size distribution (when applicable), zeta potential (when applicable), chemical composition, surface characteristics, etc. When possible, both in vitro and in vivo interactions with biologic systems should be considered and included when relevant.

The physical characteristics that should be addressed for nanomaterials that may enter the body via different potential routes of exposure include size and size distribution, composition, structure and morphology, surface chemistry, macromolecular weight, surface area, porosity, solubility, surface charge density, purity, sterility, and stability.

In vitro characteristics that should be considered and determined as warranted include:

- Binding to plasma, cells, and tissues.
- Pharmacological effects in isolated cells and tissues.
- Blood contact properties including hemocompatibility and pyrogenicity.
- Cellular uptake, distribution, metabolism, cytotoxicity, and mutagenicity.

In vivo assessment of medical device nanomaterials that are in contact with blood and/or tissues would typically include when possible:

- Route and rate of absorption.
- Pharmacokinetics/toxicokinetics (descriptive rather than modeled).
- Protein binding.
- Tissue distribution.
- Metabolism.
- Route and rate of elimination/excretion.
- Safety.
  - For route and indication.
  - Using appropriate species.
  - Exposures with appropriate route and duration.
  - Relevant toxicological endpoints.

10.4 Engineered Nanomaterial Biological Interactions and Fate

There are several challenges in conducting studies on the biological fate and effects of nanomaterials. These challenges include (1) inconsistency in the quality of nanomaterials available for study (e.g., lot-to-lot variations in important characteristics such as size, shape, and surface properties and deviations from the labeled description); (2) limitations in quantities available for study; (3) uncertainty as to the proper dose metric (e.g., mass vs surface area vs particle concentration); (4) difficulty detecting and quantifying nanomaterials in tissues; and (5) changing chemical and

physical properties of nanomaterials with time, handling, and in biological environments (Hobson et al. 2016).

As nanomaterials associated with medical devices that are in contact with the tissues, there is often an immediate, chemical environmental change that can alter nanomaterial properties that may influence biological activity, including size, shape, and surface properties such as charge, catalytic properties, and adsorbed materials (i.e., formation of a lipid and proteinaceous "corona" coating around the nanoparticle) (Hobson et al. 2016). Changes in the dynamic environment of the tissues may occur and include pH, ionic strength, and composition of body fluids, killing of microflora, and alterations in cellular functions. Contact surfaces may also change even if the nanomaterial is bound to the medical device depending on contact surface area, interaction of body tissues with the nanomaterial, and resilience of tissue-protective barriers (Hobson et al. 2016).

10.5 Use of Nanomaterials in Medical Devices

Nanomaterials have increasing utilization in the design and composition of many different types of medical device products such as those listed below.

- Topical medical device products (wound and burn dressings, etc.)
- Intravascular devices including antimicrobial catheters, coated stents, etc.
- Neurodegenerative drug delivery devices.
- Biosensors.
- Implants (biofilm issue).
- Dental materials including crowns, cements, fillers, and composites.
- Osteopathy materials disease nano.
- Diagnostic imaging agents.
- Drug/medical device combinations.

Many different forms of nanotechnology have already been used and are currently available or are emerging for a variety of applications in the design and development of nano-enabled medical devices, and the list is growing continually. It is, therefore, possible to identify different types of medical devices that incorporate nanotechnology and to describe how the evaluation safety can be effectively addressed. It is important to the advancement of new and potentially significant nanoproducts that the utilization of nanotechnology in their designs never be presumed to be inherently safe or potentially harmful. Instead, it is best to embrace the concept of "safety by design" which is best facilitated with a prudent, case-by-case, approach to safety and risk assessment. In fact, most regulatory bodies that evaluate medical device risk relative to the benefits of use ascribe to this type of approach. This is why most government regulations that address the testing and evaluation of device safety provide "guidance" on how to approach the assessment rather than prescribe in detail how to design and conduct safety tests. Therefore, it is generally best to have an experienced nanomaterial characterization scientist as well as a

board-certified, toxicologist that have worked with nanotechnology safety assessments involved in designing and planning the test and data analysis plan for each nano-enabled device on a case-by-case basis building on past experience and ever advancing the state of the art for effective as well as efficient safety evaluation.

There are many reasons why incorporation of nanomaterials in the design and function of medical devices is desirable. Most of these are related to nanomaterial-enhanced mechanical, electrical, magnetic, optical, thermal, biological, and/or chemical properties. Some examples of desirable uses of nanotechnology in specific types of medical devices are provided below with references for further, reading.

1. Antimicrobial nanomaterials such as silver nanoparticles which are among the first uses of nanotechnology in medical devices including various types of wound dressings (Fong and Wood 2006).
2. Nanostructured surfaces of medical devices for various advantages (Aninwene II and Webster 2013).
3. Miniaturization of medical devices and biosensors making them smaller, lighter, and more "wearable" (e.g., "smart" contact lenses and "stretchable" electronics and biosensors) (Park et al. 2018; Chu et al. 2016; Lee et al. 2013; Kim et al. 2015; Besteman et al. 2003).
4. Nanoencapsulation of drugs into or onto medical devices (e.g., controlled release, targeted delivery, triggered release applications) (Kumari et al. 2014; Yoo et al. 1999; Pamornpathomkul et al. 2017; Borhani et al. 2018; Yin et al. 2014).
5. Advanced structural materials for improved durability, strength, lower weight, flexibility, etc. (Aninwene II and Webster 2013; Park et al. 2018; Chu et al. 2016; Lee et al. 2013; Kim et al. 2015).

There are also undesirable aspects to incorporation of nanomaterials in medical devices when they may be unnecessary. These are generally related to the uncertainty of different types of risks including environmental, toxicological, undesirable biological effect, applicable exposure limits, and incomplete knowledge of the mode of action of the nanomaterial component. It is, therefore, important that careful consideration be made in the design phase when nanomaterials are being considered and to apply safety by design concepts to best assure success.

There is now an emerging history with the development of nanotechnology-enabled medical devices that show areas where there has been success as well as where there are significant challenges and obstacles to safe and effective commercialization (Borhani et al. 2018; Hobson 2009, 2016). This literature should be examined carefully to arrive at a state of the science perspective and approach to safety by design.

The development of nanomaterial-coated stents provide a current and valuable example of an often complex device-drug combination that provides substantial insight into the issues involved in commercialization and regulation of nanomaterials associated with medical devices. This is because in designing a stent, biological responses to the components of the stent must always be considered very carefully due to the intimacy in which they are in contact with the body as well as where they

may be applied (Borhani et al. 2018). The stent surface, in particular, plays a critical role in the success of stent implantation (Borhani et al. 2018). In developing advanced nanomaterials for application to stent surfaces, there are several coating technologies for polymer-free stents including direct coating, crystallization of the drug, nano- and microporous surfaces, inorganic porous coating, macroporous drug reservoirs, coating of nanoparticles, and self-assembled monolayers (Chen et al. 2015).

All cardiovascular implants, including stents, must provide a biocompatible surface with essential functions of anticoagulation, anti-hyperplasia, anti-inflammation, and pro-endothelialization after the implantation (Li et al. 2017). To further elaborate, the surface must meet at least three important requirements:

First, inhibit the inflammatory reaction for impeding the thrombosis formation, inhibit excessive smooth muscle cell (SMC) proliferation, and prevent intimal hyperplasia.

Second, fast endothelialization from the early time of implantation to promote the formation of endothelial layer on the stent surface within 1 month. A fast endothelialization process is essential to decrease the risk of thrombosis to the least amount.

Third, avoidance of adverse material-tissue interface interactions as it is necessary for the surface to be biocompatible, especially after complete drug elution from the surface coating (Liu et al. 2014).

There is currently a significant need for the improvement of coating technology to supply stent manufacturers to optimize drug-loading conditions (Borhani et al. 2018). Many novel techniques have been employed for improving drug release from stent coatings such as utilizing grooves and cavities on the stent struts, constructing nanocarriers like nanopores, nanofibers, and nanoparticles, and taking advantage of bioresorbable stent materials with specific drug molecules and utilizing gene NPs to specifically inhibit proliferation of vascular SMCs (Hu et al. 2015; McGinty et al. 2014; Nakazawa et al. 2008). Nanotechnology has been widely utilized to manipulate materials for developing current treatments and appears to be applicable in stents and/or within the coating layer of stents. Nanometer drug carriers with excellent biological and physicochemical properties are considered as efficient delivery tools to be used in cardiovascular stents. A drug may be taken up or be covalently bonded to the surface of nanocarriers or even be wrapped into the nanocarriers. Applying nanocarriers, as a stent-modified coating, enhances the localized drug delivery to injured locations because the cellular absorption of nanosized drugs can sometimes be more efficient and improve the duration as well as tissue distribution of the drug over other forms of administration. In addition, nanoparticles with much larger surface area in comparison to bulk or microstructures can achieve an effective slow drug release (Hu et al. 2015). In the light of developing nano-modified stents, the ability of peptide amphiphile-based nanomatrix coating for stents under physiological flow conditions in vitro has been evaluated, and the results indicate the capability of the nanomatrix-coated stent for reendothelialization, reducing neointimal hyperplasia, reducing restenosis, preventing thrombosis, and alleviating inflammatory response (Alexander et al. 2017). In an attempt to enhance the efficacy of drug-eluting stents,

highly oriented nanotubes were grown vertically on a titanium-based alloy stent platform that showed potential application for a self-expandable stent (Saleh et al. 2017). These self-grown nanotubes showed potential as a powerful tool for surface modification to enhance endothelial proliferation, to prevent vascular smooth muscle cell (VSMC) proliferation, and also as drug reservoirs (Saleh et al. 2017). Polymer nanofibers have been recently attracted considerable attention and have turned to a hotspot research focus. The properties of this new coating technique include small pore size, high porosity, large surface area, superior mechanical properties, and the relative easy surface functionality compared to other forms of coating (Morie et al. 2016). High surface/volume ratio increases drug-loading capability of the polymer nanofiber as well as cell attachment and drug diffusion (Hu et al. 2015). A heparinized titanium coating, heparin/poly-l-lysine (Hep/PLL) nanoparticles, were immobilized on a dopamine-coated titanium surface (Liu et al. 2014). The study aimed to reach a coated surface with time-ordered (three-phase) biological function. Through this study Hep/PLL concentration ratio was optimized to control both the Hep-immobilized density and release behavior, two deciding factors. The advantages of this functionalized surface were its high anticoagulant activity, selective inhibition of VSMC, and vascular endothelial cell (VEC) proliferation.

Animal studies have also demonstrated the predicated time-ordered biofunction to selectively direct an intravascular biological response using an electrospun composite nanofiber made from poly-l-lactic acid (PLLA), chitosan, and paclitaxel (PTX) (Liu et al. 2014). Nanotube-covered stents have, therefore, been suggested as a platform for restoring a functional endothelium and impairing VEC proliferation (Hemshekhar et al. 2016; Li et al. 2015a, b; Zhou et al. 2014). In the presence of nanotube (NT) arrays, it was found that there was an increased proliferation and motility of VECs (Hemshekhar et al. 2016; Li et al. 2015a, b; Yang et al. 2015). On the other hand, VSMCs showed decreased proliferation and motility in comparison to the control (Hemshekhar et al. 2016; Li et al. 2015a, b; Yang et al. 2015).

Mesoporous silica nanoparticles (MSNs) have been regarded as excellent drug carrier for polymer-free stents for the following properties: first, tunable pore size; second, high specific surface area; third, large pore volume; and finally, their biocompatibility to the tissue (Popat et al. 2011; Vivero-Escoto et al. 2010; Zhang et al. 2008). MSNs are a promising and novel drug vehicle due to their unique mesoporous structure that preserving a level of chemical stability, surface functionality, and biocompatibility ensure the controlled release and target drug delivery of a variety of drug molecules (Bharti et al. 2015). The unique mesoporous structure of silica facilitates effective loading of drugs and their subsequent controlled release of the target site. The properties of mesoporous, including pore size, high drug loading, and porosity as well as the surface properties, can be altered depending on additives used to prepare MSNs. Active surface enables functionalization to changed surface properties and link therapeutic molecules. They are used widely in the field of diagnosis, target drug delivery, bio-sensing, cellular uptake, etc. in the biomedical field. This review aims to present the state of knowledge of silica containing mesoporous nanoparticles and specific application in various biomedical fields (Bharti et al. 2015).

10.6 Regulation of Nanodevices (US FDA, EU, Asia Regulations and International Standards)

Many if not most all international regulatory bodies have at least some experience with products that contain nanomaterials. In regulating medical devices that incorporate nanomaterials, it is always important to determine initially whether the product is a device or drug-device combination. This determination may be challenging without consultation with the regulatory body because differentiation as to whether a given product is a drug or device is not always easily determined for some products where the primary mode of action (and metabolism) is more biochemical, biological, or physiological in nature (as typical for a drug substance) or is the primary mode of action physical with essentially no drug substance type action. Combination products may include those components that may have been previously approved as well as components that have no previous regulatory approval. There are several combination possibilities that will define the regulatory path for different types of nano-enabled combination products such as:

- Novel device + novel drug.
- Approved device + novel drug.
- Novel device + approved drug.
- Approved device + approved drug.

Some regulatory bodies, such as the US FDA, have established early stage guidance mechanisms to help determine the regulatory path and applicable regulations for combination products. The US FDA currently has a procedure and process where the sponsor of a new combination product can submit a formal request for designation (RFD) to the agency. Submission of a RFD is always a good first step for initial contact with the regulatory body and will help to improve the efficiency and content of the development plan, most especially for nano-enabled combination products.

10.6.1 United States of America

The US FDA has published guidance documents to address risk assessment needs for the use of nanomaterials in FDA-regulated drugs, cosmetics, food, dietary supplements, tobacco products, and medical devices. These guidance documents address a variety of different types of product that may contain nanomaterials and include the following: (1) “Drug Products, Including Biological Products, that Contain Nanomaterials,” (2) “Safety of Nanomaterials in Cosmetic Products,” (3) “Considering Whether an FDA-Regulated Product Involves the Application of Nanotechnology,” (4) “Assessing the Effects of Significant Manufacturing Process Changes, Including Emerging Technologies, on the Safety and Regulatory Status of Food Ingredients and Food Contact Substances, Including Food Ingredients that are Color Additives,” (5) “Use of Nanomaterials in Food for Animals,” and (6) “Use of

International Standard ISO 10993-1, "Biological evaluation of medical devices - Part 1: Evaluation and testing within a risk management process" (U.S. Department of Health and Human Services Food and Drug Administration 2017; U.S. Food and Drug Administration 2014; U.S. FDA 2014a, b, 2015, 2016).

US FDA guidance for the development of medical products (drugs, devices, and drug-device combinations) that use nanomaterials consistently emphasizes the importance of adequately describing and analyzing nanomaterials and their attributes in the finished product. The guidance suggests manufacturers include a description of the nanomaterials in the product, covering size, charge, composition, complexation, and morphology, in order to ensure consistent quality. The agency suggests that functionality may be included as well. Only nanomaterial properties that could affect a product's performance, safety, or quality must be defined, alongside potential risks posed from any changes. For example, a nanomaterial of a different size or shape could be seen as a batch consistency issue if it affects the quality, efficacy, or safety of the product.

Through years of evaluation of new medical technologies that contain nanomaterials, it is evident that the US FDA generally understands and accepts that the application of nanotechnology may result in product attributes that differ from those of conventionally manufactured products and thus may merit particular examination. All products that involve the application of nanotechnology are not categorically judged as intrinsically benign or harmful. The agency considers its current framework for safety assessment sufficiently robust and flexible to be appropriate for a variety of materials, including nanomaterials, and in fact has evaluated and cleared many medical products containing nanomaterials with success for about two decades. The US FDA maintains a product-focused, science-based regulatory policy and expects that technical assessments will be product-specific, taking into account the effects of nanomaterials in the particular biological and mechanical context of each product and its intended use. It is expected that the particular policies that are applicable for each product area, both substantive and procedural, will vary according to the statutory authorities and relevant regulatory frameworks of the agency. This regulatory policy allows for tailored approaches that adhere to applicable legal frameworks and reflect the characteristics of specific products or product classes evolving technology and scientific understanding (U.S. Food and Drug Administration 2014).

10.6.2 European Union

On April 5, 2017, the European Parliament (EP) approved a regulation on medical devices, as well as a regulation on in vitro diagnostic medical devices (European Union 2017). This medical devices regulation is intended to ensure that medical devices are traceable and comply with European Union (EU) safety requirements and address the use of nanomaterials in medical devices. The critical factor in classifying devices incorporating or consisting of nanomaterials is the potential for nanomateri-

als to be in contact with membranes inside the body. Devices presenting a high or medium potential for such contact will be placed in the highest risk class and will be subject to the most stringent conformity assessment procedures. This regulation included a 3-year transition period for in vivo applied medical devices and a 5-year transition period for the regulation on in vitro diagnostic medical devices to allow manufacturers and authorities time to implement the regulations.

Current EU general medical device regulations MDD (93/42/EEC), AIMDD (90/385/EEC), and IVDD (98/79/EC) have no explicit requirements in relation to nanomaterials (European Union 1990, 1993, 1998). Generally, however, all risks must be assessed and eliminated/reduced as far as possible; material toxicity, tissue compatibility, contaminants, residues, and leachables must be evaluated; and risk of injury must be minimized in connection with device physical features and dimensions.

Internationally recognized standards that apply to medical device nanomaterials include a series of International Organization for Standardization (ISO) 10,993 standards for evaluating biocompatibility of medical devices (International Organization for Standardization 2018). ISO 10993 includes globally harmonized standards for the biological evaluation of medical devices. This includes standards that have been issued by ISO and are used throughout Europe and the US FDA version of ISO 10993-1 that is used in the United States (International Organization for Standardization 2018). Even though ISO 10993 has been formally accepted in Japan, the "Japanese Guidelines for Basic Biological Tests of Medical Materials and Devices" favors certain test methods to evaluate specific categories of biological effects that may or may not include all ISO 10993 standards (International Organization for Standardization 2018).

ISO 10993-22 describes considerations for the biological evaluation of medical devices that are composed of or contain nanomaterials (ISO/TR 10993-22 2017). In addition, this guidance can also be useful for the evaluation of nano-objects generated as products of degradation, wear, or from mechanical treatment processes (e.g., in situ grinding, polishing of medical devices) from (components of) medical devices that are manufactured not using nanomaterials and includes considerations on the characterization of nanomaterials; sample preparation for testing of nanomaterials; release of nano-objects from medical devices; toxicokinetics of nano-objects; biological evaluation of nanomaterials; presentation of results; risk assessment of nanomaterials in the context of medical device evaluation; biological evaluation report; and nanostructures on the surface of a medical device, intentionally generated during the engineering, manufacturing, or processing of a medical device.

Specifically excluded from ISO-10993-22 are natural and biological nanomaterials, as long as they have not been engineered, manufactured, or processed for use in a medical device; intrinsic nanostructures in a bulk material; and nanostructures on the surface of a medical device, generated as an unintentional by-product during the engineering, manufacturing, or processing of a medical device.

Under its ISO/TC 229 nanotechnologies technical committee, the ISO has 66 published standards and 42 under development. However none of the standards under ISO/TC 229 are specific to medical devices even though they may be generally useful and applicable in many cases to resolve terminology, characterization, and testing issues in an internationally acceptable manner (ISO 2008, 2010).

The ISO standards are applied in general by most all countries including the USA, EU countries, Japan, China, India, etc. that are ISO members within each country's established framework for regulation of medical device safety.

At the present time, the UK generally follows the EU regulatory model for medical devices; however, in future the UK may decide to modify the EU or even adopt aspects of the US FDA approach. Regardless, it is reasonable to expect that regulation of medical devices that incorporate nanomaterials in their designs in the UK would be regulated in a manner harmonized with the EU and USA.

10.6.3 Asia (China, Japan, and India)

As noted above, the use and monitoring of medical device enabling nanomaterials in Japan, China, and India essentially follows the same ISO standard approach set forth by ISO 10993 standards and those of ISO/TC 229 and is generally harmonized with the US and EU requirements. While standards for quality control may differ in each country, the overall standard for safety to include devices that incorporate nanomaterials is essentially the same for all countries worldwide as proposals to improve safety assessment continue to advance and are adopted (JSA – JIS T 14971 2012; Limaye et al. 2014; Igami and Okazaki 2007).

10.7 Risk from Exposure to Medical Device Nanomaterials

The highest potential for adverse effects of medical device nanomaterials is generally associated with devices: (1) in which the nanomaterial is intended to be released, (2) that are composed of free nanomaterials, (3) which contain nanomaterials that are not chemically bound or are loosely bound to the device, (4) where there is release/loosening of nanomaterials present within or on the device surfaces during normal use, or (5) if there are chemical breakdown or wear-and-tear processes due to (bio)degradation of medical devices and (6) with nanomaterials that are released or become bioavailable are sequestered within the body and are not significantly eliminated or excreted. Devices that are grinded, polished, or shaped during application such as dental and prosthetic materials may produce significant amounts of free nanoparticles. In general, nanomaterials that are strongly chemically bound to the device surfaces are significantly less of a concern for potential adverse effects than those that are loosely bound to the device.

For most patients, the potential inhalation (including nasal administration), parenteral, dermal (topical, intradermal, and subcutaneous), mucosal, oral, ocular, and otic exposure routes for nanomaterials should always be considered. Exposure pathways where a nanomaterial may enter systemic circulation and achieve a more widespread distribution within the body should always be toxicologically considered when there is a possibility of exposure to nanomaterials involved.

10.8 Engineering Safe Nanoparticles for Medical Device Applications and Therapeutics

As noted above, nanomaterials are increasingly being incorporated into medical devices for a variety of purposes. Some new medical devices (e.g., stents, etc.) are drug-device combinations with small molecule or biologic drug substances incorporated with a device and are being used for the treatment of many diseases. Many of these treatments typically require sustained release characteristics or are designed to allow for repeated dose administration usually by injection, but other routes such as dermal, oral, etc. are also possible.

Nanotechnology is a field of research that has been stressed as a very valuable technical tool with which to approach the prevention and treatment of different human health disorders (Hobson et al. 2016; Hobson 2009, 2016; Limaye et al. 2014). Significant progress has been made in the development of nanotechnology-enabled medical devices in recent years as well as for methods of diagnostic testing that involve nanotechnology-enabled medical test devices for an array of pathophysiological situations (Hobson et al. 2016; Fong and Wood 2006; Park et al. 2018; Lee et al. 2013; Besteman et al. 2003; Kumari et al. 2014; Pamornpathomkul et al. 2017; Borhani et al. 2018; Hobson 2009, 2016; Alexander et al. 2017; Morie et al. 2016; Limaye et al. 2014; Chen et al. 2012).

The design of new medical devices that incorporate engineered nanoparticles (especially in the size range of 1–100 nm) has a broad-spectrum potential for useful applications in electronics, chemistry, environmental protection, and medicine. In the field of biomedicines, the use of nanoparticles has been growing exponentially, and there has been some concern for potential adverse effects (Roy et al. 2014). To increase the treatment efficacy and to reduce the side effects, nano-based drug delivery systems are being applied and include polymeric, solid lipid, hydrogels, gold, silver nanosystems, etc. (Riasat et al. 2016; Sengupta et al. 2014; Yang et al. 2014). Metal nanoparticles are currently used in the medical and food industry such as iron, cobalt, copper, zinc, and silica (Sengupta et al. 2014). These types of nanoparticles are physiologically important due to their different synthesis route and different physical and chemical properties. Although there are many advantages of the nano-based drug carrier systems, but toxicity parameters cannot be overlooked as there are various toxicological routes associated with the exposure of nanoparticles in human being and environment, while toxicological effects on human beings are still ambiguous (Roy et al. 2014; Riasat et al. 2016; Sengupta et al. 2014; Yang et al. 2014). Apart from this there are many types of safe nanotechnology applications that have been and are being used in the design of medical devices. These nano-enabled devices have promising and effectual prospect toward advancing the field of nanomedicine.

Various parameters such as shape, size, dose, surface characteristics, and translocation are known to play very distinguishing role in the nanoparticles toxicity. These parameters are not still fully understood in vivo. Therefore it is important to have knowledge of toxicologically significant nanomaterial interactions with the

intended biological tissue and/or system prior to final nanomaterial synthesis and medical device fabrication. These parameters can influence delivery efficiency and distribution of drugs. In designing nanoparticles a few key points should be considered such as the whether nanoparticle surface charge and any exposed surface ligands on the particle are compatible with elements of the blood and with circulatory system tissues. Also, nanoparticles can escape from clearance mechanisms and antibody opsonization can occur causing drug resistance and impact to other related metabolic parameters. Therapeutic applications of nanoparticles are dependent on these parameters when interacting with the biological system (Borhani et al. 2018; Hobson 2009; Roy et al. 2014; Riasat et al. 2016; Sengupta et al. 2014).

When a nanomaterial meets a biological system, it is being increasingly understood that there are many potential outcomes ranging from no effect to a profound biological response depending on the nanomaterial characteristics as well as the anatomical contact location (Roy et al. 2014; Frohlich and Roblegg 2012; McGill and Smyth 2010; Jachak et al. 2012; Cedervall et al. 2007; Wang et al. 2011; Faunce et al. 2008). As an example of the latter, the effects of development of a biofilm on or a protein corona around nanosurfaces and nanoparticles, respectively, vary with biological systems and locations within these systems. In some cases, cytotoxicity is reduced, perhaps by decreasing cellular nanoparticle uptake or by mitigating cell membrane damage (Lundqvist et al. 2008; Monopoli et al. 2011; Lynch et al. 2007; Jiang et al. 2010; Casals et al. 2011; Safi et al. 2011; Ge et al. 2011). With respect to the potential for protein corona surface modifications to synthetic materials, Lundqvist et al. (Lundqvist et al. 2008) found that the protein corona on 50 nm carboxyl-modified polystyrene nanoparticles varied in relation to particle size and surface modification. Highly abundant plasma proteins such as inter-alpha-trypsin inhibitors, serum albumin, clusterin, and vitronectin appear to be common to all coronas, while less abundant proteins appear to vary with nanoparticle size and surface characteristics (Lundqvist et al. 2008). Using polystyrene nanoparticles in human plasma, Zhang et al. (Zhang et al. 2011) classified nanoparticles with respect to protein coating based on size and surface properties of the parent particles. The protein corona was also demonstrated to be at equilibrium within 5 minutes of nanoparticle exposure.

The use of nanomaterials in medical device and implant design also may be useful in the prevention of biofilms (Naik et al. 2015; Zhao et al. 2009). Clearly, variable coronal protein and biofilm composition may lead to some of the different biological effects (useful as well as potentially unwanted) of nanoparticles and nanosurfaces that otherwise seem identical (Lai et al. 2013; Naik et al. 2015; Zhao et al. 2009). These studies further emphasize the need for adequate characterization of nanoparticles both before and after exposure to culture media or biological fluids (Lundqvist et al. 2008; Lai et al. 2013; Naik et al. 2015).

Animals and human studies have shown that after intravenous, inhalation and through oral exposure, nanoparticles are primarily distributed to the liver, lungs, heart, spleen, and brain. They may be sequestered in the tissues or be excreted via the urine or feces depending on whether the nanoparticle has characteristics favorable for such elimination (Hagens et al. 2007; Bahadar et al. 2016; Oberdörster et al.

2005). Thus it is clear to see why unbound nanomaterials may pose a greater risk of systemic toxicity as opposed to nanomaterials that remain tightly bound to a medical device. Also important is the possibility that during clearance of nanoparticles from the body, the components of the immune system are activated and should be a consideration in the use of nanomaterials in medical device design (Hagens et al. 2007; Bahadar et al. 2016; Oberdörster et al. 2005; Garnett and Kallinteri 2006; Takenaka et al. 2001; Zolnik et al. 2010; Dobrovolskaia et al. 2016).

10.9 Safety Evaluation of Nano-Enabled Medical Devices

The approach to safety evaluation of nanodevices varies with (1) the intended use of the device, (2) design of the device, and (3) the duration of use of the device.

Based on prior research, it is now quite clear that there can be unique properties associated with submicron (<1 micron) or nanotechnology components such as aggregation, agglomeration, immunogenicity, or toxicity (Kunzmann et al. 2011). Therefore, while safety testing of nano-enabled medical devices should proceed as much as possible similarly to medical devices without incorporated nanomaterials, it is clear that devices with submicron components may require additional, specialized techniques for characterization and biocompatibility testing (ASTM International 2014, 2018; Rivera et al. 2010). Limitations as well as additional characterization may apply when using chemical leachates-based ISO 10993-12 test conditions for the analysis of submicron component biocompatibility assessments. It is therefore imperative that the sponsor consults relevant literature and standards during the development of test protocols for device-specific submicron or nanotechnology component biocompatibility assessments and contacts the respective regulatory review division for the country of registration prior to initiation of any tests to best assure applicability and acceptability.

When conducting biocompatibility assessment of devices with submicron components, the following should be considered: (1) careful and method appropriate characterization of the test article; (2) selection of extraction conditions (e.g., solvent type) that avoid testing artifacts or significant destruction of the medical device; and (3) assurance that the test article used is representative of the device that is intended to be used clinically.

For test selection, the following items are important:

- Consideration of standard biocompatibility tests in the context of contemporary literature regarding the validity of individual tests for assessment of devices with submicron components.
- Assurance that the submicron components will not interfere with the conduct of the chosen test.
- Consideration of any additional toxicity issues that might be relevant to submicron particles, such as absorption, distribution, and accumulation into organs, potential metabolism, and elimination, since there are greater concerns associated

with submicron particles that cannot be readily detoxified and/or eliminated from the body.

As an example, copper oxide nanoparticles are often used in semiconductors, antimicrobial reagents, heat-transfer fluids, and intrauterine contraceptive devices and are generally considered safe if systemic exposure to unbound copper nanoparticles is controlled and maintained at a toxicologically negligible level (Aruoja et al. 2009). Experimentally, copper nanomaterials have been documented to possess toxic effects on the liver and kidney in experimental animal models (Lei et al. 2008; Chen et al. 2006; Meng et al. 2007). After oral administration and interaction with gastric juice, highly reactive ionic copper is formed, which is then accumulated in the kidney of exposed animals (Chen et al. 2006; Meng et al. 2007). In one in vitro study, copper oxide nanoparticles (50 nm) have been reported as being genotoxic and cytotoxic along with disturbing cell membrane integrity and inducing oxidative stress (Ahamed et al. 2010). Nevertheless, copper remains a valuable and useful nanomaterial in the design and fabrication of medical devices including those in intimate contact with the body as a recent publication evaluating the potential safety of copper/low-density polyethylene nanocomposite (nano-Cu/LDPE) materials used in intrauterine devices (IUD) in rodent exposure models where the nanomaterial was intravenously, acutely, and subchronically administered or implanted at levels tenfold higher than the clinical exposure for up to 13 weeks (Hu et al. 2018). Overall, the results from these studies indicated that the nano-Cu/LDPE IUD did not induce systemic toxicity under experimental conditions of the recommended standard practices, suggesting that the novel material IUD is safe and feasible for future contraceptive applications.

As with all medical devices, the testing required to demonstrate safety depends on the intended use of the device and should be considered and developed on a case-by-case basis. The testing of stents, IUDs, and catheters to demonstrate the safety of nano-enabled designs can serve as useful examples upon which to develop a testing strategy appropriate to a specific medical device. Other chapters in this work address many of these considerations and are useful in determining what tests should be included in safety assessments for different, nano-enabled device types.

10.10 Summary and Conclusions

Nanomedicine is an application of nanotechnology in the areas of healthcare and disease diagnosis and treatment and prevention of disease that includes pharmaceutical agents and medical devices that incorporated nanomaterials. Nanomedicine has been defined as “the monitoring, repair, construction and control of human biological systems at the molecular level, using engineered nanodevices and nanostructures” (Freitas 1999; Limaye et al. 2014). By common definition, the nanoscale ranges from 1 nm to 100 nm, however, in practice, nanomaterials may go beyond this scale and often go above 100 nm up to 999 nm or occasionally, below 1 nm. Biologically, including toxicologically, what is most important with respect to

safety assessment is whether or not a given nanomaterial imparts a characteristic that is quantal and differentiable from the biological and toxicological properties of the molecular composition of the nanomaterial. In other words, the toxicity of a nanomaterial is most accurately described with its composition, structure, and size in consideration rather than just its chemical composition. This is the essence of the emerging subdiscipline of toxicology that is termed "nanotoxicology" that seeks to understand why nanoscale materials may be toxicologically differentiable in their effects from the individual chemical components used in their fabrication (Monteiro-Riviere and Tran 2016).

Naturally occurring nanomaterials are generally not as great a concern as engineered nanomaterials which are the products of the rapidly advancing field of nanotechnology that includes nano-enabled products in most every sector of commerce including medical devices.

At present, in the regulation of nanomedical products, there is no uniform, internationally accepted definition of nanotechnology, and this may be an area of controversy for global harmonization of nano-enabled medical device testing. Fortunately, there is an ongoing effort to develop harmonized standards by internationally recognized and accepted organizations such as the International Organization for Standardization whose international technical committees have been producing an increasing number of finalized standards for nanotechnology.

The incorporation of nanomaterials into the design of a medical device must be considered carefully as current research demonstrates that:

- Nanomaterials tightly bound to a medical device may pose less of a systemic exposure risk than unbound materials.
- When systemically bioavailable, the persistence of some nanoparticles (long-term retention in target tissues) is a critical factor in determining whether nanoparticles penetrate cells and are retained long enough to provide the continuing inflammatory and immunological signals as well as other biological stimuli which may lead to toxicologically significant pathological changes.
- Nanomaterials incorporated in medical devices can provide beneficial therapeutic effects or induce adverse events such as cellular oxidative stress and the initiation of apoptosis or cellular necrosis or both.
- Nanoengineered modification of the surface activity of otherwise hydrophobic nanoparticles (e.g., formation of a protein "corona" and/or biofilm) may provide benefit or enhance their toxicity potential.
- There currently is no clear pattern of structure-activity relationships (SAR) which would permit predictability of nanomaterial (or nanoparticle) toxicity.
- There is a lack of useful chronic exposure studies with nanomaterials, especially those which clearly demonstrate dose-response toxicological effects and provide for the determination of well-characterized no-observed-adverse-effect levels (NOAELs).
- Engineered nanomaterials that may be incorporated into a medical device are rapidly developing, and even seemingly similar nanomaterials may have different biological as well as toxicological properties such that both efficacy and safety testing must be done on a case-by-case basis.

- Adequate characterization of nanomaterials used in nano-enabled medical devices administered in either in vitro or in vivo experimental toxicological studies remains critical to the interpretation of the results and comparison of findings from these studies.

In summary, there is a rapidly developing body of scientific literature that addresses the safety of emerging nanomaterials for use in the fabrication of medical devices. This literature shows that the incorporation of nanomaterials into some devices can provide significant benefit as well as create toxicological concerns. Accepting and understanding these concerns are proven valuable in the process of "safety by design" for nano-enabled medical devices, and promises increased success as the knowledge base for nanomedicine advances. Toxicological issues are being actively investigated for a growing number of engineered nanotechnologies that may have been use in the fabrication of nano-enabled medical devices and evidence a clear potential for safety. The incorporation of nanotechnology into medical device design is most certain to have major impacts toward advancing both our knowledge of the utility of nanotechnology in medicine and toward improving the quality of life of those with a wide range of afflictions that need better therapeutics as well as hope for a better future.

References

Ahamed, M., Siddiqui, M. A., Akhtar, M. J., Ahmad, I., Pant, A. B., & Alhadlaq, H. A. (2010). Genotoxic potential of copper oxide nanoparticles in human lung epithelial cells. *Biochemical and Biophysical Research Communication, 396*(2), 578–583.

Alexander, G. C., Hwang, P. T., Chen, J., Kim, J., Brott, B. C., Yoon, Y.-S., & Jun, H.-W. (2017). Nanomatrix coated stent enhances endothelialization but reduces platelet, smooth muscle cell, and monocyte adhesion under physiologic conditions. *ACS Biomaterials Science & Engineering, 4*, 107–115.

Aninwene, G. E., II, & Webster, T. J. (2013). Nanostructured and nanoparticulate metals: Redefining the field of medical devices. *Journal of Powder Metallurgy and Mining, 2*, 1–3.

Aruoja, V., Dubourguier, H. C., Kasemets, K., & Kahru, A. (2009). Toxicity of nanoparticles of CuO, ZnO and TiO2 to microalgae Pseudokirchneriella subcapitata. *Science of the Total Environment, 407*(4), 1461–1468.

ASTM International. (2014). ASTM F1904–14 "Standard Practice for Testing the Biological Responses to Particles in vivo." ASTM International, West Conshohocken, PA, 2014, www.astm.org

ASTM International. (2018). ASTM F1903-18. "Standard Practice for Testing for Biological Responses to Particles In Vitro." ASTM International, West Conshohocken, PA, 2018, www.astm.org

Auffan, M., Rose, J., Bottero, J. Y., Lowry, G. V., Jolivet, J. P., & Wiesner, M. R. (2009). Towards a definition of inorganic nanoparticles from an environmental, health and safety perspective. *Nature Nanotechnology, 4*, 634e641.

Bahadar, H., Maqbool, F., Niaz, K., & Abdollahi, M. (2016). Toxicity of nanoparticles and an overview of current experimental models. *Iranian Biomedical Journal, 20*(1), 1–11.

Besteman, K., Lee, J. O., Wiertz, F. G. M., Heering, H. A., & Dekker, C. (2003). Enzyme-coated carbon nanotubes as single-molecule biosensors. *Nano Letters, 3*, 727–730.

Bharti, C., Nagaich, U., Pal, A. K., & Gulati, N. (2015). Mesoporous silica nanoparticles in target drug delivery system: A review. *International Journal of Pharmaceutical Investigation, 5*(3), 124–133.

Borhani, S., Hassanajili, S., Tafti, S. H. A., & Rabbani, S. (2018). Cardiovascular stents: Overview, evolution, and next generation. *Progress in Biomaterials, 7*, 175–205.

Boverhof, D. R., & David, R. M. (2010). Nanomaterial characterization: Considerations and needs for hazard assessment and safety evaluation. *Analytical and Bioanalytical Chemistry, 396*(3), 953–961.

Boverhof, D. R., Bramante, C. M., Butala, J. H., Clancy, S. F., Lafranconi, M., West, J., & Gordon, S. C. (2015). Comparative assessment of nanomaterial definitions and safety evaluation considerations. *Regul Toxicol Pharmacol, 73*(1), 137–150.

Casals, E., Pfaller, T., et al. (2011). Hardening of the nanoparticle-protein corona in metal (Au, Ag) and oxide (Fe(3) O(4), CoO and CeO(2)) nanoparticles. *Small, 7*(24), 3479–3486.

Cedervall, T., Lynch, I., et al. (2007). Understanding the nanoparticle-protein corona using methods to quantify exchange rates and affinities of proteins for nanoparticles. *Proceedings of the National Academy of Sciences, 104*(7), 2050–2055.

Chen, Z., Meng, H., Xing, G., Chen, C., Zhao, Y., Jia, G., Wang, T., Yuan, H., Ye, C., Zhao, F., Chai, Z., Zhu, C., Fang, X., Ma, B., & Wan, L. (2006). Acute toxicological effects of copper nanoparticles in vivo. *Toxicology Letters, 163*(2), 109–120.

Chen, W., Weng, S., Zhang, F., et al. (2012). Nanoroughened surfaces for efficient capture of circulating tumor cells without using capture antibodies. *ACS Nano, 7*(1), 566–575.

Chen, W., Habraken, T. C., Hennink, W. E., & Kok, R. J. (2015). Polymer-free drug-eluting stents: An overview of coating strategies and comparison with polymer-coated drug-eluting stents. *Bioconjugate Chemistry, 26*, 1277–1288.

Chu, H., Jang, H., Lee, Y., Chae, Y., & Ahn, J. H. (2016). Conformal, graphene-based triboelectric nanogenerator for self-powered wearable electronics. *Nano Energy, 27*, 298–305.

Dobrovolskaia, M. A., Shurin, M., & Shvedova, A. A. (2016). Current understanding of interactions between nanoparticles and the immune system. *Toxicology and Applied Pharmacology, 299*, 78–89.

Donaldson, K., & Poland, C. A. (2013). Nanotoxicity: Challenging the myth of nanospecific toxicity. *Current Opinion in Biotechnology, 24*, 724e734.

European Union. (1990). Council Directive on the approximation of the laws of the Member States relating to active implantable medical devices (90/385/EEC). 20 June 1990, pp. 1–17.

European Union. (1993). Council Directive 93/42/EEC concerning medical devices. 14 June 1993, pp. 1–60.

European Union. (1998). Directive 98/79/EC of the European Parliament and of the Council of on in vitro diagnostic medical devices. 27 October 1998, pp. 1–30.

European Union. (2017). European Parliament. Regulation EU 2017/745 on medical devices, amending Directive 2001/83/EC, Regulation (EC) No 178/2002 and Regulation (EC) No. 1223/2009 and repealing Council Directives 90/385/EEC and 93/42/EEC. 5 April 2017, pp. 1–200

Faunce, T. A., White, J., et al. (2008). Integrated research into the nanoparticle-protein corona: A new focus for safe, sustainable and equitable development of nanomedicines. *Nanomedicine (London, England), 3*(6), 859–866.

Fong, J., & Wood, F. (2006). Nanocrystalline silver dressings in wound management: A review. *International Journal of Nanomedicine, 1*(4), 441–449.

Freitas, R. (1999). Nanomedicine [Internet]. [cited 2019 Feb 18]. Available from: http://www.foresight.org/Nanomedicine

Frohlich, E., & Roblegg, E. (2012). Models for oral uptake of nanoparticles in consumer products. *Toxicology, 291*(1–3), 10–17.

Garnett, M. C., & Kallinteri, P. (2006). Nanomedicines and nantoxicology: Some physiological principles. *Occupational Medicine (London), 56*(5), 307–311.

Ge, C., Du, J., et al. (2011). Binding of blood proteins to carbon nanotubes reduces cytotoxicity. *Proceedings of the National Academy of Sciences of the United States of America, 108*(41), 16968–16973.

Hagens, W. I., Oomen, A. G., de Jong, W. H., Casse, F. R., & Sips, A. J. (2007). What do we need to know about the kinetic properties of nanoparticles in the body? *Regulatory Toxicology and Pharmacology, 49*(3), 217–229.

Hamburg, M. A. (2012). FDA's approach to regulation of products of nanotechnology. *Science, 336*, 299e300.

Hemshekhar, M., Thushara, R. M., Chandranayaka, S., Sherman, L. S., Kemparaju, K., & Girish, K. S. (2016). Emerging roles of hyaluronic acid bioscaffolds in tissue engineering and regenerative medicine. *International Journal of Biological Macromolecules, 86*, 917–928.

Hobson, D. W. (2009). Commercialization of nanotechnology. *Wiley Interdisciplinary Reviews: Nanomedicine and Nanobiotechnology, 1*(2), 189–202.

Hobson, D. W. (2016). The commercialization of medical nanotechnology for medical applications. pp. 405–449. In A. Prokop & V. Weissig (Eds.), *Intracellular delivery III. Fundamental biomedical technologies*. Springer International Publishing Switzerland, 2006.

Hobson, D. W., Roberts, S. M., Shvedova, A. A., Warheit, D. B., Hinkley, G. K., & Guy, R. C. (2016). Applied nanotoxicology. *International Journal of Toxicology, 35*(1), 5–16.

Holdren, J. P., Sustain, C. R., & Siddiqui, I. A. (2011, June 9). *Policy principles for the U.S. decision-making concerning regulation and oversight of applications of nanotechnology and nanomaterials*. Memorandum for the heads of executive departments and agencies. https://obamawhitehouse.archives.gov/sites/default/files/omb/inforeg/foragencies/nanotechnology-regulation-and-oversight-principles.pdf. Accessed: 9 May 2018.

Hu, T., et al. (2015). Controlled slow-release drug-eluting stents for the prevention of coronary restenosis: Recent progress and future prospects. *ACS Applied Materials & Interfaces, 7*, 11695–11712.

Hu, L.-X., Hu, S.-F., Rao, M., Yang, J., Lei, H., Duan, Z., Xia, W., & Zhu, C. (2018). Studies of acute and subchronic systemic toxicity associated with a copper/low-density polyethylene nanocomposite intrauterine device. *International Journal of Nanomedicine, 13*, 4913–4926.

Igami, M., & Okazaki, T.. (2007) *Capturing nanotechnology's current state of development via 7 analysis of patents*. OECD Science, Technology and Industry Working Papers; 2007.

International Journal of Nanotechnology and Allied Sciences. (2018). Instructions for authors. http://psmpublishers.org/journal/international-journal-of-nanotechnology-and-allied-sciences. Accessed 31 Dec 2018.

International Organization for Standardization. (2018). ISO 10993-1: Biological evaluation of medical devices – Parts 1 through 12: Evaluation and testing within a risk management process. 2018, 41pp.

ISO, International Organization for Standardization. (2008). Technical specification: Nanotechnologies terminology and definitions for nano-objects, nanoparticle, nanofibre and nanoplate. ISO/TS 80004-2.

ISO, International Organization for Standardization. (2010). Nanotechnologies vocabulary part 1: Core terms. ISO/TS 80004-1.

ISO/TR 10993-22: 2 Biological evaluation of medical devices – Part 22: Guidance on nanomaterials. 2017, 61pp.

Jachak, A., et al. (2012). Transport of metal oxide nanoparticles and single-walled carbon nanotubes in human mucus. *Nanotoxicology, 6*(6), 614–622.

Jiang, X., Weise, S., et al. (2010). Quantitative analysis of the protein corona on FePt nanoparticles formed by transferrin binding. *Journal of the Royal Society Interface, 7*(Suppl. 1), S5–S13.

JSA – JIS T 14971. (2012). Medical devices – Application of risk management to medical devices. 1 March 2012, 94pp.

Kim, J., Lee, M. S., Jeon, S., Kim, M., Kim, S., Kim, K., Bien, F., Hong, S. Y., & Park, J. U. (2015). Highly transparent and stretchable field-effect transistor sensors using graphene–nanowire hybrid nanostructures. *Advanced Materials, 27*, 3292–3297.

Kumari, A., Singla, R., Guliani, A., & Yadav, S. K. (2014). Nanoencapsulation for drug delivery. *EXCLI Journal, 13*, 265–286.

Kunzmann, A., et al. (2011). Toxicology of engineered nanomaterials: Focus on biocompatibility, biodistribution and biodegradation. *Biochimica et Biophysica Acta, 1810*(3), 361–373.

Lai, X., Blazer-Yost, B. L., et al. (2013). Protein expression profiles of intestinal epithelial co-cultures after low level exposure to functionalized carbon nanotubes. *International Journal of Biomedical Nanoscience and Nanotechnology, 3*(1–2), 1–36.

Lee, M. S., Lee, K., Kim, S. Y., Lee, H., Park, J., Choi, K. H., Kim, H. K., Kim, D. G., Lee, D. Y., Nam, S., & Park, J. U. (2013). High-performance, transparent, and stretchable electrodes using graphene–metal nanowire hybrid structures. *Nano Letters, 13*, 2814–2821.

Lei, R., Wu, C., Yang, B., Ma, H., Shi, C., Wang, Q., Yuan, Y., & Liao, M. (2008). Integrated metabolomic analysis of the nanosized copper particle-induced hepatotoxicity and nephrotoxicity in rats: A rapid in vivo screening method for nanotoxicity. *Toxicology and Applied Pharmacology, 232*(2), 292–301.

Li, J., Zhang, K., Wu, F., He, Z., Yang, P., & Huang, N. (2015a). Constructing bio-functional layers of hyaluronan and type IV collagen on titanium surface for improving endothelialization. *Journal of Materials Science, 50*, 3226–3236.

Li, J., Zhang, K., Wu, J., Zhang, L., Yang, P., Tu, Q., & Huang, N. (2015b). Tailoring of the titanium surface by preparing cardiovascular endothelial extracellular matrix layer on the hyaluronic acid micro-pattern for improving biocompatibility. *Colloids and Surfaces, B: Biointerfaces, 128*, 201–210.

Li, J., Wu, F., Zhang, K., He, Z., Zou, D., Luo, X., Fan, Y., Yang, P., Zhao, A., & Huang, N. (2017). Controlling molecular weight of hyaluronic acid conjugated on amine-rich surface: Toward better multifunctional biomaterials for cardiovascular implants. *ACS Applied Materials & Interfaces, 9*, 30343–30358.

Limaye, V., Fortwengel, G., & Limaye, D. (2014). Regulatory roadmap for nanotechnology based medicines: Review. *International Journal of Drug Regulatory Affairs, 2*(4), 33–41.

Liu, T., Zeng, Z., Liu, Y., Wang, J., Maitz, M. F., Wang, Y., Liu, S., Chen, J., & Huang, N. (2014). Surface modification with dopamine and heparin/poly-l-lysine nanoparticles provides a favorable release behavior for the healing of vascular stent lesions. *ACS Applied Materials & Interfaces, 6*, 8729–8743.

Lundqvist, M., Stigler, J., et al. (2008). Nanoparticle size and surface properties determine the protein corona with possible implications for biological impacts. *Proceedings of the National Academy of Sciences, 105*(38), 14265–14270.

Lynch, I., Cedervall, T., et al. (2007). The nanoparticle-protein complex as a biological entity; a complex fluids and surface science challenge for the 21st century. *Advances in Colloid and Interface Science, 134–135*, 167–174.

McGill, S. L., & Smyth, H. D. (2010). Disruption of the mucus barrier by topically applied exogenous particles. *Molecular Pharmaceutics, 7*(6), 2280–2288.

McGinty, S., McKee, S., McCormick, C., & Wheel, M. (2014). Release mechanism and parameter estimation in drug-eluting stent systems: Analytical solutions of drug release and tissue transport. *Mathematical Medicine and Biology, 32*(2), 163–186.

Meng, H., Chen, Z., Xing, G., Yuan, H., Chen, C., Zhao, F., Zhang, C., & Zhao, Y. (2007). Ultrahigh reactivity provokes nanotoxicity: Explanation of oral toxicity of nanocopper particles. *Toxicology Letters, 175*(1–3), 102–110.

Monopoli, M. P., Walczyk, D., et al. (2011). Physical-chemical aspects of protein corona: Relevance to in vitro and in vivo biological impacts of nanoparticles. *Journal of the American Chemical Society, 133*(8), 2525–2534.

Monteiro-Riviere, N. A., & Tran, C. L. (Eds.). (2016). *Nanotoxicology: Progress toward nanomedicine* (2nd ed... 514pp). Boca Raton: CRC Press.

Morie, A., Garg, T., Goyal, A. K., & Rath, G. (2016). Nanofibers as novel drug carrier—An overview. *Artificial Cells, Nanomedicine, and Biotechnology, 44*, 135–143.

Naik, K., Srivastava, P., Deshmukh, K., Monsoor, M. S., & Kowshik, M. (2015). Nanomaterial-based approaches for prevention of biofilm-associated infections on medical devices and implants. *Journal of Nanoscience and Nanotechnology, 15*(12), 10108–10119.

Nakazawa, G., Finn, A. V., Ladich, E., Ribichini, F., Coleman, L., Kolodgie, F. D., & Virmani, R. (2008). Drug-eluting stent safety: Findings from preclinical studies. *Expert Review of Cardiovascular Therapy, 6*, 1379–1391.

National Nanotechnology Initiative. (2018a). What is nanotechnology? https://www.nano.gov/nanotech-101/what/definition. Accessed 31 December 2018.

National Nanotechnology Initiative. (2018b). Standards for nanotechnology. https://www.nano.gov/you/standards. Accessed 31 December 2018.

Nature Nanotechnology. (2018). Instructions to authors. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.187.5261&rep=rep1&type=pdf. Accessed 31 2018.

Nikalje, A. P. (2015). Nanotechnology and its applications in medicine. *Medicinal Chemistry, 5*, 081–089.

Oberdörster, G., Oberdörster, E., & Oberdörster, J. (2005). Nanotoxicology: An emerging discipline evolving from studies of ultrafine particles. *Perspectives Environmental Health, 113*(7), 823–839.

Obiweluozor, F. O., Emechebe, G. A., Tiwari, A. P., Kim, J. Y., Park, C. H., & Kim, C. S. (2018). Short duration cancer treatment: inspired by a fast bio-resorbable smart nano-fiber device containing NIR lethal polydopamine nanospheres for effective chemo-photothermal cancer therapy. *International Journal of Nanomedicine, 13*, 6375–6390.

Pamornpathomkul, B., Wongkajornsilp, A., Laiwattanapaisal, W., Rojanarata, T., Opanasopit, P., & Ngawhirunpat, T. (2017). A combined approach of hollow microneedles and nanocarriers for skin immunization with plasmid DNA encoding ovalbumin. *International Journal of Nanomedicine, 12*, 885–898.

Paradis, J. (2012). Claiming nanotechnology: Improving USPTO efforts at classification of emerging nano-enabled pharmaceutical technologies. *Northwestern Journal of Technology and Intellectual Property, 10*(3), 169–208.

Park, J., Kim, J., Kim, S. Y., Cheong, W. H., Jang, J., Park, Y. G., Na, K., Kim, Y. T., Heo, J. H., Lee, C. Y., Lee, J. H., Bien, F., & Park, J. U. (2018, January). Soft, smart contact lenses with integrations of wireless circuits, glucose sensors, and displays. *Science Advances, 1–11*, 24.

Popat, A., Hartono, S. B., Stahr, F., Liu, J., Qiao, S. Z., & Lu, G. Q. M. (2011). Mesoporous silica nanoparticles for bioadsorption, enzyme immobilisation, and delivery carriers. *Nanoscale, 3*, 2801–2818.

Powers, K. W., Carpinone, P. L., & Siebein, K. N. (2012). Characterization of nanomaterials for toxicological studies. *Methods in Molecular Biology, 926*, 13–32.

Riasat, R., Guangjun, N., Riasat, Z., Aslam, I., & Sakeena, M. (2016). Effects of nanoparticles on gastrointestinal disorders and therapy. *Journal of Clinical Toxicology, 6*, 4.

Rivera, G. P., et al. (2010). Correlating physico-chemical with toxicological properties of nanoparticles: The present and the future. *ACS Nano, 4*(10), 5527–5531.

Roy, R., Kumar, S., Tripathi, A., Das, M., & Dwivedi, P. D. (2014). Interactive threats of nanoparticles to the biological system. *Immunology Letters, 158*, 79–87.

Safi, M., Courtois, J., et al. (2011). The effects of aggregation and protein corona on the cellular internalization of iron oxide nanoparticles. *Biomaterials, 32*(35), 9353–9363.

Saleh, Y. E., Gepreel, M. A., & Allam, N. K. (2017). Functional nano-architectures for enhanced drug eluting stents. *Scientific Reports, 7*, 40291.

SCENIHR, Opinion on the Appropriateness of the Risk Assessment Methodology in Accordance with the Technical Guidance Documents for New and Existing Substances for Assessing the Risks of Nanomaterials. 21e22 June 2007. Scientific Committee on Emerging and Newly-Identified Health Risks. 2007.

Sengupta, J., Ghosh, S., Datta, P., Gomes, A., & Gomes, A. (2014). Physiologically important metal nanoparticles and their toxicity. *Journal of Nanoscience and Nanotechnology, 14*, 990–1006.

Takenaka, S., Karg, E., Roth, C., Schulz, H., Ziesenis, A., Heinzmann, U., Schramel, P., & Heyder, J. (2001). Pulmonary and systemic distribution of inhaled ultrafine silver particles in rats. *Environmental Health Perspectives, 109*(4), 547–551.

U.S. Department of Health and Human Services Food and Drug Administration. (2017, December). *Drug products, including biological products, that contain nanomaterials guidance for industry*. 24pp.

U.S. FDA. (2014a) Guidance for industry: Safety of nanomaterials in cosmetic products. 1–16, June 2014.
U.S. FDA. (2014b). Guidance for industry: Assessing the effects of significant manufacturing process changes, including emerging technologies, on the safety and regulatory status of food ingredients and food contact substances, including food ingredients that are color additives, 1–29, June 2014.
U.S. FDA. (2015). Guidance for Industry: Use of Nanomaterials in Food for Animals. 1–11, August 2015.
U.S. FDA. (2016). Use of International Standard ISO 10993-1, "Biological evaluation of medical devices – Part 1: Evaluation and testing within a risk management process." Guidance for Industry and Food and Drug Administration Staff. June 16, 2016, pp. 1–65.
U.S. Food and Drug Administration. (2014) Considering whether an FDA-regulated product involves the application of nanotechnology June 2014. https://www.fda.gov/RegulatoryInformation/Guidances/ucm257698.htm. Accessed 31 December 2018.
Vivero-Escoto, J. L., Slowing, I. I., Trewyn, B. G., & Lin, V. S. Y. (2010). Mesoporous silica nanoparticles for intracellular controlled drug delivery. *Small, 6*, 1952–1967.
Wang, Y. Y., Lai, S. K., et al. (2011). Mucoadhesive nanoparticles may disrupt the protective human mucus barrier by altering its microstructure. *PLoS One, 6*(6), e21547.
Yang, Y. X., Song, Z. M., Cheng, B., Xiang, K., Chen, X. X., et al. (2014). Evaluation of the toxicity of food additive silica nanoparticles on gastrointestinal cells. *Journal of Applied Toxicology, 34*, 424–435.
Yang, Y., Pengkai, Q., Yonghui, D., Maitz, M. F., Zhi, L. Y., Qiu, F. T., Kaiqin, X., Yang, L., & Nan, H. (2015). A biocompatible and functional adhesive amine rich coating based on dopamine polymerization. *Journal of Materials Chemistry B, 3*, 72–81.
Yin, R. X., Yang, D. Z., & Wu, J. Z. (2014). Nanoparticle drug- and gene-eluting stents for the prevention and treatment of coronary restenosis. *Theranostics, 4*(2), 175–200.
Yoo, H. S., Oh, J. E., Lee, K. H., & Park, T. G. (1999). Biodegradable nanoparticles containing doxorubicin-PLGA conjugate for sustained release. *Pharmaceutical Research, 16*, 1114–1118.
Yoshida, Y., Churei, H., Takeuchi, Y., Wada, T., Uo, M., Izumi, Y., & Ueno, T. (2018). Novel antibacterial mouthguard material manufactured using silver-nanoparticle-embedded ethylene-vinyl acetate copolymer masterbatch. *Dental Materials Journal, 37*(3), 437–444.
Zhang, L., Shizhang, Q., Yonggang, J., Huagui, Y., Sandy, B., Frances, S., Zifeng, Y., Xiaolin, W., Zhengping, H., & Max, L. (2008). Fabrication and size-selective bioseparation of magnetic silica nanospheres with highly ordered periodic mesostructured. *Advanced Functional Materials, 18*, 3203–3212.
Zhang, H., Burnum, K. E., et al. (2011). Quantitative proteomics analysis of adsorbed plasma proteins classifies nanoparticles with different surface properties and size. *Proteomics, 11*(23), 4569–4577.
Zhao, L., Chu, P. K., Zhang, Y., & Wu, Z. (2009). Antibacterial coatings on titanium implants. *Journal of Biomedical Materials Research. Part B, Applied Biomaterials, 91*(1), 470–480.
Zhou, Z., Jiang, C., Li-jie, X., Ying, X., Ping, Y., Jingan, L., Jue-jue, W., & Nan, H. (2014). Fabrication of 3D TiO2 micromesh on silicon surface and its effects on platelet adhesion. *Materials Letters, 132*, 149–152.
Zolnik, B. S., González-Fernández, A., Sadrieh, N., & Dobrovolskaia, M. A. (2010). Nanoparticles and the immune system. *Endocrinology, 151*(2), 458–465.

Chapter 11
Integrated Safety Assessment of Medical Devices

Shayne C. Gad

Abstract All of the preceding chapters examine how we select, perform, and evaluate the results of both biologic and chemical tests, as well as (Q)SAR models and literature/database sources. This chapter seeks to examine how we can best utilize all the available data by performing integrated assessments. The basic concepts of the read-across approach are expanded to support the use of data across tests species, routes, and related test data sets.

Keywords (Q)SAR methods · Duration of patient exposure · Hierarchical approval · Liquid · Model · Read-across · Risk assessment · Routes · Siloing · Species · Specifics · Tolerable exposure (TE)

11.1 Introduction

The preceding chapters have been written to provide (with critical review) the armamentarium of tools available to assess the components of biological safety of medical devices ("biocompatibility"). These tools range from in vitro and in vivo testing to computer modeling, chemical analysis, and use of stylized calculation of safe (tolerable) levels of exposure (TEs) for identified chemical entities to which patients may (in the worst case) be exposed. But as should be clear by now, there is no single prescribed approach for putting all these pieces together into an assessment that provides guidance to answer the question – "When used as intended, is this going to be acceptably safe for patients?"

The more we know, the more difficult it is to answer this question. Figure 11.1 provides a flow chart for determining safety based on identification, quantification, and evaluation of the chemical entities potentially released into patients' bodies from a device, but it does not allow for use of ISO 10993 test data in this assessment. This chapter is aimed at designing such an integrated approach.

S. C. Gad, PhD, DABT (✉)
Gad Consulting Services, Raleigh, NC, USA
e-mail: scgad@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_11

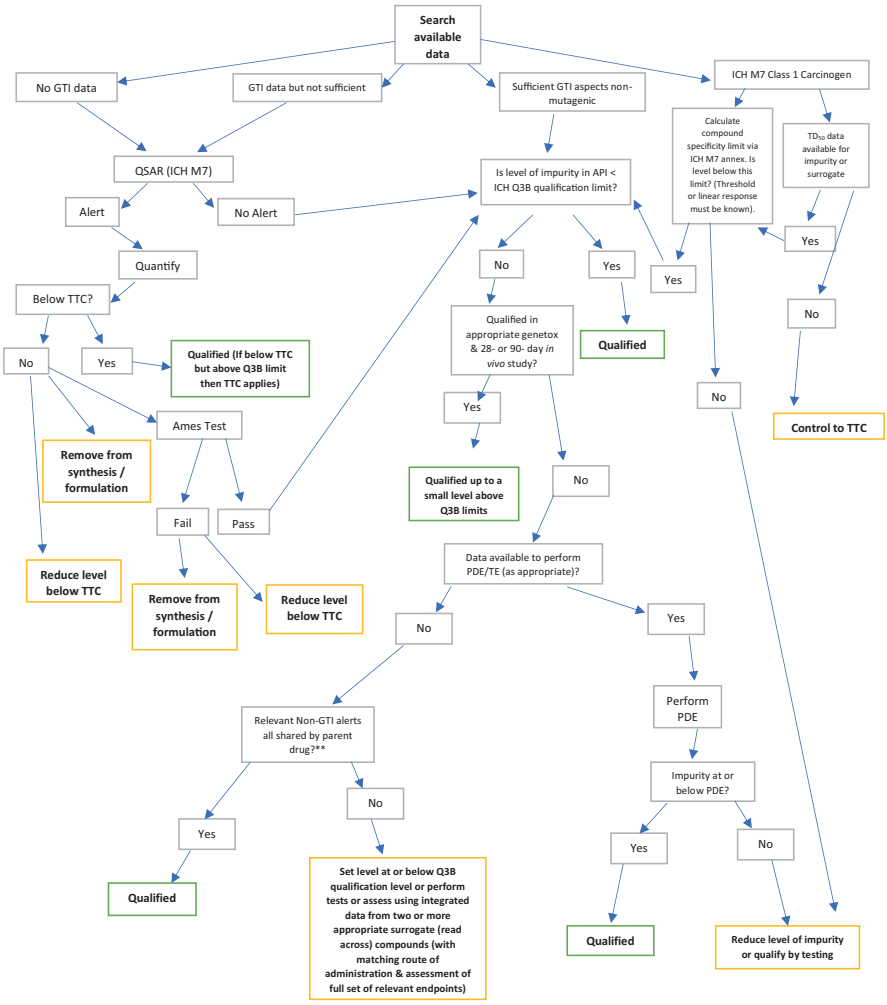


Fig. 11.1 Hierarchical approval to qualification (Q3B)

11.2 Underpinning Assumptions in Risk Assessment

Risk (or safety) assessments all proceed through an ordered set of safety steps, each fairly clearly defined until the end:

1. Identify potential risks.
2. Quantify the potential source(s) of risk.
3. Evaluate the probability of each risk component being relevant, and if possible, determine a threshold below which risk is either not present or at an acceptable level.

4. Integrate all the risk components to establish the limiting (most conservative) risk factor.

The use of the most conservative lower bound of risk is a fundamental assumption of risk assessment, but for therapeutic intervention (such as drugs and medical devices), there always must be a consideration of the potential benefit of the treatment. That is, at what point does the benefit outweigh any one risk and when must cumulative risk be considered. Notice that this requires, for each risk or risk compound, its own profile. How big/important is the risk to the patient across a range of "exposure spectrums"?

This causes us to consider a set of potential complications for evaluating a relative risk:

1. What is the route of patient exposure (source of the exposure)?
   How relevant is the data we are using (both in terms of route of exposure and the model (we have used) to evaluate the risk? How do we translate risk arising from one route (where we have available data) to another (where the patient exposure occurs)? Chap. 6 has sought to address this question.
2. What is the duration of patient exposure? That is, what is the "chronicity" of the risk?
   The most important unstated principle of both toxicity testing and risk assessment is that "good data" is rewarded. Greater precision, relevance, and reproducibility of data result in smaller uncertainty factors (or, as they used to be called, "safety factors"). The determinants of quality (and relevance to our objective) lead to a set of numerical (scalar) factors which further reinforce our results being conservative. That is, these factors support that in all cases we prefer a false positive result to a false negative.

   So, better data should be preferred to otherwise, *except* that science, particularly regulatory science, is conservative. Reliability is a major consideration. Though we are continuously presented with new techniques and methods, we default to sticking with old ones until there is adequate validation and comfort with understanding the new methods (Gad and Schuh 2018).
3. What is the model (animal species) that the data originally from (Thiel et al. 2016)?
   As devices become more complex and (especially as there are an increasing number and variety of resorbable devices either as therapeutic delivery systems or as transitory/organ/tissue support system) it is more difficult to characterize exposure in terms of what (substance), for how long, and at what concentration gradient becomes more problematic (Ronssselle et al. 2019).

A major part of the problem is "siloing" with each of the main disciplines looking at just their portion local tolerance, *innate and adaptive* immune responses, histopathology, and analytical chemistry-based risk assessment, rather than seeking to understand and utilize the integration of the individual parts. We talk of "read across" (RAAF, read-across assessment framework as per the ECHA 2017, but even

at this basic level, we rarely actually try to put the pieces together – even within each of the above-listed "fields." Indeed, for some, RAAF consists of grouping data sets to simplify the task rather than seek to use all the data present to improve the results. In this case, it is rather aimed at gap filling.

For many current regulatory applications of read-across approaches, the intent is to provide a means to "fill in the gaps" testing for the assessment of compounds (Patlewicz et al. 2019). In this chapter, however, the intent is rather to utilize all of the available data to understand the relevance of existing data to patient risks and be able to truly assess such risks.

Furthermore, true integrated assessment of all available data should allow for better understanding and use of compounding factors, in the available preclinical data (Marton et al. 2019).

So the question starts with a series of (1) is there an effect, then, if no, (2) is it adverse, then, if no, (3) what dose/exposure is it associated with, then (4) is it reversible, and (5) is it associated with other effects?

Do we see the effect in other studies? Other species? Is there a NOAEL for each species? Is there a NOAEL that would apply to all species?

Could exposure be modified to make it NOAEL? What human aspects would not be seen? What aspects could/would be seen in animal models, but not in humans?

Are there relevant (Q)SAR models? Are there relevant biomarkers that could help clarify uncertainty?

11.3 Integration Across Routes

The assessment of patient safety (and potential risks) for medical devices arises from a different prospective and history than that of drugs. ISO 10993-1 calls for three initial "endpoints" to be determined for all devices – cytotoxicity, sensitization, and irritation. All three of these recognize that the shared greatest concern for devices arises from interaction with the immediate surface of the device and the surrounding tissue of a patient (and such systemic interaction that follows). These three tests are not specific for mechanism and do not require any identification of what actual chemical entities may be released by the device at its surface and into the patient thereafter. Recognizing this limitation, the test systems are simple bioassays and are designed to be conservative (particularly the case for cytotoxicity) – that is, they "overpredict" potential risks. But these results provide little illumination as to mechanism or no means of evaluating the relevance of finding to potential patient risks. Further required tests (as per guidelines) tend to be both more specific as to potential mechanisms of identified toxicities and more focused on specific

patient exposure. Emphasis on identifying and quantitating specific chemical entities that are (at least) potentially released from a device into the body has grown, with the expectation that potential risks can then be assessed using the available data on such chemical entities.

Several difficulties limit this approach. First, no data may be available on the subject chemical entity. The use of (Q)SAR methods, presented in an earlier chapter in this book, seeks to address this issue. Less obvious in the (Q)SAR approach is that it can also serve to identify structurally similar compounds for which there may be data (such compounds may also be called "surrogate compounds"). Such surrogate compound information should be integrated into a read-across assessment. Second, the information available on identified chemical structures quite frequently is for routes (and durations) of exposure that are markedly different than those that occur with the device. Such differences can lead to both quantitative (difference in levels of systemic exposure) and qualitative differences in biological responses. These are discussed both in an earlier chapter in this volume and in Gad and Spainhour (2017). Using the methodologies discussed in these, the relevance of data assessing potential patient risks in a read across can be significantly increased and potential false conclusion avoided.

11.4 Integration Across Species

Which species are not (or most) relevant for humans? The best data for use in assessing the relevance of reported risk to patient is, of course, human data. Such is also one of the least available data sets. Rather we are much more likely to have data from other species, most commonly rodents (rats and mice) and then other common laboratory animal species. Here, assessment of the relevance of data requires an in-depth understanding of the specifics and particularly of the animal models in question. The uncertainty factors used in deriving PDE and TE values reflect the broad guidance for this (expressed as smaller uncertainty factors for larger non-rodent species – dogs, pigs, and primates – and larger factions for rabbits, rats, and mice). Table 11.1 below presents the HED approach to scaling the potential differences in species sensitivity based on metabolic rates.

11.5 Integration Across What We Know (Table 11.2)

Table 11.1 Conversion of animal doses to human equivalent doses based on body surface area

Species	To convert animal dose in mg/kg to dose in mg/m^2, multiply by k_m	To convert animal dose in mg/kg to HED[a] in mg/kg, either:	
		Divide animal dose by	Multiply animal dose by
Human	37	–	–
Child (20 kg)[b]	25	–	–
Mouse	3	12.3	0.08
Hamster	5	7.4	0.13
Rat	6	6.2	0.16
Ferret	7	5.3	0.19
Guinea pig	8	4.6	0.22
Rabbit	12	3.1	0.32
Dog	20	1.8	0.54
Primates:			
Monkeys[c]	12	3.1	0.32
Marmoset	6	6.2	0.16
Squirrel monkey	7	5.3	0.19
Baboon	20	1.8	0.54
Micro-pig	27	1.4	0.73
Mini-pig	35	1.1	0.95

[a]Assumes 60 kg human. For species not listed or for weights outside the standards ranges, HED can be calculated from the following formula: HED = animal dose in mg/kg × (animal weight in kg/human weight in kg)$^{0.33}$

[b]This k_m value is provided for reference only since healthy children will rarely be volunteers for phase 1 trials

[c]For example, cynomolgus, rhesus, and stump-tail

11.6 Failure Modes

While for the actual biocompatibility test, the most common results which lead to regulatory questions or rejection are (1) not passing the cytotoxicity tests, (2) not having adequately designed (or any) subchronic toxicity studies, and (3) improper preparation of extract solutions.

11.7 Endnote

The CDRH has prepared for some years that “liquid” devices (joint lubricants, dermal fillers, and metabolites) should be treated as drugs and therefore that regulatory oversight should be transferred to CDER. This issue has lain dormant for a number

Table 11.2 Annotated read-across table for medical devices

Specifics of potential patient exposure (route and duration) must be considered as starting place	
Cytotoxicity	Meant as a general nonspecific screen for local tissue effects of extractables from a device. Proponent/residuals in cytotoxicity with serial dilution serve to confirm local effect. In the absence of effects in vivo. Studies' relevance to potential patient risk is unlikely First principles: (i) Data trumps modeling. In vivo data trumps in vitro data. Human data always trumps data from other species (ii) Absence of dose or concentration response correlation should always be considered an indication of incomplete or faulty data (iii) In the case of medical devices, dose response should usually be expressed in terms of severity of effects decreasing with increased distance from the surface of a device and increasing with duration of patient device contact
Irritation	Is the concentration-dependent local tissue response and thus (at least as an acute response) should not be considered in the interpretation of systemic responses
Sensitization	As measured here is a Coombs and Gell Class 4 delayed contact hypersensitivity response: as such it will have both local and systemic aspects, though primarily the former
Acute systemic toxicity	As performed in accordance with ISO-10993 (or USP) guidelines is, while indicative of a prompt and severe systemic response, otherwise different to relate to other systemic toxicity responses
Pyrogenicity	Is a systemic innate acute response which can have severe adverse influences on organ systems throughout the body as is the case with irritation and sensitization; these results are *not* limited by species or route
Genotoxicity	Is a specific endpoint effect without direct relevance to specific systemic or target organ effects
Subacute/subchronic toxicity	Is intended to be broad-scope evaluation of potential adverse effects on the model species. Within any one study, it is critical not just to evaluate individual parameters as indications of an adverse effect but rather to remember that there are constant interactions between the organ/physiologic systems Read-across principles apply here first to individual studies – one must evaluate all of the individual data sets (organ weights, clinical chemistry, hematology, and histopathology) together for both individual organs but also across organ sets (e.g., liver and kidneys) This same concept of evaluating all the results identified in a study applies also to considering, if performed, separate studies of different durations, as a continuum, with longer duration serving to evaluate the same biologic response processes but with different duration of interaction (exposure) If multiple species are utilized to evaluate the prolonged (more than acute) exposure (i.e., subacute and subchronic – and potentially chronic) toxicity of exposure to a device, then one should also consider the indications of systemic toxicity across species A limitation with the interpretation of acute and subchronic device studies is that we rarely know the acute duration or kinetics of exposure of the test animals (or of patients) to the substances eluted from a device. This limitation will appear again when leachable/extractable studies and risk assessments are considered

(continued)

Table 11.2 (continued)

Specifics of potential patient exposure (route and duration) must be considered as starting place	
Implantation	Studies serve several purposes. The traditional ISO study (derived from the original USP guideline studies) assesses the direct immunological interactions between the device and test animal (host) immediate local tissue. The short-term implant studies (Ratner 2015) – that is, 7 days or so – reflect innate immune response; studies of 14–28 days will reflect adaptive immune responses. Longer implantation periods come to reflect the development and resorption of foreign body responses. Such local tissue responses are not influenced by species differences (except the model species must be large enough to provide an adequate tissue bed to physically accommodate the implant. However, which tissue that a device is implanted in does make a difference – the immediate subdermal/subcutaneous tissue is more prone to yield a robust immune response than deeper body tissues such as muscular. Accordingly, understanding the relevance of multiple implantation studies requires attention to both the duration of implantation and the tissue that devices were implanted in (Bradylak 2015) Implantation of complete devices may also serve to achieve a clinically relevant exposure route for evaluating longer-term systemic toxicology evaluations, particularly chronic (26 weeks or longer – 2 years in larger non-rodents) and carcinogenicity studies. Here species differences are important to consider
Hematocompatibility	A complete ISO-compliant evaluation of hematocompatibility requires three separate assessments: (i) Hemolysis, an evaluation of a local tissue tolerance effect – the interaction of a device and/or its eluates with red blood cells. As such it is a relevant finding across species including humans (ii) Thrombogenicity, the formation and distribution of clots through the circulatory system. The actual causation is usually a local tissue effect – at the surface of the device. But the potential adverse effects are generally systemic and applicable to all host species (including humans) (iii) Immunological, may also be the most serious complement activation The complement system is a part of the immune system that helps, or complements, the ability of antibodies and phagocytic cells to clear pathogens from an organism. It is part of the innate immune system but can be recruited and brought into action by the adaptive immune system as shown in Fig. 11.2 (Beutler et al. 2011a, b) The complement system consists of 50+ fluid phase and membrane-bound proteins found in the blood and mainly synthesized by the liver (though other organs/tissues contribute) and is organized into three pathways: • The classical pathway • The alternative pathway • The lectin pathway This is not species-dependent but is more commonly seen in non-rodents (dogs, rabbits, pigs, primates, and humans)

(continued)

Table 11.2 (continued)

Specifics of potential patient exposure (route and duration) must be considered as starting place	
Chemical characterization	Should be considered particularly in conjunction with the results of any repeat-dose exposure studies. Here the intent is to identify and quantify the chemical moieties that are potentially released from a device and potentially subsequently distributed throughout the body. The second step in such assessments is to perform a formal risk assessment of each identified chemical entity (or, in some cases, of groups of structurally related chemical entities). This is intended to be a conservative, for it is based on the worst case of quantitatively all of the potentially toxic molecule that may be eluted and systemically distributed. Assessment of potential toxicity and more sensitive than any of the bioassay-based methods (i.e., everything mentioned to this point). Its limitations are: (i) Many of the identified structures have no available (literature) data. (Q)SAR bridges the gap here for potential genotoxicity but not adherence (ii) There are no associated chronological kinetics or for exposure or information as to what organ systems are exposed (iii) The levels detected are frequently so low as to (unless genotoxic) automatically be discounted as to be a relevant indicator of potential risk. And yet, this information when properly integrated with systemic toxicity and implantation study results to provide much better understanding of what are relevant risks for patient exposure to devices and what aren't
Resorbable devices	Relatively new (and more frequent) medical devices are those that are resorbable – that is, that after implantation the physical device degrades/disappears at a planned rate. The purpose here may be either to be removed from the body/organ region (such as in the case with dermal fillers, joint lubricants, and resorbable cardiovascular stents), to deliver over time a therapeutic entity (drug) to a specific target tissue or organ (or, in some cases, systemically), or to provide longer-term exposure at a lower (less toxic) level There are difficulties with evaluating and establishing biocompatibility adequately for regulatory agencies. A significant issue (especially for longer-term resorbables) is where accurately predicting resorption rates may be problematic and therefore assessing local tissue effects (which requires termination of animal cohorts to be able to examine tissue responses) a long time-to-completion issue (potentially extending years if the device takes that long to be resorbed-FDA currently wants at least 60% resorption of the device prior to study termination), difficult to plan and such as to requiring a lot of animals in the study As resorbable device decreases in size, its mass relative to surface area and therefore the concentration of potential active substances at the device/tissue interface increase. FDA also believes that for long-term resorbables, this can lead to a late phase 1 exposure of more severe toxicity made from a resorbable device

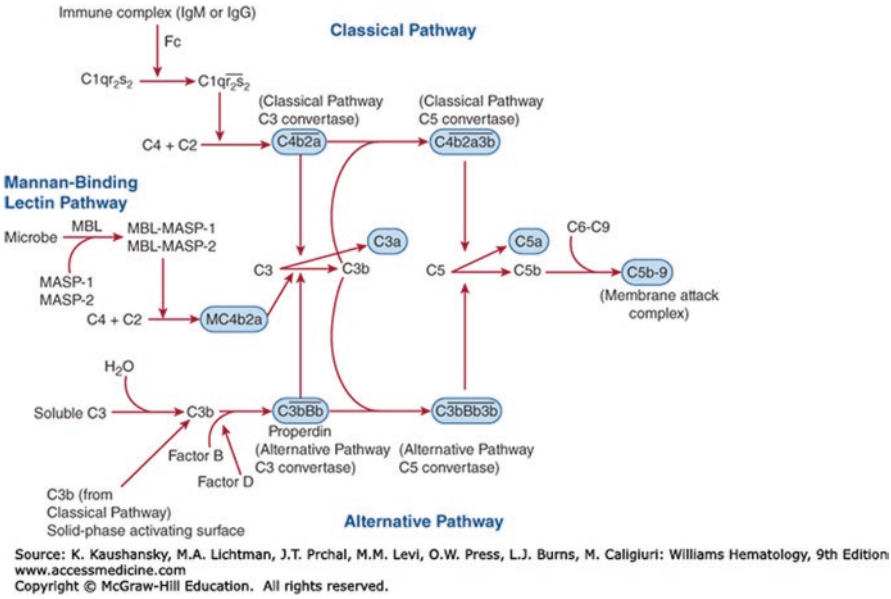


Fig. 11.2 Pathways in compliment activation

of years, but in December of 2018 published a notice preparing consideration of this change for hyaluronic acid intra-articular products (Federal Register 2018). As such devices fit the statutory definition of devices (they do not act on or alter the body's function by chemical or metabolic means), the purpose of the change is clearly not due to misclassification. Rather it would seem that the standards for establishing safety and efficacy are to be made much more strict – and the process of approval to be made much longer. It should be remembered that imaging and contrast agents likewise do not act by chemical or metabolic means but also regulate as drugs.

References

Beutler, E., Kaushansky, K., Kipps, T., Lichtman, M., Prchal, J., & Seligsohn, R. (2011a). The antiphospholipid syndrome. In J. Shanahan & H. Lebowitz (Eds.), *Williams hematology* (Vol. 21, 8th ed., pp. 2145–2162). New York: McGraw-Hill.

Beutler, E., Kaushansky, K., Kipps, T., Lichtman, M., Prchal, J., & Seligsohn, R. (2011b). The inflammatory response. In J. Shanahan & H. Lebowitz (Eds.), *Williams hematology* (Vol. 21, 8th ed., pp. 251–260). New York: McGraw-Hill.

Bradylak, S. F. (2015). *Host response to biomaterials*. San Diego: Elsevier.

ECHA (European Chemistry Agency). (2017). *Read-across assessment aramework.* Helsinki, Finland.

FDA. (2018). *Document No. 2018-27351. Intent to Consider the Appropriate Classification of Hyaluronic Acid Intra-Articular Products Intended for the Treatment of Pain in Osteoarthritis of the Knee Based on Scientific Evidence.*18 December 2018. Silver Spring, MD.

Gad, S. C., & Schuh, J. C. L. (2018). Toxicologic pathology forum opinion paper: Considerations for Toxicologic pathologists evaluating the safety of biomaterials and finished medical devices. *Toxicologic Pathology, 46*(4), 366–371.

Gad, S. C., & Spainhour, C. B. (2017). *Nonclinical drug administration*. Boca Raton: CRC Press.

Marton, L. D., Sanders, M., Reagon, W. J., Crabbs, T. A., McVean, M., & Funk, K. A. (2019). Compounding factors in the interpretation of preclinical studies. *International Journal of Toxicology, 38*, 228–234.

Patlewicz, G., Lizarraga, L. E., Rua, D., Allen, D. G., Daniel, A. B., Fitzpatrick, S. C., Garcia-Reyero, N., Gordon, J., Hakkinen, P., Howards, A. S., Karmaus, A., Matheson, J., Mumtaz, M., Richarz, A., Ruiz, P., Scarano, L., Yamada, T., & Kleinstreuer, N. (2019). Exploring current read-across applications and needs among selected U.S. Federal Agencies. *Regulatory Toxicology and Pharmacology, 106*, 197–209.

Ratner, B. D. (2015). The biocompatibility of implant materials. In S. F. Bodylab (Ed.), *Host responses to biomaterials* (pp. 37–52). San Diego: Academic Press.

Ronssselle, S. D., Ramat, U., Nyska, A., & Jackson, N. D. (2019). Pathology of bioabsorbable implants in preclinical studies. *Toxicologic Pathology, 97*, 358–478.

Thiel, C., Hoffman, U., Dhallah, A., Hengstler, J. G., & Hanepfex, L. (2016). Towards knowledge-driven cross species extrapolation. *Drug Discovery Today: Disease Models, 22*, 21–26.

Chapter 12
Toxicity of Common Extractables and Leachables of Medical Devices

Kandice M. Nappier and Sherrie Smith

Abstract The most commonly used materials in medical devices are polymers, glass, ceramics, and metals. These materials can be used in topical or internal applications, as well as in short- or long-term treatment regimens. Polymeric materials can be found in devices such as syringes, catheters, plastic bottles, molds for dental materials, membranes used in dialysis or filtrations, tubing, pumps, adhesives, balloons, gaskets, and valves. Glass is often used for storage of drug product in bottles, syringes, and vials. Ceramics are primarily used for bone and tooth replacement implants and repairs in the form of cements. Metals can be used as high-pressure containers, syringe needles, components of implantable devices such as screws and springs, and tearable aluminum foil packaging. These materials may contain chemicals that can leach out and potentially cause biological interactions in the body. For this reason, the FDA has created standards for testing these materials to address this potential toxicity. This testing includes assessment of extractables and leachables. This chapter will be a beginner's guide to the types of compounds that can be expected from extractables and leachables testing as well as the most serious toxicities associated with these compounds.

Keywords Polymers · Elastomers · Glass · Ceramic · Metals · Leachables · Extractables

12.1 Background

Medical devices are composed of different types of materials, most commonly polymers, glass, ceramics, or metals. Each of these materials can be used for topical or internal applications in short- or long-term treatment regimens. Medical devices have been sold since the 1700s, often with exaggerated or false claims of their health

K. M. Nappier, MPH (✉) · S. Smith, MS
Gad Consulting Services, Raleigh, NC, USA
e-mail: kandice.cross@gadconsulting.com; sherrie.smith@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7_12

benefits. It wasn't until the 1900s that regulations to prevent fraud and misbranding as well as control contamination began to appear in the United States (Hutt 1989). After many additions and changes to the Federal Food, Drug, and Cosmetic Act throughout the 1900s, the Medical Devices Amendment of 1976 was enacted, which allowed for a specialized regulatory approach to oversee medical devices (Hutt 1989). The FDA has since approved additional standards and has made significant strides toward establishing international standardization of medical devices through the International Conference on Harmonization (ICH) (Gad and Gad-McDonald 2016). These standards include a group of tests that must be conducted to prove the safety of these devices. Extractables and leachables testing are included in this group. Extractables are recognized by the FDA as compounds that can be extracted from the surfaces of manufacturing or storage components of pharmaceutical products in the presence of an extraction solvent or under other laboratory conditions (higher temperature and humidity conditions). Leachables are compounds that migrate from the surfaces of manufacturing or storage components of pharmaceutical products into the drug product under normal conditions. Leachables and extractables can be organic or inorganic (Lewis 2011). Depending on the extraction conditions, the solvents used, and the materials being tested, there could be numerous extractables detected (Feilden 2011). Ideally, the extractables testing is a good estimate of potential leachables. If the assessment of the extractables reveals no concerning chemicals, leachables testing may not be necessary at all (ISO 2018). Or, if leachables testing is conducted, only a small subset of the compounds detected during the extractables testing are detected (Feilden 2011; ISO 2018). Therefore, the rest of this chapter will assume one of these is the case and discuss possible results from extractables testing, assuming leachables would either be a smaller subset or not happen at all.

In the previous chapters, a guide to medical device testing strategies is provided followed by a systematic approach to evaluating the risk of the compounds detected. For medical devices, it is important to not only establish the most conservative limiting factor for the extractable in question; it is also important to consider other factors that can either change cumulative risk or change the importance of that risk. Those factors can include route of administration, duration of treatment, and what animal model data is being used to determine an exposure limit. A challenge in adequately covering all end points of toxicological concern when assessing a compound is identifying relevant data. The previous chapter offers an approach that utilizes all available data to most accurately assess risks to patients.

This chapter will be a beginner's guide to what to expect from extractables testing of the most commonly used materials in medical devices: polymers, metals, glass, and ceramic materials. This chapter will point out some classes of compounds that should be expected and identifies any serious toxicities associated with these classes. This is not intended to be an in-depth guide, but rather a starting point for those that may be new to this type of testing. The information in this chapter will introduce the reader to the types of results that could be expected from extractables testing, which in turn can be used in combination with the information provided in previous chapters to select appropriate materials to use for a device and/or assess the toxicity of the chosen materials.

12.2 L&E Profile by Material

This following chapter will be broken down into the most common materials (polymers, metals, glass, and ceramics). For each material, the common extractable profile will be discussed along with any serious toxicities associated.

12.2.1 Polymers

The term polymer covers a broad range of materials than can take on varying physical properties. Polymers (sometimes called macromolecules) are large molecules that are composed of several smaller molecules (monomers) bonded together. Polymers used in medical devices can be obtained from natural sources, such as plants and animals or synthesized through a manufacturing process. Natural polymers can be in the form of starches, cellulose, pectins, seaweed and vegetable gums, casein, albumin, keratin, and rubber (Gad and Gad-McDonald 2016). Synthetic polymers are created via a process called polymerization. There are several types of polymerization mechanisms. The most common are addition (or chain) polymerization, condensation (or step-growth) polymerization, and ring-opening polymerization (Shrivastava 2018). Examples of synthetic polymers include nylon, polyvinyl chloride, polystyrene, polypropylene, fluorocarbon resins, polyethylene phenolics, and polyesters (Gad and Gad-McDonald 2016). Polymers may also be semisynthetic or regenerated. Polymers of this type are derived from natural materials and chemically modified. Examples include vulcanized rubber, cellulose acetate, and rayon. Polymers can be made from a single repeated monomer, termed a homopolymer, or multiple monomers, termed a copolymer. Copolymers are often designed to mix the properties of various monomers (Shrivastava 2018). When full polymers are physically melted together, they are called polymer blends. This method saves time compared to creating a copolymer, and polymer blends can have properties that are superior to those of the individual polymers (McKeen 2014). Polymers can also generally be classified as either fibers, elastomers, or plastics. Fibers are typically plastics that can be softened by heating and spun into materials similar to natural fibers (Ouellette and Rawn 2015). Elastomers are polymeric materials that have rubberlike qualities and can regain their original shape after being put under stress (e.g., stretching). Plastics are polymers that harden after cooling and can be further classified as thermoplastics or thermoset plastics (Gad and Gad-McDonald 2016; Ouellette and Rawn 2015). Thermoplastics are those that can be heated and remolded into various shapes and forms while warm, over and over again. Thermoset plastics cannot. These plastics are created by chemical reactions forming highly crosslinked networks, rendering rigid materials (Gad and Gad-McDonald 2016; Ouellette and Rawn 2015; Sastri 2013). Elastomers can also be either thermoplastic or thermoset (Gad and Gad-McDonald 2016; McKeen 2014). Materials made from polymers are very often used in medical devices as well as packaging and components

integrated in the manufacturing process (Stults et al. 2015). Polymeric materials can be found in devices such as syringes, catheters, plastic bottles, molds for dental materials, membranes used in dialysis or filtrations, tubing, pumps, adhesives, balloons, gaskets, and valves (Jenke 2009).

Out of the hundreds of possible polymers that could be synthesized for use in medical devices, around 10–20 are the most commonly used (Lee et al. 2007). The most common are polyethylene, polypropylene, polyvinylchloride, and polystyrene (Czuba 2014). Polymers are chosen over other materials due to their availability, low cost, flexibility, light weight, ease of processing, and manufacturability (Lee et al. 2007; Sastri 2013). Table 12.1 below lists some of the common polymers used in medical devices along with some of their applications.

In general, polymers are considered to be of low toxicity, mainly due to their high molecular weight which does not allow for absorption into the body. Most polymers are also not typically irritating or skin sensitizers (Gad and Gad-McDonald 2016). However, in many cases polymers are not used as they are synthesized (or as they are naturally if that is the case). In order to improve the usability, performance, and quality of the polymer, additives are often employed (McKeen 2014; Shrivastava 2018). In general, additives can be classified into three categories: fillers and reinforcements, property modifiers, and processing aids (Shrivastava 2018). Fillers and reinforcements can be used in large amounts, but property modifiers and processing aids are typically used in very small quantities. Materials such as metals, glass fibers or spheres, carbon fiber, carbon black, graphite, mineral powders, and other natural

Table 12.1 Commonly used polymers and their applications

Polymer	Common applications
Polyethylene (PE)	Flexible containers, packaging, pouches, orthopedic implants, tubing, bottles
Polypropylene (PP)	Hypothermic syringes, packaging, sutures, nonwoven fabrics, bottles, containers, connectors, membranes
Polyvinylchloride (PVC)	Packaging, solution and blood bags, blister packs, connectors, tubing, containers
Polystyrene (PS)	Roller bottles, laboratory ware, vacuum canisters, containers, packaging
Polyethersulfone (PES)	Fluid handling couplings/fittings
Polysulfone (PF)	MPU connectors, membranes for dialysis, or ultrafiltration
Polymethylmethacrylate (PMMA)	IV adaptors, blood pumps and reservoir, catheter accessories, dental materials, membranes for dialysis or ultrafiltration
Polyethylene terephthalate (PET)	Containers, bottles, artificial vascular graft, sutures, meshes
Polytetrafluoroethylene (PTFE)	Medical implants
Silicone rubber	Prostheses, artificial organs, contact lenses, catheters
Polycarbonate (PC)	High-pressure syringes, glucose meters, pumps, insulin pens
Polybutylene terephthalate (PBT)	Packaging, dental instruments, high-temperature caps
Polylactic acid (PLA)	Sutures, dental implants, drug-delivery devices
Polyamide (nylon)	Containers, packaging, drug release systems

Sources: Jenke 2009; Lee et al. 2007; McKeen 2014

Table 12.2 Most common additives used in polymers for medical devices

Common additives used in polymeric materials	
Plasticizers	Blowing agents
Lubricants	Antistatic agents
Slip agents	Antimicrobials
Acid scavengers	Colorants
Antiblocking agents	UV stabilizers
Coupling agents	Antioxidants
Release agents	Curing and cross-linking agents
Flame retardants	Catalysts
Heat stabilizers	Impact modifiers

Sources: McKeen 2014; Shrivastava 2018

fibers may be used as fillers or reinforcements to enhance the physical or mechanical properties of the polymer (McKeen 2014; Shrivastava 2018). Most of these materials are generally considered to be of low toxicity. The metals that could present a toxicological concern are aluminum, lead, nickel, silver, and cadmium (Gad and Gad-McDonald 2016; Shrivastava 2018). Except for lead, these metals will be discussed later in this chapter along with the metals. A list of common property modifier and processing aid additives are presented in Table 12.2.

Besides additives there are also residual monomers and contaminants that may be present in the final polymer. All three of these entities may leach into the body and cause significant toxicity (Gad and Gad-McDonald 2016). The mechanism of polymerization can also affect how the polymer degrades. Therefore, when conducting an extractables study, it is important to understand how the polymer was created and what other chemicals may be present, in order to understand the extractable profile. The rest of this section will go over some common compounds detected in extractables testing of polymeric materials and the toxicity associated with them. These common extractable compounds will be grouped by their origin (e.g., antioxidant vs plasticizer).

As mentioned above, polymers are generally considered to be of low toxicity. However, there are monomers and other starting materials that are associated with serious toxicity. These toxic chemicals include vinyl chloride, acrylonitrile, formaldehyde, ethylene oxide, styrene, benzene, and bisphenol A (BPA). Chronic exposure to vinyl chloride in humans is associated with hepatic angiosarcoma (a rare liver cancer), narcotic effects, Reynaud's phenomenon, acroosteolysis, skin changes, and hepatocellular alterations. Vinyl chloride has been classified as a human carcinogen by the Department of Health and Human Services, International Agency for Research on Cancer (IARC), and Environmental Protection Agency (EPA) (ATSDR 2006). Acrylonitrile is considered a probable human carcinogen by EPA and IARC. Increases in tumors of the central nervous system, ear canal, mammary glands, and gastrointestinal tract have been consistently observed in rat studies. Increases in tumor incidence have also been observed in epidemiological studies, although inconsistently. In addition, acrylonitrile is a skin, eye, and respiratory tract

irritant (WHO 2002). The International Council for Harmonization (ICH) has calculated a lifetime acceptable intake of 6 μg/day for acrylonitrile (ICH 2017). Formaldehyde is also considered a probable human carcinogen by EPA and IARC (ATSDR 1999). Increases in nasal tumors have been observed in animal studies. Occupational exposure to formaldehyde has also been associated with increased incidences of lung and nasopharyngeal cancer (ATSDR 1999; EPA 2000). Ethylene oxide is classified as a human carcinogen by EPA and IARC (EPA 2018; WHO 2003). In animal studies ethylene oxide has caused lymphoid cancer and other tumors in the brain, lungs, connective tissues, uterus, and mammary gland (EPA 2018; WHO 2003). In occupational studies, there have been increased incidences of lymphoid cancer and breast cancers in females after exposure to ethylene oxide (EPA 2018). IARC has labeled styrene as a possible human carcinogen. There are several epidemiologic studies where workers showed an increased incidence of lymphatic and hematopoietic cancers; however, the data is inconclusive. Styrene is a known neurotoxin. Other effects that have been observed in animals include damage to the nasal olfactory epithelium, liver necrosis, testicular injury, and developmental effects (ATSDR 2010b). Benzene is a known human carcinogen. Occupational exposure and epidemiological studies have demonstrated an association between benzene exposure and acute myelogenous leukemia (AML). Benzene also causes aplastic anemia (ATSDR 2007b). ICH lists a concentration limit of 2 ppm as a residual solvent (ICH 2018). The biggest concerns for BPA are the effects on the brain, behavior, prostate gland, mammary gland, and changes in the onset of puberty in fetuses, infants, and children. BPA is an endocrine disruptor, a chemical that may interfere with the production or activity of hormones in the human endocrine system. Some studies have demonstrated that prenatal exposure to BPA can increase the risk of mammary cancer (HSDB 2019; NTP 2008).

Antioxidants are used to prevent thermal and thermooxidative degradation during processing. The most common antioxidants are hindered phenols, phosphites, and thioesters (Keck-Antoine et al. 2016; Shrivastava 2018). Some common antioxidants are butylated hydroxytoluene (BHT), as well as chemicals under the trade names include Irganox and Irgafos (Keck-Antoine et al. 2016). These compounds can also be used as UV stabilizers (Hahladakis et al. 2018; McKeen 2014). In animal studies, BHT caused reduced body weight as well as hepatotoxicity (Lanigan and Yamarik 2002). Some Irganox compounds have been observed to cause effects on body weight gain, organ weights, as well as reproductive and developmental toxicity (Johnson Jr. et al. 2018; Neal-Kluever et al. 2015). Reproductive and developmental toxicity has been noted in some studies with Irgafos 168 (OECD 2009). Mercaptobenzothiazole (MBT) has also been used as an antioxidant and stabilizer in polymers as well as a vulcanization accelerator in rubber products. IARC has classified MBT as probably carcinogenic to humans. Occupational exposure to MBT has been associated with an increased incidence of urinary bladder tumors. In animal studies, pituitary, adrenal gland, and hepatocellular tumors have been observed. MBT has also been reported to cause skin sensitization reactions in humans (IARC 2018).

Plasticizers are added to polymers to maintain and enhance the material's flexibility (McKeen 2014; Shrivastava 2018). They may also be added to control viscosity, particulate dispersion in a polymer matrix, lubrication, and mold release (Shrivastava 2018). Historically, the most commonly used plasticizers were orthophthalates including di-2-ethylhexyl phthalate (DEHP), diisodecyl phthalate (DIDP), and dibutyl phthalate (DBP) (SCENIHR 2008). Some other plasticizers include tricresyl phosphate and triphenyl phosphate, both of which are also used as flame retardants (IPCS 1990). Ortho-phthalates commonly demonstrate reproductive and developmental effects as well as liver and/or kidney toxicity in animal studies (CPSC 2010). Neurotoxicity and reproductive toxicity are the major concerns with tricresyl phosphate (IPCS 1990). Neurotoxicity may also be a problem with triphenyl phosphate (IPCS 1991).

UV stabilizers are added to polymers to combat photooxidative degradation that may occur upon exposure to UV radiation. Common UV stabilizers are phenolics, sterically hindered amines, and phosphates (McKeen 2014). Sterically hindered amines are available under the trade names Tinuvin, Chimassorb, and Cyasorb (Hahladakis et al. 2018; Keck-Antoine et al. 2016). Mixed results on skin sensitization have been reported on some of these materials. Other reported toxicities in animal studies include decreases in body weight gain and kidney effects (EC 2017; EPA 2006).

Slip agents and lubricants are used to reduce friction and/or adherence between the polymer molecular chains and between the polymer and metal surfaces or fillers. The most common lubricants are fatty acids, especially stearic acid compounds (Hahladakis et al. 2018; Keck-Antoine et al. 2016; Shrivastava 2018). Stearic acid compounds are also used as acid scavengers (Teasdale et al. 2015). Fatty acids are normal components of the body and food and are commonly used in foods and pharmaceuticals. At high doses, the concerns with these compounds are similar to that of excess fat in the diet. Parenteral administration may also be a source of additional concern (CIR 1987). The most common slip agents are erucamide and oleamide, both of which are primary fatty acid amides (PFAMs) (Keck-Antoine et al. 2016). PFAMs are an endogenous class of molecule that are important in various signaling pathways, including sleep, locomotion, and angiogenesis (Farrell et al. 2012). Both erucamide and oleamide are considered to be of low toxicity (Health Canada 2018). Paraffin waxes may also be used as slip agents (Hahladakis et al. 2018; Shrivastava 2018). These materials are also generally considered to be of low toxicity (ECHA Registration Dossier for Paraffin Waxes and hydrocarbon waxes 2019c).

Heat stabilizers are used to improve the thermal stability of the polymer. High temperatures may be encountered during processing, storage, or product use. Common heat stabilizers are metal salt blends, organometallic compounds, and epoxy stabilizers, such as epoxidized soybean oil (ESBO). ESBO may also be used as a plasticizer and acid scavenger (Teasdale et al. 2015). Common salt blends used as heat stabilizers in medical device polymers include those of barium, zinc, and calcium (Hahladakis et al. 2018; McKeen 2014). Zinc and calcium are essential nutrients, while barium is not. The most significant toxicity from barium is kidney toxicity (ICH 2018). Cadmium, tin, and lead organic compounds are the most common

organometallic heat stabilizer compounds (Hahladakis et al.; Keck-Antoine et al. 2016). Metal compounds and zinc and calcium stearates may also be used as acid scavengers (Shrivastava 2018). Human exposure to dibutyl and tributyl tin compounds has most commonly resulted in loss of memory, insomnia, irritation, and skin lesions. Neurotoxicity and hepatotoxicity have been reported following exposure to other organotin compounds (Magos 1986). Lead exposure can lead to adverse neurological, reproductive, developmental, immune, cardiovascular, and renal effects. These effects are more severe in utero and in children (ATSDR 2007c). ESBO toxicity can vary by the product. Some have been found to affect the liver, kidney, testis, and uterus in rat studies (Fankhauser-Noti et al. 2006). Others have caused a decrease in body weight (Larson et al. 1960). Cadmium is discussed below, along with some other metals that may present a toxicological threat.

Coupling agents are used to improve the adhesion of polymers with fillers and/or metals (McKeen 2014; Shrivastava 2018). The most common coupling agents are silanes (Shrivastava 2018). Silanes used as coupling agents have four substituents attached to a silicon atom. The most common agents have three alkoxy groups and one other functional group (e.g., amine, epoxy, methacrylate, alkane, halogen, vinyl), although there are some with fewer alkoxy groups (Pape 2011). The toxicity of these compounds will vary depending on the associated functional group(s). However, there is some information on the toxicity of siloxanes, which may also be extracted. Siloxanes are compounds with a backbone consisting of silicon atoms linked via oxygen. Each silicon atom also has other functional groups bonded to it. Several siloxane materials have been observed to cause liver toxicity in animal tests. Some have also caused reproductive effects (DMEEPA 2014).

Curing or cross-linking agents are added to polymers to improve the cohesion, adhesion, and durability (Shrivastava 2018). Common curing/cross-linking agents include 4,4′-methylenedianiline (MDA) and 2,2′-dichloro-4,4′-methylenedianiline (MOCA or MBOCA) (Hahladakis et al. 2018). MDA is classified as possible human carcinogen by IARC. In rat studies, increases in liver and thyroid tumors have been observed. Humans exposed to contaminated food containing this chemical also experienced liver damage. MDA is also irritating to the skin and eyes (ATSDR 1998). MBOCA is classified as a human carcinogen by IARC. In animals, liver, blood, and mammary gland tumors have been observed. In occupational studies renal toxicity and carcinogenicity have also been observed (ATSDR 2017).

Colorants are added to give color to a polymer (McKeen 2014). However, these additives may also exhibit other functions, such as UV stabilization (Keck-Antoine et al. 2016; McKeen 2014). Metals, metal oxides, and azocolorants are commonly used as colorants in polymers (Hahladakis et al. 2018; Shrivastava 2018). Some toxic metals are discussed in Sect. 12.2.4 below.

Antimicrobials are added to polymers to prevent mold, mildew, algae, and yeast from damaging untreated polymers (Shrivastava 2018). Silver salts, ions, and complexes are commonly used as antimicrobials in polymers (Gad and Gad-McDonald 2016; Polivkova et al. 2017; Shrivastava 2018). The use of silver nanoparticles is also picking up (Polivkova et al. 2017). The toxicity of silver will be discussed later in this chapter in the metals section.

Along with all of the chemicals listed above, their impurities and degradation products may also be observed as extractables. These products may be of similar toxicity as the parent compound or could display increased or decreased toxicity. Branched and linear hydrocarbons are common extractables that may originate from the polymer itself (Stults et al. 2015). Other examples include 2-ethylhexanol and benzoic acid monoethylhexanoate ester, both of which come from DEHP (as either impurities or degradation products) (Teasdale et al. 2015). Epoxidized fatty acids, resulting from ESBO and fatty acids used as additives are also common (Stults et al. 2015; Teasdale et al. 2015). In addition, there may be compounds that are detected due to the extraction process. For example, if isopropanol is used as a solvent, it may form isopropyl ester compounds with some of the entities during extraction. Different extraction processes may also affect the nature of the results, as some processes may cause more degradation than others (Teasdale et al. 2015). Therefore, it is important to keep these things in mind when it comes time to assess the risks of extractables test results.

12.2.2 Glass

Glass is manufactured from silica, the main component of sand, through a heating and cooling process that results in a network of silicon atoms surrounded by four oxygen atoms that are also covalently bonded to neighboring silicon atoms. This matrix of tetrahedrally surrounded silicon atoms is glassy in appearance and can undergo various processes to possess a wide range of properties (Shand 1958). Glass can be heated, cooled, shaped, and manipulated to be used for many applications. For pharmaceutical applications, several types of glass are designed based on the chemical durability required to meet specifications (Ball et al. 2012). Typical uses for glass are for bottles, tanks, vials, and syringes for storage of drug products (Jenke 2009).

Glass bottles and containers used for storage (such as prefilled syringes) of pharmaceuticals presents the greatest risk of exposure as the components of glass can leach out and migrate into the drug product. Extractable compounds can also come into contact with drug product through ion exchange, precipitation, glass dissolution, surface layer exfoliation, and corrosion (Borchert et al. 1989). The extracted elements from glass are typically aluminum, boron, barium, calcium, cadmium, cobalt, iron, potassium, silicon, sodium, and magnesium (Jenke 2009; Sumitra et al. 2016). Elemental silicon exposure has been known to cause irritation of the eyes, skin, and upper respiratory system (Pohanish 2012). A study in patients with idiopathic pulmonary fibrosis revealed an association between the occurrence of the disease and elevated levels of silicon and aluminum in pulmonary hilar lymph nodes (Kitamura et al. 2007).

Through the oral route of administration, barium can exhibit toxicities in the liver, while inhalation of barium oxide can elicit adverse respiratory effects in the form of bronchitis, cough, shortness of breath, and phlegm (CICAD 2001; NTP

1994). Boron can cause some irritation of the nose and throat, dryness of the mouth, nose, and throat, and potentially reduced sperm count and sperm motility upon inhalation in animals and humans. Irritation of the eyes after acute exposure to sodium borate dust in an occupational setting has been reported (ATSDR 2010a). Toxicities of boron species such as boric acid and borax can cause hepatic, renal, dermal, gastrointestinal and neurological effects in humans after oral exposure (ATSDR 2010a; Wong et al. 1964). Reproductive and developmental effects after oral exposure were reported in rats, mice, and dogs; however, no effects were reported in several epidemiological studies within subpopulations exposed to higher levels of boron (ATSDR 2010a; Chang et al. 2006; Liu et al. 2006a, b; Sayli 2003).

Arsenic is another compound used in the processing of glass materials. Inorganic arsenic trioxide and arsenic acid are used to remove color and clarify components used to make glass (Carapella 1992). Epidemiological studies conducted on subpopulations exposed to inorganic arsenic revealed cardiological (increased incidence of vasospasticity), dermal (excessive pigmentation, keratinization, and wart formation), neurological (decreased nerve conduction velocity), and respiratory (lung cancer) toxicities (ATSDR 2007a; Lagerkvist et al. 1986; Enterline and Marsh 1982).

Calcium, potassium, sodium, and magnesium are all essential elements needed for growth, development, and maintenance of various tissues in living organisms. There is little concern of toxicity of the body upon exposure to these elements if leached from glass into a drug product. Calcium rapidly reacts in the presence of water and air (forming calcium oxide and calcium dihydroxide) and can cause severe burns of the skin, throat, gastrointestinal tract, and eyes (TOXNET/HSDB 2019). Potassium and sodium rapidly reacts, spontaneously combusts and decomposes in the presence of water, and is highly corrosive to the skin (ECHA Registration Dossier for Potassium 2019a; ICSC: 0717 2006). Magnesium is not expected to be irritating to the eyes or a sensitizer of the skin (ECHA, Registration Dossier for Magnesium 2019b).

Aluminum, cadmium, cobalt, and iron can also be extracted from other medical device materials and will be discussed in Sect. 12.2.4 below.

12.2.3 Ceramics

Ceramic materials are made through milling of nonmetallic and nonorganic raw materials such as silicon, silicon oxides, aluminum oxides, and silicates to a very fine powder which are then strategically packed together and heated in a process call sintering to form structured devices and device components. The density, strength, and ductility of ceramics are dependent on the particle size of the milled materials, the pore size formed between the packed particles, and the heating and cooling regimen applied in the processing (Turner 2009). The various physical structures that ceramics can take on include polycrystalline, glass, a combination of glass and multicrystals and single crystals. Ceramics are selected as components of medical devices

for their hard, heat-resistant, and corrosion- and chemical-resistant properties (Gad and Gad-McDonald 2016). The largest use of ceramics is in bone and tooth replacement implants and repairs composed of hydroxyapatite synthetic $Ca_{10}(PO_4)_6OH_2$ but may also be used in hip joints composed of alumina (Griss and Heimke 1981) for the purpose of storage in vials, or drug delivery in tubing (Sastri 2013).

Alumina (Al_2O_3) and zirconia (ZrO_2) (the two most widely used materials in solid ceramic implants) are components of the ball and socket portion hip replacement implants as well as components of dental implants. Hydroxylapatite and bioresorbable calcium phosphate materials are used as metallic implant coatings and as fillers for defects in bone and alveolar ridges (Helmus et al. 2008). Compared to metals and polymers, ceramic materials have consistently demonstrated high wear resistance over time under as well as excellent biocompatibility (Turner 2009). However, in vivo studies in rabbits have shown that calcium phosphate (Ca-P) biomaterials can release degradation products postimplantation. This "breakdown" of material can occur through dissolution and removal of soluble in interstitial fluids or disintegration into fine particles or fragments that can migrate to other tissues (Lu et al. 2002).

Cytotoxicity studies have demonstrated some correlation between the particle size and concentration of bioresorbable calcium phosphate, alumina, and zirconia and cell mortality. Particles larger than 2 μm in diameter, and higher concentrations of particles tended to cause more macrophage mortality (Lu et al. 2002). The degradation products of akermanite (Ca2MgSi2O7), a silicate-based bioceramic material used in bone regeneration, have been evaluated in rats and shown to not cause any adverse systemic affects (Ma et al. 2019). The local toxicities associated with ceramic materials such as opaque porcelain, mineral-based porcelain, lithium-containing ceramic, silicon oxide ceramic, and aluminum oxide have been found to be minimal in many in vitro biocompatibility tests. Limited testing to understand the mutagenic and carcinogenic potential of ceramics has been studied as few *Salmonella typhimurium* strains have been exposed to ceramic materials and observed. Crystalline silica and alumina (major components of ceramics), however, have induced some DNA damage in mammalian cells. Systemic toxicity of ceramic materials used in therapeutic applications has been deemed extremely low (Elshahawy 2011).

The primary concern of ceramic materials in medical devices is not necessarily extractable compounds. The focus is more so on dissolution, disintegration, and migration of particles and ions into the tissues which are identified and quantitated under ISO 10993-14 guidance (ISO 2001).

12.2.4 Metals

Metals are processed by various methods, mixed with different elements to make alloys that exhibit varying properties. Many metal materials used in pharmaceutical applications are coated with organic or inorganic substances to prevent corrosion (Ball et al. 2012). The most widely used metals are tinplate, aluminum, and stainless

steel (Ball et al. 2012). Metals are often used in the manufacture (vessels used for mixing of raw materials and drug products that can withstand extreme temperatures and pressures), storage (gas tanks, closures of vials, foil packages), and administration (syringe needles) of pharmaceuticals as well as components of medical devices (screws, springs, metal implants, and surgical tools) (Ball et al. 2012; Jenke 2009).

Orthopedic implants such as fracture plates, rods, wires, stents, catheters, and joint replacement prostheses are made of either stainless steel, cobalt-chromium alloys, titanium, nitinol, or titanium alloys (Hansen 2008). Implantable devices present the greatest risk of exposure to degradation and corrosion products to patients and must be evaluated under ISO 10993 guidance (Brown et al. 2015; Hansen 2008). Compounds do not leach from metals over time as observed in materials made of polymers, and they are often used in medical devices because of their inertness, strength, biocompatibility, longevity, and malleability (Khan et al. 2014). However, through corrosion, metal ions and fragments can be released from the devices. The Table 12.3 below briefly summarizes the most common mechanisms of corrosion that can occur in metal implants and lead to the release of metal ions and debris.

Compromised metallic implanted devices can release metal ions and particles that may cause local and systemic adverse effects. The Table 12.4 below lists the three main metal materials used in medical device along with their respective components. In the rare instance that corrosion occurs, these elements can be released from the device and migrate throughout the body. The potential and known toxicities of the components listed in the table will be discussed below.

The major component of stainless steel is iron. Iron can cause irritation to the respiratory and gastrointestinal tract upon inhalation which would occur in an occupational setting (TOXNET/HSDB 2019). Iron is an essential element that is needed for hemoglobin synthesis. Long-term studies of population in which iron is consumed in excess have not revealed significant adverse effects (JECFA 1986).

Table 12.3 Mechanisms of corrosion

Corrosion mechanism	Cause/definition
Intergranular	Nonuniform composition of alloy castings → grain boundary precipitation → depletion of alloy element → susceptibility to corrosion and crack perpetuation
Pitting	Breakdown of the protective, passivating oxide film on the surface → forms cavities
Fretting	Friction between two materials as a result of motion or vibration → abrasion of fine particle fragments → oxidation of fragments → destruction of metal surface
Crevice	Depletion of oxygen in interface between two surfaces → shift in pH → formation of chloride ion species → corrosion
Galvanic	Dissimilar metals and alloys (typically at the joint) in electrical contact in a corrosive electrolyte → less noble metal corrodes at a faster than normal rate
Stress	Presence of tensile force or pressure → cracking of material

Sources: Hansen 2008; Manivasagam et al. 2010

Table 12.4 Main metals used in medical devices, their uses, and composition

Material	Medical devices (types and uses)	Composition
Titanium alloys (nitinol, Ti-6Al-4 V, Ti-5AL-2.5 Fe, Ti-6Al-7Nb)	Cardiovascular and gastrointestinal stents, dental implants and wires, stents, ablation catheters	Nickel, titanium, aluminum, vanadium, niobium
Cobalt-chromium-molybdenum alloys	Dental castings, joint replacement parts, spinal rods, fracture plates	Chromium, cobalt, molybdenum, nickel
Stainless steel	Surgical tools, wire sutures, needles, screws, joint replacement parts, cranial and fracture plates, stents, catheters	Iron, chromium, nickel, molybdenum

Sources: Hansen 2008; Khan et al. 2014

The most common metals released from implantable devices are aluminum (Al), nickel (Ni), cobalt (Co), and chromium (Cr) (Brown et al. 2015). Fine particles (or nanoparticles) released from defective or damaged devices may be engulfed by macrophages and phagocytized or carried to other tissues through the lymphatic system to reside in the lymph nodes, liver, bone marrow, or spleen (Sansone et al. 2013). Recipients of metal orthopedic implants (especially metal-on-metal implants) have demonstrated elevated lymphocyte reactivity to serum cobalt and nickel. The potential local toxicities associated with released ions as a result of corrosion or physical damage of metal implants include inflammation, deterioration of bone tissue, and pseudotumors, a complication typically seen in recipients of total hip arthroplasty (Brown et al. 2015).

The potential systemic effects of metal ions have been demonstrated in in vitro and in vivo studies. Cobalt has been found to cause adverse neurological, cardiological, hematological, and endocrine effects, leading to the development of tinnitus, vertigo, deafness, blindness (Hallab et al. 2001; Rizzetti et al. 2009), convulsions, heart disease, elevated hemoglobin in the blood, anemia, and hypothyroidism. Exposure to chromium can lead to toxicities of the kidney, liver, and reproductive organs of both male and females. Titanium dioxide caused adverse effects in the lungs of rats after intra-articular injection. Patients that experience osteolysis at the implantation site tend to develop skin hypersensitivity to metals, shortening the lifespan to the implant. (Manivasagam et al. 2010; Sansone et al. 2013). Aluminum exposure and accumulation has demonstrated toxicities within the neurological system causing memory loss and proximal muscle weakness, as well as inhibition of bone remodeling, dialysis osteodystrophy, and osteomalacia (Jeffery, E. H. et al. 1996; Sansone et al. 2013). Aluminum has also been linked to neurological disorders such as Parkinson's disease, Alzheimer's disease, and amyotrophic lateral sclerosis (ALS) as well as other adverse effects such as dialysis encephalopathy and involuntary movement (Manivasagam et al. 2010; Sansone et al. 2013). Molybdenum (Mo) is considered to display low toxicity, high resistance to corrosion, and good biocompatibility and is non-genotoxic (RIVM 2001), hence why it is often used in hip resurfacing surgery (Ribeiro et al. 2016). Niobium oxide (Nb) exhibited mild cell degeneration in the livers of Swiss mice after a single intraperitoneal dose.

Niobium is a nontoxic element often used as a component of titanium alloys (Dsouki et al. 2014).

Chromium, cobalt, nickel, vanadium, aluminum, and titanium ions have all displayed genotoxic and cytotoxic effects in tissue culture cells. Oxidative stress and chromosomal damage can be caused by the oxygen species of these ions which are reactive in the nucleus, proteins, and lipids of cells leading to apoptosis and necrosis (Sansone et al. 2013). Intramuscularly implanted nickel alloys have prompted sarcoma growth in rats while other metal alloys such as cobalt-molybdenum and a titanium alloy did not promote an increase in tumor growth after 24 months (Brown et al. 2015; Lewis et al. 1995). However, IARC has classified implanted films consisting of cobalt, nickel, and nickel/chromium/iron alloy as Group 2B potentially carcinogenic to humans. Vanadium has also been classified as a possible human carcinogen (IARC 2012). However, it has not been concluded if these materials are carcinogenic when used in implanted metallic devices (Brown et al. 2015; IARC 1999). IARC has also determined that some nickel compounds and some soluble cobalt compounds may be carcinogenic to humans. After prolonged dermal contact with nickel, many humans become sensitized to it (ICH 2019). Dermal contact with cobalt or chromium may also lead to skin sensitization (ATSDR 2012b; ICH 2019). Nickel ingestion may cause stomach pain, depression of body weight, and effects on the blood and kidneys, while long-term inhalation may result in inflammation in the lungs and nasal cavity (ICH 2019).

Silver (Ag) is commonly used as an antibacterial in implantable devices, wound dressings, catheters, bone cements, dental devices, and cardiovascular devices and as a coating on polymeric devices. It has been used in compounds such as silver nitrate, but silver nanoparticles are picking up in use (Lansdown 2010; Maillard and Hartemann 2013). Silver was found to be highly toxic when implanted in rabbit brains causing necrosis and damage of cells (Stensaas and Stensaas 1978). It is not mutagenic and is not expected to be carcinogenic. A well-known toxicity of silver exposure through the oral, respiratory, and intravenous route is argyria and argyrosis, a blue/gray discoloration of the dermis and eye, respectively (Lansdown 2010; ICH 2019). Exposure to silver nanoparticles can lead to oxidative stress which causes red blood cell damage and hemolysis (Chen et al. 2015).

Other metals that may be detected in extractables include arsenic and cadmium. Arsenic as mentioned above may cause lung cancer or other effects such as cardiological, dermal, and neurological toxicities (ATSDR 2007a; Lagerkvist et al. 1986; Enterline and Marsh 1982). Cadmium has been found to be genotoxic and a carcinogen in humans (based on epidemiologic studies) but not mutagenic. It has exhibited renal and skeletal toxicities upon inhalation and oral ingestion (ATSDR 2012a; IARC 2012).

To mitigate the risk of exposure to released ions and fragments, the surfaces of metal and metal components are coated with other materials that (1) are more biocompatible with the tissues surrounding the implants and (2) protect the implants from corrosion and release of ions (Hansen 2008).

Several metals are common extractables in multiple medical device materials. These metals are shown in the Table 12.5 below.

Table 12.5 Leachables and extractables from multiple materials

Leachable/extractable	Medical device material
Silver	Polymers, metals
Aluminum	Polymers, glass, metals
Cobalt	Glass, metals
Nickel	Polymers, metals
Iron	Glass, metals
Cadmium	Polymers, glass, metals

12.3 Summary

Extractables testing is an important step in the development of medical devices. Depending on the type of device and its construction, different types of results should be expected. The most common materials used in medical devices are polymers, metals, glass, and ceramics. This chapter provided a brief introduction to the types of results one could expect from the testing of these materials. Polymers are not typically considered to be toxic; however, finished polymers have several additives that could potentially cause toxicity. Residual monomers, contaminants, and degradants may also present a potential source of toxicity. These types of compounds are what are typically detected in extractables testing. Unlike polymers, compounds do not usually leach from metals over time; however, corrosion can cause metal ions and fragments to be released from the devices. As for glass materials, the most common extractables are elemental ions. Ceramics do not typically raise a concern for extractable compounds. Dissolution, disintegration, and migration of particles and ions into the tissues are typically the focus for these materials.

This chapter also provided a brief introduction to the toxicities that may be associated with extracted compounds. As for polymers, there are some very toxic chemicals that could potentially be observed upon extractables testing. These include vinyl chloride, acrylonitrile, and formaldehyde all of which present risks for carcinogenicity. Potential toxic compounds from glass and metal devices include arsenic, cobalt, and cadmium, which may also present risks for carcinogenicity. Some of the most concerning compounds that could be detected along with the concerning toxicities are listed in the Table 12.6 below. Especially in the case of polymers, it is important to consider how the testing itself may affect the results when designing the tests (such as picking solvents that won't cause too much degradation of the material).

This chapter is not intended to be an all-inclusive guide to the potential profile or toxicity associated with extractables testing of medical devices. Instead, the reader should use this knowledge as a starting point for what may come out of testing. Upon determining the results of testing, this information along, with the information provided in previous chapters should be used to conduct full assessments to determine the potential risks associated with a new medical device.

Table 12.6 Summary of the most serious toxicities discussed

Material	Chemical or class	Associated toxicities
Polymers	Vinyl chloride	Carcinogenicity
	Acrylonitrile	Carcinogenicity
	Formaldehyde	Carcinogenicity
	Ethylene oxide	Carcinogenicity
	Styrene	Carcinogenicity
	Benzene	Carcinogenicity
	Phthalates	Reproductive/developmental toxicity
	Mercaptobenzothiazole	Carcinogenicity
	Organotins	Neurotoxicity and hepatotoxicity
	4,4′-Methylenedianiline	Carcinogenicity
	2,2′-Dichloro-4,4′-methylenedianiline	Carcinogenicity
	Lead	Developmental toxicity
Glass	Arsenic	Carcinogenicity
	Barium	Hepatotoxicity
	Boron species	Hepatotoxicity, renal and gastrointestinal toxicities
Ceramic	Alumina (Al_2O_3)	Cell death in surrounding tissues
	Zirconia (ZrO_2)	Cell death in surrounding tissues
Metals	Chromium	Renal, hepatoxicity, and reproductive/developmental toxicity
	Vanadium	Carcinogenicity
Polymers, glass, and metals	Aluminum	Neurotoxicity
	Cadmium	Carcinogenicity
Polymers and metals	Silver	Hematological toxicity
	Nickel	Carcinogenicity
Glass and metals	Cobalt	Hematological and endocrine toxicity, severe neurotoxicity

References

ATSDR. (1998). *Toxicological profile for Methylenedianiline*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.

ATSDR. (1999). *Toxicological profile for formaldehyde*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.

ATSDR. (2006). *Toxicological profile for vinyl chloride*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.

ATSDR. (2007a). *Toxicological profile for arsenic*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.

ATSDR. (2007b). *Toxicological profile for benzene*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.

ATSDR. (2007c). *Toxicological profile for lead*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.
ATSDR. (2010a). *Toxicological profile for boron*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.
ATSDR. (2010b). *Toxicological profile for styrene*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.
ATSDR. (2012a). *Toxicological profile for cadmium*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.
ATSDR. (2012b). *Toxicological profile for chromium*. Agency for Toxic Substances and Disease Registry, Public Health Service, U.S. Department of Health and Human Services, Atlanta, GA.
ATSDR. (2017). *Toxicological profile for 4,4′-Methylenebis(2-chloroaniline) (MBOCA)*. Agency for Toxic Substances and Disease Registry, public health service, U.S. Department of Health and Human Services, Atlanta, GA.
Ball, D. J., Norwood, D. L., Nagao, L. M., & Stults, C. L. (Eds.). (2012). *Leachables and extractables handbook: Safety evaluation, qualification, and best practices applied to inhalation drug products*. Hoboken: Wiley.
Borchert, S. J., Ryan, M. M., Davison, R. L., & Speed, W. (1989). Accelerated extractable studies of borosilicate glass containers. *PDA Journal of Pharmaceutical Science and Technology, 43*(2), 67–79.
Brown, R. P., Fowler, B. A., Fustinoni, S., & Nordberg, M. (2015). Toxicity of metals released from implanted medical devices. In *Handbook on the toxicology of metals* (pp. 113–122). Academic Press, Oxford, UK.
Carapella, S. C. (1992). Arsenic and arsenic alloys. In: Kroschwitz JI, & Howe-Grant M, (Eds.). *Kirk-Othmer encyclopedia of chemical technology*. New York, NY: John Wiley and Sons, vol 3:624–633.
Chang, B. L., Robbins, W. A., Wei, F., et al. (2006). Boron exposure in China: Exploring work and lifestyle factors to boron exposure. *AAOHN Journal, 54*(10), 435–443.
Chen, L. Q., Fang, L., Ling, J., Ding, C. Z., Kang, B., & Huang, C. Z. (2015). Nanotoxicity of silver nanoparticles to red blood cells: Size dependent adsorption, uptake, and hemolytic activity. *Chemical Research in Toxicology, 28*(3), 501–509.
CICAD. (2001). Barium and barium compounds. Concise International Chemical Assessment Document 33. World Health Organization, Geneva.
CIR (Cosmetic Ingredient REview). (1987). Final report on the safety assessment of oleic acid, Lauric acid, palmitic acid, Myristic acid, and stearic acid. *Journal of the American College of Toxicology, 6*(3), 321–401.
CPSC (Consumer Product Safety Commission). (2010). Overview of phthalate toxicity. Available Online at: https://www.cpsc.gov/s3fs-public/phthalover.pdf
Czuba, L. (2014). Application of plastics in medical devices and equipment. In K. Modjarrad & S. Ebnesajjad (Eds.), *Handbook of polymer applications in medicine and medical devices* (pp. 9–19). Elsevier, Chadds Ford, PA, USA.
DMEEPA (Danish Ministry of the Environment Environmental Protection Agency). (2014). Siloxanes (D3, D4, D5, D6, HMDS). Evaluation of Health Hazards and Proposal of a Health-Based Quality Criterion for Ambient Air. Accessed July 26, 2019. https://www2.mst.dk/Udgiv/publications/2014/01/978-87-93026-85-8.pdf
Dsouki, N. A., de Lima, M. P., Corazzini, R., Gáscon, T. M., Azzalis, L. A., Junqueira, V. B. C., & Fonseca, F. L. A. (2014). Cytotoxic, hematologic and histologic effects of niobium pentoxide in Swiss mice. *Journal of Materials Science: Materials in Medicine, 25*(5), 1301–1305.
EC (European Commission). (2017). Substance Evaluation Conclusion as Required by REACH Article 48 and Evaluation Report for 2-(2H-Benzotriazol-2-yl)-p-Cresol.
ECHA (2019a). Registration dossier for ECHA registration dossier for potassium. https://echa.europa.eu/registration-dossier/-/registered-dossier/26030
ECHA (2019b). Registration dossier for ECHA Registration Dossier for Magnesium. https://echa.europa.eu/registration-dossier/-/registered-dossier/15545
ECHA (2019c) Registration dossier for ECHA registration dossier for paraffin waxes. https://echa.europa.eu/registration-dossier/-/registered-dossier/15503

Elshahawy, W. (2011). Biocompatibility. In C. Sikalidis (Ed.), *Advances in ceramics: Electric and magnetic ceramics, bioceramics, ceramics and environment*. BoD-Books on Demand, Janeza Trdine 9, 51000 Rijeka, Croatia.

Enterline, P. E., & Marsh, G. M. (1982). Cancer among workers exposed to arsenic and other substances in a copper smelter. *American Journal Epidemiology, 116*(6), 895–911.

EPA. (2000). Formaldehyde fact sheet. Accessed July 25, 2019. https://www.epa.gov/sites/production/files/2016-09/documents/formaldehyde.pdf

EPA. (2006). Reassessment of two exemptions from the requirement of a tolerance for 2-(2'-Hydroxy-5′-methylphenyl)benzotriazole and Octyl Epoxytallate. July 25, 2019. https://www.epa.gov/sites/production/files/2015-04/documents/benzotriazole.pdf

EPA. (2018). Ethylene oxide fact sheet. July 25, 2019. https://www.epa.gov/sites/production/files/2016-09/documents/ethylene-oxide.pdf

Fankhauser-Noti, A., Fiselier, K., Biedermann-Brem, S., & Grob, K. (2006). Assessment of Epoxidized Soy Bean Oil (ESBO) migrating into foods: Comparison with ESBO-like epoxy fatty acids in our normal diet. *Food and Chemical Toxicology, 44*(8), 1279–1286.

Farrell, E. K., Chen, Y., Barazanji, M., Jeffries, K. A., Cameroamortegui, F., & Merkler, D. J. (2012). Primary Fatty Acid Amide Metabolism: Conversion of Fatty Acids and an Ethanolamine in N18TG2 and SCP Cells. *Journal of Lipid Research, 53*(2), 247–256.

Feilden, A. (2011). *Update on undertaking extractable and leachable testing*. Shropshire: iSmithers.

Gad, S. C., & Gad-McDonald, S. (2016). *Biomaterials, medical devices, and combination products*. Boca Raton, FL: CRC Press.

Griss, P., & Heimke, G. (1981). Five years experience with ceramic-metalcomposite hip endoprostheses—I clinical evaluation. *Archives of Orthopaedic and Trauma Surgery, 98*, 157–164.

Hahladakis, J. N., Velis, C. A., Weber, R., Lacovidou, E., & Purnell, P. (2018). An overview of chemical additives present in plastics: Migration, release, fate and environmental impact during their use, disposal and recycling. *Journal of Hazardous Materials. 344*(15), 179–199.

Hallab, N. J., Mikecz, K., Vermes, C., Skipor, A., & Jacobs, J. J. (2001). Differential lymphocyte reactivity to serum-derived metal–protein complexes produced from cobalt-based and titanium-based implant alloy degradation. *Journal of Biomedical Materials Research, 56*(3), 427–436.

Hansen, D. C. (2008). Metal corrosion in the human body: The ultimate bio-corrosion scenario. *The Electrochemical Society Interface, 17*(2), 31.

Health Canada. (2018). Draft Screening Assessment Fatty Amides Group. Accessed July 25, 2019. https://www.canada.ca/en/environment-climate-change/services/evaluating-existing-substances/screening-assessment-fatty-amides.html

Helmus, M. N., Gibbons, D. F., & Cebon, D. (2008). Biocompatibility: Meeting a key functional requirement of next-generation medical devices. *Toxicologic Pathology, 36*(1), 70–80.

Hutt, P. B. (1989). A history of government regulation and misbranding of medical devices. *Food Drug Cosmetic Law Journal, 44*(2), 99–117.

IARC. (1999). *Monographs on the evaluation of carcinogenic risks to humans, vol. 74. Surgical Implants and Other Foreign Bodies*. International Agency for Research on Cancer, World Health Organization, Lyon.

IARC. (2012). *Arsenic, metals, fibres, and dusts: A review of human carcinogens. Monographs on the evaluation of carcinogenic risks to humans*. International Agency for Research on Cancer, World Health Organization, Lyon.;100C.

IARC. (2018). *Some industrial chemicals. Monographs on the evaluation of carcinogenic risks to humans*. International Agency for Research on Cancer, World Health Organization, Lyon.; 115.

ICH. (2017). M7(R1) Assessment and Control of DNA Reactive (Mutagenic) Impurities in Pharmaceuticals to Limit Potential Carcinogenic Risk. Accessed July 25, 2019. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Multidisciplinary/M7/M7_R1_Addendum_Step_4_2017_0331.pdf

ICH. (2018). Q3C(R7). Impurities: Guideline for Residual Solvents. Accessed July 25, 2019. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q3C/Q3C-R7_Document_Guideline_2018_1015.pdf

ICH. (2019). Q3D(R1) Guideline for Elemental Impurities. Accessed July 25, 2019. https://www.ich.org/fileadmin/Public_Web_Site/ICH_Products/Guidelines/Quality/Q3D/Q3D-R1EWG_Document_Step4_Guideline_2019_0322.pdf

ICSC: 0717 (2006). ILO International Chemical Safety Cards (ICSC). Sodium. Accessed July 25, 2019. https://www.ilo.org/dyn/icsc/showcard.display?p_version=2&p_card_id=0717.

IPCS. (1991). Environmental health criteria 111. Triphenyl Phosphate. Accessed July 25, 2019. http://www.inchem.org/documents/ehc/ehc/ehc111.htm#PartNumber:1

IPCS (International Programme on Chemical Safety). (1990). Environmental health criteria 110. Tricresyl Phosphate. Accessed July 25, 2019. http://www.inchem.org/documents/ehc/ehc/ehc110.htm#PartNumber:1

ISO (2001). ISO 10993-14 identification and quantification of degradation products from ceramics.

ISO. (2018). Draft ISO 10993-18 chemical characterization of medical device materials within a risk management process.

JECFA. (1986). WHO food additives series 18: Iron.

Jeffery, E. H., Abreo, K., Burgess, E., Cannata, J., & Greger, J. L. (1996). Systemic aluminum toxicity: effects on bone, hematopoietic tissue, and kidney. *Journal of Toxicology and Environmental Health Part A, 48*(6), 649–666.

Jenke, D. (2009). Compatibility of pharmaceutical products and contact materials: Safety. In *Considerations associated with extractables and leachables*. Wiley, Hoboken, NJ, USA.

Johnson, W., Jr., Bergfeld, W. F., Belsito, D. V., Hill, R. A., Klaassen, C. D., Liebler, D. C., & Heldreth, B. (2018). Safety assessment of Pentaerythrityl Tetra-Di-t-Butyl Hydroxyhydrocinnamate as used in cosmetics. *International Journal of Toxicology, 37*(Suppl 3), 80S–89S.

Keck-Antoine, K., Lievens, E., Bayer, J., Mara, J., Jung, D.-S., & Jung, S.-L. (2016). Additives to design and improve the performance of multilayer flexible packaging. In J. Wagner Jr. (Ed.), *Multilayer flexible packaging* (2nd ed., pp. 53–76). Elsevier, San Diego, CA, USA.

Khan, W., Muntimadugu, E., Jaffe, M., & Domb, A. J. (2014). Implantable medical devices. In *Focal controlled drug delivery* (pp. 33–59). Boston, MA: Springer.

Kitamura, H., Ichinose, S., Hosoya, T., Ando, T., Ikushima, S., Oritsu, M., & Takemura, T. (2007). Inhalation of inorganic particles as a risk factor for idiopathic pulmonary fibrosis—Elemental microanalysis of pulmonary lymph nodes obtained at autopsy cases. *Pathology-Research and Practice, 203*(8), 575–585.

Lagerkvist, B., Linderholm, H., & Nordberg, G. F. (1986). Vasospastic tendency and Raynaud's phenomenon in smelter workers exposed to arsenic. *Environmental Research, 39*, 465–474.

Lanigan, R. S., & Yamarik, T. A. (2002). Final report on the safety assessment of BHT. *International Journal of Toxicology, 21*(Suppl 2), 19–94.

Lansdown, A. B. (2010). A pharmacological and toxicological profile of silver as an antimicrobial agent in medical devices. *Advances in Pharmacological Sciences, 2010*, 16.

Larson, P. S., Finnegan, J. K., Haag, H. B., Smith, B. R., Jr., & Hennigar, G. R. (1960). Chronic toxicity studies on two epoxidized soybean oils in the rat and dog. *Toxicology and Applied Pharmacology, 2*, 649–658.

Lee, H. B., Khang, G., & Lee, J. H. (2007). Polymeric biomaterials. In J. Wong & J. Bronzino (Eds.), *Biomaterials*. Boca Raton: CRC Press. 3:1–3:24.

Lewis, D. B. (2011). Current FDA perspective on leachable impurities in parenteral and ophthalmic drug products. AAPS Workshop on Pharmaceutical Stability – Scientific and Regulatory Considerations for Global Drug Development and Commercialization. https://www.fda.gov/downloads/AboutFDA/CentersOffices/CDER/UCM301045.pdf.

Lewis, C. G., Belniak, R. M., Plowman, M. C., Hopfer, S. M., Knight, J. A., & Sunderman, F. W., Jr. (1995). Intraarticular carcinogenesis bioassays of CoCrMo and TiAlV alloys in rats. *The Journal of Arthroplasty, 10*(1), 75–82.

Liu, P., Hu, W., Wei, F. S., et al. (2006a). Effect of boron on human semen quality. *Chinese Journal Public Health-Shenyang, 22*, 940–942.

Liu, P., Hu, W., Gooping, W., et al. (2006b). Effect of occupational exposure to boron on the sperm quality of males. *Chinese Journal Industrial Hygiene and Occupational Diseases., 24*, 167–169.

Lu, J., Descamps, M., Dejou, J., Koubi, G., Hardouin, P., Lemaitre, J., & Proust, J. P. (2002). The biodegradation mechanism of calcium phosphate biomaterials in bone. *Journal of Biomedical Materials Research, 63*(4), 408–412.

Ma, N., Ma, B., Zhou, Y., Zhu, H., Zhou, Y., Huan, Z., Wang, P., & Chang, J. (2019). In vivo evaluation of the subchronic systemic toxicity of Akermanite bioceramic for bone regeneration following ISO standard methods. *RSC Advances, 9*(31), 17530–17536.

Magos, L. T. (1986). In: Handbook on the toxicology of metals (2nd ed.), edited by L. Friberg, G. F. Nordberg and V. Vouk. New York: Elsevier, p. 568–593.

Maillard, J.-Y., & Hartemann, P. (2013). Silver as an antimicrobial: Facts and gaps in knowledge. *Critical Reviews in Microbiology, 39*(4), 373–383.

Manivasagam, G., Dhinasekaran, D., & Rajamanickam, A. (2010). Biomedical implants: Corrosion and its prevention-a review. Recent patents on corrosion science. In R. Narayan (Ed.), *Biomedical materials*. Chapel Hill, NC: Springer Science & Business Media.

McKeen, L. W. (2014). Plastics used in medical devices. In K. Modjarrad & S. Ebnesajjad (Eds.), *Handbook of polymer applications in medicine and medical devices* (pp. 21–53). Elsevier, Chadds Ford, PA, USA.

Neal-Kluever, A. P., Bailey, A. B., & Hatwell, K. R. (2015). Safety assessment for Octadecyl 3-(3,5-di-tert-butyl-4-hydroxyphenyl)-propionate (CAS Reg. No. 2082-79-3) from use in food contact applications. *Food and Chemical Toxicology, 86*, 176–190.

NTP. (1994). Technical report on the toxicology and carcinogenesis studies of barium chloride dihydrate (CAS No. 10326–27-9) in F344/N rats and B6C3F1 mice (drinking water studies). National Toxicology Program, Public Health Service, U.S. Department of Health and Human Services, Research Triangle Park, NC. NTP TR 432.

NTP. (2008). NTP-CERHR monograph on the potential human reproductive and developmental effects of bisphenol A. National Toxicology Program, Public Health Service, U.S. Department of Health and Human Services, Research Triangle Park, NC.

OECD (Organization for Economic Cooperation and Development). (2009). SIDS Initial Assessment Profile Summary Conclusions of the SIAR for Tris(2,4-di-tert-butylphenyl)phosphite. Accessed July 25, 2019. https://hpvchemicals.oecd.org/UI/handler.axd?id=cfda1bb7-ec51-4010-bc84-2cee9dd819ed

Ouellette, R. J., & Rawn, J. D. (2015). *Organic chemistry study guide: Key concepts, problems, and solutions*. Elsevier, Waltham, MA.

Pape, P. G. (2011). Adhesion promoters: Silane coupling agents. In M. Kutz (Ed.), *Applied Plastics and Engineering Handbook: Processing and Materials* (pp. 503–517). Elsevier, Chadds Ford, PA, USA.

Pohanish, R. P. (2012). *Sittig's handbook of toxic and hazardous chemicals and carcinogens* (Vol. Vol 2, 6th ed.). Waltham, MA: Elsevier, Inc. 2359.

Polivkova, M., Hubacek, T., Staszek, M., Svorcik, V., & Siegel, J. (2017). Antimicrobial treatment of polymeric medical devices by silver nanomaterials and related technology. *International Journal of Molecular Sciences, 18*(2), 419–442.

Ribeiro, A. M., Flores-Sahagun, T. H., & Paredes, R. C. (2016). A perspective on molybdenum biocompatibility and antimicrobial activity for applications in implants. *Journal of Materials Science, 51*(6), 2806–2816.

RIVM. (2001) RIVM Report 711701025: Re-evaluation of Human-Toxicological Maximum Permissible Risk Levels. Ruksinstutuut Voor Volksgezondheid En Milieu (National Institute of Public Health and the Environment).

Rizzetti, M. C., Liberini, P., Zarattini, G., Catalani, S., Pazzaglia, U., Apostoli, P., & Padovani, A. (2009). Loss of sight and sound. Could it be the hip? *The Lancet, 373*(9668), 1052.

Sansone, V., Pagani, D., & Melato, M. (2013). The effects on bone cells of metal ions released from Orthopaedic implants. A review. *Clinical Cases in Mineral and Bone Metabolism, 10*(1), 34.

Sastri, V. R. (2013). *Plastics in medical devices: Properties, requirements, and applications*. Waltham, MA: William Andrew.

Sayli, B. S. (2003). Low frequency of infertility among Workers in a Borate Processing Facility. *Biological Trace Element Research, 93*, 19–29.

SCENIHR (Scientific Committee on Emerging and Newly-Identified Health Risks). (2008). Opinion on the Safety of Medical Devices Containing DEHP-Plasticized PVC or Other Plasticizers on Neonates and Other Groups Possibly at Risk. Accessed July 25, 2019. https://ec.europa.eu/health/archive/ph_risk/committees/04_scenihr/docs/scenihr_o_014.pdf

Shand, E. B. (1958). *Glass Engineering Handbook*. McGraw-Hill, New York.

Shrivastava, A. (2018). *Introduction to Plastics Engineering*, William Andrew, Oxford, UK.

Stensaas, S. S., & Stensaas, L. J. (1978). Histopathological evaluation of materials implanted in the cerebral cortex. *Acta Neuropathologica, 41*(2), 145–155.

Stults, C. L., Ansell, J. M., Shaw, A. J., & Nagao, L. M. (2015). Evaluation of extractables in processed and unprocessed polymer materials used for pharmaceutical applications. *AAPS PharmSciTech, 16*(1), 150–164.

Sumitra, A. P., et al. (2016). Pharmaceutical glass interactions: A review of possibilities. *Journal of Pharmaceutical Sciences and Research., 8*(2), 103–111.

Teasdale, A., Jahn, M., Bailey, S., Feilden, A., Taylor, G., Corcoran, M. L., Malick, R., Jenke, D., Stults, C. L. & Nagao, L. M. (2015). Controlled extraction studies applied to polyvinyl chloride and polyethylene materials: conclusions from the ELSIE controlled extraction pilot study. *AAPS PharmSciTech, 16*(3), 664–674.

TOXNET/HSDB (Hazardous Substances Data Bank). (2019). U.S. National Library of Medicine online database.

Turner, I. G. (2009). Ceramics and Glasses. In R. Narayan (Ed.), *Biomedical materials*. Chapel Hill, NC: Springer Science & Business Media.

WHO. (2002). Concise international chemical assessment document 39. Acrylonitrile. Accessed July 25, 2019. https://www.who.int/ipcs/publications/cicad/cicad39_rev.pdf

WHO. (2003). Concise international chemical assessment document 54. Ethylene Oxide. Accessed July 25, 2019. https://www.who.int/ipcs/publications/cicad/en/cicad54

Wong, L. C., Heimbach, M. D., Truscott, D. R. & Duncan, B. D. (1964). Boric acid poisoning: Report of 11 cases. *Canadian Medical Association Journal, 90*(17), 1018.

Appendix A: Biocompatibility CROs for Medical Devices

Shayne C. Gad

The number of contract laboratories serving device "land" is different from that primarily serving the pharmaceutical industry, while there are approximately 200 organizations performing testing for pharmaceuticals (see Gad, S.C, Spainhour, C.R and Serata, D. Contract Research Organizations: Their History, Selection, and Utilization).

Laboratories supporting device development fall into three broad categories:

1. Conduct GLP compliant baseline ISO-10993 studies. "Traditional" CROs do not operate in this range as their operations do not allow them to economically compete.
2. Conduct of extraction and analytics studies to provide quantitative and quantitative data for leachable and extractable (L&E) assessments. There is actually a small subset here that focuses on the analytical aspects for respiratory devices.
3. Conduct of long-term toxicity studies or studies requiring surgical implantation or replacement of devices.

Group 1: ISO 10993 studies

Lab name (affiliated labs/ past names)	Location(s)	Phone #(s)	Website
American Preclinical and Devices	Minneapolis, MN	(763) 717-7990	www.americanpreclinical.com
BD Biosciences	San Jose, CA	(408) 432-9475	www.bdbiosciences.com
BioSafety Research Center (BSRC)	Shizuoka, Japan	+81 538 58 3572	www.anpyo.or.jp
BTS	10665 Sacramento Valley Rd. San Diego, CA	(858) 605-5882	www.btsresearch.com
CBSET	500 Shire Way Lexington, MA	(781) 541-5555	www.cbset.org

S. C. Gad, PhD, DABT
Gad Consulting Services, Raleigh, NC, USA

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7

Lab name (affiliated labs/ past names)	Location(s)	Phone #(s)	Website
Covance (formerly Harlan, then Envigo)	Israel	(888) Covance	www.covance.com
CXR Biosciences	Dundee, UK	+44 (0)-1382-432163	www.cxrbiosciences.com
	Scotland, UK		
Eurofins	Multiple sites		www.eurofins.com
Geneva Laboratories	Elkhorn, WI	(262) 753-9955	www.genevalabs.com
ICP Firefly	P.O. Box 6198 Alexandria NSW Sydney, Australia	91-2-9310-3899	www.icpfirefly.com
In Vitro Technologies	Baltimore, MD	(410) 455-1245	www.invitrotech.com
NAMSA	Toledo, OH	(419) 666-9455	www.namsa.com
	Northwood, OH	(866) 666-9455	
	France	33-4-78-07-92-34	
	Irvine, CA	(949) 951-3110	
	Kennesaw, GA	(770) 427-3101	
Nelson Labs	Salt Lake City, UT	(800) 826-2088	www.nelsonlabs.com
Pacific Biolabs	Hercules, CA	(510) 565-9000	www.pacificbiolabs.com
Product Safety Lab (PSL)	Dayton, NJ	(732) 438-5100	www.productsafetylabs.com
STS duo TEK, Inc.	Rush, NY	(800) 836-4850	www.stsduotek.com
Toxikon Corp.	15 Wiggins Ave. Bedford, MA 01730	(781) 275-3330 (p) (781) 271-1136 (f)	www.toxikon.com
WuXi App Tec	Oakville, ON	(866) 337-4500	www.wellspringpharma.com

Group 2: Analytical support

Lab name (affiliated labs/past names)	Location(s)	Phone #(s)	Website
American Preclinical and Devices	Minneapolis, MN	(763) 717-7990	www.americanpreclinical.com
Avomeen Analytical Services	4840 Venture Dr. Ann Arbor, MI 48108	(734) 222-1090	www.avomeen.com
BASI	West Lafayette, IN Europe	(800) 845-4246 (765) 463-4527 44(0) 247 663 9574 (EU)	www.basinc.com
Chemic Labs	Canton, MA	(781) 821-5600	www.chemiclabs.com
EAG Laboratories (Eurofins)	Columbia, MD	(573) 474-8579	www.abclabs.com

Lab name (affiliated labs/past names)	Location(s)	Phone #(s)	Website
EKG Labs[a]	4633 World Parkway Circle Saint Louis, MO	(810) EKG-LABZ	www.ekglabs.com
Hall Analytical Labs	Millbrook Business Centre, Floats Rd, Wythenshawe Manchester M23 9YJ, UK	+44-161-286-7889	www.hallanalytical.co.uk
Impact Analytical[a]	1940 North Stark Road Midland, MI 48642	(855) 207-5894	www.impactanalytical.com
Intertek Labs	London, UK	(800) 967-5352	www.intertek.com
Jordi Labs[a]	Mansfield, MA	(508) 966-1301	www.jordilabs.com
Mass Spec Labs	18226 McDurmott West, Suite F Irvine, Ca 92614	(949) 872-2724	www.massspeclab.com
NAMSA	Toledo, OH	(419)666-9455	www.namsa.com
	Northwood, OH	(866)666-9455	
	France	33-4-78-07-92-34	
	Irvine, CA	(949)951-3110	
	Kennesaw, GA	(770)427-3101	
Nelson Labs	Salt Lake City, UT	(800)826-2088	www.nelsonlabs.com
Pacific Biolabs	Hercules, CA	(510) 565-9000	www.pacificbiolabs.com
Piper Medical[a]	1053 Village Ln Chico, BA 95926	(530) 588-6119	www.pipermedical.com
Smithers Labs	425 West Market Street Akron, Ohio 44303	+1 330-762-7441	www.smithersrapra.com
STS duo TEK, Inc.	Rush, NY	(800) 836-4850	www.stsduotek.com
Toxikon Corp.	15 Wiggins Ave. Bedford, MA 01730	(781)275-3330 (p) (781)271-1136 (f)	www.toxikon.com
VR Analytical	63020 Lower Meadow Dr. Bend, OR 97701	(541) 388-1253	www.vranalytical.com
WuXii App Tec	Oakville, ON	(866) 337-4500	www.wellspringpharma.com

[a]Respiratory devices

Group 3: Chronic and surgical

Lab name (affiliated labs/past names)	Location(s)	Phone #(s)	Website
American Preclinical and Devices	Minneapolis, MN	(763) 717-7990	www.americanpreclinical.com
CBSET	500 Shire Way Lexington, MA	(781) 541-5555	www.cbset.org
Charles River Labs	54943 N. Main Street Mattawan, MI	(269) 668-3336	www.criver.com
Covance	Princeton, NJ	(888) Covance	www.covance.com
Frontage	Concord, OH	(888) 763-4797	www.ricerca.com
Geneva Laboratories	Elkhorn, WI	(262) 753-9955	www.genevalabs.com
ICP Firefly	P.O. Box 6198 Alexandria NSW Sydney, Australia	91-2-9310-3899	www.icpfirefly.com
NAMSA	Toledo, OH	(419)666-9455	www.namsa.com
	Northwood, OH	(866)666-9455	
	France	33-4-78-07-92-34	
	Irvine, CA	(949)951-3110	
	Kennesaw, GA	(770)427-3101	
Pacific Biolabs	Hercules, CA	(510) 565-9000	www.pacificbiolabs.com
Sinclair	562 State Road DD, Auxvasse, MO	(573) 387-4400	www.sinclairreserach.com
Southern Research Institute	Birmingham, AL	(888)322-1166 (205) 211-7472	www.southernreasearch.com
SRI International	Menlo Park, CA	(650)859-2000 (866)451-5998	www.sri.com
Toxikon Corp.	15 Wiggins Ave. Bedford, MA 01730	(781)275-3330 (p) (781)271-1136 (f)	www.toxikon.com
WuXi App Tec	Oakville, ON	(866) 337-4500	www.wellspringpharma.com

Reference

Gad, S. C., & Spainhour, C. B. (2011). *Contract research and development organizations: Their role in global product development*. New York: Springer. p. 137

Appendix B

Jerry L. Bettis Jr.

Table 1 Global regulatory agencies and their major regulatory documents (Adapted from Schuh and Funk, 2019)

Regulatory agencies	Major biocompatibility, safety, and efficacy documents	Website URL
European Union—European Commission	ISO, IMDRF, OECD, GLP, Country-Specific Authority Documents and Pharmacopeia	http://ec.europa.eu/growth/sectors/medical-devices/guidance_en
United Kingdom—Medicines and Healthcare Products Regulatory Agency	ISO, IMDRF, OECD, GLP, British Pharmacopeia	https://www.gov.uk/topic/medicines-medical-devices-blood/medical-devices-regulation-safety
	Guidance on Legislation: Clinical Investigations of Medical Devices—Biological Safety Assessment, November 2013	https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/376937/Biological_safety_assessment.pdf
United States of America—Center for Devices and Radiologic Health	ISO, IMDRF, GLP, CFR Title 21 Subchapter H, USP-NF, ASTM International 10993-1 Usage Guidance (Replaces Blue Book Memorandum G95-1)	http://www.fda.gov/MedicalDevices/default.htm https://www.fda.gov/downloads/medicaldevices/deviceregulationandguidance/guidancedocuments/ucm348890.pdf
Canada—Health Canada Drugs and Health Products—Medical Devices Bureau	ISO, IMDRF, GLP, CSA Group	http://www.hc-sc.gc.ca/dhp-mps/md-im/index-eng.php
Japan—Pharmaceutical and Medical Devices Agency	ISO, IMDRF, GLP, JSA	http://www.pmda.go.jp/english/index.html

(continued)

J. L. Bettis Jr., PhD
Gad Consulting Services, Raleigh, NC, USA
e-mail: jerry.bettis@gadconsulting.com

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7

Table 1 (continued)

Regulatory agencies	Major biocompatibility, safety, and efficacy documents	Website URL
Australia—Therapeutic Goods Administration	ISO, IMDRF, OECD, GLP	http://www.tga.gov.au/medical-devices-ivds
New Zealand—New Zealand Medicines and Medical Devices Safety Authority (MedSafe)	ISO, IMDRF, OECD, GLP	www.medsafe.govt.nz
Mexico—The Federal Commission for the Protection against Sanitary Risk (Comisió n Federal para la Protecció n contra Riesgos Sanitarios)	ISO	http://www.cofepris.gob.mx/AS/Paginas/Registro%20de%20Dispositivos%20Medicos%20por%20Equivalencia/Registros-de-Dispositivos.aspx (Online translation to English available at the top of the linked page)
Brazil—The Brazilian Health Regulatory Agency (Ageˆ ncia Nacional de Vigilaˆncia Sanitária)	ISO, ABNT MERCOSUR/GMC/Res. No 40/00	http://portal.anvisa.gov.br/ (Limited English)
Argentina—The National Administration of Drugs, Foodstuffs and Medical Technology (La Administració n Nacional de Medicamentos, Alimentos y Tecnoló gıa Mé dica)	ISO	http://www.anmat.gov.ar/ (Limited English)
China—China Drug Administration (CDA)	ISO, AHWP, APEC	http://eng.sfda.gov.cn/WS03/CL0770/
Hong Kong—The Department of Health Medical Device Administrative Control System	ISO, IMDRF, AHWP, APEC [GN-01] Overview of the Medical Device Administrative Control System	http://www.mdco.gov.hk/eindex.html http://www.mdco.gov.hk/tc_chi/mdacs/mdacs_gn/files/gn_01.pdf
India—Central Drug Standards Control Organization	ISO Medical Devices Rules 2017 Implemented January 1, 2018; some devices previously regulated as drugs	https://cdscomdonline.gov.in/NewMedDev/Homepage
Singapore—Health Sciences Authority (HSA)	ISO, AHWP, Standards, Productivity, and Innovation (SPRING) Board GN-16-R2 Guidance on Essential Principles for Safety and Performance of Medical Devices (HSA Regulatory Guidance)	http://www.hsa.gov.sg/content/hsa/en/Health_Products_Regulation/Medical_Devices/Overview.html

(continued)

Table 1 (continued)

Regulatory agencies	Major biocompatibility, safety, and efficacy documents	Website URL
Russia—The Federal Service for Control over Healthcare and Social Development (Roszdravnadzor)	IMDRF, Gosudarstvennyy standart (GOST) Standards Procedure for Medical Device Conformity Assessment in the Form of Technical Trials Toxicological Studies and Clinical Trials for Medical Device State Registration Purposes (2014)	http://www.roszdravnadzor.ru/en http://en.imeda.ru/netcat_files/105/103/order_no_2n_from_09_01_2014.pdf

Note: (EU) European Union; (UK) United Kingdom; (ISO) International Organization for Standardization; (IMDRF) International Medical Device Regulators Forum; (OECD) Organization for Economic Co-operation and Development; (GLP) Good Laboratory Practices; (CFR) Code of Federal Regulations; (USP-NF) US Pharmacopeia and National Formulary; (CSA) Group Canadian Standards Association Group; (JSA) Japanese Standards Association; (ABNT) Brazilian Association of Technical Standards; (AHWP) Asian Harmonization Working Party; (APEC) Asia-Pacific Economic Cooperation; (GOST) Gosudarstvennyy standart

Table 2 International Organization for Standardization (ISO) documents applicable to biological testing of biomaterials and medical devices (Paid Access at https://www.iso.org)

ISO series[b]	Part	Year published/ revision or update	Title of standard	Current status (as of March 2018)
Technical Committee ISO/TC 194 Biological and Clinical Evaluation of Medical Devices				
10993	1	2018	Biological Evaluation of Medical Devices—Part 1: Evaluation and Testing within a Risk Management Process	Revision to 2009/Cor[c] published August 2018
10993	2	2006	Biological Evaluation of Medical Devices—Part 2: Animal Welfare Requirements	Reviewed and confirmed in 2015
10993	3	2014	Biological Evaluation of Medical Devices—Part 3: Tests for Genotoxicity, Carcinogenicity and Reproductive Toxicity[d]	Published. See TR 10993-33:2015 genotoxicity supplement
10993	4	2017	Biological Evaluation of Medical Devices—Part 4: Selection of Tests for Interactions with Blood	Published
10993	5	2009	Biological Evaluation of Medical Devices—Part 5: Tests for In Vitro Cytotoxicity	Reviewed and confirmed in 2017
10993	6	2016	Biological Evaluation of Medical Devices–Part 6: Tests for Local Effects After Implantation	Published

(continued)

Table 2 (continued)

ISO series[b]	Part	Year published/ revision or update	Title of standard	Current status (as of March 2018)
10993	7	2008/Cor 1:2009	Biological Evaluation of Medical Devices—Part 7: Ethylene Oxide Sterilization Residuals	Reviewed and confirmed 2016. ISO 10993-7:2008/ DAmd[c] 1 under development
10993	8		Withdrawn—Selection of Reference Materials	Not applicable
10993	9	2009	Biological Evaluation of Medical Devices—Part 9: Framework for Identification and Quantification of Potential Degradation Products	Replacement ISO/ DIS[b,c] 10993-9 under development
10993	10	2010	Biological Evaluation of Medical Devices—Part 10: Tests for Irritation and Skin Sensitization	Reviewed and confirmed in 2016. Replacement ISO/ AWI[c] 10993-10 under development
10993	11	2017	Biological Evaluation of Medical Devices—Part 11: Tests for Systemic Toxicity[d]	Published
10993	12	2012	Biological Evaluation of Medical Devices—Part 12: Sample Preparation and Reference Materials	Replacement ISO/ AWI[c] 10993-12 under development
10993	13	2010	Biological Evaluation of Medical Devices—Part 13: Identification and Quantification of Degradation Products from Polymeric Medical Devices	Reviewed and confirmed in 2013
10993	14	2001	Biological Evaluation of Medical Devices—Part 14: Identification and Quantification of Degradation Products from Ceramics	Reviewed and confirmed in 2013
10993	15	2000	Biological Evaluation of Medical Devices—Part 15: Identification and Quantification of Degradation Products from Metals and Alloys	Reviewed and confirmed in 2013 Replacement ISO/DIS 10993-15 under development
10993	16	2017	Biological Evaluation of Medical Devices—Part 16: Toxicokinetic Study Design for Degradation Products and Leachables	Published
10993	17	2002	Biological Evaluation of Medical Devices—Part 17: Establishment of Allowable Limits for Leachable Substances	Reviewed and confirmed in 2016 Replacement ISO/ AWI 10993-17 under development

(continued)

Table 2 (continued)

ISO series[b]	Part	Year published/ revision or update	Title of standard	Current status (as of March 2018)
10993	18	2005	Biological Evaluation of Medical Devices—Part 18: Chemical Characterization of Materials—Check New Version under Development	Reviewed and confirmed in 2013 Replacement ISO/DIS 10993-18 under development
TS[e] 10993	19	2006	Biological Evaluation of Medical Devices—Part 19: Physicochemical, Morphological and Topographical Characterization of Materials	Replacement ISO/ DTR[b,c] 10993-19-8 under development
TS 10993	20	2006	Biological Evaluation of Medical Devices—Part 20: Principles and Methods for Immunotoxicology Testing of Medical Devices	Replacement ISO/ NP[b,c] TS 10993-20 under development
TR[e] 10993	22	2017	Biological Evaluation of Medical Devices—Part 22: Guidance on Nanomaterials	Published
WD[b,c] 10993	23		Biological Evaluation of Medical Devices—Part 23: Determination of Skin Irritation of Medical Device Extracts Using Reconstructed Human Epidermis (RhE)	Under development
TR 10993	33	2015	Biological Evaluation of Medical Devices—Part 33: Guidance on Tests to Evaluate Genotoxicity—Supplement to ISO 10993-3	Published
NP TR 10993	55		Round Robin on Cytotoxicity—Part 55: (No Title)	Under development
TR 37137		2014	Cardiovascular Biological Evaluation of Medical Devices—Guidance for Absorbable Implants	Replacement ISO/ DTR 37137-2 under development
NP TS 37137	1	2014	Biological Evaluation of Medical Devices—Part 1: Guidance for Absorbable Implants	Under development
TR 37137	2	2014	Cardiovascular Biological Evaluation of Medical Devices—Guidance for Absorbable Implants. Part 2: Standard Guide for Absorbable Metals	Under development
DTS[b,c] 21726			Biological Evaluation of Medical Devices—Application of the Threshold of Toxicological Concern (TTC) for Assessing Biocompatibility of Extractable Substances from Medical Devices	Under development

(continued)

Table 2 (continued)

ISO series[b]	Part	Year published/ revision or update	Title of standard	Current status (as of March 2018)
TR 15499		2016	Biological Evaluation of Medical Devices—Guidance on the Conduct of Biological Evaluation Within a Risk Management Process	Published
CD 22442	1		Medical Devices Utilizing Animal Tissues and Their Derivatives—Part 1: Application of Risk Management	Under development; previously 22442-1:2015
NP 22442	2		Medical Devices Utilizing Animal Tissues and Their Derivatives—Part 2: Controls on Sourcing, Collection and Handling	Under development; previously 22442-2:2015
Technical Committee ISO/TC 150 Implants for Surgery				
5840	1	2015	Cardiovascular Implants—Cardiac Valve Prostheses—Part 1: General Requirements	Published
5840	2	2015	Cardiovascular Implants—Cardiac Valve Prostheses—Part 2: Surgically Implanted Heart Valve Substitutes	Published
5840	3	2015	Cardiovascular Implants—Cardiac Valve Prostheses—Part 3: Heart Valve Substitutes Implanted by Transcatheter Techniques	Under development
TS 17137		2014	Cardiovascular Implants and Extracorporeal Systems—Cardiovascular Absorbable Implants	Replacement ISO/NP TS 17137 under development
25539	1	2017	Cardiovascular Implants—Endovascular Devices—Part 1: Endovascular Prostheses	Published
NP 25539	2	2012	Cardiovascular Implants—Endovascular Devices—Part 2: Vascular Stents	Under development
25539	3	2011	Cardiovascular Implants—Endovascular Devices—Part 3: Vena Cava Filters	Published
14708	5	2010	Implants for Surgery—Active Implantable Medical Devices—Part 5: Circulatory Support Devices	Replacement ISO/CD 14708-5 under development
7197		2006/Cor 1:2007	Neurosurgical Implants—Sterile, Single-Use Hydrocephalus Shunts and Components	Published
17853		2011	Wear of Implant Materials—Polymer and Metal Wear Particles—Isolation and Characterization	Reviewed and confirmed in 2016

(continued)

Table 2 (continued)

ISO series[b]	Part	Year published/ revision or update	Title of standard	Current status (as of March 2018)
21534		2007	Nonactive Surgical Implants—Joint Replacement Implants—Particular Requirements	Reviewed and confirmed in 2016
Technical Committee ISO/TC 172 Ophthalmic Optics and Instruments				
11979	5	2006	Ophthalmic implants—Intraocular Lenses—Part 5: Biocompatibility	Replacement ISO/WD 11979 under development
11979	8	2017	Ophthalmic Implants—Intraocular Lenses—Part 8: Fundamental Requirements	Published
9394		2012	Ophthalmic Optics—Contact Lenses and Contact Lens Care Products—Determination of Biocompatibility by Ocular Study with Rabbit Eyes	Reviewed and confirmed 2017
16671		2015/Amd 1:2017	Ophthalmic Implants—Irrigating Solutions for Ophthalmic Surgery	Published
16672		2015	Ophthalmic Implants—Ocular Endotamponades	Replacement DIS 16672 under development
15798		2013/Amd 1:2017	Ophthalmic Implants—Ophthalmic Viscosurgical Devices	Published
Technical Committee ISO/TC 106 Dentistry				
TS 22911		2016	Dentistry—Preclinical Evaluation of Dental Implant Systems—Animal Test Methods	Published
7405		2008/Amd 1:2013	Dentistry—Evaluation of Biocompatibility of Medical Devices Used in Dentistry	Replacement ISO/FDIS 7405 under development
22803		2004	Dentistry—Membrane Materials for Guided Tissue Regeneration in Oral and Maxillofacial Surgery— Contents of a Technical File	Published
Technical Committee ISO/TC 84 Devices for Administration of Medicinal Products and Catheters				
10555	6	2015	Intravascular Catheters—Sterile and Single-use Catheters—Part 6: Subcutaneous Implanted Ports	Published
Technical Committee ISO/TC 121 Lung Ventilators and Related Equipment				
18562	1	2017	Biocompatibility Evaluation of Breathing Gas Pathways in Healthcare Applications—Part 1: Evaluation and Testing Within a Risk Management Process	Published

(continued)

Table 2 (continued)

ISO series[b]	Part	Year published/ revision or update	Title of standard	Current status (as of March 2018)
Technical Committee ISO/TC 210 Quality Management and Corresponding General Aspects for Medical Devices				
14971		2007	Medical Devices—Application of Risk Management to Medical Devices	Replacement ISO 14971 and DTR 24971 under development
DTR 24971			Medical Devices—Guidance on the Application of ISO 14971	
TR 13121		2011	Nanotechnologies—Nanomaterial Risk Evaluation	Published
TR 16197		2014	Nanotechnologies—Compilation and Description of Toxicological Screening Methods for Manufactured Nanomaterials	Published
TS 80004	5	2011	Nanotechnologies—Vocabulary—Part 5: Nano/Bio Interface	Under development

[a]Documents are available in paper and PDF form, with some documents becoming available in ePUb and e-book formats and available for Kindle and Apple devices

[b]Standardized notation for referencing the ISO standards numerically should include the specifics of the document series, number, and year, for example, ISO 10993-6:2016 or ISO/TR 10993-33:2015. These standards are generally reviewed every 5 years

[c]ISO Abbreviations: (Amd) amendment; (AWI) approved work item; (CD) committee draft; (Cor) corrigenda; (DAmd) draft amendment; (DIS) draft international standard; (DTR) draft technical report; (DTS) draft technical specification; (FDIS) final draft international standard; (NP) new project; (WD) working draft

[d]ICH guidelines are also frequently consulted for these assays

[e]ISO Technical Subject (TS) and Technical Reports (TR) provide technical information with no conformance required

Extracted from Schuh, J. C. L., & Funk, K. A. (2019). Compilation of international standards and regulatory guidance documents for evaluation of biomaterials, medical devices, and 3-D printed and regenerative medicine products. *Toxicol Pathol*, *47*(3), 344–357.

Appendix C: Selected Regulatory and Toxicological Acronyms

510(k)	Premarket notification for change in a device
AALAS	American Association for Laboratory Animal Science
AAMI	Association for the Advancement of Medical Instrumentation
ABT	American Board of Toxicology
ACGIH	American Conference of Governmental Industrial Hygienists
ACT	American College of Toxicology
ADE	Acceptable Daily Exposure
ADI	Allowable Daily Intake
AIDS	Acquired Immune Deficiency Syndrome
AIMD	Active Implantable Medical Device
ANSI	American National Standards Institute
APHIS	Animal and Plant Health Inspection Service
ASTM	American Society for Testing and Materials
CAS	Chemical Abstract Service
CBER	Center for Biologic Evaluation and Research (FDA)
CDER	Center for Drug Evaluation and Research (FDA)
CDRH	Center for Devices and Radiological Health (FDA)
CFR	Code of Federal Regulations
CFAN	Center for Food and Distribution (FDA)
CIIT	Chemical Industries Institute of Toxicology
CPMP	Committee on Proprietary Medicinal Products (United Kingdom)
CPSC	Consumer Product Safety Commission
CSE	Control Standard Endotoxin
CSM	Committee on Safety of Medicines (United Kingdom)
CTC	Clinical Trial Certificate (United Kingdom)
CTX	Clinical Trial Certificate Exemption (United Kingdom)
CVM	Center for Veterinary Medicine (US Food and Drug Administration)
DART	Development and Reproduction Toxicology
DHHS	Department of Health and Human Services
DIA	Drug Information Associates

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7

DIC	Disseminated Intramuscular Coagulation
DMF	Device (or Drug) Master File
DOE	Department of Energy
DOT	Department of Transportation
DSHEA	Dietary Supplement Health and Education Act
EEC	European Economic Community
EM	Electron Microscopy
EMA	European Medical Administration
EPA	US Environmental Protection Agency
EU	European Union
FCA	Freund Complete Adjuvant
FDA	US Food and Drug Administration
FDCA	Food, Drug, and Cosmetic Act
FDLI	Food and Drug Law Institute
FHSA	Federal Hazardous Substances Act
FIFRA	Federal Insecticides, Fungicides, and Rodenticides Act
GCP	Good Clinical Practices
GMP	Good Manufacturing Practices
GLP	Good Laboratory Practices
GPM	Guinea Pig Maximization Test
HEW	Department of Health, Education, and Welfare (no longer in existence)
HIMA	Health Industry Manufacturer's Association
HSDB	Hazardous Substances Data Bank
IARC	International Agency for Research on Cancer
ICH	International Conference on Harmonization
id	Intradermal
IDE	Investigational Device Exemption
IND(A)	Investigational New Drug Application
ip	Intraperitoneal
IRAG	Interagency Regulatory Alternatives Group
IRB	Institutional Review Board
IRLG	Interagency Regulatory Liaison Group
ISO	International Organization for Standardization
IUD	Intrauterine Device
iv	Intravenous
JECFA	Joint Expert Committee for Food Additives
JMAFF	Japanese Ministry of Agriculture, Forestry, and Fishery
LA	Licensing Authority (United Kingdom)
LAL	Limulus Amebocyte Lysate
LD_{50}	Lethal dose 50: The dose calculated to kill 50% of a subject population, median lethal dose
LOEL	Lowest Observed Effect Level
MAA	Marketing Authorization Application (EEC)
MD	Medical Device
MHW	Ministry of Health and Welfare (Japan)
MID	Maximum Implantable Dose

MOE	Margin of Exposure
MOU	Memorandum of Understanding
MRL	Maximum Residue Limits
MSDS	Material Safety Data Sheet
MTD	Maximum Tolerated Dose
NAS	National Academy of Science
NCTR	National Center for Toxicological Research
NDA	New Drug Application
NIH	National Institutes of Health
NIOSH	National Institute for Occupational Safety and Health
NK	Natural Killer
NLM	National Library of Medicine
NOEL	No-Observable-Effect Level
NTP	National Toxicology Program
ODE	Office of Device Evaluation
OECD	Organisation for Economic Co-operation and Development
PDI	Primary Dermal Irritancy
PDN	Product Development Notification
PEL	Permissible Exposure Limit
PhRMA	Pharmaceutical Research and Manufacturers Association
PL	Produce License (United Kingdom)
PLA	Produce License Application
PMA	Premarket Approval Application
PMOA	Principal Mode of Action
PMN	Premanufacturing Notice
po	Per os (orally)
PTC	Points to Consider
QAU	Quality Assurance Unit
RAC	Recombinant DNA Advisory Committee
RCRA	Resources Conservation and Recovery Act
RTECS	Registry of Toxic Effects of Chemical Substances
SARA	Superfund/Amendments and Reauthorization Act
sc	Subcutaneous
SCE	Sister chromatic exchange
SNUR	Significant New Use Regulations
SOP	Standard Operating Procedure
SOT	Society of Toxicology
SRM	Standard Reference Materials (Japan)
STEL	Short-Term Exposure Limit
TLV	Threshold Limit Value
TSCA	Toxic Substances Control Act
USAN	US Adopted Name Council
USDA	US Department of Agriculture
USEPA	U.S. Environmental Protection Agency
USP	United States Pharmacopoeia
WHO	World Health Organization

Index

S. C. Gad, *Integrated Safety and Risk Assessment for Medical Devices and Combination Products*, https://doi.org/10.1007/978-3-030-35241-7