Chapter 12
ICH S8: History and Perspectives

Kenneth L. Hastings

Abstract An important aspect of drug safety evaluation is determination of potential adverse effects on immune function. Drug-induced immune dysfunction can present as increased susceptibility to infections and tumors (especially virally induced), hypersensitivity reactions such as drug allergy and autoimmunity, and various inflammation-like phenomena. Although immunotoxicity test methods have been developed to assess environmental chemicals, these had not been applied systematically in drug development prior to promulgation of guidance documents by EMA and FDA. EMA and FDA guidances/guidelines differed in certain important respects, and ICH S8 was written to resolve these somewhat conflicting approaches. The key issue resolved in ICH S8 was whether functional immunotoxicity assays should be conducted routinely or when there was a cause for concern. An important result of ICH S8 is that drug developers can no longer ignore signs of compound-related adverse effects on immunity. ICH S8 provides a systematic approach to determining the need for immunotoxicity testing and includes discussion on appropriate methodology. Based on current experience with ICH S8, the issue of including immune function parameters in standard toxicity testing remains unresolved and may be addressed in future revisions of the document. In addition, guidance on unintended immunostimulation may be needed based on recent experiences in clinical drug development.

12.1 Introduction

Although immunotoxicology was first identified as a distinct specialty in toxicology by Vos (1977), the study of adverse effects on immune function parallels the emergence of immunology. Richet and Portier first described anaphylaxis in 1902, and Auer, in 1911, made the crucial discovery that this reaction required previous

K.L. Hastings (✉)
Sanofi SA, Bethesda, MD, USA
e-mail: kenneth.hastings@sanofi-aventis.com

J.W. van der Laan and J.J. DeGeorge (eds.), *Global Approach in Safety Testing*, AAPS Advances in the Pharmaceutical Sciences Series 5, DOI 10.1007/978-1-4614-5950-7_12,

exposure to the causative substance (Portier and Richet 1902; Auer 1911). These observations became important in early toxicology with the near concurrent discovery of penicillin and Landsteiner's groundbreaking work with what are now called haptens. Often overlooked when considering the discovery of what was (rightly) considered the miracle drug penicillin was an all-too-common side effect: anaphylaxis (Feinberg et al. 1953). Meanwhile, Landsteiner (1945) demonstrated that small molecular weight chemicals could bind irreversibly to proteins and, when injected into animals, could induce an immune response. Landsteiner's work would be utilized in what should be considered the first immunotoxicology assay—the Draize test for allergic contact dermatitis (ACD) (Draize et al. 1944). In the 1950s, Ovary and colleagues established the link between hapten–protein formation and induction of anaphylaxis. The first method useful in studying (if not predicting) anaphylaxis was developed by Ovary: the passive cutaneous anaphylaxis assay (PCA) (Ovary 1958).

Immune-mediated hypersensitivity reactions are now considered a type of immunotoxicity, but unintended *immunosuppression* has long been considered the more important immunotoxic effect. There were two causes identified early in toxicology: ionizing radiation and certain highly reactive chemicals. Both seemed to share a common mechanism: bone marrow toxicity (Auerbach 1958). That both could be toxic to rapidly dividing cells led to experimental therapies for cancer and to prevent rejection of transplanted organs. Thus, the link between cell proliferation and immunity was understood before many of the discoveries we now take for granted.

By the 1970s, many toxicologists and other biomedical scientists understood that the immune system was like any other: susceptible to insult which would lead to impaired function. But systematic investigation of immunotoxicity began, as is true for many subjects of interest to toxicologists, with a disaster. In 1973, an industrial accident in Michigan led to contamination of milk and milk products with polybrominated biphenyls (PBBs). Subsequent toxicology studies demonstrated PBBs to cause adverse effects on immune function in animals and humans (Bekesi et al. 1978). Other chemicals (such as aflatoxin) were found to have similar adverse effects on immune function (Thaxton et al. 1974). This was the beginning of immunotoxicology as a distinct specialty.

Food contamination is considered to be essentially an environmental problem (with the important exception of intentional adulteration), and it is thus not a surprise that immunotoxicology was developed by scientists working for regulatory agencies such as the US Environmental Protection Agency (US EPA). But it is often forgotten that pharmaceutical scientists were also interested in the subject. In 1978, the US Food and Drug Administration held a meeting on adverse effects of drugs on immunity. One paper from this meeting is of particular interest: Nelson Irey, a pathologist at the Walter Reed Armed Forces Institute of Pathology, for what appears to be the first time, grouped several types of adverse drug reactions into a single category as "immunotoxicity." Irey included penicillin-induced anaphylaxis, α-methyl dopa-induced autoimmunity, radiation and cancer chemotherapy-induced susceptibility to infections and tumors, azathioprine-induced lymphoma-like lymphoproliferative disease in renal transplant patients, and vaccine-related hypersensitivity reactions in this category (Irey 1978). For perhaps the first time, apparently unrelated pathologies were understood to have a common basis: immune system impairment.

Irey's observations did not result in a systematic approach to evaluation of xenobiotics for immunotoxic potential. This would be accomplished by Vos and his colleagues in the Netherlands, and somewhat later by scientists at US EPA, the National Institute of Environmental Health Science (NIEHS), and a few universities and associated research institutes (House and Luebke 2007). The close association between classic immunology and the new science of immunotoxicology is evident in the assays that were developed to detect immunotoxic chemicals. Virtually all of the assays are adaptations of methods that had long been used by immunologists to study basic immunobiology. For example, Cunningham (1965) reported an assay that used sheep red blood cells (SRBC) to study immune responses in rodents. This method, now known as the *plaque assay*, was used by Jerne and colleagues to study mechanisms of immune specificity (Jerne and Nordin 1963; Forni et al. 1980). The plaque assay was adapted by Vos and others to study immune impairment by xenobiotic exposure (Dean et al. 1982; Vos 1977). Just as the sheep red blood cell (SRBC) plaque assay was pivotal to understanding the evolution of immune specificity, so to it would prove to be the most useful general assay for detection of xenobiotic immunotoxicants (Putman et al. 2002; Van der Laan et al. 1997).

The second method of importance is what is generally referred to as the "host-resistance assay." Essentially, rodents are exposed (by various routes depending on the challenge agent) to either infectious agents or tumor cells. In classic immunobiology, these models have been used to study immune responses to infections and cancers. Immunotoxicologists adapted these models to study the effects of xenobiotics on the immune response to these challenge agents (Burleson and Burleson 2008).

The SRBC plaque assay is now generally referred to as the T-dependent antibody response (TDAR) assay and combined with host-resistance assays, constitutes the "gold standard" in immunotoxicology (Luster et al. 1993). There are two important points to consider: both assays are relatively nonspecific (they do not "predict" a specific human health effect) and both are assays for impaired *function*. Both are also cornerstone assays recommended in ICH S8.

12.2 Immunotoxicology and ICH

In the early 1980s, a group of scientists from several institutions developed a series of assays to detect xenobiotic immunotoxicity. The nexus of this effort was the US National Toxicology Program (NTP), and the suite of tests is commonly referred to as the "Tier Assays" (Luster et al. 1988). The overall methodology is familiar to toxicologists: one or more relatively nonspecific, but sensitive, assays are used to screen for effect(s) of concern with follow-on tests to confirm and understand targets of toxicity. In the case of immunotoxicity, a series of studies was conducted to determine "concordance," that is, which assays were most useful in identifying known immunotoxicants (Luster et al. 1992, 1993). Fifty chemicals was evaluated using adaptations of several methods commonly used by immunologists. These included xenobiotic effects on immune system organ weights, cellularity, and histology;

certain clinical pathology parameters (blood cell counts, flow cytometry, serum immunoglobulin concentration); and a set of functional assays (TDAR, delayed-type hypersensitivity, cell-mediated immunity, NK cell activity, and host resistance). The most concordant assay was the TDAR and combined with flow cytometry, detected around 90% of the 50 immunotoxicants (House and Luebke 2007). Two important points should be made: the study lacked *negative* controls, and few drugs were included. But this was an important beginning for immunotoxicology: a reasonable approach to evaluation of potential immunotoxicants could now be recommended.

In 1989, the National Institute of Public Health and the Environment in The Netherlands began what was probably the first systematic evaluation of drug immunotoxicity, including assessment of some not considered likely to be immunomodulatory (e.g., verapamil). An important result of these studies was demonstration that methods developed for immunotoxicity evaluation of environmental chemicals also were useful for drugs (De Waal et al. 1995, 1996, 1997, 1998; Van der Laan et al. 1995, 1996). Studies conducted by the US NTP to evaluate the potential immunotoxicity of drugs for treatment of AIDS produced some evidence of concern (Luster et al. 1991). One, zalcitabine, was found to suppress T-helper cell numbers in cynomolgus monkeys (Taylor et al. 1994). This often overlooked study demonstrated two important points: flow cytometry, conducted in a non-rodent species, could yield clinically relevant results (perhaps explaining why a potent antiretroviral drug could reduce mortality without an increase in the accepted surrogate marker for efficacy), and that an immunotoxicity end point could be obtained in a study that was not designed as a stand-alone immunotoxicology assay.

Immunotoxicology assays, demonstrated by Vos and colleagues to be useful for drugs, would be standardized for evaluation of environmental contaminants and food additives, but not for pharmaceuticals. There are several reasons for this. In drug development, flexibility in study design and end points was considered more appropriate, and protocols developed by the Organization for Economic Cooperation and Development (OECD) were used for general guidance (OECD 1995, 2008). For many drugs, it was difficult to distinguish immunotoxicity from exaggerated pharmacodynamics (e.g., transplant drugs, anti-inflammatory drugs, and cancer chemotherapeutics). In fact, it can be argued that the distinction is arbitrary: unlike environmental chemicals, drugs are evaluated in the context of benefit/risk for a given therapeutic indication (Piccotti et al. 2009). In a study conducted with another HIV drug, didanosine, it was shown that immunotoxic effects could be demonstrated with much longer exposure (6 months) than recommended under EPA guidelines (Phillips et al. 1997). Finally, the most important immunotoxic effect associated with drugs appeared to be *hypersensitivity reactions*, not unintended immunosuppression. The EPA tier system was not designed to detect enhanced immune responses that appeared to be the mechanism of what is commonly referred to as "drug allergy."

Nevertheless, it was understood that immunotoxicity, including unintended immunosuppression, was a potentially important adverse drug reaction. Probably, the first to address this issue was the pharmaceutical industry itself. In a 1988 white

paper, the Pharmaceutical Manufacturers Association (PMA), the forerunner of the US industry organization PhRMA, presented a rationale for incorporation of immunotoxicity end points into drug safety evaluation but cautioned that the EPA tier system was not useful in this context (PMA 1988). Two meetings sponsored by the Drug Information Association (DIA) were important for obtaining consensus on the need for guidance on immunotoxicity evaluation of new drugs (Arlington, Virginia, in 1995 and Montreux, Switzerland, in 1996). Consensus was reached in Montreux on a general approach to immunotoxicity testing with emphasis on use of the TDAR as the best general assay (Van der Laan et al. 1997).

US FDA had included an adaptation of the EPA tier system in the first edition of what is commonly referred to as the "Red Book," but these recommendations applied to new food additives, not drugs (Hinton 2000). The first FDA guidance for immunotoxicity evaluation of drugs came from the Division of Antiviral Drug Products in the Center for Drug Evaluation and Research and was developed to address concerns that AIDS patients should not be given drugs that were immunosuppressant (FDA 1993; Hastings 1996). However, it was clear that the guidance was insufficient and that a more formal document should be written.

Concurrent with FDA guidance development, health authorities in Europe were also concerned that potential immunotoxicity was not being appropriately evaluated as part of drug development. It was the divergence of opinion on a specific point that would lead to ICH S8. The European perspective was that *functional* assays should be conducted to evaluate investigational drugs as part of routine safety evaluation (Putman et al. 2003; Vos and Van Loveren 1998). This position was consistent with the approach taken by EPA: the important parameter was potential adverse effects on immune function and should be the basis for any policy recommendation. This position was clearly justified by the available scientific evidence.

The position taken by FDA was that dedicated immunotoxicity studies might not be necessary if the totality of data from nonclinical (and clinical) studies were properly evaluated. The approach advocated by CDER/FDA was that signs of unintended immunosuppression could be observed and only then would dedicated immunotoxicity studies be useful (Hastings 2002).

There was much discussion in the 1990s on the issue of dedicated versus cause-for-concern studies for immunotoxicity evaluation of investigational drugs. The proximal cause for ICH emerged from these discussions. In 2000, the European Agency for the Evaluation of Medical Products (predecessor to the European Medicines Agency; EMA) published a note for guidance on 28-day rodent toxicity studies, which included an appendix that called for dedicated immunotoxicity studies (EMEA 2000). Draft guidance on immunotoxicity evaluation of drugs by the Japanese Ministry of Health, Labor, and Welfare (JMHLW) advocated a tiered approach consistent with EMA: a functional assay should be conducted as part of standard nonclinical safety assessment. FDA/CDER promulgated a guidance that advocated a cause-for-concern approach (US FDA 2002). Clearly, there was a divergence of opinion—resulting in the need for ICH negotiations.

12.3 Writing a Guidance

ICH S8 is important because it illustrates how scientists with divergent and strongly held opinions can examine existing evidence and produce a document that appropriately addresses a safety issue. At the initial EWG meeting in London (October 2003), all of the parties involved agreed that immunotoxicity was an important issue to be addressed. At the time, CPMP (now CHMP) guidance was dominant: dedicated immunotoxicity studies were needed, whereas for CDER/FDA they *might* be needed depending on available data.

The first task of the immunotoxicology expert working group (EWG) was to determine the approaches in use at the time by the pharmaceutical industry to screen drug candidates for immunotoxic potential. This survey found that although there was considerable variability within industry, most relied on standard nonclinical toxicology studies to detect signs of immunotoxicity (Weaver et al. 2005). Some companies conducted immune function studies such as TDAR but almost always if signs suggestive of immunotoxicity had been observed in nonclinical toxicology studies or if there were other causes for concern. The most important finding from both the survey of industry practices, as well as experience by the regulatory agencies, was that signs of immunotoxicity were often either ignored, considered not relevant to clinical use, or were due to "stress." Thus, the problem did not appear to be lack of immunotoxic signs but failure to appropriately evaluate these. Concern over this particular point was important in formulation of specific guidance in the resulting document.

The second task was to determine the need for dedicated *functional* immunotoxicity studies as part of routine drug development. Although the CPMP NfG *seemed* to require either a TDAR or a combination of flow cytometric analysis of immune cells combined with natural killer (NK) cell activity as part of a 28-day repeat-dose toxicology study in rats, this may have been a false interpretation. In fact, the NfG strongly recommended including immune function end points unless there was a compelling reason no to. In effect, FDA and EMA guidances on immunotoxicology differed in recommended approach, not in whether such determinations were needed. FDA/CDER recommended follow-on immunotoxicity testing if there was a cause for concern, whereas EMA and JMHLW recommended dedicated testing unless there was *no* cause for concern. Thus, there was a basis for consensus: both regulatory authorities agreed on the need for immunotoxicity testing, and both agreed that tests such as the TDAR could be recommended.

Finally, the EWG had to consider the *scope* of the guidance. The FDA Guidance on Immunotoxicology Evaluation of New Drugs included an extensive discussion of phenomena generally referred to as "drug allergy" (US FDA 2002). Many types of drug-associated immunopathies are included in this category, but few test methods could be recommended to determine the potential of investigational drugs to cause these adverse effects. There are many methods to determine the ability of a drug administered by dermal application to cause allergic contact dermatitis (ACD), but these are accepted as adequate by all parties in the EWG. Given the relatively narrow

scope of the issue and the absence of discordance on acceptable methods, the issue of testing for ACD potential was omitted. As for other types of drug allergy, no methods could be recommended. The issue of anaphylaxis was particularly difficult since Japanese regulatory authorities had long required PCA and a related test, active systemic anaphylaxis (ASA), be conducted as part of routine drug evaluation (Udaka 1992). Aside from the issue of whether nonclinical anaphylaxis assays were useful, there were actually very few available data upon which to make recommendations. There are animal models that can be used to determine if adverse reactions consistent with anaphylaxis are in fact immune mediated, but these have not undergone sufficient validation to recommend. Finally, biologic drugs were not considered in discussions. Primarily, this decision was taken because many biologic drugs are either recombinant immune system proteins such as cytokines or are intended to modulate immune function by some other mechanism. The case-by-case approach that forms the basis of ICH S6 was considered adequate.

One issue was considered important: signs of unintended immunostimulation. The reason for including this topic was that such signs could be, and often were, seen in either nonclinical or clinical studies (Pieters 2008; Rock et al. 2010). Although such signs *could* be due to drug-specific antibody or cell-mediated mechanisms, there are other possible causes. The important point was that *any* sign of unintended immunomodulation should be evaluated when observed, whether consistent with immunosuppression or immunostimulation.

Finally, it should be noted that consensus on the issue of dedicated functional immunotoxicity assays was never achieved. The first problem was the dataset key to determining adequacy of current industry practices (Weaver et al. 2005). Although results of standard nonclinical toxicology studies (STS) accurately predicted immunotoxicity for ~90% of evaluated drugs, the actual number (42) was small. Data were inadequate for evaluation of 12 drugs, and 7 were cytotoxic oncolytics judged inappropriate for inclusion in the analysis. Most troubling was the fact that STS did not detect signs of immunotoxicity discovered with six drugs in dedicated immunotoxicity studies (primarily TDAR). Clinical data were not available for evaluating this most important measure of concordance. Thus, although agreement was achieved on the cause-for-concern approach, there was a risk that drug-induced unintended immunosuppression could be undetected.

12.4 ICH S8: The Essentials

The ICH S8 guidance document was negotiated for about 2 years—a remarkably short period of time compared to other safety topics. Immunotoxicity was accepted as an ICH topic in Osaka (November 2003), the first draft was produced in McLean, Virginia, in June 2004, and the pivotal Step 2 document was finalized in Yokohama in November 2004. The Step 4 document was signed in Brussels in May 2005, and the final guidance was published August 23, 2005. FDA promulgated the guidance in April 2006. The objectives of the final document are simple: to recommend

methods to evaluate immunotoxic potential of investigational drugs and to provide a scientifically based algorithm to determine the circumstances in which dedicated nonclinical immunotoxicity testing would be needed. Two linked methods form the structure of the guideline: a cause-for-concern paradigm which informs a weight-of-evidence determination of need for further studies.

The phrase "cause for concern," although not used in the guideline, captures the approach to evaluate need for specific immunotoxicity testing. The following factors should be considered (1) findings in nonclinical toxicology studies, (2) pharmacology of the drug, (3) indication, (4) potential structure–activity relationship(s), (5) pharmacokinetics, and (6) relevant observations in clinical use. This is a holistic approach: *all* relevant data should be evaluated for signs of test article immunomodulation.

The first cause for concern is important because there is no reliance on a specific toxicology study. Consider this in contrast to the requirements promulgated in the US EPA Office of Prevention, Pesticides, and Toxic Substances (OPPTS) *Health Effects Test Guidelines*: *OPPTS 870.7800 Immunotoxicity* (US EPA 1998). The EPA Guideline is very specific: the TDAR should be conducted in a 28-day repeat-dose oral administration rodent study, with the high-dose group given the maximum tolerated dose (MTD). For determination of potential test article effects on NK cells, the length of exposure should be 90 days using the same method of administration. ICH S8, in contrast, recommends that results obtained in *all* nonclinical studies be evaluated. There is a trade-off that should be understood: rather than rely on a single data-dense method (TDAR) in a single species (usually rats), the approach given in ICH S8 relies on signals from multiple nonclinical studies in both rodents and non-rodents. This methodology could be considered "data sparse" in comparison to the EPA approach, but vigilance in study analysis should correct this potential problem. Thus, ICH S8 provides an extensive list of relevant observations that could indicate potential immunotoxicity (including recommendations on histopathology).

Signs of immunotoxicity include alterations in immune tissue weight, cellularity, and histologic appearance, blood immunoglobulin changes, and increased incidence of infections and tumors. Anatomic and biochemical changes can suggest immunosuppression or immunostimulation: either should be evaluated. For example, increased numbers of lymph node and splenic germinal centers may be taken to suggest adverse immunoenhancement, but in combination with increased incidence of infections could, in fact, be indicators of immunosuppression. Tumor findings in rodent carcinogenicity bioassays could suggest immunosuppression if there are no other known relevant mechanisms (such as genotoxicity or hormonal activity).

The usefulness of histopathology was of particular concern. Immunotoxicologists have long debated whether immunosuppression can occur in the absence of histologic changes. S8 emphasizes an approach that toxicologic pathologists have called "enhanced histopathology." In addition, S8 provides a list of tissues that should be specifically evaluated for immune effects. Combined with published "best practices" and the fact that many nonclinical studies would be evaluated, this issue should be considered somewhat resolved (Haley et al. 2005; Kuper et al. 2000; Maronpot 2006). Certainly, there are examples where reliance on histopathology alone, especially when conducted using tissues obtained in a 1-month rodent study, would

fail to detect some immunosuppressant compounds, but this is not the holistic approach given in S8.

The second cause for concern is the pharmacology of the drug. There are at least three categories of drugs that could demonstrate signs consistent with immunotoxicity (1) cancer chemotherapeutics, (2) transplant drugs, and (3) anti-inflammatory drugs. In all three cases, immunotoxicity is likely to be exaggerated pharmacodynamics, and dedicated immunotoxicity studies might not provide useful information. However, this should not be a default assumption. Many standard cancer chemotherapeutic drugs are bone marrow toxins, and medical practice has long taken this into consideration (such as isolating patients to minimize infection risk). However, newer chemotherapeutics with novel molecular targets may not have obvious immunosuppressive activity (e.g., Yang and Moses 2008). In this case, even in the absence of concerning signs in nonclinical toxicology studies, determination of immunotoxic potential could be advisable. Drugs developed to prevent organ transplant rejection are often considered to be "obvious immunotoxins," and end points obtained in dedicated immunotoxicity studies have typically been captured in pharmacology studies. Once again, however, this may not be the case for drugs with unique pharmacodynamic properties. Immunotoxicology studies could be useful to separate wanted pharmacodynamics from unintended immunosuppressive effects.

Anti-inflammatory drugs constitute a special category for consideration. This can be illustrated by an often forgotten episode in the early development of steroid drugs. Clinical trials were conducted in patients with tuberculosis: the potent anti-inflammatory (as well as anabolic) effects of corticosteroids resulted in remarkable resolution of symptoms (Shubin et al. 1959). These effects were, of course, temporary, and patients soon developed serious, often fatal, recurrence of active tuberculosis. Pharmacodynamic activity resulted in fatal immunotoxicity. This episode in drug development was accidentally reproduced when the first anti-inflammatory biologics were used to treat rheumatoid arthritis: patients with occult tuberculosis infection sometimes developed fatal active disease (Dixon et al. 2010). The essential issue in evaluating anti-inflammatory drugs is to determine the therapeutic ratio based on immune system parameters: there is an overlap between immunopharmacology and immunotoxicology. Although therapeutic ratio is an issue with virtually any drug, the issue appears to be especially complex for anti-inflammatory drugs. Steroids can be used in combination with antituberculosis drugs for certain manifestations of the disease where inflammation is an important pathologic feature (Cunha 1995). Antitumor necrosis factor alpha (TNFα) monoclonal antibodies can be used for effective treatment of chronic inflammatory diseases such as rheumatoid arthritis and Crohn's disease if the patient does not have tuberculosis infection (CDC 2004). Adverse immunomodulation has been observed with other anti-inflammatory drugs (especially biologics), and nonclinical methods may be useful on a case-by-case basis (Gourley and Descotes 2008).

Indication is a cause for concern if the drug will be given to patients with impaired immune function (HIV patients, children with congenital immunodeficiency, elderly patients). In this context, immunotoxicity studies are likely to identify hazard, but risk would be determined in clinical trials. There was considerable

debate on the issue of indication and patient population. Although some drugs indicated for treatment of HIV infection had been demonstrated to have immunosuppressive effects in both STS and dedicated immunotoxicity studies, these findings had little, if any, impact on product labels or clinical use. The EWG recognized that infants and children could be especially vulnerable to unintended immunosuppressive effects, but did not make specific recommendations on methods for evaluating this possibility. In fact, other than including possible developmental immunotoxicity as a cause for concern (including in utero exposure), the guidance is otherwise relatively silent on best practices for addressing the issue.

Although structural similarity to known immunotoxic drugs is a cause for concern, this is a complex issue. If a drug exhibits a structural alert when analyzed using an in silico method, it is unlikely that immunotoxicity studies would be needed in the absence of signals seen in in vivo studies. Conversely, if signs of immunotoxicity are observed in toxicology studies and there is a structural alert as well, follow-on immunotoxicity studies should be considered.

The most important pharmacokinetic parameter that could indicate a cause for concern is disposition. If a drug and/or a metabolite accumulates in immune system tissues, this would not be stand-alone cause for concern. However, if there are other findings in toxicology studies consistent with such immune tissue accumulation (such as histopathologic alterations), this would be a cause for concern.

Clinical trial data may indicate cause for concern. There are many clinical findings that could indicate the need for nonclinical immunotoxicity studies. Often these studies would be needed to help establish a link between clinical findings such as increased incidence of pneumonia or urinary tract infections and immunosuppression due to the drug. This is not a rare event and has been seen with both drugs and biologics. For example, proton pump inhibitors appear to increase risk for pneumonia and *Clostridium difficile* infections (Gulmez et al. 2007; Linsky et al. 2010). Anti-adhesion molecule monoclonal antibodies may increase the risk of active JC virus encephalopathy (Bloomgren et al. 2012). More complex are issues such as the potential association between acetaminophen and risk of asthma (Eyers et al. 2011). Some types of immunotoxicity appear to decrease vaccine efficacy (Gelinck et al. 2008; Grandjean et al. 2012).

Finally, the issue of stress was extensively discussed. As noted previously, there is a long and troubling history of drug-associated immune impairment being dismissed as stress-related and not relevant to clinical safety. The complexity of the issue perhaps can be best understood as the conundrum of toxicity-induced stress. If thymic atrophy, for example, is observed in animals demonstrating evidence of toxicity not related to immune function, should this be considered immunotoxicity? There are no simple answers to this question, but the EWG concluded that far too often stress is the default explanation for observed signs of immunotoxicity and that this was not acceptable. Thus, there is the statement in the guidance that if the claim is made that signs of immunotoxicity are due to stress, *compelling* evidence must be provided to support this conclusion. Although the guidance is not specific about what should be considered compelling evidence, the implication is that a simple statement of stress causality would not

be sufficient. The guidance recommends that doses used in STS should be less than MTD in order to minimize potential for stress. The Appendix includes a thorough discussion of stress-related effects which could inform interpretation of STS findings.

The Appendix also includes an extensive discussion of specific immune function assays. Thus, although the guidance does not provide a "recipe" for conduct of studies (i.e., specific requirements), numerous useful points to consider are provided. In this respect, the guidance is somewhat unique. Flexibility in study design, based on various considerations, is recognized as an important factor.

12.5 Maintenance

ICH recognizes that scientific advances influence conduct of studies and that there will be a need to update guidance documents. There are several issues that may necessitate maintenance of ICH S8.

ICH S8 does not address the issue of drug allergy. Although there are few methods that can be recommended, the murine local lymph node assay (LLNA) should be considered appropriate to evaluate the safety of dermal drugs. This assay is validated and generally accepted (Gerberick et al. 2005). Although there are few published data using the LLNA to evaluate the potential of drugs to produce ACD, it is unlikely that inclusion in a revised guidance would be controversial.

Although developmental immunotoxicity is recognized as an important cause for concern, specific guidance is not provided in ICH S8. Since promulgation of ICH S8, there have been important advances in developmental immunotoxicology, and this issue should be addressed as part of maintenance (Holsapple et al. 2005). Although controversy is likely on some key aspects of both study design and need for studies, these issues could be successfully addressed in negotiations. Consideration should also be given to evaluation of the immunotoxic potential of drugs intended for use in the elderly.

Biologic drugs are not in the scope of ICH S8, but consideration should be given to this issue. ICH S6(R1) defers to ICH S8 on some important aspects of drug evaluation: especially important is the issue of infections and tumors associated with biologic immunomodulators. Although recommendations are made in the Appendix of ICH S8 on host-resistance assays, this section could be greatly expanded and could provide useful guidance applicable to biologic drugs.

Advanced techniques such as genomics have been applied in immunotoxicology: it is unclear, however, if guidance is needed. However, this is a rapidly changing area of drug safety evaluation, and consideration should be given to whether certain issues should be addressed. For example, immunomics is a technique that could be useful in assessing biologic drugs for adverse immunogenicity (e.g., autoimmune reactions) (Grainger 2004). It is possible that certain epitopes can be identified for which an induced immune response would be a significant hazard.

Adverse immunostimulation is addressed in ICH S8, but no specific guidance is provided on methods for assessment. As part of maintenance, some methods might be worthy of consideration. For example, in vitro methods such as the minimum acceptable biological effect level (MABEL) assay could be recommended to evaluate the safety of agonist immunomodulators (Horvath and Milton 2009; Stebbings et al. 2007). Genomic techniques could also be useful in this context: identified haplotype risk factors could be used to determine potential of test article to produce adverse effects such as "cytokine storm" and "sterile sepsis" (Luebke et al. 2006).

Finally, the original database used in ICH S8 negotiations should be greatly expanded. Useful data are undoubtedly available, and the question of whether immune function assays should be part of standard drug safety assessment can be reexamined. One of the issues that confounded negotiations on this point was whether immunogen challenge could be incorporated into STS without complicating study interpretation. This issue should be considered resolved: the most important issue that remains is optimum parameters for immunogen challenge (e.g., appropriate dose of KLH). There have been examples of unintended immunosuppression with serious clinical consequences (e.g., proton pump inhibitor association with increased risk of pneumonia and *Clostridium difficile* infection, discussed previously). If these adverse immune effects can be modeled in animals (especially as an addition to STS), a strong recommendation could be made that would be a benefit to public health. In addition to TDAR, potential adverse effects on T-independent antibody response and innate immunity should be considered.

ICH S8 should be considered a success: the drug development process has benefited from guidance provided. It is unusual today for signs consistent with immunotoxicity to be ignored or dismissed as "stress" irrelevant to clinical use. Methods to predict drug allergy are needed, and unintended immunostimulation has emerged as a significant problem. But advances in both areas continue to be reported: immunotoxicology is a vibrant field of research with much promise.

References

Auer J (1911) Lethal cardiac anaphylaxis in the rabbit. J Exp Med 14:476–496

Auerbach C (1958) Radiomimetic substances. Radiat Res 9:33–47

Bekesi JG, Holland JF, Anderson HA, Fischbein AS, Rom W, Wolff MS, Selikoff IJ (1978) Lymphocyte function of Michigan dairy farmers exposed to polybrominated biphenyls. Science 199:1207–1209

Bloomgren G, Richman S, Hotermans C, Subramanyam M, Goelz S, Natarajan A, Lee S, Plavina T, Scanlon JV, Sandrock A, Bozic C (2012) Risk of natalizumab-associated progressive multifocal leukoencephalopathy. N Engl J Med 366:1870–1880

Burleson GR, Burleson FG (2008) Animal models of host resistance. In: Herzyk DJ, Bussiere JL (eds) Immunotoxicology strategies for pharmaceutical safety assessment. Wiley, Hoboken, NJ, pp 163–177

Centers for Disease Control and Prevention (CDC) (2004) Tuberculosis associated with blocking agents against tumor necrosis factor-alpha – California, 2002–2003. Morb Mortal Wkly Rep 53:683–686

Cunha BA (1995) Pulmonary tuberculosis and steroids. Chest 107:1486–1487
Cunningham AJ (1965) A method of increased sensitivity for detecting single antibody-forming cells. Nature 207:1106–1107
De Waal EJ, Timmerman HH, Dortant PM, Kranjc MAM, Van Loveren H (1995) Investigation of a screening battery for immunotoxicity of pharmaceuticals within a 28-day oral toxicity study using azathioprine and cyclosporin A as model compounds. Regul Toxicol Pharmacol 21:327–338
De Waal EJ, De Jong WH, Van der Vleit H, Verlaan B, Van Loveren H (1996) An immunotoxicity screening study on salmeterol in rats. Int J Immunopharmacol 18:523–528
De Waal EJ, Van Loveren H, Vos JG (1997) Practice of tiered testing for immunosuppression in rodents. Drug Info J 31:1317–1323
De Waal EJ, Van der Laan J-W, Van Loveren H (1998) Effects of prolonged exposure to morphine and methadone on in vivo parameters of immune function in rats. Toxicology 129:201–210
Dean JH, Luster MI, Boorman GA, Lauer LD (1982) Procedures available to examine the immunotoxicity of chemicals and drugs. Pharmacol Rev 34:137–148
Dixon WG, Hyrich KL, Watson KD, Lunt M, Galloway J, Ustianowski A, Symmons DPM (2010) Drug-specific risk of tuberculosis in patients with rheumatoid arthritis treated with anti-TNF therapy: results from the British Society for Rheumatology Biologics Register (BSRBR). Ann Rheum Dis 69:522–528
Draize JH, Woodward G, Calvery HO (1944) Methods for the study of irritation and toxicity of substances applied topically to the skin and mucous membranes. J Pharm Exp Ther 61:377–390
European Agency for the Evaluation of Medical Products (2000) Committee for Proprietary Medicinal Products. Note for guidance on repeated dose toxicity. European Agency for the Evaluation of Medical Products, London
Eyers S, Weatherall M, Jefferies S, Beasley R (2011) Paracetamol in pregnancy and the risk of wheezing in offspring: a systematic review and meta-analysis. Clin Exp Allerg 41:482–489
Feinberg SM, Feinberg AR, Moran CF (1953) Penicillin anaphylaxis, nonfatal and fatal reactions. J Am Med Assoc 152:114–119
Forni L, Coutinho A, Köhler G, Jerne NK (1980) IgM antibodies induce the production of antibodies of the same specificity. Proc Natl Acad Sci USA 77:1125–1128
Gelinck LBS, Van der Bijl AE, Visser LG, Huizinga TWJ, Van Hogezand RA, Rijkers GT, Kroon FP (2008) Synergistic immunosuppressive effect of anti-TNF combined with methotrexate on antibody responses to the 23 valent pneumococcal polysaccharide vaccine. Vaccine 26:3528–3533
Gerberick FG, Ryan CA, Kern PS, Schlatter H, Dearman RJ, Kimber I, Patlewicz GY, Basketter DA (2005) Compilation of historical local lymph node data for evaluation of skin sensitization alternative methods. Dermatitis 16:157–202
Gourley I, Descotes J (2008) Bridging immunotoxicology to clinical drug development. In: Herzyk DJ, Bussiere JL (eds) Immunotoxicology strategies for pharmaceutical safety assessment. Wiley, Hoboken, NJ, pp 375–384
Grainger DJ (2004) Immunomics: principles and practice. IRTL Rev 2:1–6
Grandjean P, Andersen EW, Budtz-Jørgensen E, Nielsen F, Mølbak K, Weihe P, Heilmann C (2012) Serum vaccine antibody concentrations in children exposed to perfluorinated compounds. J Am Med Assoc 307:391–397
Gulmez SE, Holm A, Frederiksen H, Jensen TG, Pedersen C, Hallas J (2007) Use of proton pump inhibitors and the risk of community-acquired pneumonia. Arch Intern Med 167:950–955
Haley P, Perry R, Ennulat D, Frame S, Johnson C, Lapointe J-M, Nyska A, Snyder PW, Walker D, Walter G (2005) STP position paper: best practices guideline for the routine pathology evaluation of the immune system. Toxicol Pathol 33:404–407
Hastings KL (1996) Regulatory issues of importance to developing anti-HIV therapeutics. Toxicol Pathol 24:278–280
Hastings KL (2002) Implications of the new FDA/CDER immunotoxicology guidance for drugs. Int Immunopharmacol 2:1613–1618

Hinton DM (2000) US FDA "Redbook II" immunotoxicity testing guidelines and research in immunotoxicity evaluations of food chemicals and new food proteins. Toxicol Pathol 28:467–478

Holsapple MP, Burns-Naas LA, Hastings KL, Ladics GS, Lavin AL, Makris SL, Yang Y, Luster MI (2005) A proposed testing framework for developmental immunotoxicology. Toxicol Sci 83:18–24

Horvath CJ, Milton MN (2009) The TeGenero incident and the Duff report conclusions: a series of unfortunate events or an avoidable event? Toxicol Pathol 37:372–383

House RV, Luebke RW (2007) Immunotoxicology: thirty years and counting. In: Luebke R, House R, Kimber I (eds) Immunotoxicology and immunopharmacology. CRC, Boca Raton, FL, pp 3–20

Irey NS (1978) Drug adverse reaction reports related to immunotoxicity. In: Asher IM (ed) Inadvertent modification of the immune response: the effects of foods, drugs, and environmental contaminants. US Food and Drug Administration, Washington, DC, pp 140–143

Jerne NK, Nordin AA (1963) Plaque formation in agar by single antibody-producing cells. Science 140:405

Kuper CF, Harleman JH, Richter-Reichelm HB, Vos JG (2000) Histopathologic approaches to detect changes indicative of immunotoxicity. Toxicol Pathol 28:454–466

Landsteiner K (1945) The specificity of serological reactions. Harvard University Press, Cambridge, MA

Linsky A, Gupta K, Lawler EV, Fonda JR, Hermos JA (2010) Proton pump inhibitors and risk for recurrent *Clostridium difficile* infection. Arch Intern Med 170:772–778

Luebke RW, Holsapple MP, Ladics GS, Luster MI, Selgrade MJ, Smialowicz RJ, Woolhiser MR, Germolec DR (2006) Immunotoxicogenomics: the potential of genomics technology in the immunotoxicity risk assessment process. Toxicol Sci 94:22–27

Luster MI, Munson AE, Thomas PT, Holsapple MP, Fenters JD, White KL Jr, Lauer LD, Germolec DR, Rosenthal GJ, Dean JH (1988) Development of a testing battery to assess chemical-induced immunotoxicity: National Toxicology Program's guidelines for immunotoxicity evaluation in mice. Fundam Appl Toxicol 10:2–19

Luster MI, Rosenthal GJ, Cao W, Thompson MB, Munson AE, Prejean JD, Shopp G, Fuchs BA, Germolec DR, Tomaszewski JE (1991) Experimental studies of the hematologic and immune system toxicity of nucleoside derivatives used against HIV infection. Int J Immunopharmacol 13(S1):99–107

Luster MI, Portier C, Pait DG, White KL Jr, Gennings C, Munson AE, Rosenthal GJ (1992) Risk assessment in immunotoxicology I: sensitivity and predictability of immune tests. Fundam Appl Toxicol 18:200–210

Luster MI, Portier C, Pait DG, Rosenthal GJ, Germolec DR, Corsini E, Blaylock BL, Pollock P, Kouchi Y, Craig W, White DL, Munson AE, Comment CE (1993) Risk assessment in immunotoxicology II: relationships between immune and host resistance tests. Fundam Appl Toxicol 21:71–82

Maronpot RR (2006) A monograph of histomorphologic evaluation of lymphoid organs. Toxicol Pathol 34:409–696

Organisation of Economic Cooperation and Development (1995) Guideline for testing chemicals. Method No. 407: Repeated dose oral toxicity – rodents: 28-day or 14-day study. OECD, Paris (Updated 2008)

Ovary Z (1958) Immediate reactions in the skin of experimental animals provoked by antibody-antigen interaction. Prog Allerg 5:459–508

Pharmaceutical Manufacturers Association (1988) Immunotoxicology Task Force, Report to the Drug Safety Subsection Steering Committee on Immunotoxicology and the Pharmaceutical Industry. PMA, Washington, DC

Phillips KE, McCay JA, Brown RD, Musgrove DL, Meade BJ, Butterworth LF, Wilson S, White KL Jr, Munson AE (1997) Immunotoxicity of 2′,3′-dideoxyinosine in female B6C3F1 mice. Drug Chem Toxicol 20:189–228

Piccotti JR, Lebrec HN, Evans E, Herzyk DJ, Hastings KL, Burns-Naas LA, Gourley IS, Wierda D, Kawabata TT (2009) Summary of a workshop on nonclinical and clinical immunotoxicity assessment of immunomodulatory drugs. J Immunotoxicol 6:1–10

Pieters R (2008) Systemic hypersensitivity. In: Herzyk DJ, Bussiere JL (eds) Immunotoxicology strategies for pharmaceutical safety assessment. Wiley, Hoboken, NJ, pp 241–256

Portier P, Richet C (1902) De l'action anaphylactique de certain venins. C R Soc Biol 54:170–172

Putman E, Van Loveren H, Bode G, Dean J, Hastings K, Nakamura K, Verdier F, Van der Laan J-W (2002) Assessment of immunotoxic potential of human pharmaceuticals, a workshop report. Drug Info J 36:417–427

Putman E, Van der Laan J-W, Van Loveren H (2003) Assessing immunotoxicity: guidelines. Fundam Clin Pharmacol 17:615–626

Rock KL, Latz E, Ontiveros F, Kono H (2010) The sterile inflammatory response. Annu Rev Immunol 28:321–342

Shubin H, Lambert RE, Heiken CA, Sokmensuer A, Glaskin A (1959) Steroid therapy and tuberculosis. J Am Med Assoc 170:1885–1890

Stebbings R, Findlay L, Edwards C, Eastwood D, Bird C, North D, Mistry Y, Dilger P, Liefooghe E, Cludts I, Fox B, Tarrant G, Robinson J, Meager T, Dolman C, Thorpe SJ, Bristow A, Wadhwa M, Thorpe R, Poole S (2007) "Cytokine storm" in the phase I trial of monoclonal antibody TGN1412: better understanding the causes to improve preclinical testing of immunotherapeutics. J Immunol 179:3325–3331

Taylor LD, Binienda Z, Schmued L, Slikker W Jr (1994) The effect of dideoxycytidine on lymphocyte subpopulations in nonhuman primates. Fundam Appl Toxicol 23:434–438

Thaxton JP, Tung HT, Hamilton PB (1974) Immunosuppression in chickens by aflatoxin. Poult Sci 53:721–725

Udaka K (1992) Cellular and humoral mechanisms of immunotoxicological tissue manifestations induced by immunotoxic drugs. Toxicol Lett 64/65:93–100

US Environmental Protection Agency (1998) Health effects test guidelines: OPPTS 870.7800 Immunotoxicity (revised 2006). US EPA, Washington, DC

US Food and Drug Administration (1993) Points to consider in the immunotoxicology evaluation of new drugs intended for the treatment of HIV infection and related disorders. Division of Antiviral Drug Products, Center for Drug Evaluation and Research, US FDA, Washington, DC

US Food and Drug Administration (2002) Guidance for industry: immunotoxicology evaluation of investigational new drugs. US FDA, Washington, DC

Van der Laan J-W, Krajnc EI, Krajnc-Franken MAM, Van Loveren H (1995) Immunotoxicological screening of morphine and methadone in an extended 28 day study in rats. Int J Immunopharmacol 17:535–543

Van der Laan J-W, Timmerman H, Van Loveren H (1996) Comparison of the in vivo effects of morphine and methadone on natural killer cell activity in spleen, peritoneal cavity, and lungs in rats. Int J Immunopharmacol 18:401–407

Van der Laan J-W, Van Loveren H, Vos JG, Dean JH, Hastings K (1997) Immunotoxicity of pharmaceuticals: current knowledge, testing strategies, risk evaluation & consequences for human health. Drug Info J 31:1301–1306

Vos JG (1977) Immune suppression as related to toxicology. CRC Crit R Toxicol 5:67–101

Vos JG, Van Loveren H (1998) Experimental studies on immunosuppression: how do they predict for man? Toxicology 129:13–26

Weaver JL, Tsutsui N, Hisada S, Vidal J-M, Spanhaak S, Sawada J, Hastings KL, Van der Laan J-W, Van Loveren H, Kawabata T, Sims J, Durham SK, Fueki O, Matula TI, Kusunoki H, Ulrich P, Nakamura K (2005) Development of the ICH guidelines for immunotoxicology evaluation of pharmaceuticals using a survey of industry practices. J Immunotoxicol 2:171–180

Yang L, Moses HL (2008) Transforming growth factor β: tumor suppressor or promoter? Are host immune cells the answer? Cancer Res 68:9107–9111