Medical Imaging and Radiotherapy Research: Skills and Strategies

Aarthi Ramlaul
Editor

Second Edition

Springer

Medical Imaging and Radiotherapy Research: Skills and Strategies

Aarthi Ramlaul
Editor

Medical Imaging and Radiotherapy Research: Skills and Strategies

Second Edition

Springer

Editor
Aarthi Ramlaul
Diagnostic Radiography and Imaging
School of Health and Social Work
University of Hertfordshire
Hatfield
Hertfordshire
UK

ISBN 978-3-030-37943-8 ISBN 978-3-030-37944-5 (eBook)
https://doi.org/10.1007/978-3-030-37944-5

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

Foreword

For diagnostic imaging and radiotherapy procedures, research remains pivotal as it provides evidence upon which patient care and management can be based. Its values are wide ranging and include risk minimisation, procedure selection and optimisation, and how patients might be cared for before, during and after a procedure.

Radiography, both diagnostic and therapy, is fairly unique in the demands they impose on practitioners as the skills and knowledge needed extend from technical-orientated through to human-related, where the latter can focus on complex emotional needs. With research in mind this requires an understanding of different paradigms, including qualitative and quantitative, together with an understanding of how research might be conducted in the basic sciences, such as physics, through to the human sciences, such as psychology and sociology. As part of an educational experience, it is reasonable that undergraduate and postgraduate students in our field are exposed to a wide range of research approaches; however, in practice they are likely to concentrate on just one (e.g. physics/quantitative).

Acquiring a basic understanding of the range of research paradigms, disciplines and techniques a radiographer needs in undergraduate and postgraduate education is challenging, given the diversity. One approach would be to read widely across a range of textbooks. In doing so a good understanding can be achieved; however, the time to do this could be excessive. An alternative approach is to select a book which contains an introduction to the range of research-related matters that are relevant to a radiographer. This book fills the gap nicely, as it serves as a valuable introductory text that can be used by itself, or as a basis for further reading. As well as addressing highly specific research-related matters such as dosimetry, this book also considers more generic matters like literature critique, research funding, ethics and how to disseminate research findings through conference and journal papers. Importantly, audit and service evaluation are included too. These are essential because they help us understand the values that research findings have had after translation into practice. This book is holistic as it considers research from 'research question development' through to 'evaluation of research outcomes in practice' and I recommend it highly to you.

I would like to end the Foreword with a personal reflection about research and a word of encouragement about being involved with it, as I know learning about and doing research can be very demanding. I qualified as a diagnostic radiographer in June 1984 and in November that same year I became involved with research. Since

then it has remained a key element of what I do. I have found being involved in research and conducting it to be a highly rewarding experience. Intrinsically, it is an enjoyable process from methods design to dissemination of findings; extrinsically it is extremely rewarding, as I have seen much of my research go on to impact positively on patients across the world through practice change. I appreciate that doing research is not for everybody but helping those who do it and valuing and using research outcomes are of immense importance because without such activity research may not get done or be translated into practice.

University of Salford, Salford, UK Peter Hogg

Preface

Radiography is a dynamic field which is rapidly advancing. Research has played a crucial role in enabling this advancement. Although challenging and often demanding, research in radiography is essential and an area that we must engage in.

As a profession we have increased the amount of research within radiography and have prioritised research activities as fundamental to changing practice. Research-related activities contribute to enhancing professional practice and providing quality care that is uncompromised and benefits the patient, who is at the heart of all considerations.

The Health and Care Professions Council and the Society and College of Radiographers require inter-alia practitioners to use evidence from research to inform their everyday practice. Higher education institutions are simultaneously providing training that encourages an interest in research activities amongst their students.

This text is directed primarily towards those undertaking research studies in radiography for the first time, i.e., undergraduate and postgraduate students. However, the information contained herein would also benefit qualified medical imaging and radiotherapy practitioners who are undertaking research studies for the first time, as well as nurses and allied health professionals, who have research interests in aspects of radiography.

Students can feel overwhelmed by the number of textbooks on their reading lists, in particular, those dealing with research that are not directed specifically to meeting their needs as student radiographers. The tools provided within this text will enable the student to develop the skills needed to undertake research in a methodical and reliable manner. The research process is dynamic. Often the direction the study takes is dictated by constraints experienced due to the nature of the study. You will find practical examples of the recommended steps to follow; use these as guides in finding your own balance with your study.

This text starts off with background information on the history of research and its context within radiography. Generic aspects of the research process, from literature searching and information management to research ethics, are described. Although generic in concept, the context within which this is set relates to medical imaging and radiotherapy practice. It is intended that this style of presenting these concepts will aid your understanding.

The main types of information or data gathered from research studies are explained and tools for collecting, testing and interpreting data are provided.

The chapter on health technology assessment is a vital tool in our profession today as it enables high-quality information on the costs and implications for use of the various technologies to be assessed. By evaluating technologies, we are using the evidence to improve patient outcomes and make efficient and effective use of healthcare resources.

The inclusion of chapters on research outcome measures, reflective practice (including reflexivity), dosimetry and clinical audit provides the reader with a comprehensive package of core requirements and aspects of our profession.

Novice researchers should benefit from the chapters giving guidance on structuring, writing, publishing and presenting research findings. Writing up assignments seems to be a somewhat daunting experience for students, whether it is a simple 1000-word essay or a 10,000-word dissertation. These chapters provide you with straightforward guidance on how to structure and present your writing. Using these tools would enable you to become a confident writer and publisher.

For those considering applications for research grants, there is a chapter on applying for research funding which provides guidance on funding streams and importantly, patient and public involvement in research. Lastly, good practice tips to bear in mind throughout the research process as well as guidance on pitfalls to avoid are provided. This chapter has been written particularly with the undergraduate and postgraduate student in mind; however, it is aimed equally at all practitioners who require guidance in developing those skilled techniques.

A glossary of terms has been provided to clarify terminology used within this text and where relevant, the authors have suggested further reading material which you will find useful.

You will notice that links have been drawn between chapters and there is some repetition of information. Repetition in this context is healthy as it reinforces understanding and fosters a deep approach to learning.

The unique way that concepts were related to examples from medical imaging and radiotherapy practice was a fundamental strength of the original text. We have used feedback to constructively enhance this attribute within this new edition. Although practices within medical imaging and radiotherapy disciplines incorporate somewhat different interaction frameworks between practitioner and patient, care has been taken to present examples from practice that would enable students from both disciplines to connect with the concepts discussed.

Where reference to ‘radiography’ is made, this is generically applicable to both diagnostic and therapeutic practice. Where reference to ‘radiographer’ or ‘practitioner’ is made, this is applicable to radiographers in diagnostic and therapeutic practice, radiologic technologists or allied health professionals.

The aim of this new edition remains to provide the necessary guidance, support and direction for the novice radiography researcher. In redeveloping this textbook from its original publication in 2010, this text has retained its structure, but the content has been updated with current developments from practice and the profession as a whole.

It is my pleasure, once again, to offer this new edition to enhance and maintain research within our valuable profession. I hope that the experiences you are about to gain from your research studies encourage and motivate you to continue the process of enquiry throughout your careers.

Hertfordshire, UK Aarthi Ramlaul
March 2020

Acknowledgements

Firstly, I wish to thank the multiple cohorts of radiography students over the years at the University of Hertfordshire for their valued suggestions on the content of this textbook.

I am extremely grateful to many of the original authors for their contribution to this new edition.

I would like to thank the authors who have contributed to the previous edition but who are not part of this new edition. Your contribution remains valued.

I would like to acknowledge the support of Ms Leonie Munro, my undergraduate radiography tutor and mentor, for her considerable contribution in proofreading the entire manuscript.

I wish to thank my family for being my pillar of support during the redevelopment of this textbook amongst other endeavours.

Contents

Abbreviations

A&E	Accident and emergency
AAPM	American Association of Physicists in Medicine
ABPI	The Association of the British Pharmaceutical Industry
AD	Anno Domini
ADL	Activity of daily living
AI	Artificial intelligence
ALARA	As low as reasonably achievable
ALARP	As low as reasonably practicable
ANOVA	Analysis of variance
AP	Anteroposterior
APA	American Psychological Association
APBI	Accelerated partial breast irradiation
ARR	Absolute risk reduction
ARSAC	Administration of Radioactive Substances Advisory Committee (UK)
AUC	Area under the curve
BC	Before Christ
BDA	British Dietetics Association
BMA	British Medical Association
BMI	Body mass index
BPA	British Paediatric Association
BSc	Bachelor of Science (degree)
CCTV	Closed-circuit television
CDSR	The Cochrane Database of Systematic Reviews
CENTRAL	The Cochrane Central Register of Controlled Trials
CER	Control event rate
CHART	Continuous hyperfractionated accelerated radiotherapy
CHI	Commission for Health Improvement
CI	Confidence interval
CINAHL	Cumulative Index of Nursing and Allied Health Literature
CIT	Critical Incident Technique
CLD	Central lung depth
CMR	Cochrane Methodology Register

COIN	Royal College of Radiologists Clinical Oncology Information Network
COMET	Core outcome measures in effectiveness trials
COREQ	Consolidated criteria for reporting qualitative research
CoRIPS	College of Radiographers Industry Partnership Scheme
COS	Core outcome set
COSMIN	Consensus-based standards for the selection of health measurement instruments
CPD	Continuing professional development
CQC	Care Quality Commission
CQUIN	Commissioning for Quality and Innovation
CR	Computed radiography
CT	Computed tomography
CTDI	Computed tomography dose index
DALY	Disability-adjusted life year
DAP	Dose area product
DARE	The Database of Abstracts of Reviews of Effectiveness
DEXA	Dual energy X-ray absorptiometry
DMC	Data monitoring committee
DMPA	Depot medroxyprogesterone acetate
DoH	Department of Health
DR	Digital radiography
EBP	Evidence-based practice
ECR	European Congress of Radiology
EER	Experimental event rate
EORTC	European Organisation for Research and Treatment of Cancer
EPID	Electronic Portal Imaging Device
EPSRC	Engineering and Physical Sciences Research Council
ESD	Entrance Surface Dose
ESTRO	European Society for Therapeutic Radiology and Oncology
EU	European Union
FN	False negative
FP	False positive
FSD	Focus to skin distance
GCP	Good clinical practice
GDPR	General Data Protection Regulation
GP	General practitioner
Gy	Gray
HCC	Health Care Commission
HCPC	Health and Care Professions Council
HDL	High-density lipoprotein
HEI	Higher education institution
HQIP	Healthcare Quality Improvement Partnership
HRA	Health Research Authority
HRQoL	Health-related quality of life

HRT	Hormone replacement therapy
HTA	Health technology assessment
ICH	International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use
ICRP	International Cancer Research Partnership
IGRT	Image-guided radiotherapy
IMRT	Intensity-modulated radiotherapy
IP	Intellectual property
IPEM	Institute of Physics in Engineering in Medicine
IR(ME)R	Ionising Radiation (Medical Exposure) Regulations
IRAS	Integrated Research Application System
IRR	Ionising Radiation Regulations
ISO	International Organization for Standardisation
ISR	Independent scientific review
ITT	Intention to treat
IVE	Immersive visualization environment
J	Joules
KSF	Knowledge and Skills Framework
kVp	Kilovoltage peak
LINAC	Linear accelerator
LSR	Least squares regression
mAs	Milliampere seconds
MCID	Minimally Clinically Important Difference
MeSH	Medical subject headings
MLC	Multileaf collimator
MRA	Magnetic resonance angiography
MRC	Medical Research Council
MRI	Magnetic resonance imaging
MSc	Master of Science (degree)
mSv	millisievert
MU	Monitor unit
NCAPOP	National Clinical Audit and Patient Outcome Programme
NHS EED	NHS Economic Evaluation Database
NHS	National Health Service
NI	Northern Ireland
NICE	National Institute for Health and Clinical Excellence
NIH	National Institutes of Health
NIHR	National Institute for Health Research
NNT	Number needed to treat
NPSA	National Patient Safety Agency
NRES	National Research Ethics Service
NSF	National Service Framework
OBI	On-board imaging
OMIs	Outcome measurement instruments
OpenGrey	The System for Information on Grey Literature in Europe

OR	Odds ratio
PA	Posteroanterior
PACS	Picture archiving and communication system
PDF	Portable document format
PhD	Doctor of Philosophy (degree)
PICO	Population/participant, intervention, comparison, outcome
PIS	Participant information sheet
PPI	Patient and public involvement
PREMs	Patient-reported experience measures
PRISMA	Preferred Reporting Items for Systematic Reviews and Meta-Analyses
PROMs	Patient-reported outcome measures
PROSPERO	Prospective register of systematic reviews
PTV	Planning target volume
QA	Quality assurance
QALYs	Quality-adjusted life years
QART	Quality assurance in radiotherapy
QC	Quality control
QMS	Quality management system
QoL	Quality of life
QUADAS	Quality assessment for diagnostic accuracy studies
QUADAS-2	A revised tool for the Quality Assessment of Diagnostic Accuracy Studies
QUOROM	Quality of reporting of meta-analysis statement
QUS	Quantitative ultrasound
R and D	Research and development
RBE	Relative biological effectiveness
RCT	Randomized controlled trial
RD	Risk difference
RDS	Research and Development Service
REC	Research Ethics Committee
RIS	Radiological information system
RISRAS	Radiation-Induced Skin Reaction Assessment Scale
ROBINS-I	Risk Of Bias In Non-Randomised Studies-of Interventions
ROC	Receiver operating characteristic
RR	Relative risk or risk ratio
RSNA	Radiological Society of North America
RTOG	Radiation Therapy Oncology Group
RTTQA	Radiotherapy trials quality assurance
SCoR	Society and College of Radiographers
SHA	Strategic Health Authority
SIGN	Scotland Intercollegiate Guidelines Network
SPSS	Statistical Package for Social Sciences
STARD	Standards for Reporting Diagnostic accuracy studies

STROBE	Strengthening the reporting of observational studies in epidemiology
Sv	Sievert
TLD	Thermoluminescent dosimeter
TN	True negative
TP	True positive
UICC TNM	International Union Against Cancer Tumour Node Metastases
UKIO	United Kingdom Imaging and Oncology
US	Ultrasound
USA	United States of America
VAS	Visual analogue scale
VAT	Value added tax
WCC	White cell count
WHO	World Health Organization

Part I

Getting Started with Research

History of Research

1

Julie Hall

1.1 Introduction

Research is a planned, systematic method of scientific enquiry, which adds to existing knowledge, by providing evidence from the acquisition and analysis of data. It can lead almost anywhere you choose. Research in healthcare provides evidence that can be used to justify new practice or to challenge existing practice. Occasionally, students of radiography question the inclusion of research and statistics material in their courses. However, practitioners rarely, if ever, dispute their responsibility to maintain their skills, and develop clinical practice, in line with an ever-changing evidence base. Research provides evidence on which to base practice. Therefore, it is clear that an understanding of the research process and data interpretation is required to ensure that practice is based upon reliable evidence obtained from good quality research. The origins of research, including a historical background on the application of its theories to current and evidence-based practice, are explored in this chapter. Also considered are some recent drivers of research in radiography and why research is an important part of the role of radiographers whatever their scope of practice.

1.2 Research and Radiography

The role of a radiographer may vary around the world, with a scope of practice ranging from practitioner to autonomous consultant practitioner, but the responsibility for continued professional development, and lifelong learning to ensure optimum professional practice, remains constant. As specialist medical imaging and radiotherapy

J. Hall (✉)
Department of Radiography, Birmingham City University, Birmingham, UK
e-mail: Julie.hall@bcu.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_1

roles become the norm, practitioners find themselves involved in research on a more regular basis; this involvement can be direct or indirect and spans a wide spectrum of clinical and academic activity. Practitioners in more advanced roles are required to carry out their own research and promote a professional culture in which research is an integral component [1]. The emergence of consultant practitioners, in particular, places an emphasis on the integration of research into practice for education, innovation, and development. Both advanced and consultant practitioners use research skills to solve complex problems and transfer knowledge within the wider multidisciplinary environment to benefit service users.

All practitioners therefore require knowledge of research skills for the following purposes.

- To participate in studies contributing to the ongoing development of their profession and practice
- To address specific problems that may arise in clinical practice
- To evaluate evidence from the research of other practitioners within radiography and the wider multidisciplinary team

It is important to consider how our underpinning knowledge base is formed and the extent to which we are open to new information. This consideration leads directly to the fundamental question: What is knowledge? As practitioners involved in research, it is important that we understand the philosophy of knowledge and its emerging paradigms because we need to apply these concepts to our daily practice.

1.3 What Is Knowledge?

Epistemology is the name given to the study of the nature of knowledge; it is essentially a philosophical issue. We not only use information to survive, but also speculate on the nature of that information and our place in relation to it. This is far from being a new preoccupation. Ancient Greece is traditionally identified as the home of the first philosophers. For example, we can recognise a modern outlook emerging in Aristotle's (384–322 BC) close observations of the natural world, and his application of logic in an attempt to explain what he saw. It is then not a difficult step to recognise that it is possible, and probably necessary, to distinguish between the world as it is and the world as we perceive it [2]. Knowledge therefore poses certain problems which research must acknowledge in its attempts to separate so-called fact from belief.

Moving forward almost 2000 years, this problem was taken up by an English philosopher, Francis Bacon (1561–1626). Bacon believed in the need for a new learning, free from the 'idols' of superstition, prejudice, and the preconceptions of the human mind [3]. He was a strong admirer of Aristotle; he however differed in his insistence that observations should drive the logical process rather than vice versa. Thus, an inductive process of building up a logical structure rooted in

observation was established as being a more reliable method than starting with the logic and then applying it to experience in a deductive, top-down fashion. This philosophy gave rise to a systematic method of enquiry which, for the first time, could be termed 'scientific'.

It would probably be contentious to try and fix a date when the scientific process truly came of age. Certainly Isaac Newton set a dramatic and definitive new standard in bringing together observations of the natural world with the theoretical model purporting to explain it. The eighteenth century in Europe, dubbed 'the Enlightenment', recognised this and was characterised by an insistence that belief and observation should be mutually consistent. Its philosophers prized intellectual progress and perceived this as a measure of the advance of reason over superstition [4]. Indeed, the scientific method has been so successful that it can be argued (and often is) that it has become a dogmatic system in its own right. The strength of the challenge it offers to existing dogmas was clearly seen in the reaction provoked by Darwin's publication of *On the origin of species* in 1859. Interestingly, well over a century later there is still fierce debate, between those who believe in creationism and intelligent design and those who believe in the scientific theory of evolution through natural selection.

Clearly, the definition of knowledge and its relation to belief are not separable from social pressures, and so in looking at the pursuit of knowledge it is necessary to be transparent in taking these pressures into consideration.

1.4 Social Context of Research: Paradigms and the Pursuit of Knowledge

Once a topic of enquiry has been conceived, an appropriate method of investigation has to be applied to it. This then requires researchers to consider the beliefs and assumptions they may already hold which could limit or distort their approach. American physicist and philosopher Thomas Kuhn (1922–1996) [5] identified that research will inevitably take place within a dominant paradigm: an overarching theoretical context or set of expectations which is socially agreed, perhaps unconscious in its effect, and rooted in culture and history. For this reason, the most significant advances in knowledge and understanding are experienced as revolutionary: the so-called paradigm shifts a term coined by Kuhn. This view was strongly endorsed by Paul Feyerabend (1924–1994) [6], who argued, citing Galileo's difficulties with the Catholic Church in the seventeenth century by way of illustration, that 'science is essentially an anarchic enterprise' and that new insights are likely to meet strong resistance.

Paradigms can provide a structure that addresses the contextual and anarchic issues associated with research to support the generation of knowledge. Researchers are more likely to produce a credible outcome if they are aware that they are working within a particular paradigm. Simply put, any method of enquiry must be consistent with the nature of the research question being addressed, and both are likely to be derived within a particular paradigm. The three most common paradigms can be described as positivism, interpretivism, and critical theoretical.

1.5 Positivism

The scientific version of knowledge has become increasingly dominant in modern times; this dominance is associated with a belief that the information that science yields is true and reliable. This popular belief is drawn from a particular paradigm that is defined as positivistic.

The assertion that a belief must be testable in observed experience is clearly a powerful driving force, and the fact that this is just an assertion can easily be forgotten. The positivist outlook tends to support the assumption that an objective and measurable reality exists and simply requires a researcher to devise a way of measuring this reality as accurately as possible. Practical experimentation and observation have been pursued with increasingly subtle ingenuity and investment in advanced technology, and society can see the material fruits of this in everyday life. We can split the atom, put a man on the moon, and decode genetic structures. Clearly, in a practical sense, this sort of science works. So long as there is a quantity to be measured and an objective observer to measure it, positivists believe that eventually the truth will be revealed. From a philosophical point of view, this is something of an oversimplification, but it does provide methodologies by which certain sorts of theory can be tested. This is because the positivistic notion of a separate reality allows for the manipulation and control of that reality with no consequent loss of validity. Thus, an experimenter can manipulate an independent variable and control confounding variables, and be reasonably confident that a dependent variable will yield a reliable and valid outcome.

For example, using a radiographic phantom it is possible to vary the kilovoltage peak (kVp) for a given exposure and measure the consequent effect on image contrast, resolution, and density. The expectation is that a researcher has control over all identifiable variables, and in particular can isolate and measure the effect of varying a specific parameter.

This method forms the basis of randomised controlled clinical trials where, typically, a sample of volunteers is randomly allocated to one of two groups: one receives a placebo drug or treatment and the other receives a new version, with all other factors being the same for both groups. An attempt is thus made to eliminate systematic bias and minimise chance variability in the expectation that any subsequent difference experienced by the two groups will arise as a consequence of the new intervention. The logic is sound, the outcome is measurable, and the system works pretty well. However, there are many other situations in healthcare that also require firm evidence but are not amenable to this sort of approach. Therefore, it is often necessary to recognise the limits of the positivist outlook and adopt a more appropriate paradigm.

1.6 Interpretivism

People and their circumstances are not easy to control or measure. It may be a simple matter to check an individual's blood pressure or record their weight, but this tells us nothing about what they are thinking or feeling. Therefore, it must be

recognised that, in certain areas, there is a problem of measurement and equally so of the role of the person doing the measuring. The appropriate paradigm for this sort of enquiry is defined as interpretivist and it differs from the positivist outlook in fundamental ways. The interpretivist paradigm works on the principle that reality is socially constructed; it emphasises subjectivity rather than objectivity and regards an observer as essentially inseparable from the phenomena under observation. It is more likely that the sort of data gathered by an interpretivist enquiry will reflect the quality of an experience rather than its quantity and will tend to be concerned with theory building rather than theory testing.

For example, in a radiotherapy setting a researcher may wish to interview patients to gain insight into their experience of the treatment process. Researchers would need to be alert to the extent to which their own expectations could influence the choice of questions put to the participants, and also to a possible similar bias in the subsequent analysis of the responses.

Historically science has fought against dogma to establish itself as a reliable source of information, but perhaps it was inevitable that at some point it would reach the limits of its applicability and risk becoming a dogma in its own right. An example of this is the progress of behaviourism as a psychological model. The model of behaviourism is based on the premise that the only observable phenomenon is outward behaviour, therefore making it impossible to comment directly on possible mental events. At the start of the twentieth century, the subject matter of psychology was consciousness, and the method of enquiry introspection, but by the latter half of the twentieth century psychology was largely given over to the behaviourist biological and operant conditioning model of learning. Psychology moved on, and the importance of consciousness and the inner experience was reasserted with the development of a variety of humanistic models and methodologies. So, for example, a positivistic view of learning as a measured change in behaviour can be compared with an interpretivist version: 'learning occurs when individuals … respond, or try to respond, meaningfully to what they experience and then seek to … integrate the outcomes into their own biographies' [7].

In order to capture the lived quality of an individual's experience, an interpretivist paradigm must be embraced. This brings with it the need for a methodology which can deal with subjectivity that is nevertheless rigorous and systematic and in that sense scientific and credible. Within the interpretivist paradigm, it is possible to identify several distinct approaches and these need to be briefly described.

1.7 Phenomenology and Hermeneutics

Typically, an interpretivist approach will involve recording someone's own account of something they have experienced. The problem is to do it in such a way that a person's words are captured and used to present a credible insight which is faithful to that experience. Phenomenology aims to achieve this. Edmund Husserl (1859–1932), usually regarded as the founder of phenomenology, believed that it was possible to delineate an individual's conscious experience by a process of 'bracketing'. This involves the deliberate attempt to identify and set aside a researcher's own

preconceptions, so that one is left with a complete yet unadorned description of the phenomenon in a respondent's own terms [8]. However, there may be difficulty in achieving the desired level of objectivity when immersed in essentially subjective material. Furthermore, it may be questioned whether such a description would be meaningful anyway, since a respondent's own terms are themselves a product of that individual's circumstances. This latter point rests at the heart of hermeneutics, a phenomenological approach developed by the German philosopher Martin Heidegger (1889–1976), in which bracketing is dismissed; a researcher aims to capture individual meaning through subjective dialogue with the material [9].

For example, diagnostic imaging and therapy practitioners both come into contact with people who present with serious illness and may wish to understand their patients' condition more thoroughly. In this example it would be appropriate to talk with willing participants and allow them to describe and discuss their personal experience in some depth. A suitable methodology here may be to conduct an extended interview, gathering as much spoken and non-verbal communication as possible, and then transcribe it faithfully. Researchers would need to immerse themselves in such material and try to make sense of it while setting aside their own biases and opinions.

This is clearly a far cry from the positivist approach which tends to ignore individuals and their social context.

1.8 Symbolic Interactionism, Grounded Theory, and Ethnomethodology

The impact of social context and the roles that we derive from it form the subject matter of symbolic interactionism. Here the sense of self is regarded as arising out of the interplay between members of a social group in which we communicate by means of words, gestures, and display. The clothes we wear, the words we choose, and the mannerisms we adopt all contribute to a social consensus within which our own identity is established with reference to other people. At the level of large groups of people or populations, this process is addressed through ethnomethodology, which focuses on socially agreed customs. Within the same sociological tradition, Glaser and Strauss [10] pioneered the approach known as grounded theory. It acknowledges that individuals constantly change, and so does research.

For example, a suitable application of grounded theory could be to explore student practitioners' experience of clinical placement. Students could be asked to maintain a journal while on placement in which they record their thoughts, feelings, and behaviour. A researcher could then look for themes in these written accounts, perhaps meet with the participants, and suggest a possible analysis of the main factors which the students themselves regarded as significant to their learning. Having developed such an analysis a researcher would need to meet again with these students to confirm the extent to which a researcher's version 'rang true'. In the light of the new participant response a researcher would need to revisit their explanatory model, iterating this consultative process until consensus is reached that the model is credible.

The method is inductive, aiming to build theory from the 'bottom up' using participants' reports, and revisiting those people to check that the result is in accordance with their experience. The process of data collection and analysis is therefore iterative and ongoing, with constant elaboration and refinement in an attempt to establish a consensus.

Thus, interpretivism seeks to understand the world, while positivism expects to predict it. The third paradigm is the critical theoretical, and it aims to change it.

1.9 Critical Theoretical

Both paradigms discussed so far incorporate an ethos that the process of research is in some way separable from the area being researched. The positivist approach takes this as axiomatic and interpretivism, although it addresses individuals' experiences within their social context, still proceeds as if that context is well defined. In contrast, the critical theoretical paradigm starts with the premise that not only is research embedded in its social context, it is actually part of it. Furthermore, because society itself is neither fixed nor well defined, the validity of the product of research is therefore called into question. Thus, research is faced with both a challenge of credibility and an opportunity to be an agent of change. Action research, for example, specifically sets out to evaluate and possibly recommend change in a system at the very same time that it is gathering data on the system. This requires a team approach and potentially offers emancipatory power to the participants, but it brings problems of its own, to do with a need for flexibility and a possible challenge to existing power relationships. In this respect, it is not difficult to see the same concerns at the heart of the standpoints on research of feminists and black people. The former approach points to the failure of traditional research to address topics of particular relevance to women and places women firmly in the role of researcher and women's issues at the focus of enquiry. Likewise, the latter approach is a response to the need for culturally sensitive and competent research with an emphasis on the impact of ethnicity and culture on life and life chances. These differing standpoints share the concern that in order to be meaningful, research must be transparent in recognising personal and societal agendas. In order to achieve this, a researcher must adopt a post-modern awareness of the complexity of how the world presents to us, and how we in turn choose to perceive it. This requires a researcher to look for the 'truth behind the truth' by deconstructing existing social terms and forms of representation.

For example, role extension provides a possible example within radiography of where the status quo might be questioned or even challenged by critical theoretical research. A practitioner's role can be defined on a spectrum ranging from protocol-driven, technical tasks to autonomous patient management at consultant level. The latter end of this spectrum particularly needs to be supported by a credible evidence base, and in acquiring such a base the issue of professional boundaries would need to be addressed. The terms, conditions, and scope of the research and a researcher's own agenda cannot now be regarded as separate from the underpinning research

aim. Existing power relationships will come into play and a researcher must recognise these, allow them to inform his or her work, and so deal with them.

Language is thus crucial to any line of enquiry. For example, discourse analysis can be used to investigate social and cultural structures. This is achieved by identifying patterns of thinking revealed by language rather than the words used; hence the context for discourse analysis is sociological and regards language as an active process that reflects meaning in society [11]. Van Dijk [12] drew attention to the multileveled nature of discourse, the strategies we employ in comprehending it, and the consequent encoding of social structures and power relationships in the very words that people use. Thus, all players in a research process have agendas and it is necessary to identify and declare these. Within the critical theorist paradigm, not only is knowledge provisional, but in the words of Habermas [13], there is a 'singular interlocking of knowledge and interest'.

Clearly the type of knowledge being sought and the methods used to seek it are interdependent, and Box 1.1 attempts to summarise this relationship.

The divisions in Box 1.1 do not necessarily indicate the order in which a researcher works. It is not wrong to start with a methodology or even a method. Often, we start with a question in mind, develop a method that seems appropriate, and only then appreciate how a paradigm can inform or constrain our research design.

It can be convenient to divide research into quantitative and qualitative approaches. The former are often associated with the positivistic paradigm; the latter are often similarly associated with the interpretivist paradigm. However, this can be an oversimplification and it may therefore be safer simply to use the terms 'quantitative' and 'qualitative' as descriptions of the methods that we use, the data that we collect, and how we analyse it. For example, consider a study designed to explore the feelings of a patient undergoing a diagnostic examination or a course of radiotherapy treatment. We may assume that the approach being taken is qualitative because we are trying to capture the nature of a patient's experience and the data collected would be in the form of words requiring interpretation. However, the distinction would become slightly blurred if our analysis then involved counting the number of times that a particular feeling was expressed, because these numbers would make our approach quantitative. Furthermore, this simplistic count could lose the context in which the feeling was experienced, considering the following the examples.

- I was anxious before my examination.
- I was expecting to be anxious before my examination.

By only counting the word anxious the meaning is lost. The important thing to get right is to choose a methodology which allows you to answer your research question.

In summary it is necessary to recognise that any particular piece of research will be limited in what it can achieve; limits are set by the world view or paradigm within which a researcher is operating. We have identified three different

paradigms, but whichever approach is adopted there is a common requirement that the process of enquiry itself should be rigorous and systematic, and it is to this that we turn next.

1.10 Secondary Research

So far the assumption has been that research is all about discovering, assessing, and comparing new data. Practice can also be informed by revisiting the research carried out by others. This can be done by combining a number of studies in order to answer a specific research question (or to summarise the findings) and usually takes the form of a systematic review. A systematic review involves the painstaking collection of all relevant studies, whether they have been published or not. A good quality systematic review applies the same rigour in the review of research evidence as should have been applied in the original production of that evidence and presents the collated evidence in an impartial and balanced way. Meta-analysis is used to combine statistical data from these combined studies in a meaningful way; it takes into account the relative sizes of the studies included in a systematic review. A reliable source of this type of research is the Cochrane Database of Systematic Reviews [14]. Cochrane reviews cover a wide range of subject areas. The database is easily searched. Two recent examples include (1) 'Antidepressants for the treatment of depression in people with cancer' (April 2018) and (2) 'Prostate MRI, with or without MRI-targeted biopsy, and systematic biopsy for detecting prostate cancer' (April 2019).

1.11 Secondary Data

Another approach to research is to use secondary data: data previously collected by someone else, possibly for some other purpose. However, care must be exercised when defining what constitutes secondary data. For example, if you compiled a new data set unique to your study from existing survey material you would be considered to be doing primary research, but if you used existing summary results or results compiled by other researchers this would be considered secondary data and so your research would also be considered as secondary.

1.12 Evidence-Based Medicine: A Systematic Approach to Knowledge

We have explored the proposition that knowledge does not arise in a vacuum. In fact, putting philosophy to one side, there is nowadays an expectation that research will lead to useful applications, not least in the field of healthcare. It is this expectation that underpins the practice of evidence-based medicine, whereby current clinical activity is constantly reviewed in the light of new research. It is accepted

therefore that knowledge is only ever provisional, and with that caveat in place a research process must be robust enough to offer a definitive version of the latest 'best practice'. This must start with a systematic and critical appraisal of what is known, what is not known, and, therefore what is needed to be known to make an evidence-based decision.

The essence of evidence-based practice is how evidence is used to inform the decision-making process to meet a clinical need. The aim is to achieve the best outcome for patients by applying professional knowledge to their particular circumstances; this 'knowledge' is based upon a critical understanding of the available research and its application. However, a decision cannot be made based on the results of an appraisal of primary and secondary research alone; professional experience and consensus will inform decision-making as will the wishes of a patient. This raises another important issue: the health information seeking behaviour of patients.

Access to research and related information is no longer restricted to healthcare professionals. Patients are more frequently carrying out their own 'research', often via the Internet or other digital platforms. This activity can be positive and improve the relationship between patient and practitioner [15]. Patients who participate in their own healthcare in this way have been termed e-patients and their number appears to be increasing. This development provides yet another incentive for practitioners to be research aware, even if they are not research active, as they will be required to justify their practice to a more informed patient group.

1.13 Conclusion

We have argued that research should be a systematic and rigorous process of collecting, analysing, and sharing data. This must be done in a way which transparently acknowledges its social context. We should recognise that the knowledge we acquire is likely to be influenced by our own interests—all the more reason to therefore derive a firm knowledge base using an appropriate methodology. We can build theories from qualitative interpretation or test them by making quantitative measurements, but the fundamental principle to observe in all cases is to develop a clear focus for a research question and allow this to inform our actions.

As practitioners we are professional people, and as such we have an obligation to maintain our clinical practice to the highest standards. We practice within the wider healthcare team and in a patient-centred manner. We should take the lead in developing our profession and continue to provide an interface, which makes the highly technical environment in which we practice accessible to other professions and the patients that we encounter. Development in healthcare proceeds on the basis of clinical evidence. The way to acquire this is through research.

Box 1.1 Summary of Research Terms

Epistemology

The study of the nature of knowledge or what is out there to know in the world around us.

Paradigms

The assumptions we make about the world which influence our expectations of what it is possible to know and how we go about knowing it.

Examples: positivism, interpretivism, critical theoretical.

Methodologies

These are the general approaches to research found within each paradigm.

For example, an experimental approach is an appropriate methodology within the positivist paradigm.

Methods

These are the particular ways of carrying out a given methodology.

Using the experiment as an example of a general approach, a particular way of conducting the experiment could be a randomised controlled trial.

References

1. SCoR. Education and professional development strategy: new directions. 2010.
2. Harré R. The philosophies of science. London: Oxford University Press; 1972.
3. Peltonen M, editor. The Cambridge companion to bacon. Cambridge: Cambridge University Press; 1996.
4. Cassirer E. The philosophy of the enlightenment. Princeton: Princeton University Press; 1968.
5. Kuhn TS. The structure of scientific revolutions. 2nd ed. Chicago: University of Chicago Press; 1970.
6. Feyerabend P. Against method. 3rd ed. London: New Left Books; 1975.
7. Jarvis P, Holford J, Griffin C. The theory and practice of learning. 2nd ed. London: Kogan Page; 2003.
8. Parahoo K. Nursing research: principles, process and issues. 3rd ed. London: Macmillan; 2014.
9. Miller S. Analysis of phenomenological data generated with children as research participants. Nurse Res. 2003;10:70–3.
10. Glaser B, Strauss AL. The discovery of grounded theory. London: Weidenfeld and Nicolson; 1967.
11. Jolley J. Introducing research and evidence-based practice for nursing and healthcare professionals. 2nd ed. London: Taylor & Francis; 2013.
12. Van Dijk TA, Kintsch W. Strategies of discourse comprehension. London: Academic; 1983.
13. Habermas J. Knowledge and human interests. London: Heinemann; 1972. p. 209.
14. Cochrane Database of Systematic Reviews. https://www.cochranelibrary.com/cdsr/reviews. Accessed 27 June 2019.
15. Swee-Lin Tan S, Goonawardene N. Internet health information seeking and the patient-physician relationship: a systematic review. J Med Internet Res. 2017;19(1):e9.

Finding a Research Question

2

Martin Vosper

2.1 Introduction

There are many opportunities for asking research questions in medical imaging and radiotherapy, ranging from great to small: all of them have potential value. The aim of this chapter is to take you through the various steps involved in finding a suitable research idea or topic and formulating it into a workable question to answer.

2.2 Research Steps

Four steps are discussed below.

2.2.1 Step One: Identifying a Question

Why is it important to have a question when starting out in medical imaging or radiotherapy research? Well, having a clear and specific question permits the following things:

- A question provides focus.
- It identifies a gap in existing knowledge, or a set of circumstances which need to be explained.
- It enables us to 'stay on track' and avoid irrelevant topic material.

M. Vosper (✉)
Department of Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: m.r.1.vosper@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_2

- It reminds us of the purpose of our research.
- It enables us to seek answers to our question and to reach conclusions.
- It helps others to decide whether our work is of interest to them.

The absence of an obvious or suitable research question is one of the reasons why research proposals are rejected by academic tutors, health research ethics committees, or funding bodies. Thus, it is a good idea to spend plenty of time developing a suitable research question.

There are some available frameworks which can help us to develop a research question. One of these uses the acronym FINER to describe a good research question.

- F: feasible—possible given the time and resources available
- I: interesting—to both the reader and the researcher
- N: novel—permitting some new data or perspective to be obtained
- E: ethical—abiding by ethical research principles
- R: relevant—having practical application to our topic area (medical imaging)

SMART is also a useful acronym and pertains to what a good research should be.

- S: specific—focused on a particular subject or matter
- M: measurable—able to produce data that can be expressed numerically or in qualitative terms
- A: achievable—possible given the time and resources available
- R: realistic—sensible, practical, based on 'real-world' considerations
- T: timely—current, topical, needed now

There are many reasons why people do research: perhaps to get a diploma or degree, to advance their careers, to benefit their patients within the healthcare system, or even to expand the boundaries of human knowledge. Their expectations will vary in line with these goals, from modest to ambitious. Very few of us will do truly revolutionary research in our lifetimes, such as splitting the atom, discovering X-rays, or developing the ultimate cure for cancer. Research questions are rarely completely ground-breaking and original. A question does not have to be 100% novel to be of value. There are always questions that others have asked already, but in different circumstances. For example, we might explore whether radiation doses to patients are different this year from the last, or the same at hospitals A and B. Surveys of stress are nothing new, but it could be that no one has ever done a stress survey of patients receiving a particular type of palliative radiotherapy or diagnostic X-ray examination at our workplace. To be worthwhile a research question does not have to be new: it does have to be worth asking.

To be worth asking, a research question should try to avoid producing what we could call the 'so what' response from other people. This consists of a mix of negative feelings that critics might voice when looking at our research question.

- ‘So what is the purpose or point of this research?’
- ‘So what new understanding or information will this research give us?’
- ‘So what benefit could this research bring to imaging or radiotherapy?’

Medical imaging and radiotherapy research is most likely to succeed when it has a clear purpose, is capable of increasing understanding, and is likely to benefit patients. For undergraduate research leading to an academic award these principles could perhaps be relaxed slightly, as ‘new understanding’ might include a student’s improved awareness of a research process. The three principles of purpose, understanding, and benefit would still apply.

This might sound daunting, but really is not. Finding a suitable research question is not out of anyone’s reach. In fact, there are questions that present themselves every day, both in clinical practice and at university or college. There are too many to be answered in a lifetime, with new questions constantly arising as clinical practice changes and new technologies emerge. All we need to do is lift our eyes up from the coal face of our routine work and be observant: ask questions and keep note on a daily basis.

What types of questions can we ask? There are too many possibilities to list fully. The following may provide some suggestions.

- Are there variations between practices, workplaces, or people?
- What is happening nationally in this type of clinical practice?
- Is ‘A’ performing better or worse than ‘B’? This question could apply to procedures, environments, information, or education systems.
- What has changed since date ‘X’?
- What are people’s feelings, knowledge, attitudes, or opinions? For example, patients, staff, students, tutors, or the general public.
- What are the effects of this technology, or service? This could address benefits and risks for patients, diagnosis, treatment, costs, waiting times, technical quality, training, and education.
- What happens if we alter this? It could be technique, procedure, and parameter.
- What is the size or extent of this?
- Why did things go wrong or right in this situation?
- What is the current ‘state of the art’ (knowledge or advancement) in this topic area?

If we are completely stuck and cannot think of any research questions, it is often a good idea to look through published journal articles in medical imaging and radiotherapy for sources of inspiration. Reading them can highlight topic areas that we would not otherwise have considered. In addition, the authors may list recommendations for further research at the end of their articles; research that we might be in a position to do.

Whatever question is chosen, a researcher needs to feel interested in it and the question should be of interest to others. Having a love (or at least a liking) for a topic

helps a researcher get through the long hours of data collection and writing-up that follows. Sometimes a research question is chosen for weak reasons, because it seems easy to answer or does not need ethics approval. Although this is understandable, the result can be a first taste of research that is boring and dissatisfying. Having completed a project and got a qualification, a student in such a hypothetical scenario then closes the book on research and vows never to open it again.

It is important to ask a research question, no matter whether we are doing primary research (gathering fresh material) or secondary research (reviewing published material). Finding a fresh question is vital if a review of existing literature is to be of interest. Since there may already be published reviews in the topic area, it is best if a different angle on the material is chosen, or different inclusion criteria are used. In addition, if the last review is not fairly recent it may be possible to ask an existing question, but with the benefit of new articles.

There is another framework, termed PICO, which is often mentioned in association with developing a research question. It focuses primarily on quantitative research, which involves interventions (such as treatments). It is of relevance to radiotherapy as well as to any interventional procedure undertaken in radiology. More broadly however, diagnostic testing may also be considered as a form of intervention and could therefore fit within the PICO framework. The acronym PICO is based on the following points.

- P: population—for example a sample of patients or people
- I: intervention—such as a treatment or test
- C: comparison—such as an alternative intervention or no intervention (control)
- O: outcome—the measurable effects or consequences of the intervention

When a suitable topic area has been chosen, then PICO can be of value in 'firming up' a research question or research design. It is especially relevant to clinical trials such as randomised trials (see Chap. 12) but is not really applicable to survey study designs and certainly not used in formulating qualitative research. The mention of measurable outcomes is a positive feature of the PICO framework. Any research must have tangible and recordable outcomes in order to provide results. Sometimes researchers decide upon a research topic area without properly considering how and if its results can be measured.

2.2.2 Step Two: Is My Question Feasible?

There are many questions to ask, but not all can be answered. Well, certainly not within the time or resources available. Research into cancer survival rates following treatment is valuable. But it might need a period of years to undertake, not the 6 months of a student project. Likewise, it would be interesting to survey whether the dose from a diagnostic X-ray procedure causes an increase in cancer rates, but as the radiation dose will be small and induced cancers few, we would need a huge

sample of patients to test our question. Questions such as these can indeed be answered, but perhaps only within the wider framework of a doctoral thesis or a large funded project.

The following points should be considered by anyone starting out on a piece of research.

- Is this research question answerable?
- How big a sample will I need to have a good chance of providing an answer?
- How much time have I got?
- What resources do I have, especially money and skills?
- Has the question already been answered, to such an extent that there is nothing new to say in this topic area?

An example of an unanswerable question might be one that depends on historical data. Let us suppose that we want to retrospectively explore whether the performance of a radiography procedure has improved with time. Past hospital or patient records might be missing or incomplete, meaning that no comparison can be made with a present-day procedure. The situation is essentially unknowable. This is often a problem with retrospective studies. Another unanswerable question might be: what are the true causes of adverse patient reactions to this radiological contrast agent at our hospital? Although reaction types and rates might be recordable, the real nature of the physiological changes taking place would only be knowable after thorough tests on each patient.

More is said about sample sizes in Chaps. 15 and 16. We should however note that in any research study we need enough data to be able to reach valid conclusions. It would be flawed, for example, to make assumptions about the national views of clinical radiographers if only ten respondents return a completed questionnaire. When writing a research question, we need to consider how feasible it will be to gather data. If we are part of a big student cohort, it should be fairly straightforward to gather plenty of student attitude information, assuming that people are available when we need them and prepared to take part. But if the research topic is, for example, based on a fairly unusual clinical condition or procedure, we might not be able to gather many cases during the research time we have available.

Time available to researchers is a real issue. Clearly more can be achieved when we have three years to undertake research instead of six months. Long-term projects can be more ambitious and more thorough. But in reality most research, whether it be for a university degree or an external funding body, is driven by tight deadlines. After deciding on a research question, time is needed to explore published literature in the topic, gather data, and then write up the work. Each of these stages can take longer than expected. It is therefore important to map out a target timetable, paying attention to the key aims and objectives for each stage. In this respect it could be argued that it is better to answer a simple question well rather than a complex question inadequately, especially when time is short. It is important, when doing

research, to ask: what is my key question? This prevents becoming side-tracked with other minor queries, no matter how interesting they might appear to be. Fuzziness or lack of focus is one of the most frequent reasons why research projects underperform in assessment.

When considering the resources available to us, one key factor is money or lack of it. The costs of doing research might include the following.

- Postage, stationery (paper and envelopes)
- Travel
- Reimbursing participants
- Photocopying and printing
- Ordering journal articles
- Setting up an online survey
- Software packages and computer hardware
- Paying research assistants or statisticians (in funded projects)
- Paying for use of clinical facilities, such as scanner time (in funded projects)

Some of these things might possibly be available on a 'no charge' basis to students, but it is wise to be realistic: for example, when proposing to undertake a national survey involving all hospitals.

Resources also include our own special talents and abilities. Someone who is especially good at statistics might be drawn to a project that involves a lot of complex data analysis; someone else who is a good communicator might be well suited to doing an interview survey. Although statistical help is often available at colleges and hospitals, many people may be so scared of statistics that they try to avoid it altogether. This is not normally a good idea as it can restrict the analysis of research findings. But of course it is not true that every piece of research must make use of statistical tests. Some numerical findings can be reported perfectly well descriptively, using tables and charts, while analysis of interviews would not usually require statistics at all. It is recommended that statistical tests should be used when they add something useful to the research findings, and not merely for the sake of trying to 'show off' when there is little justification for the test's inclusion.

There are some research questions that have already been answered so completely; thus there is little scope for fresh discoveries in the topic. As an example, it is well known that giving a caudal beam angulation for PA chest radiography can slightly reduce the dose to a patient's thyroid gland. Published dose data has been available for many decades, but the question is still a popular choice for undergraduate research proposals in radiography. There can be little real justification for proposals of this type, although the measurements involved might form a useful exercise within an undergraduate science course. If the answer to a research question is already well known from standard textbooks on the subject, there is little point in asking it again, unless a new approach can be used.

2.2.3 Step Three: Types of Questions

People could ask research questions for different reasons in radiography such as the examples below.

- To test a theory or idea. This is often the case in experimental research. An experiment sounds rather lab-based, but can be much broader than this, including social science, not just physics. Testing the effects of two radiotherapy regimes in clinical practice or bringing in a new patient information leaflet could be considered to be experimental studies, as could dosimetry research using a phantom. We might be making observations and trying to create a theory to explain what we see.
- To measure a position or situation. This could take place in surveys. For example, to determine the state of current practice in digital imaging at hospital sites or radiographer attitudes to a government policy.
- To improve clinical practice. Asking questions about what we do as clinical radiographers, making use of research findings, and altering our procedures accordingly is known as evidence-based practice or evidence-informed practice. Such questioning also forms a part of the process of reflection.
- To explore feelings, viewpoints, and attitudes. These might include the anxieties of patients receiving a clinical procedure, or the factors that influence job satisfaction amongst staff.
- The types of questions we can ask could be categorised as being directed towards the following. Explaining (e.g. testing theories and ideas)
- Describing (e.g. measuring positions and situations, making observations)
- Exploring (e.g. looking at feelings and attitudes)
- Controlling (e.g. implementing change, improving quality)

These types of questions can be found within the two main approaches of healthcare research, namely quantitative and qualitative. More is said about these approaches in Chaps. 15 and 16, but quantitative research is more likely to be based on numerical data and seeks to ask questions like ‘What?’ or ‘How much?’ Qualitative research, on the other hand, is more likely to be based on attitudes and opinions and asks the question ‘Why?’

Quantitative research gathers numerical data and is based on scientific or empirical philosophies, tests, and theories. In contrast, qualitative research collects human views and perspectives for developing new theories. Either approach can create and answer research questions. Quantitative research questions are usually more specific and may be associated with testing relationships between cause and effect or differences between groups of individuals. Exploratory questions are often associated with qualitative research.

The research tradition in medical imaging and radiotherapy has mostly been quantitative, linked to measuring a statistical performance of diagnostic tests, cancer therapies, and associated equipment. However, qualitative work has a vital role

to play in exploring the attitudes, beliefs, and opinions of patients and staff. Research can often bring together both quantitative and qualitative enquiry, using a mixed methods approach, which gathers both statistical data and human insights into the underlying reasons for why things are as they are. An example might be a study of quality performance in radiography that looks at numerical trends and also staff attitudes.

Should we ask a single key question when doing a piece of research, or several? Should we stay fixed and focused on a single issue or be responsive to the findings that we make during the research process? Researchers' views on this generally differ according to their research background. Quantitative research tends to address a single question (or perhaps several closely related questions), while qualitative research may be more free ranging in its explorations and might regard a single research question as an unhelpful restraint. Qualitative research needs to be free to find and develop new research questions as people's verbal comments are explored. Putting these differences aside for a moment, it may however be advisable for anyone who is submitting a research proposal for a degree project or external funding to opt for a single principal question initially in terms of the reasons given earlier in this chapter: feasibility and focus, for example.

Does research have to have a hypothesis? The answer is no, not always. A hypothesis is a research question which aims to test a theory. Traditionally we mean scientific theory here, which is based on experimental observations, and attempts to make sense of the world by explaining these in terms of proposed explanations. A hypothesis is only associated with quantitative research. The term hypothesis could be used in research which asks explaining questions, but not in research that uses describing or exploring questions. It would not generally be found in qualitative research. As an example, it would be quite normal to have a hypothesis when doing a physics experiment in an X-ray lab, but not when exploring the belief systems of student radiographers.

People often get confused when the terms null hypothesis and alternative hypothesis are mentioned. Hypotheses generally test theories whether relationships between measurable things (variables) are real or not. There are two possibilities: either one variable will influence another or it will not. For example, we could propose a theory that the size of the X-radiation dose to living tissues is a factor influencing an irradiated person's health. Rather confusingly, this relationship would be termed the alternative hypothesis, which is given the symbol H_1. A proposal that there is no relationship at all between X-ray dose and health would be called the null hypothesis, which is given the symbol H_0. (See Sect. 15.6 in Chap. 15).

If there are two possible relationships between variables, where for example one variable might vary inversely or directly with another, we can have two alternative hypotheses. The first alternative hypothesis, symbol H_1, could here describe the more likely inverse relationship between radiation dose and health (as dose increases, health decreases), while the second alternative hypothesis, symbol H_2, could describe a less likely direct relationship (as radiation dose increases, health

increases). We all accept these days that increased X-ray doses may reduce health and would thus reject hypothesis H_2, but this was not at all obvious to the early radiation pioneers. Indeed, a theory called radiation hormesis, in which small doses of radiation might actually be beneficial in some respects, has been proposed (but is not accepted by the majority of scientists).

2.2.4 Step Four: Answering Our Question

When we have decided on a research question, we then have to take the work forward by developing a structure for the project. It is important to create aims and objectives; these are normally an expected part of a research project proposal.

It is common for students to get the terms aims and objectives mixed up. An aim is a defined goal or aspiration of a research and is much easier to write if we have a clear research question. It should be stated in broad terms and be directly related to what you hope to achieve by the end of your research study. In general, there should not be too many aims, or else the project is at risk of becoming confused in its intentions, over-ambitious, cumbersome, and not feasible. Depending on the nature of a research study, one or two aims are generally sufficient and accepted for an undergraduate research project.

Objectives are step-by-step measurable outcomes of a research study. Objectives are directly aligned to achieving the aim/s of your study and must be written using positive statements that clearly communicate your intent. These are often arranged within a timeline and set out the practical means by which a project's aim/s will be achieved. There will normally be more objectives than aims.

Let us consider an example of a survey project which asks: does the general public know less about radiography than about physiotherapy or nursing? A possible aim for a project like this might be as follows.

- To determine the general public's awareness of radiography, physiotherapy, and nursing

Possible stepwise objectives for this project might include the following.

- To explore the available published literature on public awareness of the health professions
- To explore data on public awareness of radiography
- To explore data on public awareness of physiotherapy
- To explore data on public awareness of nursing
- To synthesise the data from these datasets and draw conclusions on public awareness of radiography in relation to physiotherapy and nursing

Aims and objectives should focus on answering the research question, regardless of whether the research is quantitative or qualitative.

2.3 Conclusion

The key aspects of choosing and formulating a research question were explained in this chapter. FINER, SMART, and PICO frameworks were presented and their helpfulness in developing a research question discussed. An explanation of aims and objectives was provided as most students get them mixed up.

Further Readings

Alvesson M, Sandberg J. Constructing research questions: doing interesting research. Los Angeles: Sage Publications; 2013.

Raich A, Skelly A. Asking the right question: specifying your study question. Evidence-Based Spine Care J. 2013;4(2):68–71.

White P. Developing research questions. 2nd ed. London: Red Globe Press; 2017.

Wyatt J, Guly H. Identifying the research question and planning the project. BMJ Emerg Med J. 2002;19(4):318–21.

Literature Searching

3

Martin Vosper and Angela Dimond

3.1 Introduction

Conducting a good literature search enables you to take a broad view and to interpret your research findings in light of existing knowledge. This chapter explains the need for searching literature sources of information and provides guidance on how to conduct a search, extract information, and how to manage this information.

3.2 Why Should Literature Searches Be Undertaken?

There are many reasons why it might be important to do a literature search, both for academic work such as essay writing and dissertations and in clinical radiographic practice. Literature searches can be used in many ways, to give the following:

- A supporting background and source of evidence to help justify arguments in an essay.
- A literature review chapter covering previous published findings within a research project (dissertation), to 'set the scene' and provide comparisons with new results.
- The content material for a systematic review. More will be said on systematic reviews in Chap. 12. A systematic review is a form of 'secondary research' which provides a methodical overview of previously published research data in order to

M. Vosper (✉)
Department of Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: m.r.1.vosper@herts.ac.uk

A. Dimond
Library and Computing Services, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: a.dimond@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_3

answer a research question. For example, what is the most effective imaging modality for staging breast cancer?
- An update of current evidence to inform best clinical practice.
- Evidence to support clinical guidance documents, recommendations, and policies.
- A background section of a journal article.

So, it can be seen that everyone, from a first-year diagnostic imaging or radiotherapy student to the head of the National Health Service, has reasons to search the literature. Literature searching is not a skill that can be packed away at the end of an undergraduate degree.

3.3 Types of Literature Sources

The sources of available literature are many and include the following:

- Journal articles
- Systematic reviews (such as Cochrane)
- Textbooks
- 'Grey literature' (unpublished material such as theses and conference proceedings)
- Media articles (newspapers, magazines, Internet sites)
- Internet sites
- Government and other official publications

In fact, there is such a wide array of available material these days that it can appear bewildering. Accessing evidence electronically, via web-based search engines and library databases, has never been easier but is important to have the right tools for the job and to screen out stuff that is not relevant.

3.3.1 Journal Articles

The majority of medical imaging and radiotherapy researchers would regard journal articles as the most important source, as they tend (usually) to be reliable, of good quality, and widely read by clinical practitioners. These journal articles are peer reviewed whereby they are written by experts in the field and blind reviewed by other experts in the field in order to assure the quality of the publication. Journal articles contain new research as well, although the key content may have been reported earlier at conferences. Useful journals in medical imaging and radiotherapy include *Academic Radiology*, *The British Journal of Radiology* (BJR), *Clinical Imaging*, *Clinical Radiology*, *The European Journal of Radiology*, *The Journal of Diagnostic Radiography and Imaging*, *The Journal of Radiotherapy in Practice*, *The International Journal of Radiation Oncology, Biology and Physics*, *Radiology*,

Radiography, and *Radiotherapy and Oncology*. There are many others. Journals are available online as e-journals and full text articles can be easily accessed via databases such as Scopus and Web of Science, especially if your university library subscribes to the journals you need. Online databases which are indexes of the published literature like PubMed (including MEDLINE) are valuable as a tool for finding journal articles and their use is highly recommended.

The advantages of using journal articles are that they are up-to-date, usually peer reviewed for quality, and generally based on new original data, except for discussion papers and commentaries, which are focused on a specific topic. Some of the disadvantages are that they may provide a narrow focus and may be subject to bias.

3.3.2 Systematic Reviews

Systematic reviews, such as those produced by the Cochrane Collaboration in the UK, see Chap. 12, provide a thorough overview of published primary research in a clinical topic area at the date of publication, with an overview of findings and full assessment of research quality. There may be some summated statistical analysis, termed meta-analysis, which is valuable for assessing the effectiveness of a therapy or test. Systematic reviews are not available for every speciality; for example, there are more existing reviews in therapies than in diagnostic testing.

The advantages of using systematic reviews are that they are authoritative; for example, Cochrane reviews are unbiased, provide thorough quality assessment of included studies, and are regarded as the highest form of evidence. The disadvantages are that they may not be up-to-date, depending on publication and may not be available in all topic areas.

3.3.3 Textbooks

Textbook content tends to lag slightly behind new research developments, due to unavoidable delays in writing and publication, but may provide a useful overview of a subject area. Popular textbooks are also more accessible to the reader than research papers and provide good background material which allows a complex topic to be better understood. There are also some ‘classic’ textbooks which contain important theories and principles, written by their original authors. It is often important to refer to these, especially in subjects such as research methods, science, psychology, and social studies. But although reading textbooks can be a useful first step in a literature review, textbooks should not be the main source of references for an undergraduate essay or dissertation. Journal articles are preferred since they provide more detailed research material.

The advantages of using textbooks are that they provide reader-friendly overviews of topics, as well as good sources of references. However, some of their disadvantages may be that they are not be up-to-date and may lack depth and detail when reporting research findings.

3.3.4 Grey Literature

Grey literature refers to items which may not reach full publication status, such as conference abstracts and proceedings, or theses. Conference material tends to be very current and often contains topical new material of interest to an audience, as well as interesting debates. These sources provide good material for a state-of-the-art essay or review in clinical imaging or radiotherapy but may only include summary data rather than full accounts and may also be hard to locate. Unpublished theses and student projects tend to provide in-depth explorations of a topic and may be good references to add to a dissertation. The OpenGrey database provides a source of European grey literature. In medical imaging and radiotherapy, it may be useful to search conference abstracts, such as those of the United Kingdom Imaging and Oncology (UKIO), European Congress of Radiology (ECR), or the Radiological Society of North America (RSNA) annual meeting.

The advantages of using grey literature are that it is often up-to-date, may be highly topical, and contain alternative perspectives. The disadvantages, however, are that it may be incompletely presented and hard to reference.

3.3.5 Media Articles

Media articles from Internet news sites, popular magazines, and newspapers are useful for providing interesting scene-setting and quotations for essays or research projects but should not be relied on too heavily for factual accuracy, since the stories they contain may be biased or only partly reported.

The advantages of using media articles are that they are topical and interesting and may provide good material for debate and quotation. However, the disadvantages may be that they are inaccurate or 'sensationalist' and contain facts which are hard to verify.

3.3.6 Internet Sites

This section concerns specialist Internet sites, not electronic journal (e-journal) sites. Although Internet sites can be used as references in some circumstances, they are not always a reliable source. This is because the material found on most personal, 'special interest,' and corporate Internet sites is not subject to the same process of academic peer review and quality assessment that takes place before publication in journals. It is thus more likely to be of a variable standard and may be subject to bias. However, some sites, such as those of manufacturers, organisations, and societies, may contain very useful information which is not available elsewhere.

The advantages of using Internet sites are that they are easy to access online, may contain a wide breadth of resources and links, and are often up-to-date and highly

topical. The disadvantages, however, may be that the content may vary a lot in terms of quality and accuracy, and they may present narrow or biased viewpoints (see Chap. 10).

3.3.7 Government and Other Official Publications

Government and other official publications, such as those produced by professional bodies and colleges, provide useful material for an essay or the introduction/background section of a dissertation or journal article. It would not be usual to include such publications in a systematic review, unless they are officially sponsored primary research reports. Some types of research reviews may be based on a study of official documents.

The advantages of using these publications are that they may contain useful reference lists, provide good background material, help present the broader context, and may be written by expert panels. The disadvantages, however, may be that they are selective in the use of source material and data and may not present detailed research findings.

3.4 How Thorough Should a Literature Search Be?

Searching the published literature within a chosen research topic is vital, both in primary research (e.g., a new experiment or survey) and secondary research (a review of existing findings), or in an essay. Research may be weak and ill-informed if it shows lack of awareness of the findings of others working in the same topic area. This runs the risk of producing narrow discussions and reaching biased conclusions, without consideration of alternative evidence. This point applies to all research work, ranging from novice to expert. Undertaking a thorough literature search enables researchers to take a broad view and to interpret their own research results fully in the light of existing knowledge. It also gives researchers more to say within the analysis of findings, as the new data can be compared with that previously gathered by others elsewhere. A very thorough search of literature would not normally be expected for an essay, but even so there should be an awareness of key and up-to-date publications in the topic area, showing a balanced and well-informed perspective. Reference lists for essays are sometimes selective and may reflect a personal viewpoint or argument, depending on the essay title that has been set. References included in the background sections of published journal articles may be limited by available space, although once again it is important to show freedom from bias and awareness of relevant previous published studies.

The breadth of a literature search depends on the task in hand. Some suggestions are given in Table 3.1 below.

Suggestions for the numbers of references in Table 3.1 apply to typical student work at undergraduate degree level. For postgraduate or funded research, the expectations would be greater. But there are always some topics, perhaps very recent,

Table 3.1 Examples of numbers of literature references typically needed within common research tasks at undergraduate level

Research task	Suggested number of references	Reasons for the literature search
1. Writing a research proposal	About 10 should suffice These should be recent key references in the topic area	To set the background To show awareness of the topic To justify the proposed research To identify issues and opportunities
2. Doing a literature review chapter within a primary research project	The number depends upon how many references are available on the topic. However, fewer than about 30–40 would be disappointing to most tutors	To set the background To identify other relevant work on the topic To provide comparisons with our own findings To inform our method, discussion, and conclusions
3. Doing secondary (a systematic review of literature)	All recent relevant research in the topic area should be included But limitations will include language of publication (if other than English), type of publication, personal resources The number depends upon how many references are available on the topic, but it should be larger than in (2) above About 80+ would typically be expected for an undergraduate review in a topic area where ample literature is available	To provide a rigorous overview of recent available research in the topic area, without major omissions To portray the current state of knowledge and/or clinical practice in the topic area To synthesise the available research evidence and reach informed conclusions regarding current issues, trends, and practice

specialist, or obscure, in which little published research is available. In such cases it would be accepted that fewer references might be used—provided that the researcher really has done a thorough literature search. It is quite common for students to report, rather despondently, that they "can't find anything on the subject", when a subsequent online search brings up several journal articles which they have missed. It is important to search widely when looking for literature and to use the right tools for the job. These issues are discussed below.

Researching in an area where little previous work is available can restrict the literature review section of a research project, but it can be a positive advantage too. There is probably more chance that the research findings will be novel and original. It may also help with publication, provided that the topic is not so obscure that it is of no interest to others.

How recent should literature references be? The pace of technological change in medical imaging and radiotherapy means that older clinical references may be outdated and no longer relevant to current practice. As a guide, it can be recommended that literature sources should be from within the last five years (normally) in any topic area which is experiencing rapid change. Examples might include computed tomography and intensity modulated radiotherapy. Even in subjects like these, there

may be earlier research which should be included because it is key to understanding in the subject or contains evidence which was, and still is, of vital importance. It is sometimes suggested that research from countries which have technically advanced healthcare systems (the USA is often quoted in this regard) may be slightly 'ahead of its time' and that this should be taken account of too when thinking of a five years cut-off point for useful sources.

As mentioned above, secondary research consisting of a full literature review makes the most extensive use of literature searching. A systematic review (see Chap. 12) is the term used for a really thorough appraisal of available and relevant research evidence in a topic area, applying a specific methodology for including, excluding, and appraising studies, in order to answer a research question. Full systematic reviews are time-consuming and would not be expected within an undergraduate degree. But a reasonably complete overview of recent published evidence (in English) is achievable within a timeframe of about six to nine months for an undergraduate review.

Medical imaging and radiotherapy research is not only about physics, biology, and technology. It is also concerned with people—patients, clients, the general public, and staff. Thus, researchers need to think laterally, also considering sources within general health, psychology, social sciences, and even economics. 'Search widely' is good advice. It is important to look at research in other fields such as nursing, physiotherapy, and general industry, if exploring topics such as manual handling, job satisfaction, or anxiety (for example), which have a huge literature but not much that is specifically about radiography. In such situations useful comparisons can be made between experiences in radiography and other professions. Of course if research is being undertaken in a very 'radiographic' area such as radiation doses, it is unlikely that there will be much relevant material outside the radiological and medical physics literature—but even in this case it may be worth looking at sources in medical health, oncology, molecular biology, epidemiology, and immunology, to name just a few.

A literature source is the original published article, book, or conference proceeding. Although people sometimes find and use abstracts (which are short summaries) of published articles, because it is quicker and easier to do so, this is never a good idea. A 200–300 word abstract cannot convey the full findings of a journal article, and although there are usually some summary results, many important details, complexities, and 'angles' will be missed. Similarly, it is best to look for an author's views within their original book or article, rather than relying on secondary quotations in other sources. Secondary quotations may be selective when using an author's words, in order to support other arguments.

3.5 Writing a Literature Search Methodology

Although a method is associated in many people's minds with traditional research or an experiment in the laboratory, methodology is a vitally important (and sometimes forgotten) part of a systematic literature search too. A good literature search

for a systematic review will include details of the search strategies used, in the same way that a primary research project will have a methods chapter. Components of the search strategy should include:

- Names of the databases used (sources of information)
- Database search terms (keywords, which should hopefully allow related articles of interest to be found)
- Inclusion and exclusion criteria (justifiable reasons for leaving literature in or out of the review)
- Data extraction (details of the types of information that we want to get from the literature)
- Numbers of results (hits)
- Quality assessment of studies

A rigorous method such as this would not be expected for a literature search in an essay but might be employed to some extent (leaving out a formal quality assessment of studies), for a literature review chapter in a research dissertation.

Literature searching needs to be planned and methodical just like other aspects of research. It needs to ask clear questions, have aims and objectives, gather selected information, and report findings.

3.5.1 Approaches Used in a Literature Search

Although it is possible to search for articles within individual printed journals or their online home pages, the most efficient approach is to use an online database which indexes the content of a wide range of journals. Some of these are freely available and others are subscribed by libraries. Electronic copies of articles can be easily downloaded, provided that the institution subscribes to the journal in question or the article is free of charge. There are a number of information databases available online, which index health and related subjects, and include the following.

- PubMed (including MEDLINE), the most widely known, covering most aspects of medical and health literature and produced by the US National Library of Medicine
- CINAHL, which includes nursing and allied health
- EMBASE, for biomedical and pharmaceutical literature which includes MEDLINE
- PsycINFO, for literature in psychology
- CancerNET UK, a good resource for oncology, from the National Cancer Institute
- International Cancer Research Partnership (ICRP), a collaborative indexing of cancer research

- The Cochrane Library, a valuable source for health interventions and therapies, especially randomised controlled trials (see Chap. 12), as well as systematic reviews in health care. It contains several individual databases, such as:
  - CDSR (The Cochrane Database of Systematic Reviews)
  - DARE (The Database of Abstracts of Reviews of Effectiveness)
  - CENTRAL (The Cochrane Central Register of Controlled Trials)
- Scopus, a large database of peer-reviewed literature across all subjects
- OpenGrey: The System for Information on Grey Literature in Europe, which covers unpublished (grey) literature, categorised by country and subject area.

The large number of databases available may seem a bit confusing and many people just use a single database, such as PubMed. But it is important to know that each database may find articles that are not indexed in the others and that no single database covers everything. Sometimes it is possible to do a combined search using more than one database, but beware that this may give duplicate "hits" for the same article via different databases, giving a very long list of references. Both PubMed and EMBASE usefully combine the MEDLINE resources.

There are other sources which can be used on the web to find research articles, including those from the journal publishers, such as:

- Blackwell Synergy
- Ingenta Connect
- SpringerLink
- Sage Journals Online
- ScienceDirect, from Elsevier
- Wiley Online Library

The content of these depends on each publisher's range of journal titles, but they are a useful extra source in many cases. The web search engine Google Scholar can also bring up useful journal articles and other material such as electronic book extracts (often available on OpenAccess) and provides Advanced Search options. The term OpenAccess refers to resources that are freely available for legal download and usage. Access to these research outputs in the form of book chapters or journal articles is open to all and unrestricted. Increasingly, even a simple Google search provides journal articles that are OpenAccess.

Table 3.2 below gives an indication of the amount of material available via the various web-based search tools and gives some indication of their potential usefulness to diagnostic imaging and radiotherapy:

Additional search approaches include visiting a specialist library, such as that of the British Institute of Radiology in London, which contains student projects and subject-specific literature. It is worth contacting the librarian of such a centre before the visit, in order to check whether there are likely to be any materials which are relevant to your research, especially if a long trip is involved. Major national libraries, like the British Library, contain journals that might not be available locally, but

Table 3.2 Numbers of "hits" for the search word "radiotherapy" using various electronic databases and journals

Database, web search engine, or electronic journal	Number of hits	Comments
MEDLINE (PubMed)	202,000	This database contains over 18 million citations, dating back to the 1950s
CINAHL Plus	10,600	Articles from over 3200 journals
PsycINFO	430	Consists of several parallel databases covering different date ranges
Allied and Complementary Medicine Database (AMED)	660	Contains articles relating to alternative and complementary therapies
International Cancer Research Portfolio	580	Search for the period 2007–2008
Cochrane Library	5400	For the Cochrane Database of Systematic Reviews
Science Direct (from Elsevier)	25,700	Articles from over 1000 journals
Radiology	2000	The journal *Radiology*
British Journal of Radiology (BJR)	820	For a search within abstracts of articles in the journal BJR

access to such institutions is restricted. Most university and hospital libraries offer an interlibrary loan service whereby electronic copies of articles can be ordered, often free of charge to students or staff.

Not all original research gets published and a thorough literature search may also include unpublished material (often referred to as "grey literature") such as master's or doctoral theses and dissertations, reports of meetings, and conference proceedings. This is often to be found in university libraries and repositories such as the British Institute of Radiology Library.

The search approaches used should all be listed in the literature review methodology.

3.5.2 Database Search Terms

Within the literature search methodology, the search terms which you have used when hunting for articles via online databases should be listed. Search terms are words and phrases which we hope will score 'hits' by bringing up relevant articles of interest. The terms are entered in a search box within the database. The precise choice of words and phrases will very much affect the number and type of 'hits' that we get—the results can be a bit surprising. The database will usually look for matches between our search terms and keywords contained in the title, abstract, and text of journal articles. Sometimes, even when we know for sure that there are articles available in a topic area, a database would not seem to retrieve them for us and this can be frustrating. Possible reasons for this include:

1. The database we are using does not include these articles (no database contains everything ever written).

2. The database does contain the articles but the search terms we are using are not recognised or are not precise enough. The search phrase "X-ray" might well be a useful one to include, but just entering it on its own might bring up a huge number of hits from non-radiographic fields such as X-ray astronomy, X-ray crystallography, general physics, and so on (depending on the type of database we are using). More advice is given on issues like this later in the chapter.

3.5.3 Inclusion and Exclusion Criteria

It is necessary to have some protocol (or set of rules) in your methodology for deciding whether each piece of published research should be included in the written literature review or excluded. Every researcher needs to produce their own protocol and there is no universal guidance for this. This is because each research topic is unique and so a set of universal rules would not work. But generally, you are likely to leave out articles which are the following.

1. Irrelevant
2. Out of date
3. Unreliable

Irrelevant evidence is that which is not applicable to your research—examples might be research from other countries (where the healthcare system is very different from your own), or studies of diagnostic imaging in children when you are researching radiography of adults.

Research from many years ago might have become out of date if there have been rapid changes in clinical practice since then. This means that circumstances have changed so that the situation presented within the old research is no longer applicable today. Each researcher needs to decide a cut-off date for inclusion, before which point the research is not to be included. The date chosen will very much depend on the topic area—human anatomy does not change over the years, but chemotherapy does.

Deciding whether research evidence is reliable is often the hardest decision to make when considering whether or not to include it. More is said about literature appraisal in Chap. 4, but it is best to exclude articles which you feel are of poor quality. Poor quality research articles (as seen by you as the reviewer) might be ones with small sample sizes, flawed methods, obvious bias, weak statistics and analysis, and so on. A randomised controlled trial (see Chap. 12) is usually regarded as the best quality clinical research evidence, but such trials are rare in medical imaging (although there are more in cancer therapies). Thus, a researcher writing a review in radiology might need to compromise a bit when considering articles for inclusion (or risk having none!). When writing a literature review it may be wise to pay more attention to the findings of research which compares the effectiveness of one diagnostic test (or cancer therapy) with

another on the same set of patients. Most research will only look at a single intervention, however.

Typically, inclusion and exclusion criteria will include the following: date range, geographical location, language, age group, and type of publication. To give an example of a set of inclusion and exclusion criteria for a literature review, let us suppose that you are doing a systematic review looking at the usefulness of magnetic resonance imaging (MRI) in diagnosing suspected adult brain tumours.

Inclusion criteria in this example might consist of articles that:

- Are written in English
- Have been published within the last 5 years
- Are original primary research
- Use commonly available MRI technologies
- Have symptomatic adults as their sample group
- Involve first presentation of disease or symptoms
- Use a sample size of at least 25 clinical cases
- Are felt to be reliable and of good quality

Exclusion criteria would mostly follow on from this and might be articles that:

- Are in languages other than English (unless we have language skills)
- Are older than 5 years
- Are reviews of other work
- Involve MRI technologies not generally available elsewhere
- Include asymptomatic adults (such as volunteers and health screening cases)
- Involve recurrence of disease or symptoms
- Have a sample group of less than 25 clinical cases
- Are felt to be unreliable as evidence, due to poor quality

3.5.4 The Data Extraction Process

One very valuable, but often omitted, part of a literature search method is the data extraction form. This lists the key information that the researcher is aiming to extract from the literature. It is a good idea to complete a form for every research article, since this gives a valuable summary of major findings and also acts as a reminder about what information was found where. Otherwise you might be left later on with a lot of articles and have no recollection about which one contained a particular important finding.

An example of a data extraction form is included below in Table 3.3, for the previously mentioned review of MRI in diagnosing adult brain tumours. This is not a rigid template, just an example—every literature review will be different.

Table 3.3 Example of a data extraction form

Key data categories	Entries
Title of article	
Authors	
Year of publication	
Country of publication	
Type of MRI scanner	
Field strength of MRI scanner	
Type of research study (randomised trial, observational study, review, etc.)	
Is MRI compared with any other test and if so with what?	
Is a "gold standard" test used?	
Number of patients in the study	
Are sensitivity and specificity data included? If so state the values	
Are cost data included?	
Is there any mention of patient outcome measures, such as survival, quality of life, alteration of treatment or diagnosis, satisfaction, etc.?	

3.6 Tips and Tactics for Doing the Literature Search

Doing a literature review can either proceed smoothly or be very frustrating. The following tips and tactics may help you.

- Consider all of the possible words and phrases that might commonly be used by authors when they are writing articles in your chosen topic area. Using these words as search terms should help you find related articles. It is often a good idea to see which words are used in the reference lists of the first articles that you find. Be broad-minded in your choice of search terms.
- When using a database, look for search terms in the abstract and body of articles, not just in their titles. Many authors use rather odd or 'catchy' phrases as article titles, which do not connect well with the actual subject area. It would be easy to miss these articles if you only searched for words in their titles.
- When using the PubMed/MEDLINE database, an initial search can often bring up loads of articles that are not connected with your subject area. You can get around this problem by using the Medical Subject Headings (MeSH) tool to see what keywords and phrases PubMed/MEDLINE recognises in your topic. A repeat search using these recognised words usually gives more 'hits'.

The following steps demonstrate how you might apply a search strategy for the review of MRI in diagnosing adult brain tumours.

- Break the question down into its three main concepts: in this case MRI, brain tumours, diagnosis.
- Note down any possible alternative keywords for the same concept, such as brain neoplasms as another term for brain tumours.

- In the chosen database, search for each concept separately and then combine searches together. This gives you more flexibility if you want to adapt your search strategy as you go, according to whether you are getting appropriate results or if you want to add other keywords from abstracts or MeSH headings.
- Use Boolean operators (the words AND, OR, and NOT) to combine your terms or searches together. Remember that OR is used to combine alternative keywords and broadens out your search. AND is used to combine more than one concept together and narrows down the search results. NOT excludes a term or concept.
- Another useful technique is to use inverted commas around multiple words to search them together as a phrase.
- This example shows how the three concepts are searched separately and then finally the three searches are combined, which narrows the results down to only those articles which contain information about all of the concepts together.
  - Search 1. MRI OR "magnetic resonance imaging"
  - Search 2. "Brain tumours" or "brain neoplasms"
  - Search 3. Diagnosis
  - Search 4. 1 AND 2 AND 3
- If you find that an author or team of authors research quite frequently in your topic area, try doing a search using that author's name.
- Do not just rely on PubMed/MEDLINE as a search engine. You may find that other databases such as Scopus or Google Scholar give you more returns. No database covers all of the available literature and each will have good coverage in some specialist areas.
- Always click on the "related articles" that appear when you do a search using a database such as PubMed/MEDLINE. These in turn will lead you to other related articles in the subject area.
- If you find that articles from your topic area often appear in the same published journal, try doing a search through all of its content for the last few years. This can be done electronically via a journal homepage. You will often find that you come across other relevant articles, editorials, and correspondence that you would otherwise have missed.
- It may seem a bit obvious but do look through the reference lists of those journal articles which you have already identified as useful. Sometimes authors will quote references which do not appear in your database searches.
- Several of the above tips are "snowballing" techniques, by which finding one reference leads to locating many others in turn and is helpful if you are finding too few results.

3.7 Problems with Literature Searches

We do not want to dwell on negatives, but it is best to be pre-warned about possible problems. It can sometimes be difficult to find all of the relevant research that has been written in your chosen topic area, even with the help of databases like PubMed/

MEDLINE. Having a thorough search method will increase your chances of success.

But is the published material an accurate picture of research in medical imaging and radiotherapy? The answer is "well, yes sometimes—but not necessarily". Often there is a tendency for research which shows positive benefits (from treatments or diagnostic tests) to get published, while negative findings may end up filed in a drawer. Also, definite or statistically significant results (whether positive or negative) may be more attractive to a publisher than findings which are null or equivocal. Although journals are not like newspapers, there may still be pressure to print material which is likely to excite the readership. This tendency is called publication bias and is present in many areas of health research, including funded work. This bias can skew clinicians' perceptions of the usefulness of treatments, and there have been situations where the effectiveness of certain therapies (e.g., certain chemotherapies) has been over-exaggerated. Since most clinical staff get their updates from journal articles and conferences, this is hard to avoid.

Definite or positive findings are not only more likely to be published—they are also more likely to be published quickly. This means that the first rush of publications in a developing clinical technique may tend to give a rosy picture, while delayed reports may be more cautious. This is called time-lag bias.

Increasingly, researchers are using social media to promote their research findings as an immediate form of communication. Data about the number of tweets, blog posts, likes, bookmarks, and so on are made available through Altmetrics, which are widely added to database citations and can skew search results.

Someone searching the literature will often find several articles on a topic which are from the same group of authors. This may sometimes be essentially one piece of research, written up in slightly different ways and presented in several journals. It can lead to "multiple publication" bias. Multiple publication can be attractive to researchers since producing more "outputs" not only increases their "street credibility" but can bring in promotions and other rewards. If the articles are from widely different years, it is more likely that each is a separate piece of research, and some authors are very prolific in producing original work, even within a single year. No one would accuse authors of cheating where multiple publication of the same findings takes place—but someone undertaking a literature review should record the findings as one piece of research evidence, not several.

In technology-driven fields like medical imaging and radiotherapy, there can also be another effect, which could perhaps be termed one-upmanship bias. This means that hospital centres with the most advanced new scanners or linear accelerators may be more likely to get their research published. It is true that cutting-edge research is more likely to be achieved using the newest equipment, and active clinical researchers are attracted to the best-funded centres (hence more publications from these sites). But someone undertaking a literature review should reflect that the technologies might be unavailable at most hospitals and might not reflect the real-world situation for most of the health service. This is an example of weak generalisability or external validity.

The above-mentioned biases will be present in the research literature, and the available evidence will be influenced to some degree. A researcher cannot escape this fact but can be aware of possible biases and reflect accordingly. It may be useful to consider the following criteria for evaluating what you find.

- Who?
- Can you identify who has written the information? What are their experience and qualifications? Who do they work for? Is the work sponsored by an organisation? What else have they produced?
- What?
- Is the information biased in favour of one view? Can you locate a counter-argument? Are there obvious omissions? Can you easily distinguish between fact and opinion? Is the emphasis of the topic appropriate for your needs?
- Why?
- What is the intended audience? Is the material at a suitable academic level?

What if you are undertaking historical research, for example looking at the development of radiography during the period from the 1920s to the 1960s? Most journal articles are not available electronically in full text for those years and this can cause a problem. In these situations, it may be necessary to visit a national collection such as the British Institute of Radiology library and do a hand search of the printed copies of journals. The problem also applies to older out-of-print books. A researcher living in a large city like London is more conveniently placed to visit specialist science and medical libraries. The limited availability of older sources is not normally a problem in medical imaging and radiotherapy research, since most (but not all) older publications are no longer relevant to current clinical practice.

To access a wide range of literature you need to use an institutional library and there can be a problem if your institution does not pay to subscribe to the electronic full text version of the key journals you need. In such cases it will be necessary to order copies of articles from elsewhere. This is possible but might have cost and time implications. It is a good idea to check the library's holdings of full text journals as soon as possible, not in the last few weeks before a review has to be submitted. Remember that databases such as PubMed/MEDLINE will help you to find the title and citation (journal, year, volume, page number) of articles needed, but may not link to the full text, unless the library subscribes to the journal or it is available free to non-subscribers (OpenAccess). Abstracts are usually available via databases; do not rely on these alone.

Once you have found relevant articles, do not forget that you will eventually have to reference any that you use in your review. Saving full citations as you do your searching will save you time at the writing-up stage. It is important to have some system for keeping and organising your references, for example creating a file or folder on your computer, or making use of one of the many online reference management tools such as EndNote, Zotero, Mendeley, etc.

3.8 Getting Support

Consult your supervisor if you are undertaking a research project, as they will be able to guide you with your literature searching. Also contact your librarian for advice on online database use and other resources. Most universities have online help tutorials available to guide your literature search.

3.9 Conclusion

There are a number of ways in which you can search for information; the most popular being the online databases. It is good practice to develop a literature search strategy or method so that this plan can guide you through your searches and keep you focused on your aims and objectives of the study. Most literature reviews will need you to submit an indication of the search strategies or methods used. This forms the methodology of the literature review and would include the main inclusion and exclusion criteria for literature searching. Searching the literature can be an exciting quest, as it always provides fresh insights to the person undertaking it themselves for the first time. No matter that hundreds of other people have done the job previously—since new things are constantly being published, each search is unique. Literature searching can be frustrating too but use of the right tools and the right method eases the journey.

Further Readings

Greenhalgh T. How to read a paper: the basics of evidence-based medicine and healthcare. 6th ed. London: Wiley-Blackwell; 2019.

Hart C. Doing a literature review: releasing the research imagination. 2nd ed. London: Sage; 2018.

Straus S, Richardson WS, Glasziou P, Haynes RB. Evidence - based medicine: how to practice and teach EBM. 5th ed. London: Elsevier; 2018.

Literature Evaluation and Critique

4

Andrew J. Scally

4.1 Introduction

In the global modern healthcare environment, there is an expectation that you, as a healthcare professional, should base your practice upon the best available research evidence. National and international professional organisations for diagnostic imaging and radiotherapy practitioners emphasise the importance of an understanding, and the implementation, of an evidence-based approach to service development [1, 2].

Given that new 'evidence' continues to emerge at a rapid rate, all health professionals must be able to evaluate findings that are relevant to their practice and judge whether to incorporate change when this is necessary. This ability to critically appraise claims from research that are published in the literature, and independently evaluate the strength of such claims, is vital to diagnostic imaging and radiotherapy.

Although the idea of evidence-based medicine, or more generally evidence-based healthcare practice, has been traced back to the nineteenth century, the quality of health research has improved steadily. This does not mean that nowadays all research is conducted in a way that ensures the robustness of the conclusions. Established best practice is not always followed by researchers, and even where it is there is still the potential for hidden biases to be present in research that cannot easily be identified, eliminated, or controlled.

Standards for best practice in healthcare research have been published within recent years. These publications are an extremely valuable resource for students and qualified practitioners to help them make informed judgements on the quality and relevance of published research. Several organisations have developed critical

A. J. Scally (✉)
Diagnostic Radiography, School of Clinical Therapies, University College Cork, Cork, Ireland
e-mail: andrew.scally@ucc.ie

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_4

appraisal tools which enable you critically to appraise research papers in a systematic way that involves the consistent application of the same relevant key questions for a given research design (see Appendix for web addresses).

4.2 Hierarchies of Evidence

A natural hierarchy of research evidence quality has emerged that is informed by the ease with which potential biases can be avoided or controlled. This is covered in useful texts by Sackett and colleagues and by Greenhalgh [3–5]. Although there are variations on the precise structure of this hierarchy, in particular to take account of qualitative research [6], it is broadly outlined in Table 4.1.

The highest level of evidence is widely considered to be a systematic review of well-designed randomised controlled clinical trials (RCTs), all of which aim to answer the same research question. However, the quality of a systematic review is necessarily constrained by the quality of the individual trials of which it is composed. Not all systematic reviews are reviews of clinical trials. Also, it is worth bearing in mind that published randomised controlled trials may be relatively uncommon in diagnostic imaging and radiotherapy. It is perfectly reasonable to perform a systematic review of observational studies, when there is little or no evidence from RCTs in a particular area of interest. The second level in the standard hierarchy is a large, well-designed RCT. It is therefore considered by some to represent the strongest kind of evidence. The third level of evidence is an observational cohort study; the fourth and fifth are observational case–control studies and cross-sectional studies (surveys) and the sixth is an observational ecological study (where individual-level exposure data are lacking, but aggregate, population-level data are available). Ecological studies are not common in medical imaging or radiotherapy, but are conducted occasionally. The lowest level of evidence

Table 4.1 The traditional hierarchy of evidence

Rank	Study design	Comment
1	Systematic review	Ideally of well-designed homogeneous RCTs. Status in the hierarchy may be relegated if RCTs are heterogeneous or if the review is of observational studies. A systematic review may or may not include a meta-analysis
2	Randomised controlled trial	Judgement required of the size and quality of the study and whether the results are definitive
3	Cohort study	A large, well-designed study may be more persuasive than a weak RCT, but a cohort study is more prone to bias
4	Case–control study	Causal inference more difficult to establish and more prone to bias than in a cohort study. Efficient design for rare conditions
5	Cross-sectional study (survey)	Causal inferences cannot be made. Provides information at a single instance of time
6	Ecological study	An observational study that uses aggregate level data, in the absence of an assessment of individual exposures
7	Case reports	Lack generalisability due to very limited sample size and their selective nature

is considered to be individual case reports, owing to their usual lack of generalisability.

Although this hierarchy reflects the reliability of the different study designs in terms of researchers' ability to eliminate or control biases within them, it should not be assumed that an observational study is always inferior to an RCT. Randomised controlled trials have their own potential problems and may not always be the most appropriate design for diagnostic imaging studies. They are not even ethical or appropriate in many situations, for example, when studying the health effects due to exposure to toxic agents, such as ionising radiation or chemical pollutants. Furthermore, the need for and appropriateness of such a hierarchy has been challenged. Wherever the design methods rank in the hierarchy, a well-designed study produces results that are more plausible than those from a poorly designed study. It has also been suggested that rigid adherence to this hierarchy has seriously misrepresented, or under-reported, the evidence supporting the more widespread use of new imaging methods in oncology [7].

In medical imaging, studies are commonly designed to measure the diagnostic accuracy of alternative imaging techniques and their combinations, or of the diagnostic performance of individuals or groups of observers/interpreters or even combinations of technology and observers. Diagnostic accuracy is commonly determined by the sensitivity of a test (the ability to detect disease when it is present), and the specificity of a test (the ability to exclude disease when it is absent). Although standards of good practice have been developed specifically for the design and reporting of such studies [8], a diagnostic accuracy study can have the characteristics of either an RCT or an observational study and so can be evaluated broadly within the evaluation framework relevant to an RCT or observational study.

4.3 Examples of Different Research Designs in Medical Imaging and Radiotherapy

Many different types of studies are published in the medical and health science literature. However, not all of them are primary or secondary evaluations of patient outcomes or, in the case of medical imaging studies, diagnostic performance. The research literature is broad and may cover many aspects of professional practice, for example, clinical audits, development of guidelines, developments in education and training, surveys of professional practice, surveys of user views and experiences, and experimental studies relating to assessment of health technologies. Although many of the principles addressed in this chapter can be applied to the appraisal of such articles, the main focus of this chapter is on research involving patient outcomes and diagnostic performance. It is from such studies that suggested changes can be made to clinical practice and improvements made in patient care.

Some examples from the medical imaging and radiotherapy/oncology literature that have used the primary research designs identified above are outlined in Table 4.2 [9–49]. No attempt here is made to appraise these studies but they could serve as helpful examples to which you could apply an appropriate critical appraisal tool from the options presented in the Appendix.

Table 4.2 Examples of the different study designs used in medical imaging and radiotherapy research

Authors	Title	Purpose
Systematic reviews		
Younger et al. [9]	Describing ionising radiation risk in the clinical setting: a systematic review	A systematic review seeking to identify and explore the techniques advocated for disclosing the risk to patients of ionising radiation from clinical medical imaging examinations
Sierinka et al. [10]	Systematic review of flexion/extension radiography of the cervical spine in trauma patients	To investigate whether flexion/extension (F/E) radiography adds diagnostic value to CT or MRI in the detection of cervical spine ligamentous injury and/or clinically significant cervical spine instability of blunt trauma patients
Gupta et al. [11]	Systematic review and meta-analyses of intensity-modulated radiation therapy versus conventional two-dimensional and/or or three-dimensional radiotherapy in curative-intent management of head and neck squamous cell carcinoma	To compare IMRT with conventional two-dimensional (2D) and/or three-dimensional (3D) radiotherapy (RT) in curative-intent management of HNSCC regarding disease-related endpoints
Harris et al. [12]	Systematic review of endoscopic ultrasound in gastro-oesophageal cancer	To review the literature about the use of endoscopic ultrasound for the preoperative staging of gastro-oesophageal cancer, especially staging performance and impact
Bryant et al. [13]	Cardioprotection against the toxic effects of anthracyclines given to children with cancer: a systematic review	To conduct a systematic review of the clinical effectiveness and cost-effectiveness of cardioprotection against the toxic effects of anthracyclines given to children with cancer
Brealey et al. [14]	Accuracy of radiographer plain radiograph reporting in clinical practice: a meta-analysis	To quantify how accurately radiographers report plain radiographs in clinical practice compared with a reference standard
Randomised controlled trials		
Gupta et al. [15]	Neoadjuvant chemotherapy followed by radical surgery versus concomitant chemotherapy and radiotherapy in patients with stage IB2, IIA, or IIB squamous cervical cancer: a randomized controlled trial	To compare the efficacy and toxicity of neoadjuvant chemotherapy followed by radical surgery versus standard cisplatin-based chemoradiation in patients with locally advanced squamous cervical cancer

Table 4.2 (continued)

Authors	Title	Purpose
Brealey et al. [16]	Influence of magnetic resonance of the knee on GPs' decisions: a randomised trial	To assess the effect of early access to MRI, compared with referral to an orthopaedic specialist, on GPs' diagnoses and treatment plans for patients with knee problems
Bartholomew et al. [17]	A randomised controlled trial comparing lateral skull computerised radiographs with or without a grid	To investigate the effect on perceived image quality of the use or non-use of a secondary radiation grid for lateral skull radiography
Harrison et al. [18]	Randomized controlled trial to assess the effectiveness of a videotape about radiotherapy	To investigate whether the provision of a videotape, in addition to the standard information booklet, reduced pre-treatment worry about radiotherapy in cancer patients
Sala et al. [19]	A randomized controlled trial of routine early abdominal computed tomography in patients presenting with non-specific acute abdominal pain	To compare the effect of initial early computed tomography (CT) versus standard practice (SP) on the length of hospital stay, diagnostic accuracy, and mortality of adult patients presenting with acute abdominal pain
Ravasco et al. [20]	Dietary counseling improves patient outcomes: a prospective, randomized controlled trial in colorectal cancer patients undergoing radiotherapy	To investigate the impact of dietary counselling or nutritional supplements on several outcome measures (nutritional intake, nutritional status, and quality of life) in colorectal cancer patients
Cohort studies		
Slaar et al. [21]	Plain radiography in children with spoke wheel injury: a retrospective cohort study	To evaluate the type of radiographs that are obtained in children with BSI, to assess in which anatomical regions fractures occur, and to evaluate on which radiographs a fracture can be detected in children with bicycle spoke injury (BSI)
Damen et al. [22]	Additional value of different radiographic views on the identification of early radiographic hip and knee osteoarthritis and its progression: a cohort study	To investigate the prevalence and progression of early radiographic osteoarthritis (OA) of the hip and knee on different radiographic views, to determine whether different radiographic views have additional value in detecting early hip and knee radiographic OA cases or progression
Trakada et al. [23]	Pulmonary radiographic findings and mortality in hospitalized patients with lower respiratory tract infections	To identify whether specific radiographic findings in patients with lower respiratory tract infections predict mortality

(continued)

Table 4.2 (continued)

Authors	Title	Purpose
Aktas et al. [24]	Concomitant radiotherapy and hyperthermia for primary carcinoma of the vagina: a cohort study	To evaluate the supplementary value of adding hyperthermia to radiotherapy in patients with primary vaginal cancer
Virtanen et al. [25]	Angiosarcoma after radiotherapy: a cohort study of 332,163 Finnish cancer patients	To evaluate the risk of angiosarcoma after radiotherapy among cancer patients in Finland
Jaremko et al. [26]	Do radiographic indices of distal radius fracture reduction predict outcomes in older adults receiving conservative treatment?	To investigate whether radiographic deformities suggesting inadequate reduction would be associated with adverse clinical outcomes
Case–control studies		
Zhang et al. [27]	Diagnostic radiography exposure increases the risk for thyroid microcarcinoma: a population-based case–control study	A population-based case–control study to investigate whether there is an association between ionising radiation-based medical imaging procedures and incidence of thyroid cancer
Darby et al. [28]	Risk of ischemic heart disease in women after radiotherapy for breast cancer	A population-based case–control study of major coronary events (myocardial infarction, coronary revascularisation, or death from ischemic heart disease) to investigate if there is an increased risk due to receiving radiotherapy for breast cancer
Sernik et al. [29]	Ultrasound features of carpal tunnel syndrome: a prospective case-control study	To examine the most adequate cut-off point for median nerve cross-sectional area and additional ultrasound features supporting the diagnosis of carpal tunnel syndrome (CTS)
Cheng et al. [30]	Yoga and lumbar disc degeneration disease: MR imaging based case control study	To identify whether lumbar disc degenerative disease was reduced in practicing yoga instructors compared to a control group
Spruit et al. [31]	Regional radiotherapy versus an axillary lymph node dissection after lumpectomy: a safe alternative for an axillary lymph node dissection in a clinically uninvolved axilla in breast cancer. A case control study with 10 years follow up	To compare disease-free survival and overall survival in patients with clinically uninvolved axilla undergoing radiotherapy or axillary lymph node dissection following lumpectomy for breast cancer
Finlay et al. [32]	Advanced presentation of lung cancer in Asian immigrants: a case-control study	To determine if Asian immigrants to the USA present with more advanced lung cancer compared to non-Asians

Table 4.2 (continued)

Authors	Title	Purpose
Cross-sectional studies (surveys)		
Nightingale et al. [33]	A national survey of current practices of preparation and management of radical prostate radiotherapy patients during treatment	To gain insight into the variation of radiotherapy practices in the UK, focusing on pre-treatment preparations, on-treatment review, and management of radical prostate cancer patients undergoing radiotherapy
Snaith et al. [34]	A UK survey exploring the assistant practitioner role across diagnostic imaging: current practice, relationships and challenges to progression	An electronic survey of individual assistant practitioners (APs) within the NHS in the UK to explore utilisation, role scope, and aspirations
Goense et al. [35]	Patient perspectives on repeated MRI and PET/CT examinations during neoadjuvant treatment of esophageal cancer	To evaluate the experienced burden associated with repeated MRI and positron emission tomography with integrated CT (PET/CT) examinations during neoadjuvant treatment for oesophageal cancer from the perspective of the patient
Lutz et al. [36]	Survey on use of palliative radiotherapy in hospice care	Hospice professionals were surveyed to assess the need for palliative radiotherapy in the hospice setting
Davies et al. [37]	Radiation protection practices and related continuing professional education in dental radiography: a survey of practitioners in the North-east of England	To survey the opinion of practitioners on the availability of related postgraduate courses in the region
Jones and Manning [38]	A survey to assess audit mechanisms practised by skeletal reporting radiographers	To survey the role of plain film reporting radiographers and the methods they employ to evaluate the quality of their performance
Power et al. [39]	Videofluoroscopic assessment of dysphagia: a questionnaire survey of protocols, roles and responsibilities of radiology and speech and language therapy personnel	To survey videofluoroscopic practice and identify the roles and responsibilities of radiology and speech and language therapy personnel
Studies of diagnostic test accuracy		
Yi et al. [40]	Detection of noncalcified breast cancer in patients with extremely dense breasts using digital breast tomosynthesis compared with full-field digital mammography	To evaluate the tumour visibility and diagnostic performance of digital breast tomosynthesis (DBT) plus full-field digital mammography (FFDM), compared to FFDM alone, in patients with noncalcified T1 breast cancer

(continued)

Table 4.2 (continued)

Authors	Title	Purpose
Wooten et al. [41]	Bedside ultrasound versus chest radiography for detection of pulmonary edema: a prospective cohort study	This study compared the sensitivity and specificity of bedside ultrasound and chest radiography in diagnosing pulmonary edema
Grisaru et al. [42]	The diagnostic accuracy of ^{18}F-Fluorodeoxyglucose PET/CT in patients with gynaecological malignancies	To compare the diagnostic accuracy of PET/CT with standard imaging (CT/MRI/US) in patients with suspected recurrence of gynaecological malignancy
Burling et al. [43]	Virtual colonoscopy: effect of computer-assisted detection (CAD) on radiographer performance	To determine whether CAD as a "second reader" improves polyp detection by trained radiographers reporting on virtual colonoscopy examinations
MERCURY Study Group [44]	Diagnostic accuracy of preoperative magnetic resonance imaging in predicting curative resection of rectal cancer: prospective observational study	To assess the accuracy of preoperative staging of rectal cancer with magnetic resonance imaging to predict surgical circumferential resection margins.
Dai et al. [45]	Does three-dimensional power Doppler ultrasound improve the diagnostic accuracy for the prediction of adnexal malignancy?	To investigate the diagnostic accuracy of 3-D power Doppler ultrasound in the differentiation between benign and malignant adnexal masses
Qualitative studies		
Nightingale et al. [46]	A qualitative analysis of staff-client interactions within a breast cancer assessment clinic	An exploration of the culture of staff–client interactions within a breast cancer assessment clinic, using an ethnographic approach: the impact upon client experience
Nagle et al. [47]	Exploring general practitioners' experience of informing women about prenatal screening tests for foetal abnormalities: a qualitative focus group study	To explore GPs' experience of informing women of prenatal genetic screening tests for foetal abnormality
Poulos and Llewellyn [48]	Mammography discomfort: a holistic perspective derived from women's experiences	To use qualitative research methods to consider discomfort from a holistic perspective of the mammography experience derived from the women themselves
Colyer [49]	The role of the radiotherapy treatment review radiographer	A qualitative study to gain an understanding of the role of the radiotherapy treatment review radiographer

4.4 Basic Concepts in Critical Appraisal

There are some common concepts of critical appraisal of research literature that are relevant to most study designs. Some key pointers to evaluating a piece of published research are indicated below.

- Are there clear aims and objectives?
- Is there a defined research question?
- Do the authors have a good grasp of previous research in this field?
- Is the study relevant to clinical practice and carried out in 'real-world' circumstances?
- Is the method clear and well reported?
- Is the sample group sufficient and representative?
- Have appropriate inclusion and exclusion criteria been defined and used?
- Is the analysis of findings (quantitative and/or qualitative) appropriate and are the results appropriately interpreted?
- Are all study participants accounted for in the analysis?
- Are there any possible sources of bias, identified or unidentified by the authors, and have these been controlled or adjusted for in the analysis?
- Are unexpected events or negative findings discussed?
- Are weaknesses in the study acknowledged by the authors?
- Are the authors balanced in their views and conclusions?
- Are there useful recommendations?

All research should be designed to produce valid results. Validity is concerned with the extent to which inferences can be drawn from a study, in particular generalisations extending beyond the study sample, having taken into account study methods and the representativeness of the study sample. Two types of study validity can be distinguished.

- Internal validity relates to the ability of a research method to show a real relationship between cause and effect, such as whether observed differences in patient outcome can be attributed to the effect of the intervention under investigation.
- External validity is concerned with how generalisable the findings from a study are to a wider population, based on the sample of patients included in the study.

Bias, confounding, and chance can all reduce internal validity and may provide alternative explanations for an observed difference between study groups. Bias is often related to faults in the study design and can arise, for example, from an unrepresentative or skewed selection of patients for a study (selection bias), or a partial or unbalanced collection of data (information bias). To prevent bias, wherever feasible and necessary, good study designs will blind (or mask) the patients, clinicians, and even the researchers so that they are kept ignorant of anything that could lead them to a change in behaviour that might affect study findings.

Confounding occurs when an apparent effect of the intervention on patient outcome is in fact due to the action of a variable other than the intervention. When confounding is known or suspected, it can be controlled for in the design (e.g. randomisation, matching) or in the analysis (e.g. multivariable analysis). The effects of unknown confounders can be reduced by randomisation, but can never be eliminated entirely. A confounder is defined as an additional variable that is related to the dependent variable (e.g. disease or other outcome), but is not a consequence of this outcome. It is also related to the independent variable under study (e.g. intervention or exposure), but is not a consequence of this variable either.

The effect of any intervention can also be explained by chance. Even a randomised trial, which protects against systematic differences between groups, does not prevent differences between samples arising by chance although this does diminish as a sample size increases. The probability of an observed difference occurring by chance when no real difference exists is demonstrated by a *p*-value. A *p*-value of, for example, $P = 0.01$, informs us that assuming there is no real difference between treatments, the probability of uneven randomisation explaining the difference is around 1 in 100. Therefore you would not expect the play of chance to explain your study findings. External validity is likely to be threatened when only a small sample of patients is obtained from a single geographical location or there is self-selection of patients into a study (e.g. volunteers). This is therefore addressed by conducting research at multiple sites, increasing the sample size, and, when possible, selecting a random sample of patients into a study so that every eligible patient has an equal chance of being selected and thus the sample should be representative of the target population. It is often very difficult, if not impossible, to obtain a truly random sample. Many studies therefore use a convenience sample of, for example, consecutive eligible patients attending a department/clinic. In such circumstances particular attention should be given to the representativeness of a sample chosen for a study.

4.5 The Typical Structure of a Research Paper

Most research articles are similarly structured, though the precise structure may vary according to the editorial policy of a journal and the design of a study. The general structure of a published research article is as follows.

- *Title*—Making clear the purpose and design of the study.
- *Authors*—Including names, qualifications, and affiliations.
- *Abstract*—Summarising the background and purpose, structure, results, and conclusions of the study.
- *Introduction*—Presenting the background to the study and its rationale, including reference to previous relevant research.
- *Methods*—Including a thorough description of the study design, an outline of the practicalities of how it was done, an explanation of how potential biases were addressed, and a description of the data analysis methods used.

- *Results*—Presentation of the results, with emphasis on the primary outcome measure identified for the study.
- *Discussion*—Interpretation of findings, recognition of any limitations of the study, the discussion of the findings in the context of what was previously known, and suggested implications for practice.

4.6 Preliminary Steps in a Critical Appraisal

When setting out to identify relevant research in an area of practice, the first task should be a systematic search of the literature using the methods discussed in Chap. 3.

From the outset, it is important to understand that there is no such thing as a perfect research study. Even the best conducted studies have potential flaws that are impossible to avoid. For example, almost all research involving patients or staff needs informed consent from the participants. If those who refuse to give that consent are over-represented in particular subgroups, such as gender, age or ethnicity, then the representativeness of the sample could be open to challenge. Also, all research is subject to logistical and economic constraints and so compromises have to be made when considering what is feasible. It is far easier to criticise the work of others than to design a study that is beyond criticism. It is thus important, when critically appraising a paper, to consider unavoidable constraints within which researchers are working and to assess whether they have implemented all measures reasonably available to them to optimise the robustness of the study. A critical evaluation of a study is not just about finding fault. We should also praise when this seems appropriate.

Another issue to bear in mind is that there is a difference between the assessment of the method and findings of a research article and the assessment of the written presentation of that article, although both are important. Students often focus too much on the presentation of a research study when evaluating it, leading to a critique which is descriptive and uncritical. Reports of studies of high inherent quality may be poorly presented by the authors, meaning that some information may be lacking and a fair assessment of study quality is hard to undertake. Conversely, a weak study could be well presented, with strong structure and great detail, and yet could contain flaws so significant that no meaningful inferences can be drawn from it.

Once you have identified a research article that may be of relevance to you and that you may wish to critically evaluate, there are a few preliminary steps and questions that should be considered before progressing further.

1. A reading of the abstract may clearly identify whether or not the paper is relevant to your purpose. If it is still unclear after reading the abstract, a quick reading of the article may be necessary before you are able to make a decision. Is the nature of, and emphasis within the study relevant to the purpose of your literature search and evaluation? Do not spend too much time on articles that are peripheral or irrelevant to your purpose.

2. Does the title accurately reflect the content of the study or is it uninformative or misleading?
   a. The title may give the impression that the study comprises fresh (primary) data, but it may in fact be a review of previously published work. In this case the article may be of help to you in appraising some of the other pieces of published research to which it makes reference, but this is no substitute for your own independent assessment of the original studies.
   b. Is the study measuring the outcome(s) it says it is measuring, or are surrogate outcome measures being used? (A surrogate outcome measure is one that is presumed—with or without good evidence—to be associated with the primary outcome of interest, but is usually easier to measure.)
3. Does the list of authors suggest that they have the relevant expertise in all important aspects of the research? You should never assume that eminence in a particular field guarantees the quality of the research, nor that an unknown author, or an author from a different discipline should not be trusted or believed. All research should be appraised on its merits, but extra vigilance in the appraisal of the robustness of the research may be suggested where certain relevant expertise may appear to be lacking, for example the absence of a medical statistician from the list of authors of a paper that utilises seemingly complex data analysis methods.
4. Is the study design what it says it is? Not all studies reported as RCTs are randomised or adequately controlled; some studies reported as cohort studies could more accurately be described as cross-sectional studies. The answer to this question is not always clear-cut and the paper may require more thorough evaluation before it can be definitively answered.
5. Has the paper been commented upon already? Peer-reviewed journals normally include a letters section, in the printed edition and/or online, within which members of the health/scientific community pass informed comment on research previously published in the journal. In the online content pages of peer-reviewed journals, letters commenting on the research are often identified adjacent to the original article. It is always worthwhile to read the published views of other commentators on a research article, though of course these comments themselves should be subject to critical appraisal.
6. In the introduction to the paper, have the authors adequately identified and summarised the available evidence in the relevant subject area and justified the need for their own study? The Declaration of Helsinki, which governs the ethics of biomedical research, requires that research involving people should be underpinned by a thorough knowledge of the scientific literature in order that research volunteers are not subject to unnecessary harm or inconvenience.

4.7 Critical Evaluation Strategies According to Design Method

We next consider the specific requirements for a critical evaluation of studies comprising the designs illustrated in Table 4.2. Some key resources have been identified to assist students, and qualified practitioners alike, in performing the

evaluation, including reference to key publications explaining the rationale for giving attention to specific aspects of the study design and an evaluation tool/checklist that provides a pro forma for a systematic evaluation. A few of the key issues for each study design are briefly outlined, but a more thorough explanation of the importance of each issue is provided in the essential resources indicated.

4.7.1 Critical Evaluation of Systematic Reviews

- *Useful resources*
- The PRISMA Statement (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) [50]; an appropriate checklist from Appendix.
- *Specific issues to consider*
  The purpose of a systematic review is to help healthcare providers and other decision-makers to make clinical decisions about best practice. Rather than reflecting the views of the authors or being based on a possibly biased selection of published literature, a systematic review involves locating all the available evidence in relation to a specific research question, appraising the quality of the evidence identified, synthesising the available evidence, and if relevant, statistically aggregating the evidence of all relevant studies. Systematic reviews, and the statistical meta-analytic methods they use, were originally developed to synthesise the results from several homogeneous randomised controlled trials. In today's healthcare environment, they have a much broader application and can incorporate more heterogeneous RCTs and observational and qualitative studies, respectively. Systematic reviews should adhere to strict scientific design in order to make them more comprehensive and to minimise the chance of bias (systematic errors) and random errors (mistakes occurring by chance), thus providing more reliable results from which to draw conclusions and make decisions. The following should therefore be considered when critically appraising the quality of a systematic review.
- *Research question*
  What question did the systematic review address? The main research question should be clearly stated and preferably describe the relationship between population, intervention (or test or exposure), comparison intervention, and outcome (PICO). Knowing the population is important to decide whether the review applies to your specific patient group. The intervention is a planned course of action and the exposure something that happens. These again need to be described in detail, as should the comparison intervention, to ensure clarity and to help you determine what contributed to the outcome. The most important outcomes, beneficial or harmful, should also be clearly defined. The title, abstract, or final paragraph of the introduction should clearly state the research question. See Chap. 2 for more guidance on finding and formulating a research question.
- *Searching*
  Is it unlikely that important, relevant studies were missed? The information sources searched should be clearly described (e.g. databases, registers, personal files, expert informants, hand-searching) and any restrictions (e.g. years considered, publication

status, language of publication). A comprehensive search for all relevant studies should include the major bibliographic databases (e.g. Medline, EMBASE, Cochrane), a search of reference lists from relevant studies, contact with experts to inquire about, in particular, unpublished studies, and the search should, ideally, not include English language only. The search strategy should be clear, explicit, and reproducible and be described in the methods section of the paper.

- *Study selection*
  Were the criteria used to select articles for inclusion appropriate? The inclusion and exclusion criteria (defining population, intervention, principal outcomes, and study design) should be clearly defined before the search is undertaken to ensure the consistent and appropriate selection of eligible studies into the review. The methods section should describe in detail these criteria.
- *Validity assessment*
  Were the included studies sufficiently valid for the type of questions asked? There should be predetermined criteria used to assess the quality (e.g. randomisation, blinding, completeness of follow-up) of each included study depending on the type of clinical question being asked. The process of assessing validity should also be described, for example, masking the reviewers to who were the authors of the study and whether two reviewers independently applied the quality criteria. The methods section should describe the quality criteria used and the process of applying the criteria. The results section should provide information on the quality of studies and, if applicable, extent of agreement between reviewers when appraising studies.
- *Study characteristics*
  Were the study characteristics similar? The type of study design, participants' characteristics, details of intervention, and outcomes should be described. Heterogeneity, or inconsistency of results across different studies, could be explained by differences in study characteristics. The possibility of heterogeneity should be explored visually through the examination of forest plots of the results of studies or, more formally, with statistical tests such as chi-square (see Chap. 15 for common statistical tests).
- *Data synthesis*
  Were the methods used to combine the findings of the relevant studies reported (if applicable)? The principal measures of effect (e.g. relative risk), method of combining results (statistical testing and confidence intervals), and a priori sensitivity and subgroup analyses should all be reported in the methods section and the findings in the results.

4.7.2 Critical Evaluation of Randomised Controlled Trials

- *Useful resources*
  The CONSORT 2010 Statement (Consolidated Standards of Reporting Trials) and associated resources [51]; an appropriate checklist from Appendix.
- *Specific issues to consider*

A randomised controlled trial is usually regarded as the strongest type of primary research study in health care, although it may not be feasible in all situations and for some types of research question it is not appropriate. For example, determination of the prevalence of a particular disease in a population requires a well-designed survey. In an RCT, subjects (usually patients) are allocated in a controlled but random way to two or more groups, receiving different interventions. An RCT is generally the best available design to test (a) whether a medical intervention works at all (e.g., a drug, surgical technique, exercise regimen, radiotherapy treatment, or diagnostic screening test), by comparing outcomes of the intervention group with the placebo or control group; (b) whether a new intervention is superior to existing treatment, by comparing outcomes with a group receiving standard care; or (c) whether a new, cheaper, or less invasive intervention is equivalent in its effect to the current expensive or invasive procedure.

- *Participants*
  What were the eligibility criteria for participants? What were the settings (primary care, secondary care, community) and geographical locations from which recruitment was made? What were the specific inclusion and exclusion criteria and were they appropriate? Were participants randomly selected (how subjects are randomly allocated is discussed later) or was a convenience sample used (e.g. all consecutive patients over a 1-month period)? Are the sample characteristics representative of the population of patients in whom you are interested, for example in terms of age, gender, ethnicity, socioeconomic characteristics, disease type, and severity? If not, caution may be advised in generalising the results to your patient population.
- *Interventions*
  Is the precise nature of the experimental and control interventions clear? How and when were the interventions administered? Did the control group receive a placebo or standard care? Was this a comparison of a new intervention compared to standard care or a new intervention in addition to standard care?
- *Outcome measures*
  Ideally there should be only one primary outcome measure, though occasionally more than one may be justified. Several secondary outcome measures may also be identified, but should be interpreted with more caution. Were the outcome measures adequately defined and accurately measured? Were they measured just once or repeated measures made over time? If the latter, then there will be important statistical issues to consider. Did the primary outcome measure evaluate the real concept of interest or were surrogate outcome measures used?
- *Sample size*
  Was an appropriate prospective sample size estimation undertaken? If so, was previous research used to estimate a likely effect size (true difference in outcomes between the human groups included in the trial) or was a judgement made regarding the minimum effect size that would represent a clinically important effect? If no sample size estimation was undertaken, then there is a serious risk of the study being overpowered (an unnecessarily large sample) or underpowered (too small a sample to make any valid findings).

- *Randomisation*
  Was the randomisation method adequately described and was it open to abuse? Were random number tables used, or a computer-generated random number sequence? Was simple randomisation used or a restricted method, for example random number blocks, stratification, or minimisation (a method used to minimise differences in baseline characteristics between groups)? Was the group allocation of all participants adequately concealed?
- *Blinding*
  Ideally, the group allocation of all participants should remain unknown to participants and to those responsible for the administration of the intervention and data collection and for their general medical care, until after the data are analysed. Sometimes this is very difficult and may at times not be logistically feasible. Lack of blinding, or its inadequacy, can in some circumstances seriously compromise the validity of a study (due to complex psychological issues affecting both patients and those responsible for their care), but in other circumstances it may be of limited importance (e.g., lack of blinding of a patient is unlikely to seriously compromise a study evaluating the diagnostic accuracy of alternative tests, since the outcome measure relates to observer interpretation of imaging signs rather than to the degree of improvement in the health status of the patient). Were all reasonable steps taken to ensure adequate blinding? What more could have been done?
- *Statistical methods*
  Were appropriate statistical methods chosen to analyse all outcome measures? For simple analyses the answer to this question should be within the scope of all readers. Although RCTs can be complex to undertake, the statistical methods chosen for their analysis (at least that of the primary outcome measure) are usually relatively simple because the groups should be fairly well balanced on all factors that may affect outcome, apart from the intervention group to which they have been assigned. More complex methods may be used for some secondary outcome measures. The primary analysis of a clinical trial should be based on 'intention to treat'. In other words, patients should be analysed within the group to which they were randomly allocated rather than according to the treatment they may actually have received.
- *Results*
  Was the flow of participants through each stage of the trial made clear? Were all important baseline characteristics of participants summarised and were they very similar between trial groups? Were all participants accounted for, with the number of dropouts evaluated and reasons given for all missing data? Were the results of statistical analyses adequately reported (effect size, confidence intervals (where possible), and statistical significance)? Were secondary and further exploratory analyses identified as such? Was an appropriate account taken of multiple analyses in determining the threshold for statistical significance?
- *Interpretation*
  Are the researchers' claims justified by their results, in the context of what is already understood from previous research? Were the limitations of the study (in

terms of inclusion criteria, uncontrolled potential biases, sample size, and precision) adequately recognised by the authors? Is the evidence presented sufficiently strong to confirm, or warrant reconsideration of, current practice?

4.7.3 Critical Evaluation of Observational Studies

- *Essential resources*
- The STROBE Statement (Strengthening the Reporting of Observational Studies in Epidemiology) [52]; an appropriate checklist from Appendix.
- Specific issues to consider are presented below.

4.7.3.1 Cohort Studies

- A cohort is a group of people with shared or common characteristics for the purpose of health research and is often followed longitudinally over time. As in the case of an RCT, groups within the cohort (sample) are compared with one another. The main difference between an RCT and a cohort study is that in the latter, subjects are not allocated at random to interventions or exposures. This lack of random allocation makes it harder to eliminate or control biases due to systematic baseline differences between cohort subgroups to be compared. Otherwise, the characteristics of cohort studies are similar to RCTs. A cohort study is usually the best available study design in situations where an RCT is either unethical or impractical. Cohort studies are not the most efficient design for studies investigating rare occurrences or diseases with long latency periods.
- *Participants*
  Settings, locations, and periods of recruitment, follow-up, and data collection should all be stated. What were the eligibility criteria for inclusion and were they appropriate? If two or more sub-cohorts were compared, might there be any other systematic differences between them (e.g., different prior information, recruited at different times)?
- *Exposure*
  What was the nature of the "exposure", how was it measured, and how did it vary across the cohort? Was it measured reliably? In most cohort studies, the exposure consists of some agent which the subject physically receives, for example, a vaccine, drug, other medical intervention, or an environmental toxin such as a radiation exposure or inhalation of some toxic chemical agent. In many medical imaging studies, such as those by Trakada et al. [23] and Jaremko et al. [26] in Table 4.2, the role of the exposure is taken by imaging findings because we want to assess the degree to which the imaging appearances can predict patient outcome.
- *Outcome measures*
  Was a primary outcome measure adequately defined and was it appropriately and adequately measured? What were the additional outcome measures? If the study was longitudinal (repeated measurements over time), was a specific time point

identified as the primary time point or was the trend over time of primary interest?

- *Other variables*
  Unlike the case with RCTs, in a cohort study we cannot be assured of reasonable balance between groups in a cohort study, so baseline differences between groups may need to be accounted for in the analysis. There may be a number of potential confounders (other variables associated with both the exposure and the outcome measure) that need to be adjusted for in the analysis. All variables of importance in a study, their method of measurement/determination, and their role (measure of exposure, outcome measure, or confounder) should be identified.
- *Sample size*
  The same considerations are applicable as for RCTs, but the methods of estimation could potentially be more complex due to a necessarily more complex statistical analysis.
- *Control of biases*
  Did the authors identify all serious potential sources of bias in the study and make all reasonable efforts to control them?
- *Statistical methods*
  In some cohort studies, the analysis methods used can be quite straightforward, but often various types of regression model are required to accommodate repeated measures on individuals and/or adjustment for confounders. The authors should explain clearly the nature of the analyses proposed.
- *Results*
  All relevant details relating to the recruitment of participants should be reported, including the total number of people eligible for participation, the numbers declining consent, any missing data, and the numbers lost to follow-up. Actual numbers, rather than just percentages, should be reported. The analysis process should be adequately described, including unadjusted and adjusted estimates, and the confounders adjusted for. Effect size and measures of uncertainty should be presented as well as statistical significance.
- *Interpretation*
  As for RCTs, but potential limitations due to uncontrolled biases require even more careful consideration.

4.7.3.2 Case–Control Studies

Case–control studies involve comparing people with a disease or characteristic (the cases) with otherwise similar people who lack that disease or characteristic (the controls). These studies have proved very useful for investigating cause and effect, for example, linking smoking with lung cancer. They are most appropriately used in situations where a disease process being investigated is rare. They are however more prone to hidden biases than cohort studies. A cohort study in such cases would need to be inordinately large to ensure that sufficient cases of disease were included in order to effect comparisons between subgroups. In a case–control study, the cases of disease are identified first; appropriate controls are then selected for comparison, and the focus is on a comparison of an exposure of interest between the two groups.

Direct inference of causation cannot be made from case–control studies because our starting point is the identification of cases that already have the disease of interest.

- *Participants*
  Particular care is required in explaining how case ascertainment was determined because misclassification is a serious potential bias in studies of this type. Suitable controls are often also problematic to recruit. A control group should be similar in all its characteristics to a case group except with regard to their disease status and, potentially, their 'exposure'. Were there equal numbers of cases and controls or are two or more controls recruited for each case? Were controls matched or unmatched to cases? If matched, what were the matching criteria?
- *Exposure, outcome, other variables, sample size, and control of biases*
  As for cohort studies
- *Statistical methods*
  As for cohort studies. An additional issue for case–control studies arises when the cases and controls are matched. In this case, the matching has to be specifically accounted for in the methods of analysis. For example, McNemar's test should be used to analyse case–control pairs, rather than a simple chi-squared test that compares groups at the aggregate level. If a logistic regression model were to be used for a cohort study, then for a matched case–control study a conditional logistic regression model, which incorporates the matching variables, should be used.
- *Interpretation*
  As for cohort studies.

4.7.3.3 Cross-Sectional Studies

Studies of this type involve a 'snap-shot' investigation of some phenomenon of interest at a particular instant or over a short period of time. In epidemiology, they are often used to ascertain the prevalence of a particular disease at a moment in time in a well-defined geographical area or subject group. Surveys are usually examples of this design and are used widely in studies involving both patients and health professional groups.

- *Participants*
  Were the eligibility criteria for inclusion clearly stated? What were the settings and locations of recruitment? What were the methods of recruitment? Are the characteristics of the sample similar to those of your population of interest? What potential biases are present in the methods of sample selection?
- *Variables*
  In epidemiological studies, this study design is often used to determine the prevalence of a disease in a population of interest. More broadly, surveys can be used to obtain information on a wide and complex range of issues using simple or complex, single or multiple questionnaires. Were all quantitative variables adequately defined and were the measures valid? If a questionnaire was used, has it been previously validated and was it suitable for the purpose for which it was used?

- *Statistical methods*
  Analyses could comprise simple evaluations of prevalence of disease (or other concept of interest), where confidence intervals should also be provided if a random sample of the population of interest is used. Commonly, surveys are based on non-random samples, in which case any statistical inferences should be treated with caution. Many surveys are essentially descriptive in nature, with assessment of responses to a large number of questions. The validity of any statistical comparisons in such circumstances is even more open to question unless efforts were made to minimise the number of formal comparisons and account for multiple testing. It is not possible to ascertain causality from cross-sectional studies.
- *Interpretation*
  As for cohort studies.

4.7.4 Critical Evaluation of Studies of Diagnostic Test Accuracy

- *Useful resources*
  The STARD (Standards for Reporting of Diagnostic Accuracy Studies) 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration [53]. QUADAS-2: A Revised Tool for the Quality Assessment of Diagnostic Accuracy Studies [54]. This is a generic tool used to appraise the quality of primary studies in systematic reviews of diagnostic accuracy. An appropriate appraisal tool from Appendix.
- *Specific issues to consider*
  Diagnostic accuracy studies are integral to the evaluation of new and existing imaging technologies and to the measurement of their ability to distinguish patients with and without the target disorder. Studies that assess the performance (or accuracy) of a medical imaging modality, such as magnetic resonance imaging of the knee, should apply the modality to a prospective and consecutive series of patients with and without the target disease, such as meniscal or ligamentous injury, and then the patients undergo a second gold standard or reference test, such as arthroscopy. The relationship between the results of the imaging modality (or index test) and disease status, as determined by the gold standard, is described using probabilistic measures such as sensitivity (correct abnormal diagnosis of patients with disease) and specificity (correct normal diagnosis of patients without disease). It is important that the results of the gold standard are close to the truth, or the performance of the imaging modality will be poorly estimated.
- *Patient selection*
  Was the setting for the evaluation described? Was the patient spectrum representative of patients who will receive the test in practice? Were selection criteria clearly described? Patient selection processes affect which patients enter a study and this can affect both its internal validity (in that a biased selection of patients could inflate the index test performance) and external validity (in that a

narrow selection of patients could limit the generalisability of the findings). The setting, such as a specialised centre, could be referred rare or problem cases which could affect the prevalence and severity of disease in a patient sample and thus study generalisability. Similarly, an appropriate spectrum of patients should be selected in terms of demographics and clinical features; a limited spectrum can considerably bias the sensitivity and specificity of a test. Predetermined selection criteria should be described to ensure the explicit and reproducible selection of patients into the study.

- *Observer selection*
  Was the effect of the characteristics of observers on test performance considered? Was observer variability determined? The characteristics of the observers involved in the interpretation of images are important in diagnostic accuracy studies of imaging modalities, as they can affect estimates of test performance and generalisability. For example, a study that includes a single, highly specialist observer is likely to have low external validity. In contrast, such an observer could help to produce the best estimates of test accuracy and so increase internal validity. Characteristics of observers that have been considered important in the appraisal of a diagnostic accuracy study include allocation of images to be read by observers; number, experience, and training of observers; profession of observers; and assessment of observer variability and examination of its effect on test accuracy. The variability of an observer, or the reproducibility with which an observer interprets an image, can be assessed as different observers interpreting the same sample of images (interobserver) or the same observers interpreting the same images on separate occasions (intra-observer). The greater the observer variability, the less reliable are the results of the imaging modality (see receiver operating characteristics (ROC) in Chap. 12).
- *Choice and application of the reference (gold) standard*
  Was the reference standard likely to correctly classify the target condition? Was the time period between reference standard and index test short enough to be reasonably sure that the target condition did not change between the two tests? Did the whole sample or a random selection of the sample receive verification using a reference standard diagnosis? Did patients receive the same reference standard regardless of the index test result? Was the reference standard independent of the index test? The reference standard is the method used to determine the presence or absence of the target condition and is assumed to be 100% sensitive and specific. In reality, every test is fallible, but if the reliability of a reference standard is high then methods can be used to account for imperfection. The choice and application of the reference standard is therefore very important in determining estimates of an index test performance. A valid reference standard should be chosen that correctly classifies the target condition and is applied within a clinically acceptable timeframe after the index test to prevent a change in the target condition explaining a difference in the results between the index test and reference standard. The same reference standard should be applied regardless of the results of the index test and preferably to the whole or at least a random sample of patients. Not applying the same reference standard to deter-

mine the definitive diagnosis in the sample of patients could also explain differences in results between an index test and reference standard and thus estimates of test performance. Nor should an index test form part of the reference standard as this too will introduce bias.

- *Independence of interpretation*
  Were the index test results interpreted without knowledge of the results of the reference standard? Were the reference standard results interpreted without knowledge of the results of the index test? Assessments that involve clinical judgement, such as the interpretation of medical images, are susceptible to bias owing to prior expectation. Therefore, the interpretation of the results of a test under evaluation should be undertaken independently, blind to the results of the reference standard. Similarly, the results of a reference standard should be interpreted blind to the results of an index test. Not avoiding this bias may lead to inflated measures of diagnostic accuracy.
- *Measurement of results*
  Were uninterpretable/intermediate test results reported? Were withdrawals from a study explained? Indeterminate index test results might arise due to factors such as technical faults or inferior image quality. Patients might also withdraw from a study before the results of either or both of an index test and reference standard are known. This could be for many uncontrollable reasons such as death, changing residency, or unwilling to continue co-operation. A study should fully report these indeterminate test results and withdrawals. If they are essentially random and not related to the true disease status, they should not introduce bias but could affect generalisability [55].

4.7.5 Critical Evaluation of Qualitative Studies

- *Useful resources*
  Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups [56]; an appropriate appraisal tool from Appendix.
- *Specific issues to consider*
  Qualitative research aims to provide an in-depth understanding of social phenomena such as people's experiences and perspectives in the context of their personal circumstances or settings. To explore the phenomena from the perspective of those being studied, qualitative studies are characterised by the use of unstructured methods which are sensitive to the social context of the study; the capture of data which are detailed, rich, and complex; a mainly inductive rather than deductive analytic process; and answering 'what is,' 'how,' and 'why' questions. It employs a variety of methods including interviews, focus groups, observations, conversation, discourse and narrative synthesis, documentary, and video analysis.
- *Sampling*

Were the criteria for selecting the sample clearly described? Was the sampling strategy comprehensive to ensure the generalisability of the analyses? Were the characteristics of the sample adequately described? As with quantitative studies, it is important for exclusion and inclusion criteria to be clearly specified. This will help you to judge whether the appropriate characteristics of participants according to age, gender, ethnicity, and other relevant demographic features were identified. Unlike quantitative research that requires the selection of a consecutive or random sample of patients that are representative of a population, qualitative research requires the selection of specific groups of people that possess characteristics relevant to the phenomena being studied. Convenience sampling might be used for pragmatic reasons and involves choosing individuals that are easiest to reach, but this might introduce bias. Alternatively, there is purposive sampling when patients/participants are deliberately selected because they possess a certain characteristic and this helps to ensure a range of viewpoints are represented. The characteristics of a sample must be described to help you judge whether an appropriate selection of patients/participants has been included.

- *Data collection*
  Were the data collection methods appropriate for the research objectives and setting? Common methods of data collection include observations, interviews, focus groups, or document analysis. Observation is used to record social phenomena directly by the investigators themselves or indirectly through audiotape or videotape recording. Direct observation requires an investigator to spend time in the social context under investigation and collect data through their nonparticipation or participation in a setting. In nonparticipant observation a researcher does not get involved in the social interactions being observed. It is therefore important to consider whether an observer is likely to be ignored or could inadvertently affect the behaviour of those observed. In participant observation a researcher is part of the social setting, but again it must be considered whether their dual role as observer and participant influences social interactions. Collecting data using interviews might include semi-structured or unstructured individual interviews or may be conducted in focus group settings. Individual interviews are more useful for evoking personal experience, in particular, on sensitive topics; focus groups use group interaction to generate data, but their public forum might inhibit candid disclosure. You should consider the rationale for the choice of a particular method of data collection and its appropriateness for the topics being studied. Finally, analysis of documents such as charts, journals, and correspondence might provide qualitative data. This can be achieved by counting specific content elements (e.g., frequency of specific words being used) or interpreting text (e.g., seeking nuances of meaning). The former rarely provides adequate information for analysis. You should consider whether multiple methods of collecting data are included. This approach can improve the rigour of a study as it allows investigators to examine subjects' perspectives and behaviour from different angles and to capture information with one method that was not possible with another.

- *Validity*
  Are the results of the study valid? This is concerned with whether the data collected truly reflect the phenomena under scrutiny. One method to achieve this is to use triangulation, which refers to the collection of data from different sources using different research methods to identify patterns of convergence. Another approach to validating data is to feed the findings back to the subjects to see if they consider the findings a reasonable account of their experience. There should also be appropriate consideration of 'negative' or 'deviant' cases by a researcher who should give a fair account of these occasions and explore reasons for why the data may vary.
- *Data analysis*
  Were the data appropriately analysed? Qualitative research begins with a general exploratory question and preliminary concepts. Relevant data are collected, patterns observed, and a conceptual framework is developed. This process is iterative, with new data being incorporated that may corroborate or challenge an emerging framework. The process should continue until the framework stabilises. Further data would thus not substantially affect the process. At this point theoretical saturation or informational redundancy is said to have been achieved. Qualitative data, and their interpretation, should be cross-referenced across multiple sources, using triangulation, in order to ensure the robustness of the analysis. Data synthesis should also, ideally, be undertaken by more than one person, and consensus agreement reached, to reduce the risk of researcher bias due to preconceived ideas about the phenomena investigated.

4.8 Conclusions

It is an expectation of all health professionals that they maintain an awareness of relevant research developments in their area(s) of practice in order to inform continuous improvement in patient care. In medical imaging and radiotherapy, rapidly evolving technology continually leads to the refinement of existing diagnostic/therapeutic techniques, and the development of new diagnostic and therapeutic methods.

Evidence-based practice requires the use of current best evidence in making decisions about patient care. This can only be achieved through (a) an understanding of research concepts; (b) an awareness of the characteristics, application, and limitations of commonly used research designs; and (c) an ability to critically appraise and evaluate research evidence in order that appropriate decisions can be made regarding when and how practice should evolve or change.

Key steps in terms of adopting a systematic approach to critical evaluation of the literature are presented. In addition internationally accepted standards, detailing best practice in research design for all commonly used research approaches and methods, are highlighted. Links are also provided to a variety of critical appraisal templates that can be applied to individual research studies, thereby aiding a consistent and systematic approach. References to several professionally relevant

examples of published research studies are provided, for each type of research design, which students, educators, and practitioners can use to practise their critical appraisal skills.

Appendix: Resources for Critical Appraisal

The following resources have been developed to assist medical and health practitioners in the critical appraisal of research appropriate to their practice. The checklists have similarities and some differences, so it is worth exploring a few of them to find a checklist that you think is best suited to your purpose:

- BestBETs (Best Evidence Topics)—critical appraisal worksheets for a wide range of study types: https://bestbets.org/links/BET-CA-worksheets.php (accessed 19 May 2019)
- Boynton PM and Greenhalgh T. Hands-on guide to questionnaire research: Selecting, designing, and developing your questionnaire. BMJ (2004);328:1312–1315.—Checklists for questionnaire design and the critical evaluation of a questionnaire based studies. Table E Critical appraisal checklist for a questionnaire study available at: https://www.bmj.com/content/suppl/2004/05/27/328.7451.1312.DC1 (accessed 19 May 2019)
- Centre for Evidence-Based Medicine (CEBM)—Critical appraisal tools for systematic reviews, RCTs, diagnostic accuracy, prognostic and qualitative studies: https://www.cebm.net/2014/06/critical-appraisal/ (accessed 19 May 2019)
- Critical Appraisal Skills Programme (CASP)—Critical appraisal tools for systematic reviews, qualitative studies, RCTs, cohort, case-control and diagnostic accuracy studies, economic evaluation studies and clinical prediction rules: https://casp-uk.net/casp-tools-checklists/ (accessed 19 May 2019)
- The Scottish Intercollegiate Guidelines Network (SIGN)—Critical appraisal notes and checklists: https://www.sign.ac.uk/checklists-and-notes.html (accessed 19 May 2019)
- The Joanna Briggs Institute (JBI)—Critical appraisal tools for a broad range of study designs: https://www.joannabriggs.org/critical_appraisal_tools (accessed 19 May 2019)

References

1. EFRS statement on radiography education. 2019. The European Federation of Radiographer Societies (EFRS).
2. European Qualifications Framework (EQF). Level 6 benchmarking document: radiographers. 2nd ed. 2018. European Federation of Radiographer Societies (EFRS).
3. Greenhalgh T. How to read a paper: the basics of evidence-based medicine and healthcare. 6th ed. Wiley Blackwell: Hoboken; 2019.
4. Sackett DL, Rosenberg WMC, Gray JAM, et al. Evidence based medicine: what it is and what it isn't. Br Med J. 1996;312:71–2.

5. Sackett DL, Straus SE, Richardson WS, et al. Evidence-based medicine: how to practice and teach EBM. 2nd ed. Edinburgh: Churchill Livingstone; 2000.
6. Daly J, Willis K, Small R, et al. A hierarchy of evidence for assessing qualitative health research. J Clin Epidemiol. 2007;60:43–9.
7. Hicks RJ. Health technology assessment and cancer imaging: who should be setting the agenda? Cancer Imaging. 2004;4(2):58–60.
8. Bossuyt PM, Reitsma JB, Bruns DE, et al. STARD 2015: an updated list of essential items for reporting diagnostic accuracy studies. BMJ. 2015;351:h5527.
9. Younger CWE, Wagner MJ, Douglas C, Warren-Forward H. Describing ionising radiation risk in the clinical setting: a systematic review. Radiography. 2019;25:83–90.
10. Sierinka JC, van Lieshout WAM, Beenen LFM, et al. Systematic review of flexion/extension radiography of the cervical spine in trauma patients. Eur J Radiol. 2013;82:974–81.
11. Gupta T, Kannan S, Ghosh-Laskar S, Agarwal JP. Systematic review and meta-analyses of intensity-modulated radiation therapy versus conventional two-dimensional and/or or three-dimensional radiotherapy in curative-intent management of head and neck squamous cell carcinoma. PLoS One. 2018;13(7):e0200137.
12. Harris KM, Kelly S, Berry E, et al. Systematic review of endoscopic ultrasound in gastro-oesophageal cancer. Health Technol Assess. 1998;2(18):1–134.
13. Bryant J, Picot J, Levitt G, et al. Cardioprotection against the toxic effects of anthracyclines given to children with cancer: a systematic review. Health Technol Assess. 2007;11(27):1–84.
14. Brealey S, Scally A, Hahn S, et al. Accuracy of radiographer plain radiograph reporting in clinical practice: a meta-analysis. Clin Radiol. 2005;60(2):232–41.
15. Gupta S, Maheshwari A, Parab P, et al. Neoadjuvant chemotherapy followed by radical surgery versus concomitant chemotherapy and radiotherapy in patients with stage IB2, IIA, or IIB squamous cervical cancer: a randomized controlled trial. J Clin Oncol. 2018;36(16):1548–55.
16. Brealey SD, DAMASK (Direct Access to Magnetic Resonance Imaging: Assessment for Suspect Knees) Trial Team. Influence of magnetic resonance of the knee on GPs' decisions: a randomised trial. Br J Gen Pract. 2007;57(541):622–9.
17. Bartholomew AL, Denton ERE, Shaw M, et al. A randomised controlled trial comparing lateral skull computerised radiographs with or without a grid. Radiography. 2004;10:201–4.
18. Harrison R, Dey P, Slevin NJ, et al. Randomized controlled trial to assess the effectiveness of a videotape about radiotherapy. Br J Cancer. 2001;84(1):8–10.
19. Sala E, Watson CJE, Beardsmoore C, et al. A randomized controlled trial of routine early abdominal computed tomography in patients presenting with non-specific acute abdominal pain. Clin Radiol. 2007;62:961–9.
20. Ravasco P, Monteiro-Grillo I, Vidal PM, et al. Dietary counseling improves patient outcomes: a prospective, randomized controlled trial in colorectal cancer patients undergoing radiotherapy. J Clin Oncol. 2005;23:1431–8.
21. Slaar A, Karsten IHCM, Ludo FM, Beenena LFM, et al. Plain radiography in children with spoke wheel injury: a retrospective cohort study. Eur J Radiol. 2015;84:2296–300.
22. Damen J, Runhaar J, Kloppenburg M, et al. Additional value of different radiographic views on the identification of early radiographic hip and knee osteoarthritis and its progression: a cohort study. Arthritis Care Res. 2017;69(11):1644–50.
23. Trakada G, Pouli A, Goumas P. Pulmonary radiographic findings and mortality in hospitalized patients with lower respiratory tract infections. Radiography. 2006;12:20–5.
24. Aktas M, de Jong D, Nuyttens JJ, et al. Concomitant radiotherapy and hyperthermia for primary carcinoma of the vagina: a cohort study. Eur J Obstet Gynecol Reprod Biol. 2007;133:100–4.
25. Virtanen A, Pukkala E, Auvinen A. Angiosarcoma after radiotherapy: a cohort study of 332,163 Finnish cancer patients. Br J Cancer. 2007;97:115–7.
26. Jaremko JL, Lambert RGW, Rowe BH, et al. Do radiographic indices of distal radius fracture reduction predict outcomes in older adults receiving conservative treatment? Clin Radiol. 2007;62:65–72.
27. Zhang Y, Chen Y, Huang H, et al. Diagnostic radiography exposure increases the risk for thyroid microcarcinoma: a population-based case–control study. Eur J Cancer Prev. 2015;24(5):439–46.

28. Darby SC, Ewertz M, McGale P, et al. Risk of ischemic heart disease in women after radiotherapy for breast cancer. N Engl J Med. 2013;368:987–98.
29. Sernik RA, Abicalaf CA, Pimental BF, et al. Ultrasound features of carpal tunnel syndrome: a prospective case-control study. Skelet Radiol. 2008;37:49–53.
30. Cheng T-C, Jeng C-M, Kung C-H, et al. Yoga and lumbar disc degeneration disease: MR imaging based case control study. Chin J Radiol. 2008;33:73–8.
31. Spruit PH, Siesling S, Elferink MAG, et al. Regional radiotherapy versus an axillary lymph node dissection after lumpectomy: a safe alternative for an axillary lymph node dissection in a clinically uninvolved axilla in breast cancer. A case control study with 10 years follow up. Radiat Oncol. 2007;2:40.
32. Finlay GA, Joseph B, Rodrigues CR, et al. Advanced presentation of lung cancer in Asian immigrants: a case-control study. Chest. 2002;122(6):1938–43.
33. Nightingale H, Conroy R, Elliott T, Coyle C, Wylie JP, Choudhury A. A national survey of current practices of preparation and management of radical prostate radiotherapy patients during treatment. Radiography. 2017;23:87–93.
34. Snaith B, Harris MA, Palmer D. A UK survey exploring the assistant practitioner role across diagnostic imaging: current practice, relationships and challenges to progression. Br J Radiol. 2018;91:20180458.
35. Goense L, Borggreve AS, Heethuis SE, van Lier ALHMW, van Hillegersberg R, Mook S, et al. Patient perspectives on repeated MRI and PET/CT examinations during neoadjuvant treatment of esophageal cancer. Br J Radiol. 2018;91:20170710.
36. Lutz S, Spence C, Chow E, et al. Survey on use of palliative radiotherapy in hospice care. J Clin Oncol. 2004;22:3581–6.
37. Davies C, Grange S, Trevor MM. Radiation protection practices and related continuing professional education in dental radiography: a survey of practitioners in the north-east of England. Radiography. 2005;11:255–61.
38. Jones HC, Manning D. A survey to assess audit mechanisms practised by skeletal reporting radiographers. Radiography. 2008;14(3):201–5.
39. Power M, Laasch H, Kasthuri RS, et al. Videofluoroscopic assessment of dysphagia: a questionnaire survey of protocols, roles and responsibilities of radiology and speech and language therapy personnel. Radiography. 2006;12(1):26–30.
40. Yi A, Chang JM, Shin SU, Chu AJ, Cho N, Noh D-Y, et al. Detection of noncalcified breast cancer in patients with extremely dense breasts using digital breast tomosynthesis compared with full-field digital mammography. Br J Radiol. 2019;92:20180101.
41. Wooten WM, Shaffer LET, Hamilton LA. Bedside ultrasound versus chest radiography for detection of pulmonary edema: a prospective cohort study. J Ultrasound Med. 2019;38:967–73.
42. Grisaru D, Almog B, Levine C, et al. The diagnostic accuracy of 18F-Fluorodeoxyglucose PET/CT in patients with gynaecological malignancies. Gynecol Oncol. 2004;94:680–4.
43. Burling D, Moore A, Marshall M, et al. Virtual colonoscopy: effect of computer-assisted detection (CAD) on radiographer performance. Clin Radiol. 2008;63:549–56.
44. MERCURY Study Group. Diagnostic accuracy of preoperative magnetic resonance imaging in predicting curative resection of rectal cancer: prospective observational study. Br Med J. 2006;333:779.
45. Dai S, Hata K, Inubashiri E, et al. Does three-dimensional power Doppler ultrasound improve the diagnostic accuracy for the prediction of adnexal malignancy? J Obstet Gynaecol Res. 2008;34(3):364–70.
46. Nightingale JM, Murphy F, Eaton C, Borgen R. A qualitative analysis of staff-client interactions within a breast cancer assessment clinic. Radiography. 2017;23:38–47.
47. Nagle C, Lewis S, Meiser B, et al. Exploring general practitioners' experience of informing women about prenatal screening tests for foetal abnormalities: a qualitative focus group study. BMC Health Serv Res. 2008;28(8):114.
48. Poulos A, Llewellyn G. Mammography discomfort: a holistic perspective derived from women's experiences. Radiography. 2005;11(1):17–25.
49. Colyer H. The role of the radiotherapy treatment review radiographer. Radiography. 2000;6(4):253–60.

50. Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA). http://www.prisma-statement.org/. Accessed 17 May 2019.
51. CONSORT 2010 Statement (Consolidated Standards of Reporting Trials). http://www.consort-statement.org/. Accessed 17 May 2019.
52. The STROBE Statement (Strengthening the Reporting of Observational studies in Epidemiology). https://www.strobe-statement.org/index.php?id=strobe-home. Accessed 17 May 2019.
53. STARD (Standards for Reporting of Diagnostic Accuracy Studies) 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. Access via the EQUATOR network (Enhancing the Quality and Transparency of Health Research). http://www.equator-network.org/reporting-guidelines/stard/. Accessed 17 May 19.
54. Whiting PF, Rutjes AWS, Westwood ME, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155:529–36.
55. Kelly S, Berry E, Roderick P, et al. The identification of bias in studies of the diagnostic performance of imaging modalities. Br J Radiol. 1997;70:1028–35.
56. Consolidated criteria for reporting qualitative research (COREQ): a 32-item checklist for interviews and focus groups. Access via the EQUATOR network (Enhancing the Quality and Transparency of Health Research). http://www.equator-network.org/reporting-guidelines/coreq/. Accessed 17 May 2019.

Further Readings

Azzam Al-Jundi A, Sakka S. Critical appraisal of clinical research. J Clin Diagn Res. 2017;11(5):JE01–5.

How to read a paper – links to articles by Trisha Greenhalgh on the BMJ website. https://www.bmj.com/about-bmj/resources-readers/publications/how-read-paper. Accessed 19 May 2019.

Williams V, Boylan A-M, Nunan D. Critical appraisal of qualitative research: necessity, partialities and the issue of bias. https://doi.org/10.1136/bmjebm-2018-111132.

Reflective Practice and the Patient Voice

5

Pauline J. Reeves

5.1 Introduction

There is a growing body of evidence to suggest that healthcare practitioners are not readily integrating research findings into their clinical practice [1], even those who hold postgraduate qualifications [2, 3]. A study of sonographers found that despite having positive attitudes towards research, practitioners did not feel that their university courses prepared them to undertake research themselves [2]. Even those in radiography consultant posts cited barriers, such as lack of time and high clinical workloads, as reasons (some acknowledged that they were excuses) for not doing research [3].

5.2 Research Methods and Their Contribution to Course Outcomes and Professional Development

In this chapter a number of skills, which link clinical radiographic practice to research utilisation and the study of research methods, are outlined. For most people, the acquisition of these skills is likely to come as part of the pursuit of particular qualifications as you train as a radiographer, then climb up the career ladder. The different levels of radiographer education, and how they relate to research methods and related skills, are summarised below. The numbers in brackets relate to Scottish Higher Education [4]. These level descriptors are outlined together with relevant extracts from the UK Society and College of Radiographers' (SCoR) Education and Career Framework [5]. It is clear that an understanding of research, and research methodology, is expected at whatever educational level you may be studying.

P. J. Reeves (✉)
Department of Medical Imaging, Sheffield Hallam University, Sheffield S10 2BP, South Yorkshire, UK
e-mail: P.Reeves@shu.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_5

- *Level 4 (7) Year 1 undergraduate/Cert HE-Assistant Practitioner*
  Have a sound knowledge of the concepts and principles underlying radiography together with the ability to evaluate both quantitative and qualitative data in order to develop lines of argument, to reflect on and learn from experience within own scope of practice [4].
- *Level 5 (8) second year undergraduate/DipHE/Foundation degree-Assistant Practitioner*
  Display knowledge and the ability to apply concepts and principles in a radiographic context. Students should be able to initiate and undertake critical analysis of information and demonstrate knowledge of the main methods of enquiry used within radiography and radiology [4].
- *Level 6 (9/10) BSc (Hons)-Radiographer (Autonomous Practice)*
  To make use of scholarly reviews and primary sources in order to extend the understanding of the complex body of knowledge in radiography. To initiate and carry out projects, by engaging in audit, research and continuous professional development. Contribute to the development of professional practice for the patient benefit [4, 5].
- *Level 7 (11) PgCert/PgDip/MSc-Advanced Practitioner*
  Understand how the boundaries of knowledge are advanced through research. Engage in audit and research and disseminate the outcomes. Use conceptual understanding to evaluate critically current research, advanced scholarship and methodologies [4, 5].
- *Level 8 (12) PhD- Consultant Radiographer*
  Initiate and lead audit and research, with dissemination of outcomes via presentation or publication. Be responsible for the creation and interpretation of new knowledge which is at the forefront of professional practice, through original research or other advanced scholarship, of a quality to satisfy peer review [4, 5]

These outline descriptors help to demonstrate the importance of study and active utilisation of research to the ability to achieve a degree qualification, whether that be at foundation (assistant practitioner), honours (autonomous practitioner), master's (advanced practitioner) or doctorate (consultant practitioner) level.

Having outlined how research skills are relevant to all levels of radiographic practice, we examine the meaning of the term 'reflective practice' and then outline a research method, the 'critical incident' technique, which can be used to pick out key events that can help and encourage you to reflect clinically, even at a very early stage in training.

We also look at the need to be critical about written research. This leads to a definition of the term 'evidence-based practice' and what this means in radiography; also covered is 'reflexivity' and we contrast it with clinical reflection. How and why it should be used in the teaching of research methods to student radiographers is discussed. These are presented as a developmental series of skills. The overall aim of this chapter is to demonstrate how knowledge of research methods and the associated skills can be the key to both your undergraduate and postgraduate development as a fully autonomous professional radiographer and, in turn, how these skills may be applied to the determination of the needs of patients and the articulation of the 'patient voice'.

5.3 Reflective Practice

Reflective practice has been identified as a skill, which is relevant at all levels of practice within the profession, from assistant practitioner to consultant radiographer. These levels and skills were outlined in the previous section. The levels of knowledge and understanding of reflective practice should be such as to allow a practitioner to use research evidence and experience as tools to aid reflection, with a view to encouraging their own professional growth [6].

The process of reflection helps you to resolve apparent inconsistencies between theoretical teaching and what occurs in clinical practice; something which troubles students, especially in the early stages of training [7]. Many undergraduate courses in the United Kingdom (UK), and overseas [6], require some form of written reflection in the form of portfolios or reflective journals. This requirement is also part of continuous professional development (CPD) after qualification [6, 8].

The stimulus for written reflection may come from any number of events: for example, changes in rotation to include a new work area, a promotion, a specific interaction with a patient or member of staff or from a news item or journal article [7]. For example, an Australian research study examined diagnostic radiographers' communication skills from patients' perspectives [9]; such an article can be used to provide a new viewpoint on a critical area of personal performance.

Reflection on practice allows you to create your own body of theoretical experiential knowledge to underpin your clinical skill base and the intuitive aspects of clinical work. One way to start out is to recall significant (critical) negative or positive events and to focus on those as a learning tool. As radiography practice develops and the working environment becomes more stressful [10], we may argue that the need to critically examine our practice by focussing on certain events becomes more pressing.

Once you have described the incident itself, try and write down your reactions and feelings about what happened. Why were things different than you expected? This may lead you to conclude that there is knowledge missing; this may be behavioural (as with the example of communication skills given above) or you may conclude that you have a gap in your theoretical knowledge, which should then lead you to investigate (research) that gap in your knowledge.

5.4 Critical Incident Technique as a Tool for Research and/or Reflection

Critical incident technique (CIT) has been defined as: 'A "focussed" overview of factors, events, behaviours or experiences that result in satisfaction/dissatisfaction with care or that promote or detract from good quality delivery of care, and of why those things influence satisfaction with or quality of care' [11].

Data may be collected via questionnaires, interviews or by observer reports [11]. In nursing research CIT has typically been used to highlight aspects of best and worst practice by asking respondents to recall and write about critical incidents themselves [11]. The advantages of focussing on specific experiences, behaviours

or activities are that atypical incidents facilitate recall and allow participants to clarify both their perceptions and feelings about the incidents identified [11]. It can be a very effective tool in this respect, but has not been commonly used in radiography studies so far. It also has the advantage of encouraging you to focus on good practice, rather than the tendency to recall only negative events.

5.4.1 Using Critical Incident Technique

CIT was first described by John Flanagan in 1954 [11]. He stated that there were five steps in the use of the technique.

1. Determining the general aim of the study.
2. Planning and specifying how data will be collected.
   This can be done using interviews; the simplest and, arguably, most effective method is by using an open questionnaire. A response sheet can be relatively simple to devise, as shown in Fig. 5.1, depending on how many incidents you wish to collect. This is one of the advantages of the technique; the response sheet is easy to design, as opposed to the use of structured questionnaires. Use of a questionnaire allows respondents to bring up difficult or sensitive incidents, which they may not have been willing to reveal in a face-to-face interview.
3. Collecting the data
   Respondents are asked to consider, and then describe, events or behaviours that help or hinder in the activity, which forms the focus of a study [11]. Such events may include both positive and negative examples, with respondents being asked to describe up to four events.
4. Analysing the data
   The data gathered in step 3 above can be analysed by grouping the different incidents together to form overarching themes, or categories, which relate to the aims of the study [11] (See Chap. 16). One way to achieve this is manually, using coloured highlighter pens to identify keywords. Various software programmes are available for those working with large volumes of data.
5. Interpreting and reporting the outcomes
   Further analysis of the overarching themes should eventually enable the development of a list of factors, which contribute to the phenomenon under study.

Good	
Good	
Bad	
Bad	

Fig. 5.1 Example of simple response sheet design

5.4.2 Discussion

Research suggests that reflection on specific events is easier than attempting to reflect on practice overall [7]. Undergraduate programmes typically require students to do a research project, or dissertation, but they are often not permitted to collect data directly from patients or qualified staff. The use of the CIT encourages individual students to reflect on practice, and is also a way to generate anonymised, clinically related information within a university setting (see Fig. 5.2). It enables participants both to reflect on practice and to link that clinical reflection to the practice of research. Critical incidents can be very powerful, and provide a means by which tacit, implied knowledge can be accessed and scrutinised [7].

At postgraduate level, with ethical approval, CIT may be used with patients as participants. At each level in the career framework outlined by the Society and College of Radiographers [5], a requirement is to provide care that is 'patient-centred'. This has recently been backed up with the publication of a document entitled *Patient public and practitioner partnerships within imaging and radiotherapy: guiding principles* [12]. Each of the core values outlined in the document is written using the patient voice and is illustrated throughout using critical incidents collected from patients regarding their quality of care. The use of reflective practice and research methods, such as CIT, can form major tools in focussing our care directly on the needs of patients. A study based in an oncology centre in Canada [13] suggested writing the phrase 'I demonstrate patient-centred care by…' at the top of a sheet of paper as a tool for reflection and orientation towards the needs and wishes of each patient.

Recent studies used CIT to look at the long-term impact of interprofessional education (IPE) on behavioural outcomes in professional practice [14]. CIT was also used to ask a small group of assistant practitioners to describe incidents in which they did, and did not, feel valued within the healthcare team [15].

Good; During an ultrasound examination, a fetal death was discovered
The radiographer dealt with it in a very sympathetic manner.
Good; I was walking through the department and I saw a therapeutic radiographer and a patient sitting and chatting and when the patient got upset the radiographer just sat and listened and held her hand. I see good patient care all the time but this was raw and unstaged. True empathy and compassion being shown to another human.

Contrast these with the following:

Bad; Failure to ask a patient if they can be touched for repositioning. The need to touch and manhandle a patient rather than actually ask them to adjust themselves. I hate seeing competent and independent patients being helped to undress. Ask patients to pull their trousers down, don't just pull and grab at them
Bad; A ward patient was brought to the department following a total hip replacement for check X-rays. However...the request form was delayed. The patient had to wait an hour before it was realised that she was still waiting for an X-ray. When her form was finally found, the department had become busy and her request form was placed at the bottom of a large pile which resulted in a further wait.

Fig. 5.2 Actual examples of positive and negative critical incidents

5.5 Critical Analysis of Research

Throughout both your course and career you will be expected to search for and appraise research papers. Unfortunately it is the case that the quality of research 'evidence' varies dramatically. Reasons for this have been cited as selective reporting (including the tendency to report positive results but not negative ones) and issues of bias, including the viewpoint of a particular medical speciality. The quality of radiology research has been argued to be poor and very variable [16]. This is partly said to be because of the pace of technologic change and the need for rapid assessment of potential new technologies. Later chapters in this book advise you how to appraise research carefully and what to look for; suffice to say it is important to be critical and not merely accept the veracity of the papers that you utilise in your studies.

5.6 Evidence-Based Practice

In its latest published research strategy document the SCoR states that its primary aim is to 'embed research at all levels of radiography practice and education' in order to 'grow and implement a high quality evidence base' for practice [17]. A study in Norway gave the following definition: 'Evidence-based radiography is informed and based on the combination of clinical expertise and the best available research-based evidence, patient preferences and available resources' [18]. The authors argued that previously radiographers relied on personal experience and tradition and did not believe that radiographic practice was a matter for investigation [18]. A study into research utilisation in ultrasound found that 72.9% (n = 218) of the sonographers surveyed claimed that they did integrate research findings into their clinical practice [2]. The researchers noted, however, that this group largely comprised higher-grade sonographers rather than younger staff and/or recently qualified sonographers.

A more recent paper states that what is actually happening in departments tends more towards practice 'creep' and 'drift' rather than anything based on research evidence [19]. Practice creep is argued to consist of gradual, small changes in practice over time; often this may be due simply to changes in personnel and the way they were taught in particular institutions. These changes may, or may not, be evidence-based and are often not even disseminated within a single department [19]. Such individual variations are clearly illustrated in an ethnographic study that exposed differences in (and reasoning behind) exposure selection amongst radiographers working in digital rooms [20].

Alternately practice drift occurs as staff become more experienced. Staff then begin to lose their primary knowledge and cut corners, especially in busy situations such as on-call. It is argued that this may become more common in diagnostic radiography, as the loss of colleague interactions (and thus quality control) in viewing areas is caused by the move to the relative isolation of working in digital rooms [19].

"Radiation protection, such as gonad shields, lead rubber sheets...were rarely used. I observed that the radiographers are very busy and stressed and they usually 'forget' to use these shields."

"During a chest X-ray on a young male patient, the radiographer did not place lead protection as it was not in sight at the time, despite me offering to go and find one; her response was 'it doesn't matter'".

Fig. 5.3 Examples of critical incidents regarding use of gonad shields

All UK NHS employees are now required to demonstrate their competence annually as part of an individual review [21]. The Health Professions Council (HPC), and the College of Radiographers, have their own mandatory requirements for continuous professional development (CPD), [8, 22] as do other professional bodies worldwide such as the Australian Society of Medical Imaging & Radiation Therapy [23]. Registrants with the HPC are required to keep a portfolio documenting their learning activities and showing that those activities are relevant to current or future practice (e.g. for someone studying to become a sonographer) [22]. These learning activities do not just include attending courses. They can include research and activities such as critical analysis of original journal articles [8].

This step forms the final link in the reflective loop. Reflection on practice begins the process and should lead to identification of gaps in knowledge or clinically related questions that need answering. This in turn leads a practitioner into the research literature, being careful to be critical of what is found in that literature. The final step is to integrate those findings into everyday practice.

This can be illustrated if we look at two similar examples of critical incidents from diagnostic radiography as shown in Fig. 5.3. Reflective practice requires a student to critically appraise the evidence base regarding the placement of gonad shielding. Students do find that there are inconsistencies between what they are taught at university and what they see in clinical practice.

5.7 Reflexivity

At the culmination of any radiography degree course, whether BSc or MSc, you are normally be required to undertake some form of research project. This may include active research but, increasingly at undergraduate level, this is likely to be an extended study or literature review (referred to as a dissertation).

There is often a lot of stress generated in deciding what topic to select for your project. Unlike a doctoral project, the research, whether for BSc or MSc, does not have to be original. The project can be a reworking of another study in a new setting or a review of existing literature. Indeed it can be argued that the actual topic chosen for a research study is, to some extent, immaterial since the aim of such a study is to allow you to demonstrate that you have the ability to carry out research (be it primary/active or secondary/documentary) and to critically analyse the findings.

The term reflexivity is taken to refer to reflection upon a research process itself. The term may be defined as: 'Reflexivity; the project of examining how the researcher and intersubjective elements impact on and transform research' [24].

The concept of reflexivity in research arose from qualitative methods, principally ethnography. Ethnographic methods utilise the concept of researcher as the instrument; a researcher immerses themselves in the setting and collects data via a variety of methods, including participant observation [24]. Many qualitative healthcare research projects may be classified as participant observation since they take place in the workplace. Reflexive critique forms part of a research report in order for a researcher to determine, and make explicit, the effect that they themselves (as a radiographer, including their preconceptions) have had upon the research, since they have been a participant within the setting. Did their participation introduce bias (contamination) into the research (especially where the setting is also a researcher's own workplace)? [25]. One author writes of having to cultivate a 'sense of strangeness' and of having to fight against the familiar by consciously adopting a fresh perspective when carrying out observations within a radiography department [25].

In the context of a student research project the inclusion of a reflective chapter at the end of a project report (dissertation) has the function of requiring you to explicitly and critically audit your own research. Guideline questions here would include: What did you learn from the research process? What would you do differently if you were to approach the research again? Can you identify any flaws in your research? This process readily leads in to recommendations for further research and to the final conclusions.

The inclusion of a reflexive section allows you to demonstrate critical analysis and to show whether you have understood the research process. It can also help those marking the project to differentiate between gradings.

5.8 Conclusions

In this chapter a range of skills, which link clinical practice with research methods, have been reviewed and an attempt made to demonstrate how these skills can help your development as a practitioner and lifelong learner in radiography.

Any radiography student who graduates in the twenty-first century would be equipped with the skills discussed above. You would have been encouraged to reflect on your practice as you developed your clinical skills. You would have been introduced to a variety of research methods and the ability to critically appraise them, as this book sets out to do. You would have been given the understanding and skills of evidence-based practice. The extent to which you use those skills once you have qualified, by choosing to make your clinical practice evidence-based and by pursuing higher qualifications, is likely to determine how far up the career ladder you progress. Mandatory continuous professional development (CPD) requires a radiographer to demonstrate lifelong-learning. However, anyone wishing to become an advanced and/or consultant practitioner needs to make a positive choice regarding the use of research methods and appraisal skills and the drive to implement the

"We had a paediatric patient attending treatment and I was so lucky to be part of such an amazing team treating him. He had the same therapeutic radiographers every day and the effort the staff went to to make him enjoy his treatment. His head shell was painted to create a superhero mask, he had a sticker chart to celebrate each days treatment, the staff would dress up and after day 1 when he was scared to be left on his own they used string for him to hold in the room and his mum held the other end outside of the room.

The whole department celebrated him ringing the end of treatment bell."

Fig. 5.4 Example of true patient-centred care

resultant knowledge into changing clinical practice. Reflective practice, based on the utilisation and dissemination of research, is one of the keys to the higher levels of the profession and to safe and effective radiographic practice centred on the needs of each patient as shown in Fig. 5.4.

Acknowledgements Thanks go to Joanna McNamara for encouraging her students to provide the therapeutic radiography incidents.

References

1. Hendricks J, Cope V. Research is not a 'scary' word: Registered nurses and the barriers to research utilisation. Nordic J Nurs Res. 2017;37(1):44–50.
2. Elliott V, Wilson SE, Svensson J, Brennan P. Research utilisation in sonographic practice: attitudes and barriers. Radiography. 2009;15(3):187–95.
3. Harris R, Paterson A. Exploring the research domain of consultant practice: experiences of consultant radiographers. Radiography. 2016;22(1):e25–33.
4. Quality Assurance Agency. The frameworks for HE qualifications of UK degree awarding bodies. London: QAA; 2014 [Cited 2019 May 3]. https://www.qaa.ac.uk/quality-code/qualifications-and-credit-frameworks.
5. Society & College of Radiographers. Education and career framework for the radiography workforce. London: SCoR; 2013 [Cited 2019 May 3]. https://www.sor.org/learning/document-library/education-and-career-framework-radiography-workforce
6. Abrahams K, Brady C. Graduate reflective practice program: were long-term objectives achieved? J Med Imaging Radiat Sci. 2013;44(4):203–8.
7. Chapman N, Dempsey SE, Warren-Forward HM. Workplace diaries promoting reflective practice in radiation therapy. Radiography. 2009;15(2):166–70.
8. CPD Now: your online portfolio and personal CPD accreditation. London: SCoR. [Cited 2019 May 8]. https://www.sor.org/learning/cpd/cpd-now.
9. Pollard N, Lincoln M, Nisbet G, Penman M. Patient perceptions of communication with diagnostic radiographers. Radiography. 2019;25:333. https://doi.org/10.1016/j.radi.2019.04.002.
10. Brown A. Professionals under pressure: contextual influences on learning and development of radiographers in England. Learn Health Soc Care. 2004;3(4):213–22.
11. Viergever R. The critical incident technique: method or methodology? Qual Health Res. 2019;29(7):1065–79.
12. Society & College of Radiographers. Patient public and practitioner partnerships within imaging and radiotherapy: guiding principles. London: SCoR; 2018.
13. Calisi R, Boyko S, Vendette A, Zagar A. What is person-centred care? A qualitative inquiry into oncology staff and patient and family experience of person-centred care. J Med Imaging Radiat Sci. 2016;47(4):309–14.

14. Graybill E, Heggs A, Truscott S, Vinoski E, Crenshaw M, Crimmins D. Using the critical incident technique to measure long-term outcomes of interprofessional education. J Interprof Care. 2017;31(4):533–6.
15. Jones C, Waller C. Professional integration of APs via critical incident technique. Br J Healthc Assist. 2012;6(12):603–7.
16. Reeves P. Research in medical imaging and the role of the consultant radiographer: a discussion. Radiography. 2008;14(1):e61–4.
17. Society & College of Radiographers. 2016-2021 research strategy. London: SCoR; 2015 [Cited 2019 May 22]. https://www.sor.org/learning/document-library/research-strategy-2016-2021/2016-2021-research-strategy.
18. Hafslund B, Clare J, Graverholt B, Nortvedt MW. Evidence-based radiography. Radiography. 2008;14(4):343–8.
19. Snaith B. Evidence based radiography: is it happening or are we experiencing practice creep and practice drift? Radiography. 2016;22(4):267–8.
20. Hayre C. Cranking up', 'whacking up' and 'bumping up': X-ray exposures in contemporary radiographic practice. Radiography. 2016;22(2):194–8.
21. NHS Employers. Simplified knowledge & skills framework. 2019 [cited 2019 May 3]. https://www.nhsemployers.org/SimplifiedKSF.
22. Health & Care Professions Council. Standards of continuing professional development. London: HCPC; 2018 [Cited 2019 May 3]. http://www.hpc-uk.org/standards/standards-of-continuing-professional-development/.
23. Australian Society of Medical Imaging & Radiation Therapy. Professional practice standards. Sydney: ASMIRT; 2018 [Cited 2019 May 13]. https://www.asmirt.org/media/371/371.pdf.
24. Finlay L, Gough B. Reflexivity-a practical guide for researchers in health and social science. Oxford: Blackwell; 2008.
25. Strudwick R. The radiographic image: a cultural artefact? Radiography. 2014;20(2):143–7.

Ethical Considerations

6

Hesta Friedrich-Nel and Aarthi Ramlaul

6.1 Brief History and Timeline of Research Ethics

The history of ethics can be traced to the time of Socrates (469–399 BC), Plato (427–347 BC) and Aristotle (384–322 BC). Many different views of how best to implement the 'good for all' emerged from BC (before Christ) to AD (Anno Domini) with the advent of Christianity and the biblical commandments. The views continued to develop and evolve through the centuries leading to multiple theories that have shaped the way we perceive ethics today. The need for ethical considerations through formal review gained ground. However, in the nineteenth century experiments on human subjects during World War II (1939–1945) created difficulties. Some of the difficulties were that many participants were not informed that they were part of a research study; they did not provide informed consent [1]. In many instances the vulnerability of a group was used to the advantage of a researcher [2]. In addition, researchers did not as a rule explain the risks associated with the research. Resnik [3] indicated that publications such as the Nuremberg code [4], the Belmont report [5] and the Declaration of Helsinki (first published in 1964) [6] addressed these difficulties by establishing guidelines to protect research participants. The Declaration of Helsinki is recognised as the most authoritative guide on ethical standards for human or clinical research and has been revised several times (1975, 1983, 1989, 1996, 2000, 2008 and 2013) [2].

H. Friedrich-Nel (✉)
Department of Clinical Sciences, Central University of Technology, Bloemfontein, South Africa
e-mail: hfried@cut.ac.za

A. Ramlaul
Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: a.ramlaul@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_6

Aligned with the Declaration of Helsinki, the World Health Organisation (WHO) [7] and the Health Research Authority (HRA) [8] published key principles that researchers need to consider when conducting clinical research. The WHO defines good clinical research practice (GCP) as a process that incorporates established ethical and scientific quality standards for the design, conduct, recording and reporting of clinical research involving the participation of human subjects [7]. Good clinical practice (GCP) is an international quality standard that is provided by the International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use (ICH). This international body defines a set of standards that governments can transpose into regulations for clinical trials involving human participants [7].

Researchers who conduct clinical research need to consider GCP. It gives the assurance that a researcher considers, respects and protects the rights, safety and well-being of research participants [7, 8]. Both the WHO and HRA regard human research as any research project involving individuals in a physical or psychological intervention, observation, collection, storage or dissemination of information. In any of the mentioned circumstances, an individual could be exposed to an unwanted risk.

Figure 6.1 provides a brief overview of the good clinical research principles, which are explained in this chapter. A researcher using human participants must be familiar with and apply these principles in a research project. Research involving human participants, participants who lack capacity, human tissue or radiation, by law will need approval from an appropriately constituted research ethics committee

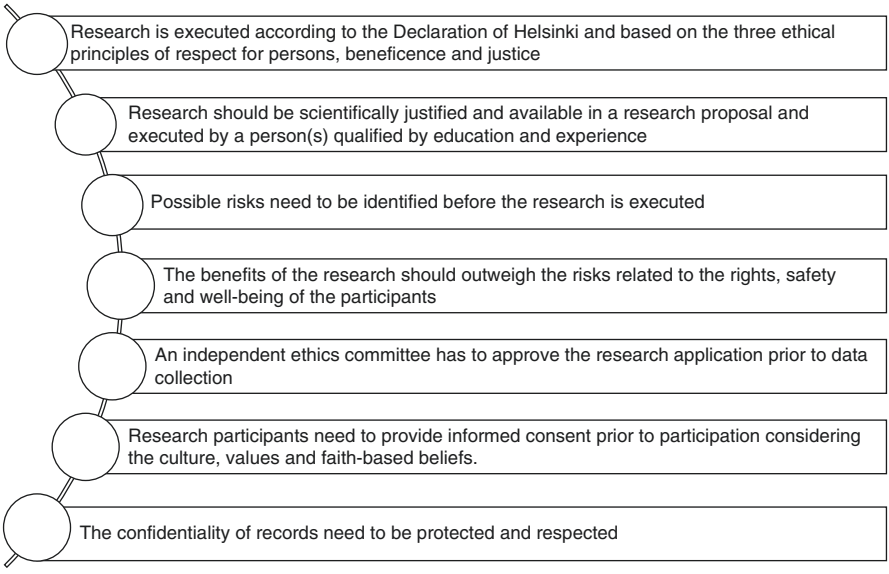


Fig. 6.1 The principles of good clinical research adapted from the European Medicines Agency [9]

(REC). In the UK, this application is made through the Integrated Research Application System (IRAS) and is discussed in Sect. 6.5.

Mainly UK sources are cited in this chapter. Researchers from outside the UK should also access the requirements of their respective countries and universities.

6.2 Ethics in Research

Ethics is a branch of philosophy that deals with making the right decision to justify a moral outcome [10]. Put simply, it means distinguishing between what is considered as 'right' and 'not right'. It deals with critical analysis and evaluation of assumptions we hold and make to decide the best way to deal with problems that arise. Medical ethics is a branch of ethics relevant to healthcare researchers due to their commitment to best practice in their professional roles, responsibilities and accountability.

Ethical decision-making is not a simple process. It involves analysing often large amounts of information and evidence to answer a complex problem. There are no right or wrong answers in solving ethical problems, and herein lies the dilemma of managing ethics related issues. A researcher must ensure that the best possible decision is made based on reasoning following a methodical and rigorous interrogation of an issue at hand. Making these decisions requires a critical thought process. There are several principles and guidelines available that a researcher can consult to help in the decision-making process.

6.3 The Principles of Ethical Conduct

Beauchamp and Childress [11] in their 1979 seminal book *Principles of biomedical ethics* devised four principles that later became known as the Georgetown mantra. These principles are autonomy, beneficence, non-maleficence and justice [11]. Several authors are critical about the reduction of research ethics and professionalism to include only these principles [12]. With this criticism in mind, it is important to note that these principles should be the core when research includes human participants. A researcher must however respond to all ethical principles in executing a research project. These principles are briefly discussed.

Respect for persons indicates that prospective research participants should be treated with autonomy. Autonomy means that individuals have the right to self-determine what happens to them. It implies a rational thought process where a person actively decides whether or not to take part. Informed consent plays a vital role in enabling research participants to exercise their autonomy. Individuals need to be able to choose whether or not they want to participate in ethically approved research. They need to be fully informed before they give consent. They should not get involved until they have granted valid informed consent.

There are two legal aspects regarding consent. The first involves the act of a researcher giving information to a prospective participant. The second involves a

participant agreeing to participate based on an understanding of what the involvement in a study would entail. It is therefore a participant's right to receive information in their own language. It is furthermore a researcher's duty to provide information in such a way so that a prospective participant can make an informed decision whether or not to take part. Researchers must create a balance between right and duty. To give consent, participants must be competent (of sound mind), sufficiently informed (getting the right amount of information) and not be subjected to coercion or influence (no prompting or manipulation). Information must be provided in a way that is comprehensible to them and devoid of technical jargon or confusing language.

There are additional ethical considerations to be made for people under the age of 16 years. For example, parental or guardian consent is required. Depending on the nature of research studies, most undergraduate research studies focus on adults over the age of 18 years because they are not considered a vulnerable group. However, some groups are classified as vulnerable. These groups include pregnant women, children, prisoners, and mentally handicapped persons. Additional measures are needed to protect the rights and welfare of these groups; the principle of 'do no harm' needs to be emphasised. All available information about the benefits and possible risks of a project needs to be communicated to prospective participants prior to them giving consent for participation in research studies.

Beneficence means to do good or prevent harm. The principle involves balancing the benefit and risk associated with the proposed research. A well-designed research ethics application, based on sound scientific and ethical guidelines, is required to ensure that this principle is upheld (see Sect. 6.5). In addition, a researcher needs to be duly qualified to undertake the proposed research in order to protect all participants. In the case of a student performing research, a research supervisor needs to have the necessary qualification with a regulatory or professional body registration.

The principle most closely linked to beneficence is non-maleficence: do no harm or having an obligation not to inflict harm. Beneficence and non-maleficence can be considered as two sides of a coin. Usually 'doing good' and 'not doing harm' often confuse people about where one stops and the other starts. Doing either should lead to good research practice.

The principle of justice means to be treated in a fair manner; a fair process is necessary to select and recruit research participants. Fairness needs to be applied in the procedures to select individuals, and in the recruitment of individuals to participate in a study. The Belmont report [5] identifies individual justice, social justice and equity in the selection of research participants. Individual justice means that the proposed research should not only benefit some patients or select 'undesirable' persons for research with a risk. Social justice refers to specific groups such as vulnerable groups, racial minorities, economically disadvantaged or any group that is easily available in a setting where the research is to be conducted. Equity refers to the fact that no group or individual should be advantaged or disadvantaged through favouritism or discrimination. One common ethical dilemma, in relation to justice,

lies in the fair allocation of resources to a population where the demand outweighs the supply. For example, in the distribution of a new treatment a researcher would have to question how to decide who to treat and who not to treat.

6.4 The Need for Ethical Considerations in Our Roles as Radiographers

Ethical and moral behaviour is an expectation of our practice as radiographers. It enables us to take a rational, coherent and consistent approach to making moral decisions. Ethics are the rules of human conduct. Our roles as student radiographers or qualified practitioners contain the rules for professional conduct. Doing what is ethical according to these rules is doing what is right. In this way ethics is a core professional attribute.

Another need for considering ethical issues is to produce a framework based on principles that can be applied universally in decision-making. Decision-making deals with a critical evaluation of assumptions and arguments. This is also evident during the review of ethics application documents when reviewers must be satisfied that a proposed research study meets ethical principles (see Sect. 6.8). This expectation is written into radiographers code of conduct and statements detailing expectations of proficiency and competence as practitioners. In clinical practice we must do what is right for our patients. In research practice we must do what is right for our research participants. When undertaking a research project or data collection exercise involving human participants, a researcher must understand the basic principles of ethics and how these may apply during a research process.

Medical practice involves scientific facts. At the point of service delivery ethics must be seamlessly blended with scientific facts giving holistic practice and delivery of patient-centred care. Our scope of practice, and our professional role and responsibilities, are set out by professional and regulatory bodies in the country that a practitioner is studying and practising in. There is an expectation of trust, reliability and accountability. A core expectation in relation to professional conduct in research is trustworthiness where participants expect researchers to be 'faithful' to their involvement; this includes the need for privacy and modesty. Participants expect practitioners to be competent. They expect practitioners to be well trained to know what they are doing and that they can be depended on to do the right thing. Trustworthiness is a character trait as it encompasses attributes such as reliability, honesty and dependability.

Radiography practitioners, as members of the allied and/or health professions, have a duty of care to report colleagues who act inappropriately, in both clinical and research capacities. An example of this is if a practitioner was investigating the number of repeated X-ray examinations in a diagnostic imaging department and found that a radiographer persistently and unnecessarily repeated X-ray examinations in order to aim for perfection even though diagnostic quality of the images was not compromised. Another example would be if a patient incorrectly received a therapeutic radiation dose of another patient. These incidences must be reported to

a senior member of staff, even if anonymity has been promised. All individuals have their own beliefs and values, their own biases and prejudices. When a student enrols to study radiography, s/he has to subscribe to the beliefs and values of the discipline. As such conducting research within that field must also be undertaken with the same frame of considerations. Guidelines alone are insufficient; the final responsibility lies with the person conducting the research. Radiographers must therefore work within ethical and legal boundaries of their scope of practice and expectation regarding their role and responsibilities.

6.5 Considerations When Applying for Ethics Approval

The first consideration that any researcher must determine is whether their proposed study is likely to require ethical review. All formal enquiry has some ethical component, even if it is only that researchers conduct themselves honestly in undertaking their study and do not deliberately influence (bias), copy (plagiarise) or even fabricate the work. However, not all clinical studies require ethical review. For example, the research governance framework within England, and to a major extent across the rest of the United Kingdom (UK), involves ethics application through the IRAS system. IRAS is a single system to apply for the permissions and approvals for health and social/community care research in the UK [13]. All applications for research within the National Health Service (NHS) are made via the IRAS and then reviewed by the various research ethics committees (RECs) linked to IRAS. A REC consists of a group of people appointed to review research applications, and to formally assess whether a proposed research adheres to ethical principles. It must conform to recognised ethical standards, which include respecting the dignity, rights, safety and well-being of participants [13]. Similar application processes for ethical approval of research studies are required in countries such as Europe, South Africa, New Zealand, Australia and the United States of America (USA).

Researchers need to follow a clear but rigorous process of determining whether their study requires ethical review. The IRAS website gives detailed guidance on this process. Within the UK, formal NHS ethical review is not required if a proposed study is deemed to be an audit and/or service evaluation and not research. This is dependent upon the intention of a researcher. For example, is the aim of the research to obtain new knowledge through rigorous and systematic approaches (research) or to measure existing practice/undertaking quality assurance? Table 6.1 outlines some of the determinants that can be used in deciding which category a proposed study may fall into and whether it requires formal IRAS ethical review.

While the basic premise of ethical review is to protect every participant from any potential harm, and at the same time respecting their dignity, rights and well-being, it may also seek to safeguard a researcher and/or institution undertaking a study. However, the process of implementation differs around the world. Researchers are encouraged to discuss the requirements for ethics approval as per the regulatory standards in their respective country. If a researcher is a student or member of staff in a university, internal review may also occur within the institution. For academic

Table 6.1 Determining the need for ethical approval

		Researcher's intent	Questions	Data collection	Allocation/ randomisation	Decision
Service evaluation		To evaluate service delivery and/or current care	What is the standard that this service achieves?	Measures current service only; can only be tested on treatments/investigations/ techniques in practice	None	May require NHS Trust R&D review but not usually ethical review
Clinical audit		To evaluate the delivery of current care against best practice	Does the service meet predetermined standard/s?	Measures service against guidance or professional standards; can only be tested on current treatments/investigations/ techniques in practice. Usually involves the analysis of existing data but may include simple interview or questionnaire administration	None	May require NHS Trust R&D review but not usually ethical review
Research	Quantitative approaches	To derive new approaches, knowledge, test hypotheses and interrogate research questions	Based on hypothesis	Evaluation or comparison of new treatments/investigations/techniques using existing or new data. Understanding the implications of new treatments/ investigations/techniques and/or relationships using existing or new data	May use both	Formal NHS Trust R&D and ethical review required
	Qualitative approaches		Research questions to identify and explore themes		May use both	

Adapted from Health Authority Research—Defining Research [8]

or non-clinical research this may be the only review required. In the case of a student or staff member wishing to undertake clinical research in a health and social care setting, an application to IRAS is required. The information in Table 6.1 should help a researcher decide how to proceed with the correct ethics application process. If the research is part of a university degree, in the case of undergraduate and postgraduate radiography courses, students can discuss this with their respective research supervisor in the first instance.

6.6 Key Ethical Considerations

This section provides guidance on the main ethical considerations when making an application for approval. It is presented in alphabetical order for ease of reference.

Anonymity refers to the identity of participants being kept unknown. Anonymity may be achieved in a variety of ways. For example, pseudonyms may be used to protect the identity of participants and/or locations. In addition, codes may be used to identify participants, with the information that relates to these codes (participants) being kept on a separate central list (key). In some cases, however, the identity of a participant is known to a researcher but anonymous to other research participants. For example, in face-to-face interviews, participants cannot be anonymous; however, their identity remains confidential in that they are unknown to others. Their respective identities and views are then 'hidden' in any subsequent report or publication.

Assent is the acceptance to be involved in a research study by a participant under the age of consent for research purposes (16 years and over in the UK). To obtain assent a participant information sheet (PIS) should be age-appropriate with respect to the language and explanations utilised. It has no legal standing; where a participant is old enough to understand what taking part in the research entails, it is good practice to interact with such a young person as an individual as well as with an adult (parent/guardian/legal representative) who gives the formal (legal) consent. It should be noted that research on children should be avoided if the data can be obtained by using only adult participants.

Coercion relates to payments in monetary terms or in goods such as gift vouchers. These are sometimes offered by researchers to thank participants for taking part in their research studies. Payment may also be made to investigators, usually by pharmaceutical companies, for their time in taking part in clinical trials. However, if the level offered is too high, this may be viewed as coercive, in that it may induce investigators to sign up numbers of participants purely for monetary return. A similar situation may happen with participants. Coercion may also take place when researchers, in whatever way, pressurise participants into taking part in research. This can inadvertently occur when researchers attempt to recruit participants for a study without allowing them time to consider the implications of their involvement in the research before they consent to take part. Usually a minimum period of 24 h after initially discussing the research should be given to potential participants to allow them time to consider whether or not they want to take part.

Coercion may also occur in respect of the nature of the relationship between researchers and participants, especially where research is being undertaken by clinical staff or by academics with their own students. It may be difficult for a patient to refuse to participate in a clinical trial if asked to consider this by a surgeon who is going to perform the operation. The same dilemma would occur if a first year student were to be invited by his/her professor to be interviewed as part of the professor's research. In circumstances where participants have a particularly dependent relationship with a researcher, consideration should be given to asking another member of the clinical/research team to take consent.

Confidentiality refers to the duty of a researcher to securely manage the information obtained from or about a research participant. Participants have the right to privacy and confidentiality; they expect professionals to keep their information safe and secure. Researchers must follow the data protection principles and use this in their judgement and decision-making. Like anonymity, confidentiality is a promise that it will not be possible to attribute/connect the findings of the research to the participants themselves, unless they gave permission for this prior to consenting to take part. If a researcher wishes to utilise anonymous direct quotations from participants, it is good practice to obtain express permission from them and to do this prior to them taking part in the research.

Conflicts of interest may arise where there is some form of relationship between various individuals or groups within a study that could possibly affect the outcome of the research through bias or coercion. Such a relationship should be declared and clearly identified in an application and, if appropriate, in the PIS (see PIS below). An example here could be the source and amount of funding provided to a researcher. Participants may not want to take part if they are unhappy about a research funding body. For example, inviting patients that have lung cancer to participate in a research study funded by a tobacco manufacturer. It is also good practice for researchers, particularly undergraduate students, to declare on the PIS if the research is to be done in fulfilment of an academic qualification. In some ways this may have a beneficial impact, with altruistic patients wanting to help students to fulfil their research, although this could also be construed as possibly being coercive.

Consent is the formal acceptance given by a participant to be involved in a research study. Any consent should, as far as possible, be fully informed and written, in that a potential participant should be made aware of what the research is about, and of methods and implications in taking part. This is normally given in the form of a written leaflet (see PIS below) which sets out the details of the proposed research. Informed consent becomes difficult when prospective participants are unable to give consent because of their age (young and old), mental capability, or physical state (e.g., unconscious). In these situations, consent should be obtained as far as possible from the individual concerned. If it cannot be obtained then another person such as a parent/guardian, carer or legal representative may be asked to give consent on behalf of the prospective participant. However, if a participant were to be only temporarily incapacitated it is important that consent be obtained from such a participant once they regain their full faculties.

Written, informed consent is taken as the standard. There are other forms of consent that may occur in research practice. Implied consent occurs when a participant does not expressly give consent, but this is inferred through their actions. In a research sense this generally occurs with survey methods utilising questionnaires, when consent is not specifically asked for by a researcher but is taken to be given (implied) if the questionnaire is returned. Once given, consent does not become permanent; participants may withdraw from a study without being required to give any reason and may also be able to ask that their data are not to be used. This must be indicated on the consent form. However, if data have been anonymised and aggregated it would be difficult for an individual participant's information to be separated out. This may also occur with data obtained from focus groups because it is the group interaction that generates the data; withdrawing one participant's data would therefore make the remaining data difficult to interpret. In cases such as these it needs to be made clear on a PIS that data collected up to the point of withdrawal have to be retained. In addition, ethics approval is only given for any one named project at a time; it does not cover future studies where the subjects being identified have not been clearly stated. Research that evolves from current work requires separate ethics approval and consent from participants at some time in the future. In other words a current authorisation will not apply. As well as the various types of consents, researchers also need to consider who is to take/obtain consent from a participant as there may be issues of coercion and a possibility of power bias, as outlined above. Conversely, a person taking the consent must be aware of the implications of a study to be able to answer any queries, thus enabling each participant, by being fully informed, to decide whether to take part.

Data are needed in research. Researchers must operate in accordance with several legislative acts governing data protection and access to medical records relevant to the country in which research is performed. Researchers should indicate how and where data are to be stored (usually in a locked cabinet or password protected computer), who will have access to the data (usually the researchers), how long data will be stored (a minimum of 3 years if data are to be published), and what will happen to the data post-study (i.e., destruction). Data that are to be sent outside the country in which a study has been done should be anonymised. As outlined in the section on anonymity, the use of codes/keys can be used to separate identifiable data.

Data monitoring committee (DMC) plays an important role in clinical trials. A DMC reviews a study while it is in progress to assess the impact of an intervention (e.g., drug/new technique) upon the participants. If it is shown that serious side-effects are beginning to occur in large numbers, then the trial should be stopped to avoid exposing future participants to harm. On the other hand, when a study shows overwhelming positive results then it might be suggested that enough data have been collected to show benefit, thus it would be unethical to inconvenience or recruit further participants and therefore the study should be stopped.

Data protection regulations apply to how researchers collect and hold information about their participants. On 25 May 2018 the General Data Protection Regulation (GDPR) came into force in the European Union (EU). It pertains to protection of personal information (data). According to the GDPR, one must have a

defined lawful basis to hold and use personal data. Researchers who will be holding and using health information, which is a special category of personal data in GDPR (most researchers producing a PIS), are also required a further condition to this lawful basis. In most cases this condition should be to support 'scientific and historical research'. GDPR also requires that a researcher should be fair and transparent about holding and using personal data. This includes all personal data used to support research. The PIS provides a large part of how to meet fairness and transparency requirements. However, the information provided in the PIS is not the only information a researcher should provide. GDPR demands that all potential research participants can access the information provided and are likely to understand it [14]. Researchers outside of the EU must adhere to the protection of personal information legislation in their country.

Participant information sheet (PIS) is arguably the most important document of a research study. The information contained within it explains and invites participants to participate in a study. It does receive scrutiny at the REC meeting. A clearly written, well defined and appropriate PIS should give participants enough information on the nature of a proposed study for them to be able to make an informed choice about whether or not to take part. Guidelines and a template are provided on the HRA website. The RECs prefer the PIS to be in a specific format, but this is not compulsory. However, if the template is not used, then researchers should make sure that the appropriate sections relevant to their study are included in whatever alternative format they utilise, such as a letter.

A PIS must be written in lay terms and in a language style that is understandable to possible participants. If assent is being sought from a participant under the age of 16/18 years, it is necessary to amend the level of reading ability. For younger children, around 8 years of age, it could be beneficial to use drawings or diagrams to explain the information. Language again may be an issue with respect to multinational studies; a PIS may have been written in another country. The language used within a PIS must be suitable to the audience in the country the research is being carried out in. Therefore, if a PIS is required for non-English speakers it must be translated. Unlike clinical practice, it is not enough to get relatives to translate or act as interpreters for participants. Participants must have the relevant information available directly to them so that they can make an informed choice; professional services should therefore be utilised. With respect to research documentation, applications and in particular a PIS, a researcher should make sure that all paperwork is devoid of errors in spelling and grammar, and that all sections have been completed correctly with the information required, otherwise the decision of the REC may be delayed.

Radiation research refers to studies involving the use of radiation. Such research generates specific sections of an ethics application form to be completed detailing the type of radiation and particularly the dose to be received. This has to be substantiated by a local radiation protection advisor who has to sign the form confirming the proposed level of radiation exposure. Researchers need to provide information on a PIS to participants about any possible radiation effects. The concept of measuring radiation dose in millisieverts (mSv) may probably not be understood by a lay

participant. Thus, it may be useful to use a comparator; the most commonly used being levels/hours of background radiation. Informing patients of the risks of radiation is highlighted within the recent updated Ionising Radiation (Medical Exposure) Regulations (IR(ME)R) 2017 [15] and IR(ME)R (NI) 2018 guidelines [16]. There is however limited guidance on what is considered as appropriate comparators so that these can be understood when explaining them to patients and/or research participants. Nonetheless, the basis of IR(ME)R lies in justification of the risks and benefits: researchers are therefore encouraged to use this as a reasonable approach in their explanation to participants.

Respondent distress/expectation refers to the potential for participants to become distressed or expect further information about topics that are highlighted because of their participation within a study. A researcher has to provide details, in an application and on a PIS, on how these situations will be dealt with. This is usually done by giving advice on access to further information and support services or, if appropriately qualified, undertake this directly themselves. In determining the suitability of a PIS an ethics committee may be concerned that a study is not artificially raising participants' expectations of a particular treatment or examination or causing unnecessary anxiety and stress by the information given. For example, a researcher may want to ask patients with prostate cancer their views on which treatment they would prefer (radiotherapy or surgery). However, while both options could be available in one hospital, only surgery may be available in another. For patients in the second hospital equity may not be apparent; it may thus be deemed unethical as the proposed research may be raising expectations that such patients may have a choice of treatment. Researchers should also consider issues like time inconvenience and the sensitivity of the matter being investigated. In addition, with the increasing prevalence of mental health disorders that participants do not need to disclose, a researcher should additionally consider any associated hazards that could arise during a study.

Researcher issues and responsibilities are important when conducting a study. The primary purpose of any research ethics system is to protect each participant. However, researchers also need to be aware that there may be times when their own actions/circumstances need to be considered within a study. This could include visiting a participant in their home or collecting data alone in a city centre. The REC looks for some indication that researchers are aware of these issues (risk assessment undertaken) and that they have put into place a mechanism to protect themselves (e.g., lone-worker policy). This is particularly important for research involving radiation. Prospective researchers must show that they are aware of the implications of their actions about any use of radiation, complying with the principles of ALARA (as low as reasonably achievable) and ALARP (as low as reasonably practicable). In addition, researchers have a responsibility to participants and to society as a whole by the very nature of what they are undertaking. They should not copy (plagiarise) or falsify data, and should act fairly (unbiased) in their approaches to all participants, for good quality findings to be obtained. Otherwise it becomes unethical to subject participants to poor research practice.

Sponsor is often erroneously understood by inexperienced applicants to refer to financial contributions to undertake a study. They are sometimes confused regarding

questions as to whether their research has a sponsor. They often answer in the negative due to their misunderstanding of a sponsor. In governance terms a sponsor is taken to be a person or company, usually an employer, who accepts responsibility for the actions of a researcher in respect of any claims for negligence or harm because of such research. In most cases the answer therefore would be in the affirmative, particularly for researchers working within the NHS. In terms of student researchers, their respective university should take responsibility as sponsor. This would be through a research supervisor who is directly employed, rather than a student.

6.7 The Process for Ethical Clearance of a Research Project

Research studies are broadly classified as qualitative and quantitative. Nonetheless, if a study involves human participants (adults and/or children) it requires ethical approval. The main reason is to protect both a participant and researcher. Each participant is protected since an ethics committee considers the risks involved in the proposed research. A researcher is protected because there will be evidence that an ethics committee approved the research project. In other words an approved project adheres to specific standards. Approval from an ethics committee needs to be obtained before data collection can commence.

An ethics committee provides guidelines on the application requirements; this is usually provided as a checklist. Figure 6.2 is based on the main points of such checklists. A researcher must use a checklist as a guideline to complete an application and to provide additional evidence to support the review and approval processes as well as for self-assessment. Submission of a research ethics application and supporting documents for the ethics committee approval is commonly done via online platforms. This means that it is easy to verify that all required documents were submitted. The committee administration can verify if an application is incomplete and notify the researcher.

Submissions can be classified as research with a risk or a minimal risk project. Research with a risk may involve prospective interventions with human subjects. A project with a minimal risk may be a retrospective study using information from a patient's records. Once submitted, an ethics committee refers the complete submission to one or two independent reviewers. The reviewers may use the principles in Fig. 6.2 to approve a project, request modifications or, in rare cases, may even reject an application. The role of an ethics committee is to ratify reviewers' reports at a meeting where a final decision is captured. A researcher receives written notice with the reviewers' consolidated feedback. Once a project is approved, a researcher is notified in writing with an ethics approval reference number linked to the project.

Ethics committees usually provide helpful resources; for example, templates and examples are usually available on their websites so that a researcher can prepare a submission for the approval process. Good practice is to access the online submission site of your university to verify the deadline date for submission of documents,

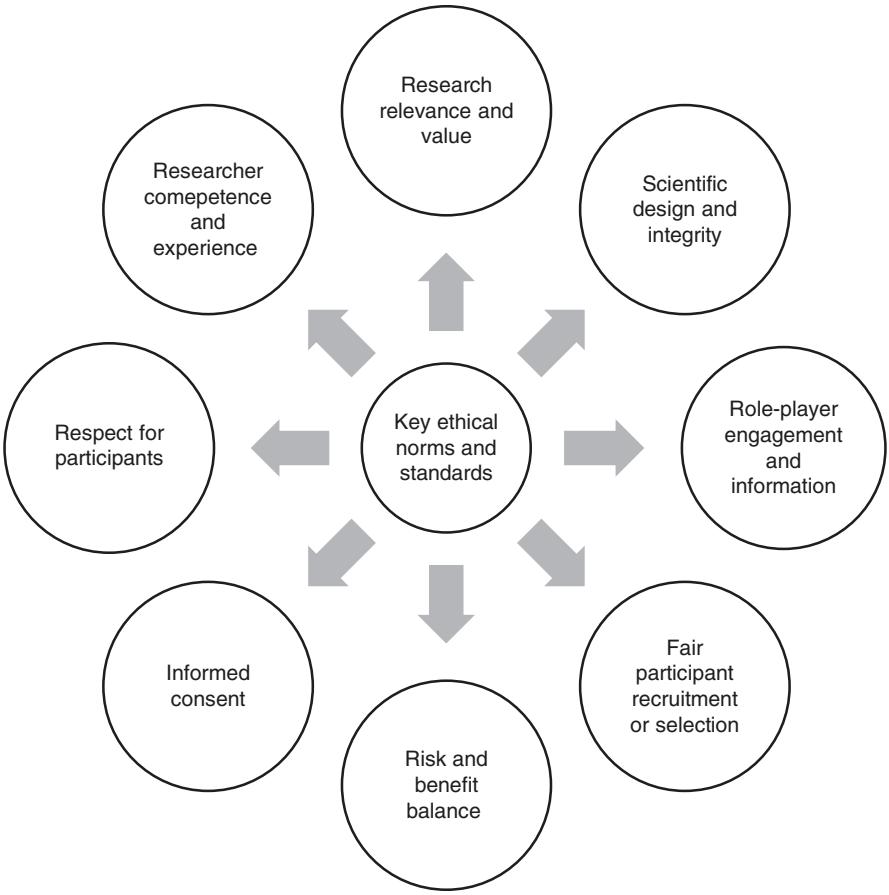


Fig. 6.2 Criteria for approval of applications

to identify the specific requirements, to make a list of the documents needed and to access templates (e.g., PIS). A formal application to the RECs is a lengthy process which takes time, so it is wise to plan ahead. As a researcher you should create a folder with all the documents ready to upload. Each ethics committee may use a different checklist, and these requirements may also be university or country specific. However, the information broadly corresponds with the criteria for approval of a project as given in Fig. 6.2.

The scientific design of a research ethics application requires sound alignment between the research title, aim, objectives, and methodology. In the selection of participants, the recruitment procedures and information to potential participants need to be clearly outlined. For example, inclusion and exclusion criteria or if any potential participants belong to a vulnerable group. The possible risks and benefits of the proposed research must be highlighted. The informed consent process must be clear and explained in terms that a lay person can understand. An ethics

application must clearly indicate privacy and confidentiality matters, and whether participants will receive compensation and the projected cost of the proposed research. A researcher must state how the participants will be informed regarding the outcome of the research, plans for the safe keeping of records and length of time of retention of such records. It should be clear that a researcher, sometimes with the guidance of a supervisor, has the required competency to perform the proposed research. One example is to indicate competency to interview research participants.

In addition, the research procedures must be clear and aligned with the research title and inclusion of participants. This includes the data collection process and analysis. An example here would be a survey of patients' experiences of a colonoscopy examination or of a radiotherapy planning session, where it was proposed to interview the patients 15 min after the end of the examination/session. Patients are unlikely to be able to answer questions at this time as they will be recovering from their examination. Such a hypothetical study might be better served by a questionnaire for each participating patient to complete in their own time, or interviewing patients after a suitable time period.

Lastly, undergraduate students often misunderstand the requirement for permission with the requirement of consent. Consent applies to research participants: once they have formally agreed to participate means they have consented. Permission on the other hand refers to approval to gain access to participants. For example, in a university setting this would need to be the course or programme leader of a student population (diagnostic radiography, radiotherapy, physiotherapy and so on) that a researcher wishes to recruit for a study. Students are required to write to their respective course/programme leader or dean of the faculty to ask for permission to access their students. In a hospital setting, students are required to write to the department manager or lead superintendent of that specific clinical area to ask their permission to approach either the radiographers and/or patients. Depending on the nature of a study, the local R and D department of the trust may also have to be consulted. A department manager and research supervisors should be able to advise students further in this regard. The term R and D is used generically to describe research and development offices or departments within either NHS organisations or universities [13].

6.8 Dealing with Reviewer Feedback

Reviewers provide written feedback on an application. This feedback is linked to the criteria for approval of projects (Fig. 6.2). It is however seldom a pleasant experience for a researcher to receive feedback from reviewers, specifically if the outcome of the ethics committee is that an application needs to be modified and resubmitted. A researcher must keep in mind that the purpose of the reviewers' feedback is to improve the submitted application, and address areas of concern so that the project responds to all ethical principles. The feedback should never be personalised. If a researcher must resubmit an application and supporting

documents to an ethics committee for final approval of the project, it is good practice to highlight the changes within the application. A cover letter to accompany the resubmission should indicate the changes and responses to the reviewers' comments as this may be helpful for them to review the changes. This step may even speed up the approval process. In some cases, a researcher, with a supervisor's support, may explain with the necessary evidence that it is not possible for some of the recommended changes to be executed as per the reviewers' feedback.

Reviewers also comment on administrative aspects such as an incomplete submission. For example, some documents not uploaded or incorrectly uploaded and/or missing permissions; a common one is where a supervisor did not give permission for the research project. Other matters relate to the inclusion and exclusion selection criteria of research participants being unclear; the PIS being unavailable in the languages spoken within the research environment, and the layman summary of the project being done in technical and academic language. Reviewers may also request revisions if the risks associated with the proposed research have not been comprehensively addressed and the data collection method and the data collection tool are not aligned with the aim of the proposed study. Lastly, reviewers may request a revision of a project's timelines as these may not be feasible to conduct the proposed study.

In the end, the most successful application is the one in which the ethical implications are carefully considered and addressed, and supported by applicable documentation.

6.9 Conclusion

The origin of research principles of good clinical research practice, why ethical approval is required, and guidance on the process involved to obtaining ethical clearance, are covered in this chapter. Differentiation is provided to help researchers determine whether their respective study will be considered as research, audit or service evaluation and the need for ethics approval. The specific ethical considerations in our roles as radiographers were presented as well as considerations that should be made when devising ethics application documents. Common pitfalls that reviewers comment on were shared with the integration of good practice tips to ensure successful outcomes during a review process.

References

1. Beecher HK. Ethics and clinical research. N Engl J Med. 1966;274:1354–60. https://doi.org/10.1056/NEJM196606162742405.
2. Dhai A. The research ethics evolution: from Nuremberg to Helsinki. S Afr Med J. 2014;104(3):178–80.
3. Resnik DB. Research ethics timeline (1932 to present). National Institute of Environmental Health Sciences; 2019. https://www.niehs.nih.gov/research/resources/bioethics/timeline/index.cfm.

4. Nuremberg Military Tribunals. The medical case. In: Trials of war criminals before the Nuremberg military tribunals under control council law no. 10, vol. 2. Washington, DC: US Government Printing Office; 1949. p. 181–2. https://www.loc.gov/rr/frd/Military_Law/pdf/NT_war-criminals_Vol-II.pdf. Accessed 23 Oct 2015.
5. National Commission for the Protection of Human Subjects of Biomedical and Behavioural Research. The Belmont report: ethical principles and guidelines for the protection of human subjects of research. 1979. https://www.hhs.gov/ohrp/policy/belmont.html. Accessed 21 Oct 2015.
6. World Medical Association. WMA declaration of Helsinki—ethical principles for medical research involving human subjects. Revised October 2013. http://www.wma.net/en/30publications/10policies/b3/index.html. Accessed 23 Oct 2015.
7. WHO (World Health Organisation). Handbook for good clinical research practice (GCP) guidance for implementation from the World Health Organization. Geneva: WHO; 2002. https://www.who.int/medicines/areas/quality_safety/safety_efficacy/gcp1.pdf.
8. Health Research Authority. Defining research. 2018. http://www.hra-decisiontools.org.uk/research/docs/DefiningResearchTable_Oct2017-1.pdf. Accessed 10 May 2019.
9. European Medicines Agency. Guidance on good clinical practice (E6). 2016. https://www.hra.nhs.uk/planning-and-improving-research/policies-standards-legislation/good-clinical-practice/. Accessed 16 May 2019.
10. Ramlaul A, Gregory T. Ethical and legal considerations in professional practice. In: Ramlaul A, Vosper V, editors. Patient-centred care in medical imaging and radiotherapy. Oxford: Elsevier; 2013. p. 257–70.
11. Beauchamp TL, Childress JF. Principles of biomedical ethics. 6th ed. New York: Oxford University Press; 2009.
12. Tsay C. Revisiting the ethics of research on human subjects. AMA J Ethics. 2015;17(12):1105–7.
13. Integrated Research Application System. Streamlining the research application process. 2019. https://www.myresearchproject.org.uk/. Accessed 10 May 2019.
14. Health Research Authority. Consent and participant information guidance. 2019. http://www.hra-decisiontools.org.uk/consent/. Accessed 10 May 2019.
15. The ionising radiation (medical exposure) regulations 2017 No. 1322. http://www.legislation.gov.uk/uksi/2017/1322/contents/made. Accessed 11 Feb 2019.
16. The ionising radiation (medical exposure) (amendment) regulations 2018 No. 121. http://www.legislation.gov.uk/uksi/2018/121/made. Accessed 30 May 2019.

Part II

Research Planning

Planning Your Research Study

7

Susan Cutler and Peter Williams

7.1 Introduction

One of the challenges of research, particularly for new researchers, is thinking about an area for research that is valid, worthwhile and researchable; this is the first step in research design. Exploration of predominantly peer-reviewed published and, to a lesser extent, unpublished literature is a good place to start. This helps you evaluate what research has already been carried out and give you an indication of any gaps in your current research area. Even if a study has been carried it may still be valid to repeat it if new research will add additional knowledge. Guidance on searching, evaluating and critiquing are covered in Chaps. 3 and 4, respectively. Once you have considered an appropriate area you then need to frame a specific question or questions. The process of developing and refining a research question is discussed in Chap. 2. A research question is fundamental to the research approach adopted, and throughout this book a number of research approaches are explored.

It is, however, important to learn the difference between a method and methodology. Method relates to the tools of data collection or analysis, such as questionnaires or interviews. Methodology refers to the approach or paradigm that underpins research. For example, an interview conducted within a qualitative approach, which seeks to explore, sayings, feelings or experiences, will have a very different underlying purpose and produce different data from that of an interview conducted within a quantitative design. For example, you may want to explore patients' experiences of a visit to a medical imaging or radiotherapy department. You could explore this

S. Cutler (✉)
Head of Department, School of Health, Teesside University, Middlesbrough, UK
e-mail: s.cutler@tees.ac.uk

P. Williams
Department of Information Studies, University College London, London, UK
e-mail: peter.williams@ucl.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_7

qualitatively or quantitatively. However, a mixed-methods approach can be undertaken, which uses both qualitative and quantitative methods in the same study. This approach is growing in popularity in view of the range of data it can collect. It is however time-consuming to conduct.

First, we will look at a qualitative approach, and then a quantitative one. "Tell me about your recent visit to the medical imaging/radiotherapy department". This open-ended question gives patients an opportunity to tell you about issues, concerns and experiences that are important to them. The narrative produced will be rich in data and may highlight issues, experiences or concerns that you had not considered. The data will be in a patient's own words. The first stage is identifying some themes. After a first read through of the data you then have to compile a list of words or phrases for each issue or concept. For example, patient A may have stated the following. "They all seemed very busy, I was worried that I might have to wait a long time for my appointment, but I was only 5 min late. I was then taken to another waiting area, where I had to get changed and then I had to wait again. I wasn't expecting that".

When you look at this verbatim transcript there are possibly a few words that you could highlight: for example, "busy", "worried", "waiting", "unexpected:. When you have explored the other patients' interviews, you may find there are similar identified that will enable you to develop categories and themes.

On the other hand should you use a quantitative approach, then you might ask 20 randomly selected patients to complete a questionnaire using a rating scale. An example of a question could be phrased as follows. "How would you rate the efficiency of your recent visit to the medical imaging/radiotherapy department?" You may want to distribute this to two different patient groups; for example, those who utilise an open access service and those with appointments (Table 7.1). You would give these patients one of five options from which they will be required to select their responses.

- 5. Excellent
- 4. Very good
- 3. Average
- 2. Poor
- 1. Very poor.

Table 7.1 Data gathered using a quantitative approach

Efficiency rating	Open access patients	Patients with appointments
5	2	5
4	4	8
3	6	5
2	7	1
1	1	1
	*n*¼20	*n*¼20

The data produced will be in a very different form. The raw data could be as follows.

- Patients with appointments: 54333454454344342551.
- Open access patients: 33343224422522334521.

As you can see, these two different approaches result in very different data; the first highlights issues that are pertinent to a patient; the second gives you numerical data that illustrate how many patients rated the department as efficient.

In order for your project to be successful you need to have clear aims and objectives and a plan of action. Making changes, as you go along, could be a recipe for disaster. You therefore need to think about the research process and the actions that you will have to undertake in a systematic and structured way. Often the ideal is superseded by what is practicable; this needs to be taken into consideration at the beginning of your project. Ironing out potential issues at the beginning of a process will reap benefits later on in your project. It is during the planning process, and selection of the appropriate tools, that a researcher must acknowledge and recognise potential pitfalls that could arise. As you will be committing a great deal of time and effort to this endeavour, we think an important consideration of undertaking research is that of choosing a research topic, which has personal interest or is of professional significance, to help motivate you. It is suggested that you consider your own skills set and those you use to conduct your research. For example, if you like talking to people you might want to consider using interviews, focus groups or observing people to maximise your skills.

Whether you realise it or not, as the researcher you have a powerful influence on your research project. What you have read in the past influences your thinking. How you phrase or ask questions influences the data collected [1], as discussed earlier. You do need to take a reflective approach to your project and should try and maintain a sceptical approach to the evidence provided by respondents and other data sources [2].

In order to plan your work, it is good practice to have a structure or a framework to work from. An example is presented below.

1. Orientation of the research
2. Review of the literature identified through a structured search
3. Research design and methodology
4. Data analysis
5. Reporting and writing up the research

7.2 Orientation of the Research

The generation of a research idea, a specific research question, aim or hypothesis, is discussed in detail in Chap. 2. However, there are some practical issues that are worth considering during this initial phase.

Time is one of the key resources available to a researcher. However, as with all the best laid plans, things can and do go wrong. Therefore, it is imperative that time is assigned in the overall plan to account for any potential problems (see Chap. 20). This is particularly important if research is undertaken as part of undergraduate or postgraduate studies because the timeline is often very tight and is frequently the main factor in determining a research design. A longitudinal study will not be feasible if a project needs to be completed in a set number of weeks. For example, it would be unrealistic to try and undertake a research project exploring practitioners' attitudes to continued professional development activities and their career projection. Such research cannot be undertaken in only 20 weeks.

A timescale affects a research question and the way the data are collected. You may only be able to use one instrument, but would have ideally liked to validate your data further with the use of another data collection instrument. The number of researchers undertaking the project also influences the focus of a project and the research question.

You may be undertaking a project on your own or working as part of a research group. Careful planning and ground rules must be identified at the planning stage in a group research project. This is necessary to ensure all group members are aware of the timescale, targets and deadlines. You need to consider the amount of time needed to undertake the data collection: conducting interviews, for example, can be very time-consuming and this may constrain the number of interviews that an individual can undertake.

The cost of the project has to be considered at this early stage. If your research is being undertaken as part of your degree, one of your priorities will probably be to keep costs to a minimum. Therefore, if you would have travel to interview participants from diverse geographical areas, as part of a qualitative project, then this would likely preclude this type of question. This will impact on your time also. Similarly, if you are considering testing a hypothesis that requires the use of equipment not available at your university, this too may exclude a particular research focus.

Whether you are a first-time researcher or one with experience, the importance of regular meetings with your supervisor cannot be overemphasised (see Chap. 20). Your supervisor is also a key resource with experience in guiding other students in a research process [3]. These meetings will give you an opportunity to discuss or clarify each stage of your project and receive constructive feedback (see Chap. 20). This is especially useful when you have collected the data; you can then discuss the implications of your findings during meetings with your supervisor. If you are a preregistration undergraduate or postgraduate then these meetings will also help you keep on schedule by ensuring you submit your work on time.

7.3 Review of the Literature

It is important to undertake a literature review to gather information from a range of sources and to contextualise what is known about a subject. A review of literature should lead to you identifying any gaps in the literature so that you can motivate

why your research will be important to conduct. The first stage in a search process involves developing a focused research question to identify relevant literature. A focused question should ensure all relevant studies will be included and that a systematic and replicable process will be used in your project.

7.4 Research Design and Methodology

This is the stage where your idea or theory moves towards a proposed concrete project. Your proposed research's aims are translated into specific questions and what instrument is to be used to answer these questions. This entails deciding what research instruments will be used to gather the data. A detailed planning of a project should be written up in a document called the "research proposal", which may be submitted as part of the ethics approval process and should be presented to your supervisors (Box 7.1). It is prudent to keep referring to this document as your project progresses. At this stage your area of research will not have been fully fleshed out. Your research proposal thus serves as your guide to completion.

Research design is an essential consideration in any research project. There are two main paradigms employed in research studies: quantitative and qualitative research designs. Depending on the nature of your aims and objectives, your study would usually fall within one of these two types of research designs. See Chaps. 8, 15 and 16.

Box 7.1 What Should You Include in Your Proposal?
The main headings in the proposal should include.

- Aims and objectives—what are you going to do, broken down into measurable objectives?
- Background—why is this topic interesting or important?
- What is the clinical relevance of the study if taking a qualitative approach?
- Methods—you will need to justify your chosen methodology and approach. How will you carry the research out? This should include a detailed description of the data you will collect, including sample size, access and if applicable, statistical tests that will be applied. How are you going to ensure the validity and reliability of the study if undertaking a quantitative study? Have you thought about credibility, transferability and trustworthiness of the research if undertaking a qualitative approach?
- Literature review—this should be a brief outline of the current literature that relates to your study. This is essential to enable you to become familiar with and take account of research already undertaken in a specific area. This should include your key references.
- Ethical issues—those related specifically to your study should be identified in this protocol. Submission of this protocol is usually an integral aspect of your ethical approval process.

- Resources—costing of staff, travel, materials. This may have to be very explicit, especially if your project is funded and you are accountable for all the monies spent. Often new researchers underestimate the costs of staff time particularly. It is easy to underestimate how long tasks will actually take. Normally in an undergraduate project there are no significant cost implications and the materials and equipment required are usually supplied by the university.
- Pilot study—if applicable, a pilot study should always be undertaken and you will need to describe in this protocol how this will be conducted and what your sample size will be.
- Timescale—this should include the important milestones such as the start and completion of the data collection process; data analysis; chapter drafts and time allocated for amendments. It is always worthwhile to include some contingency time, as well as a completion date.
- Dissemination—how are you going to inform others, including participants, of the findings? [4].

Quantitative research is usually structured to test a hypothesis. It aligns with a positivist viewpoint: collecting precise, measureable "scientific" data results in high levels of data reliability. This requires a researcher to analyse the data in order to look at correlation or possible cause and effect. It is deemed less subjective than qualitative research. A hypothesis consists of a suggested explanation for a phenomenon or a reasoned proposal suggesting a possible correlation between multiple phenomena. It is a methodology that aims to determine a relationship between one thing (independent variable) and another (dependent variable) in a specific population. Quantitative research is often described as being reductionist [4]. It is used to collect a range of numerical data in an effort to answer a research question. Though this may not always be the case, you may be seeking to describe a specific set of circumstances or characteristics of a study sample or target population [5]. The designs used in this paradigm are experimental and non-experimental: surveys, epidemiology and quasi-experimental designs, for example.

7.4.1 Experimental Designs

In experimental design investigators deliberately control and manipulate the conditions that determine the events in which they are interested [6]. An experiment involves making a change in the value of one variable, and is often undertaken in a laboratory. Healthcare experiments often need to be undertaken in a hospital or clinical setting; this then allows for a randomised controlled trial (RCT) to be undertaken. A RCT is an experimental design that aims at assessing clinical effectiveness (see Chap. 12). It is quite often used to ascertain the effectiveness of new drugs. RCTs provide strong evidence for efficacy of healthcare treatments or interventions

due to their low probability of bias. A population in a RCT is defined by the researchers. A population, for example, could be men over the age of 60 with prostatic disease, in a RCT pertaining to this pathology. A RCT sample would then be selected from this population. A randomised assignment procedure is used to allocate participants to a group. After the RCT intervention, the outcomes are then measured.

7.4.2 Non-experimental Designs

In non-experimental designs surveys are often used. In medical imaging and radiotherapy practice they can be employed to ascertain the attitudes, opinions and beliefs of people using the services we provide. The participants could be patients; doctors, dentists or other healthcare practitioners who refer patients; or radiographers. A questionnaire is often used to collect data. A survey questionnaire could be undertaken, for example, to give us an overview of the use of imaging services or radiotherapy in a given community.

Epidemiology is an approach used particularly in public health. It is concerned with a population as a whole or a particular group within it. This approach can use comparison studies, comparing responses of different cultural groups to an intervention, for example. This approach can be used to undertake correlation studies, which are used to identify interrelationships. Epidemiological methods are discussed in Chap. 9. Quasi-experimental designs resemble experiments, but the participants are not randomly assigned. Also, such research may include time-series designs. These are when a sample may have an intervention allocated and then observed over a period of time. For example, some patients may be recruited to this type of experiment and have radiotherapy while another group may have radiotherapy in addition to a new drug.

Qualitative research is used to study human behaviour and to interpret how people conceptualise the world and find meaning in their experiences. The use of this type of research does not usually test hypotheses. It is concerned with understanding personal meaning through specific questions that aim to guide an investigation. This method is used to seek to make sense of, or interpret, phenomena in terms of the meanings people bring to them. Qualitative research is used to help us understand how people feel and why is it that they feel as they do (see also Chap. 8). Samples tend to be much smaller compared to quantitative projects. The latter tend to include larger samples. A good qualitative study addresses a clinical problem through a clearly formulated question and often uses more than one research method, known as triangulation. Analysis of qualitative data can and should be done using explicit, systematic methods and should be reproducible [7]. Interviews, focus groups and observations are common techniques used in qualitative research. These are discussed in Chap. 16. Qualitative methodology is a useful method to assess individual feelings, experiences and knowledge. Some examples are presented below.

- How do student practitioners perceive the role of a personal tutor?
- What do students feel about distance learning while on clinical placement?

- What are patients' experiences of imaging/interventions/treatments?
- What are radiographer's experiences of imaging patients with a hearing impairment, or who do not have English as their first language?

An interview is defined as a two-person conversation initiated by an interviewer with the specific purpose of gathering information relevant to a research objective. Interviews allow for rapport to be developed between an interviewer (researcher) and interviewee (respondent). This allows a researcher to probe and explore complex issues. Interviews can be conducted face-to-face or telephonically. The former allows a researcher to observe respondents' body language and emotions thus providing additional information. Telephone interviews reach respondents over a wide geographical area. They are cost-effective and a more accessible approach to consider. The success of an interview depends upon the skill of an interviewer and how articulate a respondent is. Some interviews may be quite formal and a series of questions are asked. The responses to the questions are recorded; a researcher is quite directive during a formal interview. Interviews could be less formal, where a researcher raises a number of issues in a conversational way and is less directive.

Focus groups can be used to generate data from a wider range of responses compared to those of individual interviews. A researcher uses a focus group to deliberately select participants to discuss a particular area of interest. A focus group facilitates a trusting relationship between participants and a researcher. A focus group is unique in that participants build on the answers of others in the group. This allows for the production of rich data through social interaction. A focus group discussion can produce new thoughts that a researcher may not have thought of; participants encourage a collective group discussion of a topic. A focus group is often quicker to use, but it does not really allow personal matters to emerge or to be explored in any great depth. Individual interviews would be appropriate to find out how diagnostic and therapeutic practitioners cope with death and dying, whereas a focus group would be useful to ascertain what support diagnostic and therapeutic practitioners need for continued professional development.

Observations allow for gathering of data from live situations and enable a researcher to look at what is happening in situ rather than second-hand. A researcher uses a structured observation to specifically know what to look at: hand washing technique undertaken by practitioners or manual handling techniques of student radiographers, for example. An unstructured observation, on the other hand, means that a researcher is less clear what to look for and has go to a situation and observe what is taking place in order to determine what is significant. For example, you could provide some patient information in your waiting room, and then you could observe whether and how patients engage with it over a period of time.

7.5 Advantages and Disadvantages of Quantitative and Qualitative Research

The advantages and disadvantages of both these approaches are listed in Table 7.2. Let us look in more detail at each approach.

Table 7.2 Advantages and disadvantages of quantitative and qualitative approaches

Quantitative approach	Qualitative approach
Advantages	
Objectivity—the elimination or reduction of subjectivity of judgement or interpretation	Interviews yield much richer data than a questionnaire might
Independent and dependent variables clearly and precisely specified	Can handle complex topics
Data reliability high—due to controlled observations, laboratory experiments, mass surveys or other forms of research techniques	Interviews help where the topic is new or unfamiliar
Representative—often more representative of a wider population and can compare similar studies more easily	
Results are usually statistically reliable	
Disadvantages	
Not appropriate for learning why people act or think as they do	Inconsistency: It is very possible to be inconsistent in the way you ask questions, what questions you ask and how you interpret answers
Questions must be direct and easily quantified	Generalisability—difficult to generalise from interviews
Participants may answer differently in a structured survey than they would in a real-life situation	Low validity/reliability—the data-gathering measure may not measure what you want it to, or may not give the same results if repeated
Very difficult to prevent or detect researcher-induced bias in quantitative research	Experience—conducting interview or managing focus groups effectively does require a high level of experience from the researcher to obtain the required information from the respondent
	It is time-consuming to immerse yourself in the wealth and volume of data generated

7.5.1 Advantages of a Quantitative Approach

- Subjectivity of judgement or interpretation is eliminated or reduced. Subjectivity affects all research to some extent. Even before a study begins it is already influenced by factors and constraints such as a researcher's desires, interests and preoccupations. Quantitative research eliminates some of these factors because it requires a clear statement of a research problem in specific and well-defined terms in order to obtain numerical data that can be statistically analysed [8].
- The independent and dependent variables under investigation are clearly and precisely specified. An independent variable can be changed or manipulated by a researcher. A dependent variable is a response that is measured. An independent variable is the presumed cause; a dependent variable is the presumed effect.
- High levels of data reliability can be achieved due to controlled observations, laboratory experiments, mass surveys or other forms of research technique.
- A more representative sample of a wider population is usually obtained. A comparison of similar studies can be made more easily because of the larger sample sizes.
- Results are usually statistically reliable.

Consider the above-mentioned example of a questionnaire which was distributed to patients for them to rate the efficiency of the department. They had to choose from one of the five categories. The numerical data collected from this question could then be analysed statistically. The example used in this illustration is a small sample, but a questionnaire could be distributed to a much larger sample comprising several 100 patients. You might collect demographic information and, providing that the participants completed this information correctly, then the number of men and women who completed the questionnaire would then be clearly defined. If you use interview data from a few participants, you will not be able to apply statistical tests.

7.5.2 Disadvantages of a Quantitative Approach

- It is not appropriate for learning why people act or think as they do.
- Questions must be direct and easily quantified. It is vital that the questions are unambiguous, or the responses would detract from the focus of a research and will reduce the validity of a study.
- Participants may answer differently in a structured survey than they would in a real-life situation. There is no opportunity to probe respondents' answers.
- It is very difficult to detect researcher-induced bias in quantitative research.

7.5.3 Advantages of a Qualitative Approach

- Rich data are collected in interviews compared to questionnaires (see Chap. 16). An in-depth analysis of phenomena may be done as there is no rigid parameters of definable variables. The example above of patients talking about their experiences allows you to think about how a patient's meaning of the issues alluded to can be explored further.
- Complex subjects sometimes make it difficult to construct a questionnaire. The issues could be too complex to encapsulate in a series of relatively closed questions. Interviews allow a comprehensive exploration of such issues.
- Unknown territory, in which a researcher/interviewer may not have a good grasp of a particular phenomenon, requires an instrument to obtain data: interviews are excellent as a way in. For example, many years ago an interview was personally used by one author of this chapter to undertake a study of the (then) new phenomenon of the world wide web, and its impact on journalistic practices.

7.5.4 Disadvantages of a Qualitative Approach

- Inconsistency is a disadvantage. It is very possible to be inconsistent in the way you ask questions, what questions you ask and how you interpret answers. Of course, if your brief is wide, and you are exploring people's wider views and experiences, then an interviewee leads to an extent, so you may not always ask the same questions, but you must try and adopt the same (disinterested) approach.

- Generalisability of interview results may be difficult if a study aimed to do so. If not then it may be possible to do so. A study, by Davies and Bath [9] of Somali women in the UK regarding interpersonal sources of health and maternity information, comprised eight participants in an exploratory focus group and five who participated in semi-structured interviews. They did not attempt to generalise beyond postnatal Somali mothers receiving maternity care in a particular city in the UK, or beyond the specific (maternity information) topic examined.
- Low validity/reliability is a disadvantage. Validity according to Hammersley [10] is "truth: interpreted as the extent to which an account accurately represents the social phenomena to which it refers". In other words, does the data-gathering measure what you want it to measure? As for reliability, this is, also according to Hammersley [11], "the degree of consistency, with which instances are assigned to the same category by different observers or by the same observer on different occasions". For example, does the data-gathering produce the same results if a study was repeated? These are consistency issues which you must be aware of when planning and delivering your interviews (see Chaps. 10 and 16).
- Being time-consuming is factor to consider when undertaking qualitative research. In order to immerse yourself in the wealth and volume of data that are generated you need a great deal of time. Trawling through the data can be labour-intensive.
- Experience is a prerequisite when conducting interviews or managing focus groups effectively in order for a researcher to obtain the required information from respondents.

Let us go back to the interview example of patients about their recent visit to the medical imaging department; patients said they were worried. It is not clear from this initial response why they were worried. An open interview would allow you to explore this further with them. They also stated there were aspects of their experience they did not expect. You might want to find out more about this. If you used a questionnaire, then the reason for a patient being worried might not be captured or explored any further.

7.6 Instrumentation

This refers to how the data will be collected. These may be questionnaires, interviews, accounts, observation and tests. Research could be based on an observation made with instruments; recording electrodes, microscopes, and standardised clinical tests, weighing scales, for example. It could however not require the use of instruments to collect the data. You should think about the practicalities of using instruments, which may include the cost to use them; time and also their reliability. You should use standardised tools. If you decide to use a new tool, then you have to establish its reliability and validity. If your study includes weighing patients before a radiographic test or therapeutic intervention, for example, then you have to ensure that the scales are calibrated and checked regularly. A feature of scientific observation is the accuracy and reliability of the equipment employed. Instrumentation

relates to a situation when the instrument changes over the period of the study thus invalidating comparison of measured results [5], and this may compromise the internal validity of a study.

7.7 Selecting Your Sample

As part of the process you have to select your sample from the subject population. Sampling is the process of selecting a representative group of participants from a population and must be justified and fulfil an explicit purpose. Sample size is determined by the methodology of a study: effective sampling ensures the external validity of quantitative research and the trustworthiness of qualitative research. Generally, a sample in a qualitative study tends to be small, whereas it tends to be larger for quantitative data collection. As discussed earlier, if you were to interview patients about their experiences you may only have the resources and opportunity to interview five or six participants, but they need to be representative of the population. If the focus of your study is the experience of adolescents, you have to define adolescent and ensure you interview participants who meet the criteria. If your design is quantitative and your instrument is a questionnaire, and you are looking at an over 60s population, then there is a potential to distribute this to much larger numbers. You want to ensure your sample size is sufficient for the purpose of the analysis you intend to perform. You must ensure your sample is representative of a population you are studying.

Before selecting your sample, it is important to find out as much as possible about your study population. Population refers to a larger group from which a sample is taken. You need to know some of the overall demographics of the selected population, such as age, gender, etc., before commencing sampling. To some extent, when undertaking a project as part of a degree, your sample size will be determined by your available population and the time frame of your study. You also have to consider how much it might cost to interview a large sample or distribute a large number of questionnaires. The context of a sample should be considered. If a sample were to be selected from a population of practitioners working within a rural setting, how applicable would it be to generalise the findings to those practitioners who work in large urban environments? A sample is a subset of a population. You could consider all practitioners registered with the Health and Care Professions Council (HCPC) in the UK and select 50 therapeutic radiographers as your sample. But the sample should be representative of the population; a biased sample does not adequately represent the key groups.

7.8 Validity and Reliability

Cohen et al. [6] describe validity as an important aspect of effective research. It is a requirement of any research paradigm (see Chap. 10). Reliability refers to the ability of a research to be replicated over time and different samples. A reliable instrument should produce similar data when similar respondents are used. Validity and

reliability of a study can be maintained by careful sampling, using the correct instruments, and the application of correct statistical tests (see Chap. 15). Conversely, using qualitative data, validity might be addressed through honesty, depth, the participants and the extent of triangulation. So it is important that a researcher has confidence in the elements of research planning, data collection, processing and analysis, interpretation and judgement.

7.9 Mixed Method Designs

This is the combination of at least one qualitative and one quantitative method in a single research project. This approach should be used if you think it will give you greater objectivity; it could help you reduce bias in your data collection [12]. It is often used when conducting social research. For example, you may wish to use a questionnaire as the initial tool, but to get more detail from the respondents who completed the questionnaire you could then interview them [13]. The advantages and disadvantages of quantitative and qualitative approaches should be considered. You should think about some practical consequences of using a mixed method approach in that it will require additional resources to collect and analyse the subsequent data. Triangulation techniques are often used to explain more fully the richness and complexity of human behaviour by studying it from more than one viewpoint, hence the use of both quantitative and qualitative methods. It is also useful to confirm concurrent validity; i.e., where the data or results of the test concur with the data or results from other tests. For example, if you were to investigate practitioners' perceptions of their role as mentors to undergraduates, a predominantly qualitative mixed method approach could be used as this would be an exploratory study rather than testing a hypothesis. A series of interviews could be conducted in the first instance in order to generate data from which a rating questionnaire could be developed. In order to triangulate the data collected a focus group could be conducted to confirm the concurrent validity.

7.10 Data Analysis

It does not matter which research paradigm you use for data collection; you will find out very quickly just how much data you generate. This information needs to be organised and you will have to account for each piece of data generated. Information such as a person's age and gender are important data points. An interview that has lasted an hour could generate 50 pages of transcribed data. So, whether the data are written or numerical, you need to manage them [14]. On a practical note, you must ensure that data are stored and kept in a secure place. Remember to back-up all data or make hardcopies. Security is particularly important if you used an audio-recorder or video to collect your data from identifiable participants. You must implement measures to ensure that signed informed consent forms are not linked to your raw data so that there is no risk of the participants being identified (see Chap. 6).

If you use an experimental approach you will have to capture your data using computer software. DePoy and Gitlin [14] suggest a staged approach: check the data collection form for omissions, and then try to address missing data. You have to label each variable and then decide on the order to you need have to enter the variables into computer software programme. Guidance on quantitative data analysis is provided in Chap. 15. When you input your data, it is good practice to carry out a systematic check to ensure you have entered the data correctly. It is very easy to hit a wrong key on your keyboard, especially if you spend long hours working at your computer. The check also ensures that you have captured all your data and that no information has been omitted.

For qualitative data, it is common to undertake analysis while sorting out the data. This approach generates enormous amounts of data; usually narrative, but could be videos. Audio-recorded information from interviews has to be transcribed verbatim; video data have to be analysed frame by frame. These data are verbal or nonverbal; quantitative data on the other hand are numerical. Analysis of data requires sorting the data into meaningful categories, taxonomies or themes that explain the meanings or underlying patterns of a phenomenon of interest [15]. The collected data have to be managed. Usually the first stage for interview data is transcribing of audio-recorded interviews. You can do this yourself or get professional transcribers to do this. It is important that interviews are transcribed verbatim.

One advantage of personally transcribing data is that you do get to listen to the audio recordings many times; this should give you a feeling of what is being said. This should help you in your analysis of the data. However, it can be very time-consuming and tedious. There is voice to text software available. It may take you approximately 6–10 h to type a 1 h interview transcription. Once the transcription is complete you then have to check that you typed the data accurately and verbatim, and ensure that there are no misspellings of words or missing data. Once this is complete you then have to immerse yourself in the data: you read or listen to it several times. When you commence generating categories of information, you should establish a way of accessing the key passages that reflect these categories. How to organise this is to a great extent up to you. You could use cards, or word processing programmes, to catalogue, store and manipulate the information. Then you will have to analyse the data, and you must allocate time to do. It is useful while you undertake this task to make notes so that you can make comments regarding the data. Once again you have to refer to your original questions and aims to ensure that your analysis addresses the focus of your research. You can often generate further questions during this process.

7.11 Reporting and Writing Up Research

This stage is the culmination of all your hard work. Your project may have a strict assigned word count. If so then you should consider carefully what you want to write about. Clearly you have to address the questions you posed, but you can also include some discussion of the strengths, limitations and constraints of your project. Your research may have highlighted some unexpected findings, and this should be

considered as a possible area for further research. Evidence-based practice does serve to decrease the uncertainty that patients and healthcare professionals experience in a complex healthcare system. Your research, if in the public domain, will be scrutinised. It is worth considering what practitioners will be looking for when critiquing your work; this should be considered when developing you protocol. Beaven and Craig [16] suggest considering the following.

- Are the results of the study valid?
- Is the quality of the study good enough to produce the results that can inform clinical decisions?
- What are the results and what do they mean?
- If your research relates to practice, you might also consider whether the results of the research can be applied in a clinical setting.

Structure and presentation of a research dissertation are covered in Chap. 17. Writing for publication in order to disseminate research findings is presented in Chap. 18.

7.12 Conclusion

Designing and carrying out a research project is a multifaceted activity. Planning a study can often take more time than conducting research. The key to planning a research project is organisation and good time management. When undertaking a research study it is important to review relevant literature to inform a proposed study's methodology and contextualise what is known about the subject area. After reviewing literature, informed decisions regarding the methodology for a proposed study can be made. Ensuring you are aware of the advantages and disadvantages of the selected methodology to collect your data is essential to increase the validity or trustworthiness of the data collected. It is very important to organise and hold regular meetings with your project supervisor to ensure your progress is monitored and deadlines are met.

References

1. Blaxter L, Hughes C, Tight M. How to research. 3rd ed. Maidenhead: Open University Press; 2006.
2. Carter S, Henderson L. Approaches to qualitative data collection in social science. In: Bowling A, Ebrahim S, editors. Handbook of health research: investigation, measurement and analysis. Maidenhead: Open University Press; 2005.
3. Marshall G, Brennan P. The process of undertaking a quantitative dissertation for a taught MSc: personal insights gained from supporting and examining students in the UK and Ireland. Radiography. 2008;14:63–8.
4. Parahoo K. Nursing research: principles, process and issues. 3rd ed. Basingstoke: Palgrave Macmillan; 2014.
5. Polgar S, Thomas SA. Introduction to research in the health sciences. 6th ed. Oxford: Churchill Livingstone; 2013.

6. Cohen L, Manion L, Morrison K. Research methods in education. 8th ed. London: Routledge Falmer; 2017.
7. Bettany-Saltikov J. How to do a systematic literature review in nursing: a step-by-step guide. Maidenhead: McGraw-Hill; Open University Press; 2012.
8. Gerrish K, Lathlean J. The research process in nursing UK. 7th ed. Hoboken: Wiley-Blackwell; 2015.
9. Davies MM, Bath PA. Interpersonal sources of health and maternity information for Somali women living in the UK: information seeking and evaluation. J Doc. 2002;58(3):302–18.
10. Hammersley M. What's wrong with ethnography? Methodological explorations. London: Routledge; 1992.
11. Bowling A. Research methods in health. 4th ed. Berkshire: Open University Press; 2014.
12. Bergman MM. Advances in mixed methods research. Los Angeles: Sage; 2008.
13. Blaxter L, Hughes C, Tight M. How to research. 4th ed. Maidenhead: Open University Press; 2010.
14. DePoy E, Gitlin LN. Introduction to research. 4th ed. London: Mosby; 2015.
15. Holloway I, Galvin K. Qualitative research in nursing and healthcare. 3rd ed. West Sussex: Wiley Blackwell; 2017.
16. Beaven O, Craig JV. Searching the literature. In: Craig JV, Smyth RL, editors. The evidence-based practice manual for nurses. 3rd ed. Edinburgh: Elsevier; 2011.

Types of Information

8

Riaan van de Venter

8.1 Introduction

The research problem emanating research questions, hypotheses, and aim and objectives of a research study primarily drives the subsequent research design and methods that a researcher would utilise to conduct a particular study. In turn, these aspects of a research process and study determine the type of information that needs to be gathered in order for a researcher to meet the underpinning objectives and aim, as well as answer the research questions and hypotheses of a study. To this end, researchers use various types of information, in a multitude of ways, to find and present answers to their research questions.

An exposition of common information types that are utilised in diagnostic imaging and radiotherapy research is presented in this chapter. The discussion covers three categories: literature, quantitative and qualitative information. The focus in this chapter is twofold: to describe the types of information and to provide some insights on when and how to use the information.

8.2 Literature

Researchers continuously need information to formulate research problems or find answers to research questions. Literature can fulfil both of these in a variety of ways. Literature refers to the plethora of sources that exist in various media: for example, digital, print, artefacts, images, text and audio recordings, related to a

R. van de Venter (✉)
Department of Radiography, Nelson Mandela University, Port Elizabeth, South Africa
e-mail: riaan.vandeventer@mandela.ac.za

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_8

topic of interest [1]. Research literature can be defined as the current available body of knowledge, made up of publications in various formats (for example, original research articles, opinion papers, editorials, legal documents, blogs, tweets and videos) related to a specific research problem of interest. This set of literature includes current theories, hypotheses, practices, evidence and recommendations for further research pertaining to the research area of interest [2]. Furthermore, literature also sets the scene for one's quest for knowledge to ultimately inform theory and practice in diagnostic imaging and radiotherapy [2, 3]. In scientific research literature can be used as a literature review or a literature control.

8.2.1 Literature Review

A literature review entails activities and processes of searching, identifying, recording, understanding, sense-making, evaluating and synthesising what is already known on a topic or subject of interest, and then presenting it in a manner that can be understood by the diagnostic imaging and radiotherapy community, as well as the broader scientific community [3, 4]. It can also be seen as an end product of a research study because it often entails various decisions and justifications by researchers regarding the inclusion of particular research literature opposed to other research literature, for example, or the specific presentation of the current status of the existing knowledge and debates on the area of interest [3]. A literature review has the following functions [1, 3–5].

- It provides a theoretical background and justification for a study to be conducted and facilitates focussing on and clarification of a research problem.
- It broadens researchers' knowledge based in their chosen subject of study.
- It provides a researcher with information on research design and methodology to follow.
- It assists a researcher to develop a theoretical or conceptual framework for a proposed study.
- It situates and contextualises a proposed study's significance in relation to what already exists.
- It enables a researcher to integrate and demonstrate how a conducted study advances the corpus of knowledge within a specific discipline.

Depending on the purpose of a study a researcher wants to conduct, a literature review can take on one of two major formats. The presentation can be in a narrative or systematic format. A narrative literature review commonly forms the introduction and/or background to research reports; it can also be a standalone product. In contrast, systematic literature reviews are research studies on their own. They are a standalone product, but can also form part of a bigger research study [3]. Table 8.1 provides an overview of the main characteristics of each literature review type [3, 6, 7].

Table 8.1 Characteristics of the two main formats of a literature review

Characteristic	Narrative literature review	Systematic literature review
Purpose	• Summarises and critiques the existing body of knowledge to provide a broad overview of what is already known on the topic of interest • Provides a background to a research problem to be studied	• Summarises and critiques existing literature on a topic of interest, by integrating either quantitative or qualitative data or both, to address a specific research question by following meticulous, rigorous methods that are spelled out in advance
Stance	• Reflexive, subjective and interpretive	• Neutral, objective and aggregative/ translative
Literature inclusion	• Non-exhaustive, and only a sample of literature is used • No explicit inclusion/ exclusion criteria • No formal critical appraisal of literature and therefore no inter-rater agreement	• Use all literature that exist in relation to the research question • Explicit inclusion/exclusion criteria • Formal critical appraisal of literature is done, and the final appraisal is subject to inter-rater agreement
Methodology	• No exact methodology is followed; the process is therefore iterative	• A precise methodology is followed with regard to research question formulation, literature search strategy and reporting of findings. Thus, the process is more linear
Sense-making of literature	• Inductive reasoning is used to interpret the literature and extract meaning	• Deductive or abductive reasoning is used to interpret and extract meaning from the literature
Rigour	• Less rigorous	• More rigorous
Variations	• General review • Theoretical review • Methodological review • Historical review	• Narrative systematic review • Meta-analysis • Meta-summary • Meta-synthesis • Rapid review • Integrative literature review

8.2.2 Literature Control

Literature control is used as a premise of departure to which one can compare and contrast one's results or findings to demonstrate the significance of one's study in relation to the existing body of knowledge in the field [8]. It locates a researcher's study findings within the existing body of knowledge to demonstrate how the findings contribute, further develop, negate or challenge the current state of knowledge related to the topic of interest [9]. It can be used to provide readers with an insight into how a researcher conceptualised a specific research issue, interpreted the research findings or chose a particular research instrument [10]. One would typically use literature control in the discussion section of a research report, thesis or journal article. For example, a researcher would discuss how a study contributes to the advancement of knowledge, theory and practice in diagnostic imaging and radiotherapy.

Diagnostic imaging and radiotherapy theory and practice cannot only be advanced with existing evidence. Therefore, there is a continuous need to explore new avenues within the profession to inform evidence-based practice. Researchers design and conduct a variety of research studies that produce and make use of quantitative or qualitative information, or both information types. Quantitative and qualitative information is explored below. Some of the relevant types used in medical imaging and radiotherapy research are presented.

8.3 Quantitative Information

Quantitative information is the final presentation after quantitative data have been interpreted and organised in a meaningful manner. Data are generally considered quantitative if they are numerical in nature. Researchers can use participant observation, surveys/questionnaires (digital, hardcopy, face-to-face, telephonic, self-reported) or experimental techniques to obtain quantitative data. These methods of data collection can consist of closed or open-ended questions, or a mixture of both. The data are interpreted in a rigid manner through statistical analysis, to either describe the characteristics of the sample/population or to make inferences (see Chap. 15). By implication, quantitative studies usually make use of large sample sizes in order to be representative of a target population. The types of data thus collected relate to measurements, scores or counts, for example [11]. Quantitative data are classified in several ways based on the amount or characteristics of the information as shown in Fig. 8.1. Data can be categorised in four levels: nominal, ordinal, interval and ratio. Nominal data provide the least amount of information and ratio data the most. These four levels can be further classified into categorical and numerical data, and even further into non-metric and metric data [12]. Let us have closer look at each of the four data levels in greater detail.

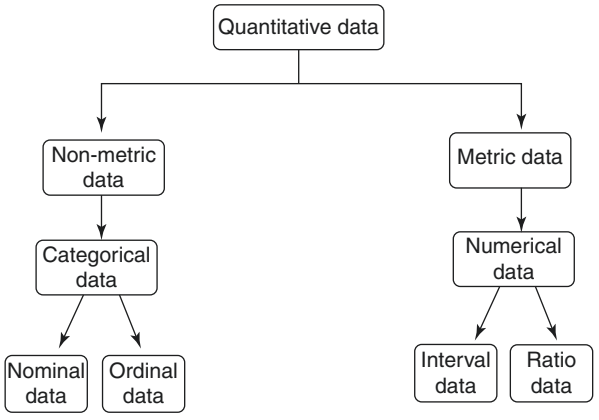


Fig. 8.1 Classification of quantitative data

8.3.1 Nominal Data

Nominal data classify or name information into categories or groups, as opposed to measuring the data. Values can be distinguished from one another by different names. A nominal scale usually has two or more classes or categories. When using nominal scales, a researcher allocates a category or class with identical features with an identical numerical value [11–13]. For example, if a researcher is collecting data by means of a questionnaire with regard to the gender of patients visiting a radiotherapy department then it will be necessary to list three options: male, female, other (please specify). During the data analysis stage a researcher would then code each of these categories numerically: male 1, female 2 and other (please specify) 3. These numerical codes would then be subjected to statistical analyses so that a researcher can then interpret and present the findings regarding the gender of patients visiting the particular radiotherapy department. Other examples of categories in this data type are marital status, tribe and even feelings. One aspect to remember is that the categories or classes of nominal data are usually mutually exclusive and there is ordinarily no order or ranking; comparisons between categories are thus not possible as they are qualitatively dissimilar [11, 12]. By implication, the questions are therefore close-ended.

8.3.2 Ordinal Data

Ordinal data allow for mutually exclusive classes or categories to be arranged in some rank order. This allows for comparisons between variables. Researchers can compare the differences in intensity or magnitude to find out whether two values of a variable are equal, lesser than or greater than the other. Surveys usually use Likert scales when working with ordinal data to group the numerical values together for a particular category. The use of a Likert scale (e.g., 5-point Likert scale or 7-point Likert scale) allows a researcher to obtain the degree of agreement of a respondent with statements in a questionnaire [11–13]. If a researcher wants to find out how satisfied a patient was with the level of care provided during a contrast medium examination in a particular radiology department by the radiographers, then use could be made of a 5-point Likert scale as shown in Table 8.2.

Respondents should mark the most appropriate answer based on their experience or opinion. Other examples of ordinal types of data include levels of education, levels of distress, ranking radiographic images in order of resolution from lowest to highest, or the level of importance a radiotherapist attaches to a particular professional skill attribute. The questions asked are usually closed-ended in nature.

Table 8.2 Example of an ordinal data type question

The level of care that radiographers provided me during my contrast medium examination was				
1	2	3	4	5
Very poor	*Poor*	*Satisfactory*	*Good*	*Very good*

8.3.3 Interval Data

Interval data allow for numerical measurements to differentiate between magnitude and quantity between different values of a variable. The units of measure are also always equal for all values attached to a variable; definite conclusions can therefore be made between the differences between value points in close-ended questions [11–13]. Examples of interval data could be the following.

- The temperature of an X-ray unit at a specific time
- The number of hours that a radiographer spends completing a particular quality control test in five general X-ray rooms
- The emotional intelligence quotient of first year radiography students
- Radiation dose
- Body mass index (BMI)
- Blood pressure

Another example could be the number of days that a patient had diarrhoea following radiotherapy treatment to the pelvic area. To obtain relevant data a researcher could include a question to determine when a patient had diarrhoea as shown in Table 8.3.

The range of options to select interval data are useful when a researcher aims to investigate any sort of relationship between variables or test a hypothesis. Interval data are commonly used in experimental studies where a researcher may want to assess the effect of kilovoltage peak (kVp) on the temperature of the anode of an X-ray tube. The kVp is then altered and the different anode temperatures noted for each kVp setting. The kVp will be the independent variable; the anode temperature will be the dependent variable. Depending on the hypothesis to be tested, the tabulated data of kVp values versus anode temperature can then be subjected to statistical analyses to establish whether the null hypothesis should be accepted or rejected.

8.3.4 Ratio Data

A ratio data scale has a true zero value. It is unique compared to the other three data types discussed above. A zero value represents the absence of the particular measurement that a variable measures. Ratio scales have equal intervals and magnitude properties. This means one should be able to assume that two units on a scale

Table 8.3 Example of an interval data type question

Following my radiaton treatment session to my pelvic area I had diarrhoea for:
☐ 1 day
☐ 2 days
☐ 3 days
☐ 4 days

represent twice the distance of one unit. Put differently, if one unit represents 10 m, then two units would represent 20 m (10 m × 2 = 20 m). Examples of ratio data and scales include physical measurements of weight (mass), length, time, volume, distance and angles [11–13]. Ratio data can be used in a similar fashion to interval data with regard to establishing relationships, associations and differences.

8.3.5 Literacy and Language Barriers and Obtaining Appropriate Data

Sometimes researchers are faced with the challenge to ensure that the data that they obtain are accurate, but the respondents may have differing levels of literacy or a language barrier may exist. From personal experience, using visual cues in a questionnaire, or survey, may assist in alleviating and countering this challenge. For example, if you want to obtain information about the respondents' biological sex or level of satisfaction with a service that received, you may formulate the questions as displayed in Table 8.4.

8.3.6 Quantitative Data Analysis and Display

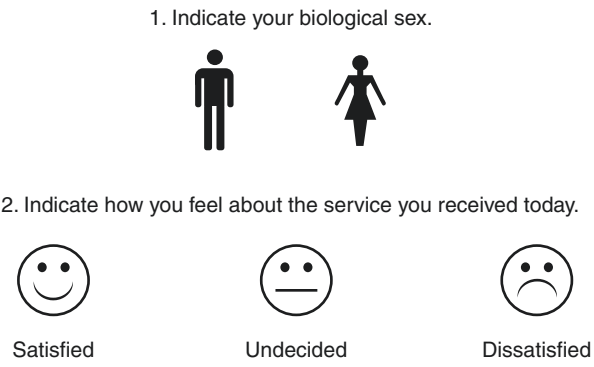
When the required data have been collected by a researcher then this is followed by data interpretation/analysis. This is required to present the information in a meaningful manner. This depends on the information available and the intended audience.

8.3.6.1 Data Analysis

Researchers make sense of quantitative data by using a variety of statistical analyses for interpretation of such data [7, 13, 14]. However, an in-depth exploration of data analysis is beyond the scope of this chapter.

Use is made of flowcharts in Figs. 8.2, 8.3 and 8.4 to depict the next step of the research process after data collection. The purpose of a research study should inform

Table 8.4 Example of formulating questions to address literacy and language barriers



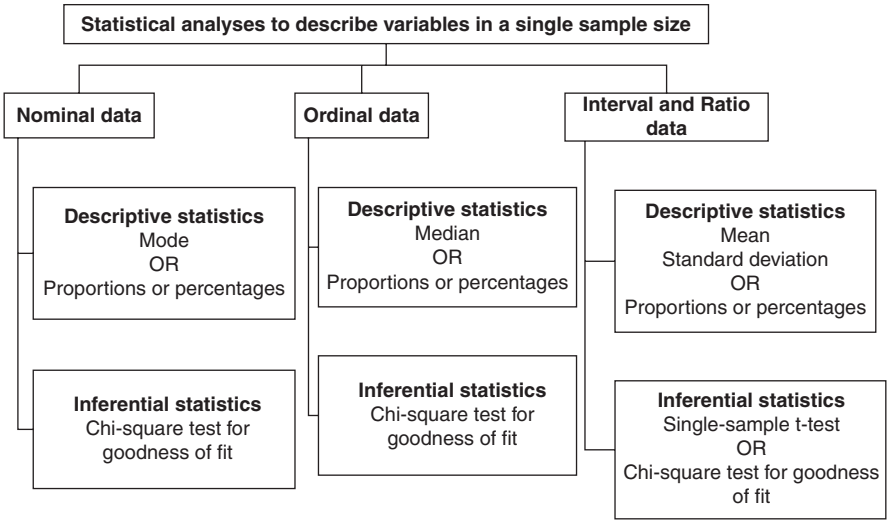


Fig. 8.2 Describing a single sample size with statistics

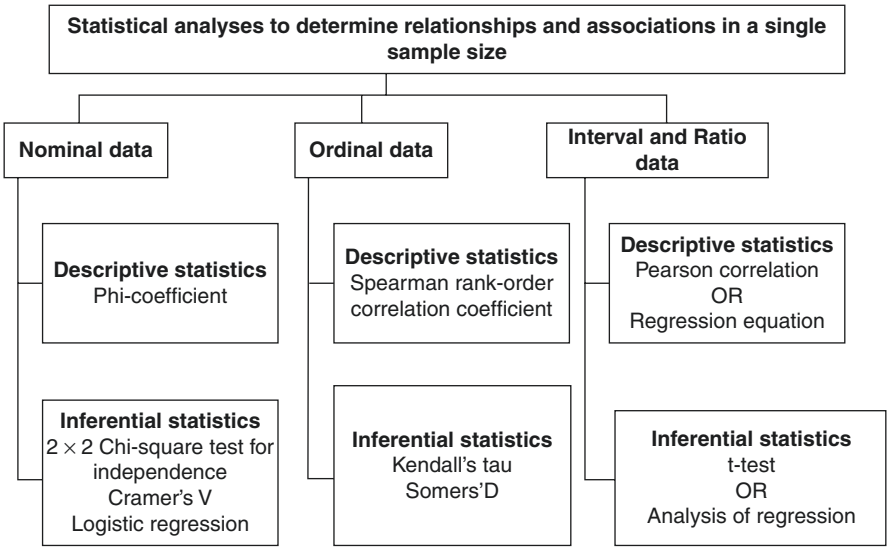


Fig. 8.3 Determining relationships or associations in a single sample size with statistics

the decisions made regarding the appropriate statistical techniques used for analysis. Figure 8.2 demonstrates the decisions of a researcher to describe variables in a study in a single sample size: one score per respondent per variable. Figure 8.3 demonstrates the statistical techniques that can be utilised when relationships or associations need to be established from a single sample size where two variables

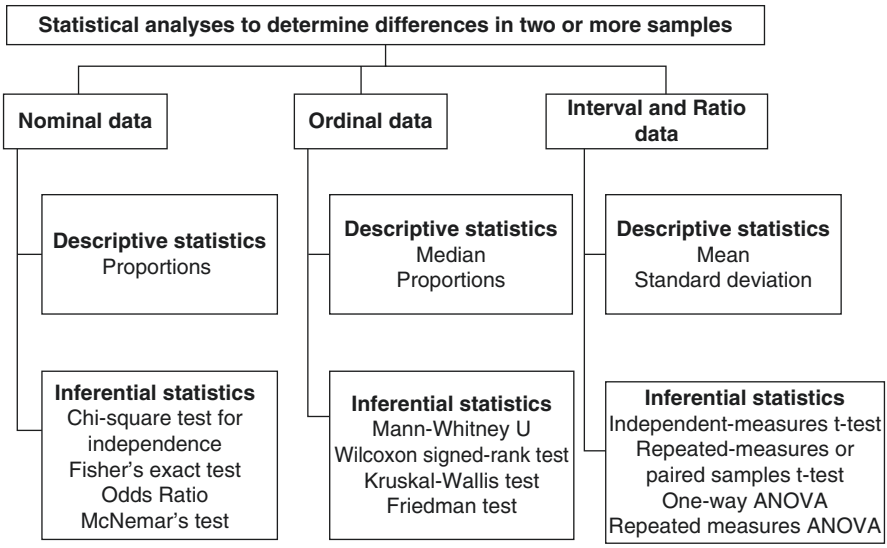


Fig. 8.4 Determining differences in two or more sample sizes with statistics

are measured for each respondent. Figure 8.4 portrays the statistical analyses that can be done to determine the differences between groups of variables for two or more different sample sizes.

8.3.6.2 Data Display

Data display involves a process of organising raw data into comprehensible and unambiguous information so that an intended audience or readership can understand the results. The choice of presenting the organised data for interpretation depends on whether one is dealing with discrete or continuous data.

Discrete data can be counted. A researcher thus needs to determine distribution frequencies for each category of a particular variable being measured. This can also subsequently be presented in distribution frequency tables, bar graphs or pie charts. Data that fall within the discrete data realm are nominal and ordinal data [2].

The use of distribution frequency tables allows a reader to see how many cases fall into a particular category. If a bar graph is used a reader can determine the number of cases in each category by looking at the height of the bar and with which value it intercepts on the y axis. A pie chart can be used to demonstrate the proportions relative to the number of cases that make up each category for a particular variable. These points are illustrated in the following example.

- A researcher is interested in determining how many patients, by sex, have undergone chest radiography at a public hospital for a period of 7 days. Using the patient register, the following information was obtained (M = male; F = female): M F M M M F F F F F F F M M M M M M F F F F.

Table 8.5 Frequency distribution of patients undergoing chest radiography for 7 days

Sex	Frequency (f)
Male (M)	10
Female (F)	13
Total (n)	23

Fig. 8.5 Bar graph

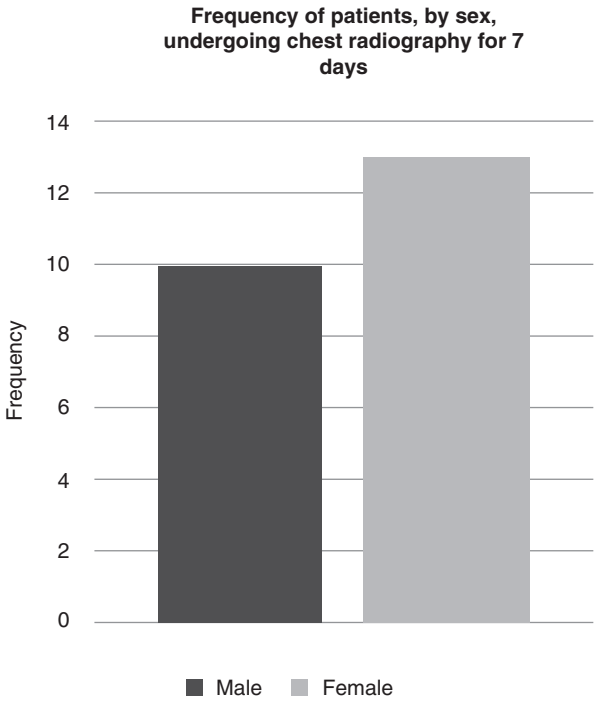


Table 8.5 illustrates presentation of the data in a distribution frequency table. The numerical data may be presented in bar graph (Fig. 8.5) or a pie chart (Fig. 8.6).

Continuous data produce real numbers and can be processed by standard mathematical rules [2]. Interval and ratio data are categorised. Similar to discrete data, a researcher could display this information by means of grouped distribution frequency tables or graphically by means of histograms, frequency polygons, line graphs or scatter plots.

One could group the number of radiographic images rejected as part of a quality audit of hospitals in a particular district. If there are ten hospitals in the district where the audit was being conducted, then a grouped distribution frequency table may look like the one represented by Table 8.6. The grouped distribution frequencies could be used to compile a frequency polygon. By looking at the information in Fig. 8.7 a reader should be able to understand frequency of the unplotted values. The trend of the number of patients visiting a department over a period of one year, for example, can be represented in a line graph (Fig. 8.8). A reader should be able to see that the most patients were seen in the department during

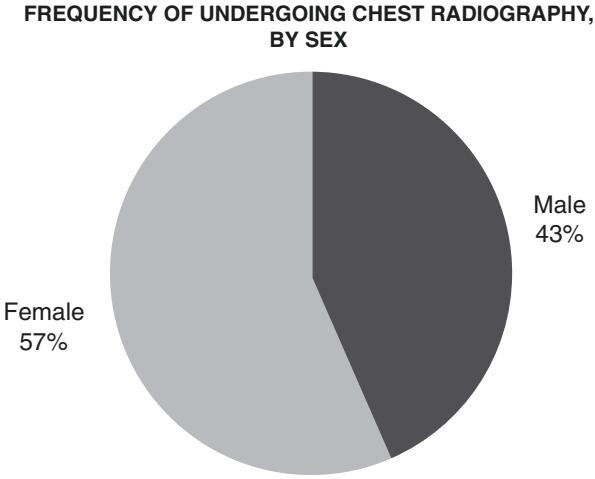


Fig. 8.6 Pie chart

Table 8.6 Grouped frequency table

Class interval	Frequency (*f*)
0–10	1
11–20	2
21–30	4
31–40	2
41–50	1
Total (*n*)	10

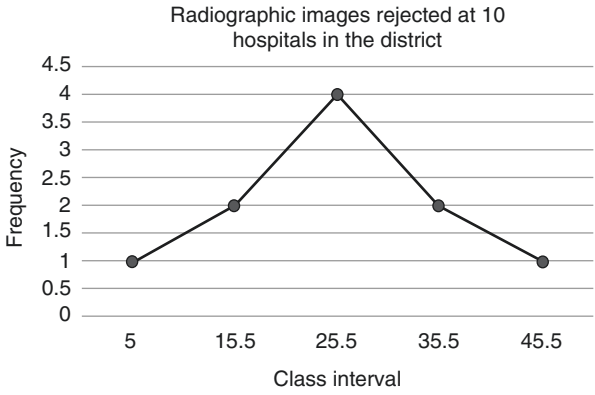


Fig. 8.7 Frequency polygon

December, and the least during March 2018. Histograms can be used to represent continuous data; they look similar to bar graphs except that the bars touch one another. Similar to the bar graphs the height of the bar represents the frequency of cases in each category or class interval. When a relationship between two variables needs to be graphically depicted, a scatter plot would be meaningful as demonstrated in Fig. 8.9.

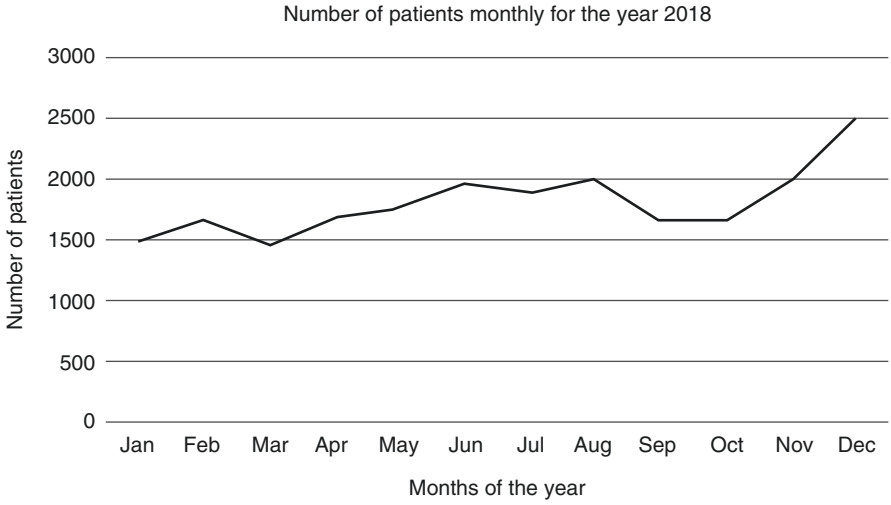


Fig. 8.8 Line graph

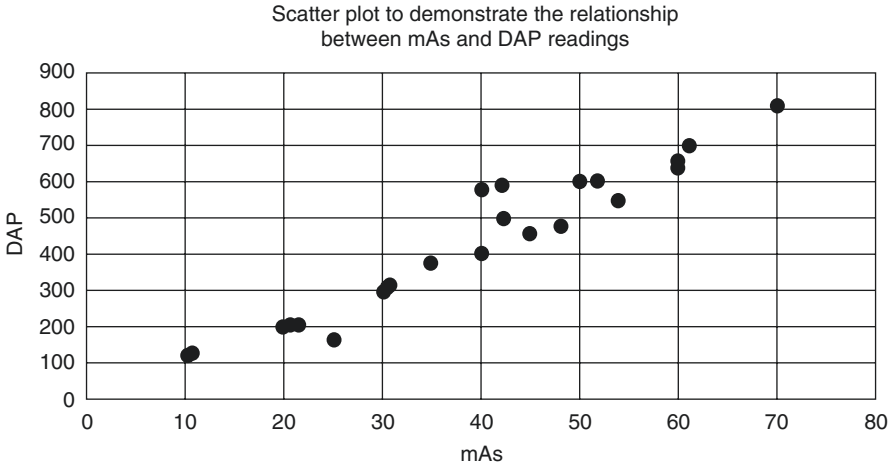


Fig. 8.9 Scatter plot

8.3.7 Accuracy, Specificity, Sensitivity

Establishing the performance of radiographers' image interpretation abilities is well covered in diagnostic imaging research. With the establishment of and continuous research into radiographer-led image interpretation practices and the effect of education and training, researchers may be interested to determine participants' performance regarding accuracy, specificity and sensitivity pertaining to image interpretation.

When applied in this context, researchers basically want to determine radiographers' ability to distinguish between normal and abnormal radiographic images. The radiographers' decisions or responses may be classified as follows.

- Correctly interpret a radiographic image as abnormal (true positive).
- Correctly interpret a radiographic image as normal (true negative).
- Incorrectly interpret a radiographic image as abnormal (false positive).
- Incorrectly interpret a radiographic image as normal (false negative).

Using these responses, the sensitivity, specificity and accuracy of radiographers' image interpretation ability can then be calculated [15–19]. This information can also be represented graphically using the receiver-operator characteristic (ROC) curve (Fig. 8.10). The closer the curve is to the upper left corner, the higher the accuracy of a radiographer with regard to discriminating between normal and abnormal radiographic images. The scale used for both the y and x axes usually ranges between 0 and 1. The greater the area is under the curve (AUC), the better a radiographer's accuracy. These measures can also be used to determine the performance of different imaging modalities in a department or different procedures [16, 20–22].

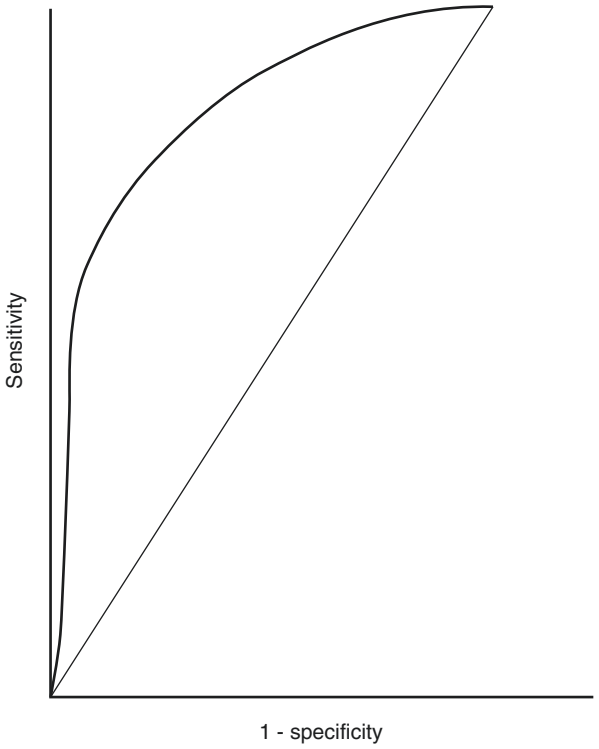


Fig. 8.10 An example of the characteristic shape of the ROC curve

8.4 Qualitative Information

The other side of the coin, relative to quantitative information, is qualitative information (see Chap. 16). This is sourced, documented, interpreted, organised and presented in a totally different manner. Raw qualitative data can be utterances of individuals, written narratives, visual artefacts (digital or physical) or audio recordings, for example. Researchers obtain this kind of data through unstructured, semi-structured or structured interviews; reflective journaling; visual methodologies; and observations [23]. The aim of qualitative studies is to understand phenomena as they occur in their natural setting and how individuals make sense of their experiences or context they find themselves in [23, 24]. Examples of qualitative topics may include: the lived experiences of therapeutic radiographers regarding workplace bullying; the particular culture in a diagnostic radiography department; or the manner in which undergraduate diagnostic radiography students experience and cope with death and dying during clinical placements. Looking at these examples one can see that they deal with rather specific and sensitive information. The elicitation of appropriate and rigorous information is imperative to ensure credible findings in the end to inform the theory and practice of diagnostic imaging and radiotherapy. Hence, data gathering methods alter depending on the focus of a research study to be conducted. Table 8.7 provides common research purposes and the appropriate qualitative data gathering approach to use as well as research designs [23].

Qualitative data need to be recorded in appropriate ways in preparation for analysis and interpretation. Recording can be done by transcriptions on paper, audio-visual recordings or photographs/collages/sketches and the like. During the analysis and interpretation phases of working with qualitative data and information, researchers need to go about this in a systematic manner in order to hold scientific grounding and to ensure trustworthiness of the information presented at the end of a research process.

Table 8.7 Research study aims and corresponding data gathering technique and research design

Research study aim/purpose	Data gathering approach	Research design
Understanding the culture or a group of people	Participant observation Document analysis Interviewing	Ethnography Grounded theory
Gaining insights into converging or diverging perspectives on a particular topic especially when exposure of participants to various perspectives are crucial to the study	Focus groups	Explorative-descriptive
Understanding a phenomenon from individual perspectives or the lived experiences of individuals	Individuals interviewing	Phenomenology Explorative-descriptive Case study

Data analysis in qualitative studies starts with coding. The process of coding entails, for example, a researcher reading all the interview transcripts generated and then assigning a meaningful unit to a segment of the transcript to capture the meaning thereof. This is done for all transcripts. Thereafter, this initial set of codes generated during this first cycle of coding is clustered into code families or groups. A researcher then assigns the code of this newly formed group of codes to the respective segments of the transcripts or keeps a thorough list of the codes clustered in the particular group. This clustering continues to form categories by way of further reducing the code groups or families. These categories are then further refined to form themes and sub-themes. In the final research report, it is then the themes and sub-themes that are discussed. A theme and sub-theme abstractly capture the participants' experiences. For example, workplace bullying in a radiotherapy department and the various manifestations related to the experiences [8]. The content of each theme and sub-theme is formed by the categories based on the data. This is known as a thematic synthesis or thematic analysis based on the content of the data, which is the most commonly used method in diagnostic imaging and radiotherapy research [6, 8, 25]. The main factor, in the data analysis process, is the research question underpinning a study. It is easy to go astray when analysing and interpreting data. Analyses and interpretations pertinent to the research questions should be the focus and included in the final presentation of the information. One can therefore appreciate that data analysis works in a reductionist fashion to make sense of the data. This allows a researcher to present data in a meaningful manner for the readers (see Fig. 8.11). The presentation of qualitative information is done in a descriptive and narrative format. Extracts from the data sources are used as substantiating evidence for the narrative or description provided. The same procedure can

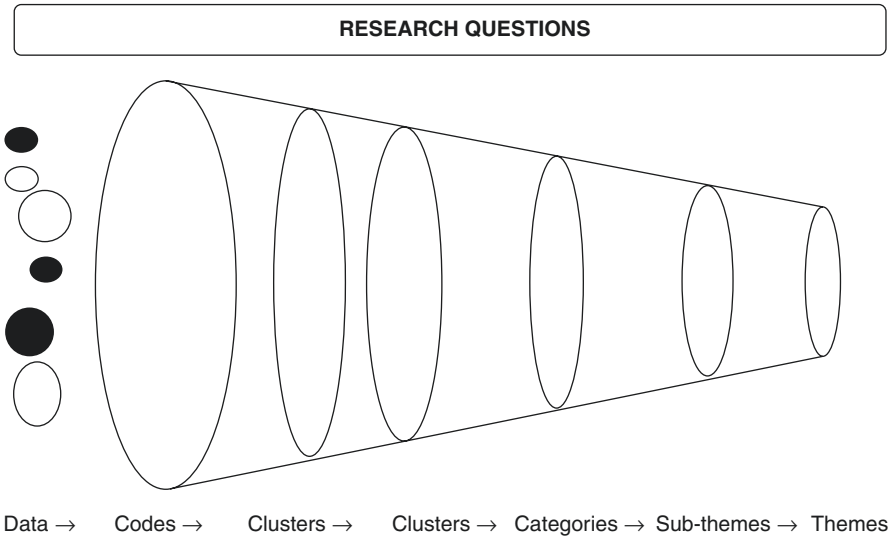


Fig. 8.11 Data reduction from data to themes underpinned by the research questions

be followed using computer-assisted qualitative data analysis software (CAQDAS) like ATLAS.ti and NVIVO [8].

A trend in qualitative research is the use of visual methodologies. An in-depth exposition of this methodology is beyond the scope of this chapter. Nonetheless an introduction to this field is deemed necessary.

8.4.1 Visual Methodologies

Visual methodology is an umbrella term for social sciences methods that use arts-based techniques to generate and represent knowledge. These methods are often used for topics where participants are required to think differently about a topic or subject and where they have to represent the unconscious mind in a conscious manner. Some people can also represent their ideas and experiences better in a visual manner opposed to in the written or spoken word. Particular methods included under visual methodologies are naïve (unplanned) sketching, photo-voice, photo-elicitation, collages, video, sculpture (for example, Mmogo-method™).

The use of visual artefacts as elicitation or transmittal media allows participants, during variations of interviewing or focus groups, to verbalise their viewpoints on the topic at hand. For example, the author personally used the Mmogo-method™ as a transmittal medium in a study to understand undergraduate diagnostic radiography students' experiences and coping with death and dying in the workplace. Another example where visual methodology may be useful is in studies focussing on resilience.

Visual methodologies almost always are combined with other forms of data gathering whether it is written text or spoken words. Visual methods add another dimension to qualitative information that is otherwise unattainable and enhance data richness [26, 27].

8.5 Conclusion

Data need to be sourced using appropriate methods/techniques. Collected data in academic writing are recorded, analysed, interpreted, organised and then presented as meaningful information to the readers. What is deemed appropriate is dependent on the approach to research that researchers take and the underpinning research questions of a study. An overview of common data and information, including visual methodologies, used in diagnostic imaging and radiotherapy, was covered in this chapter. The examples and discussion are by no means an exhaustive body of literature. The information used in research is grouped into three classes: literature, quantitative and qualitative information. The type of information that a researcher ends up with should be guided by the purpose of the study; there is no one size fits all. A clearly formulate research problem, research question(s), aim and objectives are crucial for any research study.

References

1. DePoy E, Gitlin LN. Introduction to research: understanding and applying multiple strategies. 5th ed. St. Louis: Elsevier; 2016.
2. Polgar S, Thomas SA. Introduction to research in the health sciences. 6th ed. Churchill Livingstone Elsevier: Edinburgh; 2013.
3. Onwuegbuzie A, Frels R. 7 steps to a comprehensive literature review: a multimodal and cultural approach. London: SAGE Publications Ltd; 2016.
4. Kumar R. Research methodology: a step-by-step guide for beginners. 5th ed. London: SAGE Publications Ltd; 2019.
5. Hofstee E. Constructing a good dissertation: a practical guide to finishing a Masters, MBA or PhD on schedule. Johannesburg: EPE; 2006.
6. Polit DF, Beck CT. Essentials of nursing research: appraising evidence for nursing practice. 9th ed. Philadelphia: Wolters Kluwer; 2018.
7. Gray JR, Grove SK, Sutherland S. Burns and Grove's the practice of nursing research: appraisal, synthesis, and generation of evidence. 8th ed. St. Louis: Elsevier; 2017.
8. Creswell JW. Research design: quantitative, qualitative & mixed methods approaches. 4th ed. Thousand Oaks: SAGE Publications Inc.; 2014.
9. Braun V, Clarke V. Successful qualitative research: a practical guide for beginners. Los Angeles: SAGE; 2013.
10. Mouton J. How to succeed in your master's and doctoral studies: a South African guide and resource book. Pretoria: Van Schaik Publishers; 2001.
11. Bless C, Higson-Smith C, Sithole SL. Fundamentals of social research methods: an African perspective. 5th ed. Cape Town: Juta; 2013.
12. Maree K, Pietersen J. The quantitative research process. In: Maree K, editor. First steps in research. 2nd ed. Pretoria: Van Schaik Publishers; 2016. p. 161–72.
13. Flick U. Introducing research methodology. 2nd ed. London: SAGE Publications Ltd; 2015.
14. Gravetter FJ, Wallnau LB. Essentials of statistics for the behavioral sciences. 8th ed. Wadsworth: Cengage Learning; 2014.
15. Gunn C. Radiographic imaging: a practical approach. 3rd ed. Edinburgh: Churchill Livingstone; 2002.
16. Manning DJ. Evaluation of diagnostic performance in radiography. Radiography. 1998;4:49–60.
17. Piper KJ, Patterson A. Initial image interpretation of appendicular skeletal radiographs: a comparison between nurses and radiographers. Radiography. 2009;15:40–8.
18. MedCalc. ROC curve analysis: introduction. https://www.medcalc.org/manual/roc-curves.php. Accessed 20 Apr 2019.
19. McConnell J, Eyres R, Nightingale J. Interpreting trauma radiographs. Oxford: Blackwell Publishing; 2005.
20. Vining DJ, Gladish GW. Receiver operating characteristic curves: a basic understanding. Radiographics. 1992;12:1147–54.
21. Obuchowski NA. Receiver operating characteristic curves and their uses in radiology. Radiology. 2003;229:3–8.
22. Park SH, Goo JM, Jo C-H. Receiver operating characteristic (ROC) curve: practical review for radiologists. Korean J Radiol. 2004;5(1):11–8.
23. Holloway I, Galvin K. Qualitative research in nursing and healthcare. 4th ed. Oxford: Wiley Blackwell; 2017.
24. Lobiondo-Wood G, Haber J. Nursing research: methods and critical appraisal for evidence-based practice. 9th ed. St. Louis: Elsevier; 2018.
25. Gray DE. Doing research in the real world. 4th ed. London: SAGE Publications Ltd; 2018.
26. Pink S. Visual methods. In: Jupp V, editor. The SAGE dictionary of social research methods. London: SAGE Publications Ltd; 2011. p. 321–2.
27. Roos V. The Mmogo-method™: an exploration of experiences through visual projections. https://repository.nwu.ac.za/handle/10394/4966. Accessed 15 Apr 2019.

Epidemiological Research Methods

9

David M. Flinton

9.1 Introduction

Epidemiology is the systematic study of the distribution of health and illness, or the investigation of factors affecting the health of populations, i.e., how often diseases occur in different groups of people. This information informs health professionals; links can be made between health and specific causes of disease in order to try and prevent illness and disease through effective health strategies and campaigns.

Epidemiology for the most part uses a binary outcome. Binary data can only take one of two values: has the disease or does not have the disease. Values can then be assigned to these two states: 1 = has the disease; 0 = does not have the disease. Summarising this variable is quite simple. It can be done in one of two ways using either proportions (also defined as risk or rates) or odds.

Measuring the disease frequency can be quite simple. For example, consider six subjects, one of whom has cancer while the others do not. In this instance we have clear binary outcomes, one person with cancer (a case) and five people who do not have cancer. If, however, we needed to measure cholesterol levels, there would be a continuum of severity rather than a binary outcome. When a continuum occurs, we could define a case as someone outside of two standard deviations from the norm, or perhaps base the cut-off on clinical importance, i.e., at a level where the risk of heart disease increases significantly. In the United Kingdom (UK) the healthy total cholesterol level in adults should be ≤5 mmol/L for healthy adults [1]. We could define a case of high cholesterol using this figure as a reference, but we are ignoring other possible figures such as the individual levels of non-HDL (high density lipoprotein) cholesterol and HDL cholesterol and the ratio. This is fine if everyone uses this definition; comparing results if researchers use different case definitions becomes problematical.

D. M. Flinton (✉)
School of Health Sciences, City, University of London, London, UK
e-mail: d.m.flinton@city.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_9

9.2 Disease Occurrence

As well as defining what a case is, it is also important to understand how disease occurrence is measured. Two common measures of a disease's occurrence are incidence and prevalence. Incidence is the rate at which new cases occur in a given time period. When the risk is roughly constant incidence is measured as:

$$\frac{\text{The number of new cases}}{\text{Population at risk} \times \text{Time during which cases were collected}}$$

If there were approximately 46,000 new cases of breast cancer in England and the population was 60 million, this would give an annual incidence of approximately 76.7 per 100,000. In fact, the rate reported in 2005 was almost exactly this with 45,947 new cases and a crude incidence rate of 76.3 per 100,000 [2]. The information this figure gives, called the crude rate, is quite basic and it can mask important information such as the different rates in males and females or the relationship of the incidence of breast cancer with age. To overcome this problem, specific rates can be used, breaking the data down into rates for specific age ranges and by gender. Standardised rates can also be produced, which allow better comparison between populations when there are differences in the populations that might affect the event we are looking at. For example, if we were interested comparing incidence rates of prostate cancer between two countries and the age structure of the two populations was very different, we could use standardised rates for comparison. The standardised rate is not an actual rate. It is the rate that a population would have if it had a 'standard' age structure. A frequently used standard is the World Health Organization (WHO) world standard population.

Prevalence is the actual number of cases at a given point in time (point prevalence) or in a certain period (period prevalence). It provides a better measure of disease burden (the impact of a health problem on a population) as it includes information on the total number of cases at/during a specific time.

The two measures, incidence and prevalence, are linked; when a new case (incident) occurs it joins the prevalence figures and stays there until the person either recovers or dies. The relationship between incidence and prevalence can therefore be expressed as:

$$\text{Prevalence} = \text{Incidence rate} \times \text{The average duration of the disease}$$

If the average duration of a disease is short, such as someone who has pancreatic cancer prevalence will be low because although new cases are constantly being added, cases will also be being removed relatively quickly through death. If the time period to recovery or death is long, as it often is with chronic illnesses such as asthma, then even a disease with a relatively low incidence rate can give rise to a high prevalence. Sometimes prevalence can change. For example, the use of lapatinib alone or in combination with trastuzumab increases the survival period in patients with HER2-positive metastatic breast cancer [3]. So, assuming the incidence rate remains the same use of lapatinib and trastuzumab will lead to a higher prevalence rate as the duration of the disease is increasing.

9.3 Disease Burden

Disease burden is a concept that measures the impact of a health problem on a given population. The impact of a health problem can be measured using different factors such as financial cost, mortality, morbidity, or other indicators. Two common measures used are: quality-adjusted life-years (QALY) and disability-adjusted life-years (DALY).

QALYs are a composite measure of disease burden that considers the gains in life expectancy and health-related quality of life. QALYs can be calculated by multiplying the duration of time spent in a health state by the health-related quality of life (HRQoL) weight (utility score) associated with that health state (years of life x utility value = QALY). A value of 1 QALY equates to 1 year in perfect health. An issue with using QALYs for cancer patients is that often the side effects of treatment can be quite severe so reducing a patient's HRQoL during the treatment phase, but there is the expectation that their life expectancy will increase, and their health improved in the long term [4].

DALYs are the sum of the years of life lost due to premature death in a population and the years lost due to disability for people living with the health condition or its consequences. When there are improvements in health this will reduce the DALY figure which is opposite to QALYs which are increased.

9.4 Study Designs

A number of different designs can be utilised in epidemiological studies most of which can be classed as observational studies with a control group. The term observation is used as a researcher's intention is to observe and not to interfere with the routine care being given.

These designs allow for large sample sizes and provide data when randomised controlled trials would be considered unethical. For example, knowing that asbestos might cause a cancer means that it would be unethical to deliberately expose individuals to this substance, but studying people who might be exposed to this agent as part of their job is possible.

The three most commonly used study designs are discussed in this section.

9.4.1 Cohort Studies

Cohort studies track people forward in time. In their simplest form they follow two groups (cohorts) from the point of exposure to the outcome of interest. The difference between the two groups is that one has been exposed to a risk factor whereas the other group has not. If on analysis a difference exists in the rate of the outcome of interest between the two groups, it can suggest an association between the exposure and the outcome.

Cohort studies are usually prospective in design, but occasionally can be retrospective. In a prospective cohort study an investigator recruits subjects and collects

baseline exposure data on all subjects. This is done before any of the subjects have developed any of the outcomes of interest. The subjects are then monitored periodically in order to record who develops the condition(s) of interest. Various methods of follow-up can be utilised: mailed questionnaires, phone interviews, face-to-face interviews, or follow-up examinations, for example. There is a clarity of temporal sequence in this form of design as the exposure precedes the outcome. It is also possible to make precise estimates of both the incidence of an outcome and the relative risk of an outcome based on exposure. A retrospective cohort study indicates that when the study was planned the data or part of it already existed. This design tends to be used when data become available that have already been collected for other purposes. The methodology is the same, but the study is performed retrospectively.

For example, if we were interested in investigating the effect of lens irradiation, as part of a CT scan in childhood, in terms of cataract formation in later life we could identify a suitable population to study such as children who have a condition/illness that requires CT scans of the head. A control group could then be established such as children attending outpatients who do not require any radiological examinations, and then both the cohorts could be followed for a number of years via the check-up clinics. At the end of each data collection period we could then compare the number of subjects with cataracts in the two groups.

This type of design is good for looking at rare exposures and can consider multiple exposures and multiple outcomes. The main advantage of this type of study is that unlike case–control studies there is no recall bias; it also avoids survivor bias (see Chap. 10). Weaknesses of this type of study are that if the outcome or event takes a long time to develop a study will take a long time to complete and therefore will be expensive; maintaining contact with individuals over long periods can sometimes be difficult. This makes such a study more prone to having a high drop-out rate and therefore very susceptible to attrition bias; those who stay with the study and those who drop out of the study, although similar at baseline, may be different at the time of follow-up.

9.4.2 Case–Control Studies

Case–control studies are usually retrospective in nature. A study starts by recruiting a sample of subjects with a particular health disorder (cases) and a group of subjects without the health outcomes (controls). We can then assess and compare both of their exposure to the risk factor in question. Therefore, in this type of study the outcome is measured before that of exposure. Since subject selection is based on exposure this makes this type of study very prone to selection bias.

For example, if we were interested in looking at lung cancer and smoking, we could identify a group of subjects with lung cancer (from the clinics) and a group of controls, patients without lung cancer from healthy staff employed at the hospital. A questionnaire could then be given to the subjects in both groups asking them about their smoking history. Matching the two groups is the main method available to a researcher for controlling the factors that might distort or confound the relationship

that is being studied. Two variables that are often matched in case–control studies are age and gender. Despite matching the data selection bias may exist as hospital workers might not be representative of the general population and it could be argued might be less likely to smoke than the general population due to their awareness of the impact of smoking on their health.

This type of study has the ability to study rare diseases. It has the advantages over cohort studies of being cheap and relatively quick to conduct. Such studies are, however, more prone to bias than cohort studies, particularly recall bias (see Chap. 10) due in part to differential recall of factors in the control group compared to the cases group.

In case–control studies because we start with the outcome it is impossible to comment on the incidence or prevalence. For example, if we had selected 50 cases and matched them with 50 controls the prevalence in the study would be 50%, but the figure is dependent on a researcher's choices not the incidence or prevalence of the disease and because of this odds ratios are reported rather than relative risk.

Other limitations of this type of study are: the design is not useful to study multiple outcomes as the cases are selected based on an outcome, so the study is only useful for the association between exposures and that particular outcome; and the temporal sequence of the exposure and outcome may be difficult to determine.

The main difference in design of cohort and case–control studies is shown in Fig. 9.1.

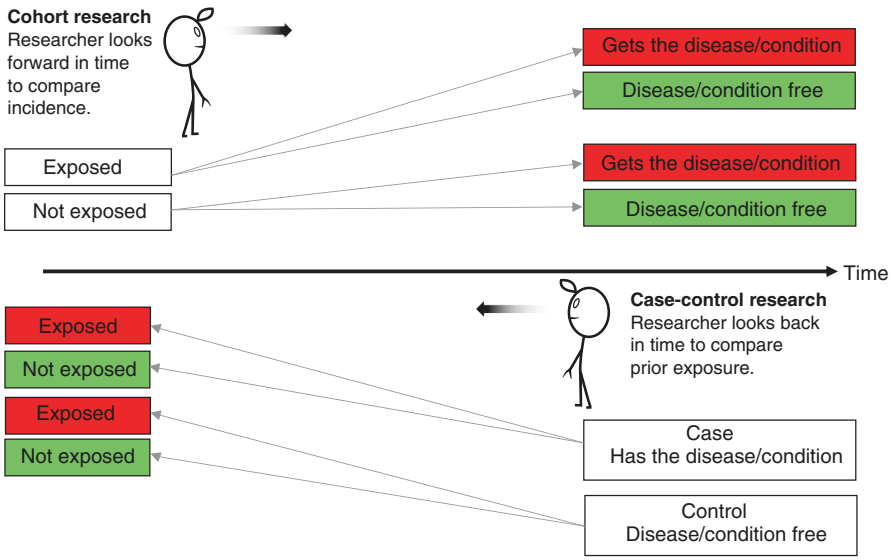


Fig. 9.1 Case study and cohort designs

9.4.3 Cross-Sectional Studies

Cross-sectional studies are surveys that determine exposure and disease at the same time or over a very short period. They are literally a 'snapshot' as all information on both the outcome and the exposure are collected at the same point in time. If we were interested in the prevalence of wrist problems in sonographers, a questionnaire could be sent to a representative sample. Depending on how detailed the questionnaire or survey was we might be able to look at the prevalence of wrist issues in sonographers and link it to parameters of the scan, such as the way they hold the probe, or the pressure exerted. What might be difficult to establish is what might be cause and effect. For example, is it holding a probe certain ways that causes wrist issues or do sonographers with wrist issues adjust the way they hold a probe?

The design of these studies range from simple to very complicated, and they are usually analysed by regression analysis. They are prone to bias and are not very suitable for rare conditions. As stated above it may be difficult to ascertain whether the exposure or the cause came first. They are however the quickest and cheapest of the studies detailed in this section, and they can compare multiple outcomes and exposures.

9.5 Bradford Hill Criteria

In 1965 Austin Bradford Hill proposed a set of nine criteria as a means of judging whether a causal relationship between a factor and a condition existed. The criteria are still widely used as a benchmark in determining a causal inference.

1. Temporal relationship
   The exposure to the factor in question must always precede the outcome. If wearing a lead rubber apron causes back pain, the wearing of it must precede the back pain. This is the only criterion that must always be met. Prospective cohort studies provide stronger evidence of this than a retrospective study or cross-sectional study.
2. Strength of association
   This is defined by the size of the risk as described in the following sections. The stronger the association, the more likely it is that the relationship is causal.
3. Consistency
   If the association is consistent across different studies in different populations and different designs, the association is more likely to be causal.
4. Biological gradient
   A dose response should be seen. Changes in disease rates should follow from a corresponding change in exposure. For example, we should see more lung cancer cases in subjects who smoke more and less in those who smoke less cigarettes.
5. Plausibility
   Does the association make sense? Is there a sound theoretical basis or one that can be postulated for making an association between a risk factor and the outcome being measured?

6. Specificity
   This is established when a single cause leads to a specific effect. This is considered by some to be a weak criterion as causality can be multiple.
7. Coherence
   The association should not contradict existing theory and knowledge. Coherence is similar to plausibility: both consider whether the cause and effect proposed makes sense.
8. Experiment
   Causation is more likely if there is evidence from randomised experiments. If we can get subjects to reduce the number of cigarettes smoked or stop completely, it will reduce the risk of developing lung cancer.
9. Analogy
   In some circumstances the effect of similar factors and their actions may be considered. In other words, if we know that one type of causal agent produces an effect, a second similar agent may cause a similar effect.

9.6 Odds and Risk

Cohort studies often report findings as a relative risk whereas odds ratios are used for case–control studies. To understand relative risk and odds ratios better, we must first look at what risk and odds are. Risk and odds are slightly different, as shown by the formulae below, but both are strongly linked with probability. The latter is covered in other chapters in this book.

$$\text{Risk} = \frac{\text{Number of outcomes resulting in the event}}{\text{Total number of possible outcomes}}$$

$$\text{Odds} = \frac{\text{Number of outcomes resulting in the event}}{\text{Number of outcomes NOT resulting in the event}}$$

If in a group of 200 subjects, only one patient presented with cancer, the risk of cancer within the group would be 1 in 200 or 0.005. If we report the odds on the same group, we could say that the odds were 1/199 or 0.005025. A small difference, and for rare events we can say that the odds and risk approximate each other. However, this difference becomes greater for common events. For example, if in a group of 200 patients, 50 had cancer, the risk of cancer within the group would be 50 in 200 or 0.25; the odds 50 in 150 or 0.33, so the risk of getting cancer is 25% and the odds 33%. Odds and risk only approximate each other for rare events; the more common the event, the greater the difference. Risk has the advantage of presenting the data in a way that is understood by most people.

Odds and risk contain the same information, so it is possible to calculate one value if the other is known.

$$\text{Odds} = \text{Risk} / (1 - \text{Risk}) \qquad \text{Risk} = \text{Odds} / (1 + \text{Odds})$$

In the above example the risk was 0.25; therefore to calculate the odds, we would substitute the following figures in the first formulae to calculate the odds. Odds = 0.25/(1–0.25) = 0.33.

9.7 Relative Risk or Risk Ratio (RR)

The relative risk is a ratio of two risks and describes the risk in one group as a multiple of the risk in a second group. If we have two groups, one that is exposed to a possible carcinogen and one that is not exposed, we can work out the relative risk by looking at the ratio of disease occurrence in both groups. Using Table 9.1 the relative risk would be calculated by looking at the risk in the exposed group called the experimental event rate (EER) [A/A + B] and the risk in the unexposed group, the control event rate (CER) [C/C + D]. The relative risk can then be calculated by dividing the EER by the CER. A resulting value of 1 would mean that the risk between the groups is identical: values above 1 represent a positive association (cause) and values below 1 a negative association (protective). The further a value is away from 1, the stronger the association between exposure and outcome.

If we look at some hypothetical data on subjects who are smokers and non-smokers shown in Table 9.2, it is possible to calculate the relative risk of lung cancer associated with this lifestyle activity.

We first need to calculate the risk for smokers, the EER, which is 288/545 or 0.53, 53%, and then the risk in the non-smokers, the CER, 164/906 or 0.18, 18%. The relative risk of cancer is calculated by dividing the EER by the CER, 0.53/0.18 = 2.9. In this sample, the relative risk of cancer in smokers is 2.9 times that of non-smokers.

9.8 Odds Ratio (OR)

To calculate the odds ratio we again use a contingency table, but labels are slightly different to that of the one used to calculate risk reflecting the different design of the study. From Table 9.3 we first have to calculate the odds by looking at the number of

Table 9.1 Contingency table for a cohort study

	Diseased	Not diseased	Total
Exposed	A	B	A + B
Not exposed	C	D	C + D
	A + C	B + D	

Table 9.2 Example data for a cohort study

	Diseased	Not diseased	Total
Smokers	288	257	545
Non-smokers	164	742	906
	452	999	

Table 9.3 Contingency table for a case–control study

	Cases	Controls	Total
Exposed	A	B	A + B
Not exposed	C	D	C + D
	A + C	B + D	

Table 9.4 Example data for a case–control study

	Cases (cancer)	Controls (no cancer)	Total
Exercise	100	500	A + B
Sedentary	250	350	C + D
	A + C	B + D	

events and non-events in each group. In the exposed group this would be A/B and the odds in the not exposed group would be C/D. From this we can work out the odds ratio by dividing the odds from each group to give the odds ratio = (A/B)/(C/D).

Table 9.4 uses hypothetical data looking at colon cancer and exercise: the exposure measurements are subjects who exercise compared with a sedentary group. First, we need to calculate the odds in the exercise group (A/B): 100/500 or 0.20. The odds in the sedentary group (C/D): 250/350, 0.71. The ratio of these odds is 0.20/0.71 = 0.29. In this instance the value is below 1 indicating that exercise is protective of colon cancer development as we see a third less cases. We could reverse this by taking the reciprocal of this figure 1/0.29 = 3.5 which tells us that people who are sedentary are 3.5 times more likely to get colon cancer than those who exercise.

Binomial logistic regression may be used to estimate the probability of an event. It is similar to linear regression but accounts for the fact that the dependent variable is binary. The process is quite similar to that already described, except that in this process we can consider other variables at the same time: a logistic or logit function is required, which is the log to the base e of the odds, in order to link the variables and allow a linear relationship to exist.

Some studies do quote *p*-values for relative risk/odds ratios and these can be calculated by a variety of tests such as Fisher's exact test, chi-square test, and Wald test, but it is also very common when looking at papers only to see confidence intervals (CI), usually the 95% CI quoted after the relative risk or odds ratio. The CI gives information on the size and the precision of a study's results; a wider CI means less precise results.

For example, a study looking at the protective effect of exercise on breast cancer reported that women who were inactive in early life, but later became active had an OR of 0.6 (0.3–1.1), whereas lifelong exercisers had an OR of 0.6 (0.4–0.8). In this instance although there is a risk reduction for late-life exercisers, it is not considered significant as the CIs include the value 1 (no difference), whereas the effect of lifelong exercise on the reduction in breast cancer was significant for lifelong exercisers (CI did not include the value 1). The result is shown below in graph form in Fig. 9.2. The odds ratio of 1 is highlighted with a blue line. As can be seen the CI for the first group who were inactive in early life includes the odds ratio value of 1 (passes through the blue line) and so is not significant, the lifelong exercisers' CI is all one side of the 1 line and so is significant.

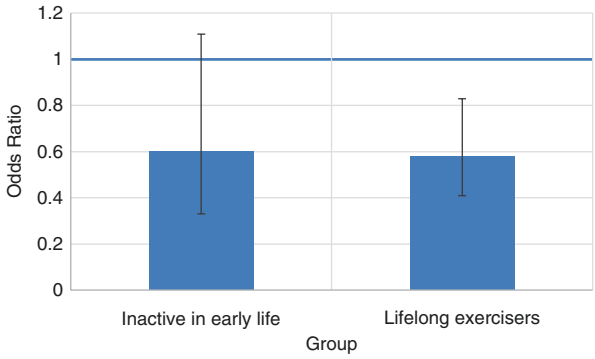


Fig. 9.2 Odds ratio result

Table 9.5 Contingency table for a cohort study

	Bad outcome	Good outcome	Total
Drug	A (117)	B (62)	A + B (179)
Control	C (114)	D (35)	C + D (149)

9.9 Other Terms Used

Another useful figure is the absolute risk reduction (ARR), which is also called the risk difference (RD). It is the absolute difference in outcome rates between the control and treatment groups (CER—EER). It is most useful in deciding whether using a certain treatment reduces the risk by a clinically meaningful amount. Table 9.5 presents a study where one group is given a new drug which is hoped will cure cancer, whereas another group receives standard treatment.

In this hypothetical study, the risk of a bad outcome (CER) in the control group was 76.5% (114 bad outcomes in a total of 149 patients) or 0.765 and in the drug group (EER) was 65.4% (117 bad outcomes in 179 patients) or 0.654. The absolute difference of a bad outcome occurring in the two groups is therefore 76.5% − 65.4% = 11.1%. In other words, taking the drug reduces the 'absolute' risk of a bad outcome by 11.1% compared to the control group.

Linked to ARR is another value called 'numbers needed to treat' (NNT). It is the reciprocal of the ARR and can be thought of as the number of patients that would need to be treated to prevent one additional bad outcome. For the above data the NNT is 1/0.111 = 9. In this instance for every nine subjects who get the new drug approximately one case of cancer would be stopped. An ideal NNT would be 1, where everyone who receives the treatment gets better and therefore the lower the NNT value, the better the treatment being considered.

NNTs are very easy to understand and to calculate and have gained popularity in recent publications. NNT data are especially useful in comparing the results of trials in which the relative effectiveness of the treatments is of interest. For example, the

Table 9.6 Contingency table for a cohort study

	Bad outcome	Good outcome	Total
Drug	A (24)	B (125)	A + B (179)
Control	C (42)	D (137)	C + D (149)

NNT for long-term survival following treatment using radiotherapy and drug A for breast cancer might be 9 whereas for radiotherapy and drug B it is 4. Clearly the use of drug B with radiotherapy is more beneficial as every four treatments given would result in a life being saved whereas with radiotherapy and drug A every nine treatments would prevent a death.

The final figure to be considered is relative risk reduction (RRR) which as with NNT is derived from the ARR. RRR estimates the baseline risk that is removed because of the new treatment and is the ARR/risk in control group (CER). In the above example the risk in the control group was 76.5% so the calculation is 11.1/76.5 = an RRR of 14.4%.

RRR therefore differs from ARRs and this difference might be quite large. Consider Table 9.6 above. The ARR = [(42/149) = 0.281] − [(24/179) = 0.134] = 0.17 or 14.7%, so for every 100 treatments about 15 bad outcomes would be avoided. The RRR = ARR/CER = 14.7/28.1 = 0.52. This means that bad outcomes were reduced by 52% in the treatment group compared with the control group.

Relative risk measures by how much the risk is reduced in the experimental group compared to a control group. If, for example, 50% of the control group died and 25% of the treated group died, the treatment would have a relative risk reduction of 0.5 or 50% and this would probably be significant. However, if 0.5% of the control group dies and 0.25% of the treated group dies, this would probably be insignificant even though the RRR would again be 50%.

9.10 Summary

The interpretation of both relative risk and odds ratios follows the same guidelines. A value of <1 indicates a negative association; >1 indicates a positive association; and 1 indicates no association between the exposed and unexposed groups. As with risk and odds, the relative risk and odds ratio approximate each other when an event is rare. For common events the odds ratio always follows the same direction as the relative risk. It may not be an accurate estimate of the risk; the odds always being overestimated compared to risk.

If the RR/OR is >1, and the CI does not include 1, then events are significantly more likely in the treatment than the control group.

If the RR/OR is <1, and the CI does not include 1, then events are significantly less likely in the treatment than the control group.

Previous tables and information have all been combined to give a summary in Tables 9.7 and 9.8.

Table 9.7 Summary of contingency tables

	Bad outcome	Good outcome	Total	
Exposed	A	B	A + B	EER = (A/(A + B))
Not exposed	C	D	C + D	CER = (C/(C + D))

Table 9.8 Summary of equations used in the chapter

Variable	Equation	
Odds ratio	(EE/EN)/(CE/CN)	(Experimental event/experimental non-event)/(control event/control non-event)
Relative risk	EER/CER	Experimental event rate/control event rate
Absolute risk reduction	CER-EER	Control event rate—experimental event rate
Relative risk reduction	(CER-EER)/CER	Absolute risk reduction/control event rate
Numbers needed to treat	1/(CER-EER)	1/absolute risk reduction

9.11 Conclusion

In this chapter the common research methods for epidemiological studies were considered. Epidemiological studies have played and continue to play a big part in helping healthcare professionals understand the factors affecting the incidence and prevalence of disease. The three main types of epidemiology study design are: cohort, case–control, and cross-sectional. Each has associated strengths and weaknesses in their use. In order to establish causality certain criteria, as defined by Bradford Hill, have to be met: the main criterion being a temporal relationship between exposure and outcome. Studies report their findings as an 'odds ratio' or 'relative risk'. These two measures are often used interchangeably and when the event being measured is rare these two figures do approximate each other. Unlike a lot of other studies, the p-value is sometimes omitted. If the 95% confidence interval for the relative risk or odds ratio is not either side of 1, then you can state that there is a significant disease exposure association.

References

1. NHS. High cholesterol. Crown copyright. https://www.nhs.uk/conditions/high-cholesterol/. Accessed Apr 2019.
2. Dewis R, Gribbin J. Breast cancer: diagnosis and treatment. An assessment of need. NICE clinical guidelines, no. 80-81S. National Collaborating Centre for Cancer; 2009.
3. Madden R, Kosari S, Peterson GM, Bagheri N, Thomas J. Lapatinib plus capecitabine in patients with HER2-positive metastatic breast cancer: a systematic review. Int J Clin Pharmacol Ther. 2018;56(2):72–80.
4. Devlin NJ, Lorgelly PK. QALYs as a measure of value in cancer. J Cancer Policy. 2017;11:19–25.

Further Readings

Bruce N, Pope D, Stanistreet D. Quantitative methods for health research: a practical interactive guide to epidemiology and statistics. 2nd ed. Chichester: Wiley; 2018.

Hulley SB, Cummings SR, Browner WS, et al. Designing clinical research. 4th revised ed. Philadelphia: Lippincott Williams and Wilkins; 2013.

Kirkwood B, Sterne J. Essential medical statistics. 2nd ed. Malden: Wiley; 2003.

Stewart A. Basic statistics and epidemiology: a practical guide. 4th ed. London: CRC Press; 2016.

Sampling Errors, Bias, and Objectivity

10

David M. Flinton

10.1 Introduction

This chapter covers much of what should be considered before you undertake your research: what the population is; how to get a sample; and why sampling is important, and probability. The different types of bias that can exist in study design are covered. Finally, there is a brief section covering power and its importance in research.

10.2 Sampling

One of the key issues of research is how to choose a sample to be studied. Sampling is inherently different in quantitative studies compared to qualitative studies. A general difference between quantitative and qualitative research is that qualitative research utilises an inductive approach, taking specific information and making a broader generalisation. It is important in qualitative research that there is detailed, rich, complex data that are set in context. Qualitative data usually pertain to 'meanings' and as such are mainly in the form of words, themes, or patterns.

Quantitative research is more commonly associated with deductive approaches where a hypothesis is tested using statistical methods. We go from the general statement, the hypothesis to the specific: the observations. Because of this fundamental difference qualitative research tends to be based on purposive sampling where a small sample is selected because of certain characteristics. Quantitative research, which this chapter focuses on, relies on a sample, a sub-set of a population, being large enough to be representative of a population; otherwise, the results will be biased and not represent the population parameter.

D. M. Flinton (✉)
School of Health Sciences, City, University of London, London, UK
e-mail: d.m.flinton@city.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_10

A parameter is a figure that is derived from the whole population. Samples are collected because it is usually not possible to collect data from the whole population, even when it is small. The term 'population' in this instance does not refer to the population as a whole, but rather is the set of individuals or items from which a sample is taken. For example, if a researcher was interested in quality of life after a heart attack in the United Kingdom (UK), the population would be all UK 'heart attack' patients; it is from this population that a sample would be taken. If we provide and report figures from a sample, they are referred to as statistics.

Parameters are very rarely known for a number of reasons. The main issue mentioned above is the inability to include all of a population due to the huge task of tracking down and collecting data on every single person in it. Another concern is the ethical issue of using a population when we know a sample would give an efficient estimate. There are also issues relating to the identification of subjects that could affect the ability to collect a whole population. Let us consider men with prostate cancer. How can we include those who die from undiagnosed prostate cancer? Finally, the transient nature of the population of interest can give problems. If we look at breast cancer, in England and Wales there are roughly 110 new cases every day and 30 deaths, so the breast cancer population changes on a day-to-day basis. Because of all these issues relating to collecting data from a population we take a smaller number of cases and assume that they are representative of the population: we sample.

The first stage of effective sampling is to define a population precisely and then construct a sampling frame. The ideal sampling frame is a list including all the items/people that you are trying to sample. So, for a study investigating radiography practitioners working in the UK, a comprehensive list of all UK based practitioners would be ideal. In practice the ideal hardly ever exists. In this case the closest we could probably get is the UK Health and Care Professions Council (HCPC) register. The list, if we could get permission to access it, would not be perfect as it would include practitioners taking a sabbatical, those who have recently retired or are maintaining their registration but working abroad.

The relationship between population, sampling frame, and sample can be seen in Fig. 10.1. Note in the diagram that the sampling frame is incomplete and does not cover the population so reflecting the discussion above. The figure on the left shows no sampling frame error as it takes evenly from the population, whereas we have a sampling frame bias on the right as we are not evenly covering the population. A sample is chosen to reflect a population from a sampling frame; from the sample we will get responses which we can analyse.

10.2.1 Sample Bias

If the data collected are not representative of a population then the estimates will not be accurate, and we say that the sample is biased. A selected sample is in some way systematically different to its population. The first and most obvious reason sampling bias may exist is if a sample size is too small and will therefore not produce a reliable

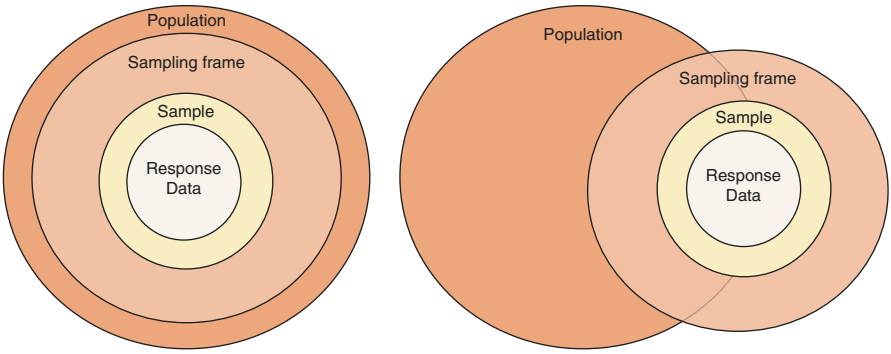


Fig. 10.1 Relationship between data and population

estimate of the population. Taking this to its extreme, my wife and I are both radiographers. If I used this sample, I might conclude that the gender ratio of radiographers is 50% male and 50% female, and 50% are therapeutic radiographers and 50% diagnostic radiographers. This obviously is not representative of the true population. This issue is explained further in the section on power calculations in Sect. 10.7.

Another factor to consider is the sampling method. There are various methods of sampling, some of which are better than others at obtaining a representative sample of a population and reducing sampling bias. In order to reduce bias a method is needed where every subject has an equal chance of inclusion/exclusion in a study. A sample is biased if certain members are under-represented or over-represented compared with others in the population; this bias can occur in the selection of both a study and control group. This may arise for several reasons some of which are detailed below.

10.2.2 Types of Sampling Bias

- Self-selection bias

  Imagine that a researcher was interested in what radiographers thought about the use of technology in the clinical teaching of students. A poster could be sent to all NHS radiography departments asking for volunteers for the study. Radiographers interested in technology or clinical teaching would probably be more likely to respond to this request, so they are effectively self-selecting themselves for the study and they are likely to differ in important ways such as age and gender from the population the experimenter wishes to draw conclusions about. This type of bias can lead to a polarisation of responses with extreme perspectives; those strongly supporting and those strongly against the issue being researched being more likely to respond than those who are more neutral.
- Selection from a specific real area.

  The above example also includes this type of bias as it only included radiographers from the NHS. Radiographers working in private hospitals might have

slightly different issues and considerations and therefore opinions compared with NHS departments. Another example of this type of bias might be if questionnaires were handed out at the front entrance of a hospital to those entering the hospital. This might give an overrepresentation of healthier individuals as the more infirm might use transport and thus might have a separate entrance, as might some separate clinics/wards. In this scenario these populations would be under-represented.

- Healthy user bias

  This form of bias occurs when a study population is likely healthier than the general population. The healthy user effect has been cited as a likely source of bias in observational studies looking at the use of hormone replacement therapy (HRT) [1]. It was postulated that women who took HRT were systematically healthier than those who did not use HRT; this implied that the benefits observed in the studies might not be due wholly to HRT.
- Berkson's bias

  This form of bias was first recognised by Dr. Berkson when looking at case–control studies where both the cases and controls were sampled from a hospital rather than from the population at large. A classic example of this was first published by Roberts et al. in 1978 [2] who, when looking at hospitalised cases, found a large positive association between the presence of both respiratory disease and locomotor disease. The association between respiratory disease and locomotor disease came about because the hospitalisation rate of patients with both these conditions was a lot higher than that for people who had only one of the conditions and even lower for patients without either disease. The observed association was therefore false; the finding would have been very different if the sample had been taken from non-hospitalised individuals.

10.3 Sampling Methods

Various sampling methods are at our disposal, some are better than others at removing sampling bias.

10.3.1 Random Sampling

The idea behind random sampling is to remove sampling bias. There are a number of ways of performing a randomisation process, some practicable, some not so practicable. Some types of sampling are detailed below, but this list is not exhaustive.

10.3.1.1 Simple Random Sampling

If we were interested in investigating patients' perception of the care received during their visit to casualty, a list of all patients who had attended casualty could be obtained from the picture archiving and communication (PAC) system during the timeframe in question. A random sample could be obtained by each name being written on a piece of paper, placed in a drum, and then randomly drawn.

58	06	96	03	51	50	09	96	67	74	08	97	06	71	22
89	08	40	54	15	03	69	94	98	91	94	21	91	29	06
91	88	08	83	54	54	13	04	94	67	70	01	31	25	18
38	66	48	14	30	31	03	96	65	30	53	43	55	20	97
24	35	69	21	18	55	71	78	54	94	80	58	47	46	48
45	60	39	34	12	91	57	51	73	08	01	18	58	92	87
74	04	28	68	68	60	67	37	34	48	22	86	73	51	53
06	57	05	72	96	97	27	78	55	27	57	77	50	08	68
24	74	76	86	46	82	64	38	07	30	42	09	48	15	05
88	81	89	45	85	68	79	50	38	10	80	74	93	23	39
55	96	15	31	08	60	04	04	98	24	21	81	45	12	83
50	42	28	55	02	16	49	48	46	14	72	41	83	08	56
50	72	79	30	45	88	47	51	44	73	31	99	76	80	18
58	84	67	26	03	86	96	77	42	59	04	01	58	99	86
48	07	34	94	44	45	14	79	40	72	48	14	01	05	92
48	58	32	58	97	87	76	42	29	20	11	83	94	89	92
48	94	21	60	13	93	48	44	82	39	74	85	68	11	13
14	78	45	22	08	11	77	20	35	75	41	43	25	31	44
35	67	95	35	86	02	03	29	35	42	87	53	10	18	46
95	44	89	14	56	52	25	47	96	79	52	04	59	73	04

Fig. 10.2 Table of random numbers

A better approach would be to use a table of random numbers. Each patient would be given a number and would be selected by matching numbers generated from the table. The table is random and so it does not matter where you start or in which direction you move. Assuming there were 820 patients available to a study, we could start at the top left corner of Fig. 10.2 and reading down move from right to left, so the first number generated would be 589, the next 891, then 008, 688 … and so on. Patients 589, 008, and 688 would then be approached to join the study. There was no patient 891 as there were only 820 patients in total so any number above 820 would be ignored. This process, although still time-consuming, is better than the first option proposed, but is still only really suitable for studies of relatively small size.

An important issue to consider is when we want to allocate subjects to groups: for example, if we want to produce two subject groups, one to receive the intervention and the other the placebo or alternative intervention. Again, we would like to remove selection bias and have a system of allocating patients to the groups which is random, and again using the random numbers table allows us to do this.

Again, let us make an assumption, this time that we want 24 subjects randomly allocated into two groups. We arbitrarily pick a starting point on the random numbers table and we have decided to move down and then along each block from left to right. The data selected are shown surrounded by a box in Fig. 10.2.

The numbers selected are then used to code which group they will belong to; odds go in the intervention group (I) and evens into the placebo group (P). Randomisation using the data would mean the allocation of subjects as shown in Fig. 10.3.

60	04	04	96	24	16	49	48	46	14	88	47	51
P	P	X	P	P	P	I	P	P	P	P	I	I

44	73	86	96	77	42	59	45	14	79	40	72
P	I	P	P	I	P	I	I	P	I	P	P

Fig. 10.3 Random numbers and sample sequences

1 I, I, P, P	2 I, P, I, P	3 I, P, P, I
4 P, P, I, I	5 P, I, P, I	6 P, I, I, P

Fig. 10.4 Block randomisation

5 P, I, P, I	6 P, I, I, P	6 P, I, I, P	3 I, P, P, I	5 P, I, P, I	1 I, I, P, P

Fig. 10.5 Sample sequence

The method is simple and random. It can give rise to uneven sizes of the groups, particularly in small trials, which can be a problem as if you calculated the sample size for the study it would have assumed equal group size. In the example above twice as many subjects were allocated to the placebo group compared with the intervention group.

10.3.1.2 Block Randomisation Sampling

Block randomisation overcomes the problem of the different number of subjects in different arms of a trial by keeping the subjects balanced throughout the study. The blocks can be of any size; they are usually a multiple of the number of treatments. If we use blocks with a size of four, we get six possible ways of assigning the two possible (placebo or intervention) treatments keeping the balance between treatments equal as shown in Fig. 10.4.

The allocation sequence is then decided by using the random numbers table to decide the sequence of the blocks. If we read horizontally on the table starting at the top left position, we get the figures 5, 8, 0, 6, 9, 6, 0, 3, 5, 1 so the selection of the 24 patients is as shown in Fig. 10.5. Note how we now have 12 subjects in each arm of the trial. This method can be further refined by varying the block length.

10.3.1.3 Stratified Sampling

Stratified sampling is a further development of block randomisation. It is used when it is important to achieve a balance between important characteristics in the subjects. A separate block randomisation is carried out for the important characteristic. If we were comparing alternative treatments for reducing stress in radiography

practitioners, for example, it might be important to stratify by gender. Each gender would have its own block randomisation, so each gender would be equally distributed between the two different treatments.

10.3.2 Purposeful Sampling

Snowball sampling is a technique for developing a research sample where existing study subjects suggest further recruits to take part in a study from among their acquaintances. It is of particular use when a researcher is studying a hidden population, a population with no sampling frame.

Judgement sampling is where a researcher actively selects subjects who are believed will be the most productive sample to answer a research question. This can be done via a framework looking at the possible variables that might influence a subject's contribution to a study.

Convenience sampling is where you sample subjects easiest to reach. It is generally considered as being the poorest way of getting a study sample, having the lowest credibility since the subjects are selected arbitrarily.

10.4 Bias and Error

One type of bias associated with sampling was discussed above. There are two other main forms: response and information bias. As we saw with selection bias a number of different variants exist and the same exists with these other two forms of bias.

Bias can be defined as an error in sampling or testing that will systematically affect the outcome of a study. If present, it infers that the findings of a study are less meaningful.

10.4.1 Response Bias

This is a type of bias that can affect the results of a study. It arises if there is a tendency for participants to respond inaccurately or falsely to questions.

10.4.2 Acquiescence Bias

Acquiescence bias occurs when respondents have a tendency to agree with all the questions in a measure. This may be made worse due to bad question design that encourages respondents to reply in a way they think the questioner wants them to answer, rather than what they actually think. An example of a badly phrased question is shown below. It leads respondents to an affirmative answer even if they disagree. Put differently the way the question is phrased makes you want to agree.

Please indicate the extent to which you agree or disagree with the following statement.

I was extremely satisfied with my radiotherapy treatment.

Agree	*Somewhat agree*	*Undecided*	*Somewhat disagree*	*Disagree*

Two ways acquiescence bias may be reduced is to 'reverse' some of the items on a questionnaire and to carefully consider question design. See the question below which asks the same question as above.

How satisfied/dissatisfied were you with your radiotherapy treatment?

Very satisfied	*Somewhat satisfied*	*Undecided*	*Somewhat dissatisfied*	*Very dissatisfied*

10.4.3 Demand Characteristics

Demand characteristics refer to a type of response bias where participants alter their response or behaviour because they become or think they become aware of what a researcher is investigating. They are trying to please the researcher (get it right) by conforming to what they see is the purpose of the experiment. Subjects might pick up on subtle cues such as body language or phrasing of the questions, which might be enough for the subjects to work out what the expectancy of a researcher is.

10.4.4 Social Desirability Bias

This form of bias occurs when a respondent provides an answer that they consider is more sociably acceptable, trying to conform to the social norm rather than revealing their own true opinions. The reason this is thought to happen is that respondents feel uncomfortable revealing their true answers. Latkin et al. 2016 [3] indicate that there are two dimensions to this: firstly altering the response to influence how the respondent is perceived by others; secondly self-deception, undertaking an action to enhance their own self-perception. Social desirability bias is more common when looking at sensitive or controversial subject matters. For example, consider what a potential respondent might think when confronted with a question that states: "indicate your level of racism on the scale below".

A number of ways of reducing this bias exist. Not revealing the purpose of a study and allowing anonymous responses help. Indirect rather than direct questions are also thought to reduce this form of bias. The use of a social desirability scale as part of the questionnaire can determine if respondents have responded with a high social desirability bias and therefore can be excluded from the study.

10.4.5 Extreme Responding

This occurs due to a respondent's tendency to pick the most extreme options from a ratings scale. It is most commonly observed in self-reporting questionnaires.

It is suggested that culture affects the amount of extreme responding and can be more pronounced in certain cultures. Culpepper and Zimmerman [4] found that Hispanic Americans evidence a lot of extreme responding. Van Herk et al. [5] noted higher levels in Spanish and Italian respondents compared to British, German, and French samples.

10.4.6 Question Order Bias

This occurs when a respondent may answer differently to questions based on the order in which questions appear. This is important when considering the design of a questionnaire or interview. This can occur as the respondents are unconsciously trying to apply meaning based on the order, i.e., a list of possibly responses might be interpreted as best on top, worse on the bottom. Another example might be if you draw attention to certain aspects in a preceding question as shown in the example below.

1. Which of the following features of your mobile X-ray unit would you most like to see improved? Battery life, physical size, limited mAs output.
2. What do you find most frustrating about your mobile unit?

10.5 Information Bias

10.5.1 Recall Bias

A subject's recall is thought to be dependent on their disease status, the exposure, even if irrelevant, being remembered better by cases than controls thus leading to exposure being under-reported by the controls. This can be exacerbated by certain issues such as a patient's preconceptions about the link between exposure and disease. These in turn can sometimes be influenced by the media, which may emphasise links between certain exposures to certain factors and the related health outcome. This problem was identified with case–control studies in the epidemiology section in Chap. 9 (see Sect. 9.4.2). It relies on subjects remembering an event in their past.

10.5.2 Observer Bias

The bias here is with the researcher(s). This occurs when researchers know the aim(s) of a study and allow this knowledge to influence their observations. In research the effect of observer bias can be removed by carrying out a specific type of study: a double-blind study. This is a study where neither the researchers nor the participants know which arm of a trial a subject is in. Note that there is also something called observer effect which is different to observer bias. The observer effect occurs when subjects change their behaviour because they know they are being watched.

This list is not exhaustive and other types of bias occur. Another area where bias exists is in scientific publications that can be used to support or refute work. Examples of bias that occur here are publication bias and reporting bias.

- Publication bias: the predisposition of journals to accept for publication studies that have a positive finding. Again, this can be exacerbated by authors, who have a tendency to only submit articles with a positive outcome.
- Reporting bias: the tendency of authors in studies that have multiple outcomes to only report the outcomes that are significant and ignore the non-significant outcomes.

10.6 Error

Error is generally considered to refer to a difference between an observed value and true value. There are two different types: random error and systematic error.

Random errors tend to mainly affect the variability around the mean; if a sample size is small, it may impact on the mean. Figure 10.6 represents two data sets: the blue line represents accurate data and the green line a repeat of the data collection with more random errors present. As can be seen in Fig. 10.6 random errors affect the precision of results as they are more spread out around the mean, i.e., have a larger standard deviation, but the mean is unaffected.

Systematic errors affect the mean value being reported, moving it higher or lower and their effect tends to be independent of the sample size. Systematic errors are often referred to as bias as they affect the accuracy of the results which are then no longer a true reflection of the population. Figure 10.7 shows the original data with a blue line. The green line represents the data recollected with systematic errors introduced.

Consider the example of an ionisation chamber being used to measure radiation dose. If the chamber was incorrectly positioned for one reading, we would have a random error and as the sample size is increased this one value would have less effect on the mean reading. If, however, the wrong chamber factor was given to a

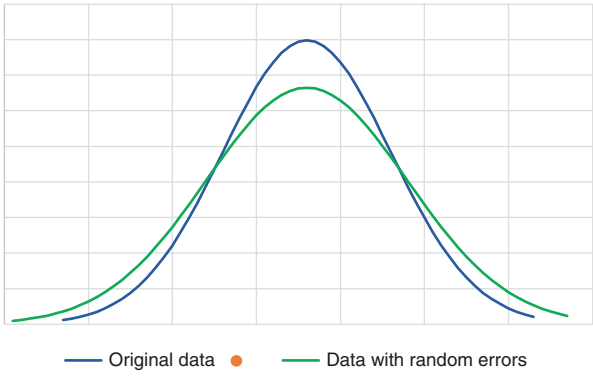


Fig. 10.6 Distribution change with random errors

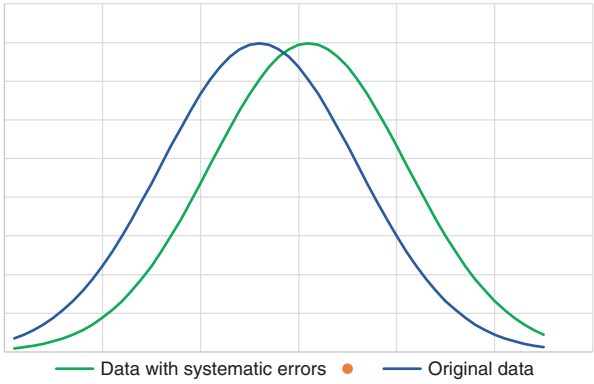


Fig. 10.7 Distribution change with systematic errors

researcher, this would affect all the readings, moving them in one direction (higher or lower), and the number of readings taken would not affect the results. With a systematic error all readings are affected to the same degree; the bias will not be apparent when looking at the data, but we might spot the random error as it might be very different to all the other readings. This is why it is important to look at the data before you start your analysis. You can simply look at the figures to see if one stands out as being different or you could do some simple plots to check the data to see if there are any outliers. This check should always be done as it highlights errors that occur when you input the data. Another way to reduce error coming in at this stage of the process is to carry out double entry of the data.

There are various types of bias that can be introduced into a study, some of which have already been mentioned. The types of bias a study is open to will depend on the type of study and how well it was performed.

10.7 Power Calculations

A frequently encountered problem in quantitative research is deciding how big a sample needs to be to find a 'reliable' result. The smaller a sample is, the less likely it is that if you repeated the sampling you would find you had the same result and the less reliable it is. On the other hand you do not want to waste time and resources collecting unnecessary data when a smaller sample would give a sufficiently reliable result. In reality this only applies to instances when it is feasible to repeat the exercise (such as handing out a questionnaire), but it is possible to extend the idea to non-repeatable samples as well in an imaginary way. So how do you measure reliability? Information on reliability and validity is given in Sect. 10.8 below.

When a decision is made to reject a null hypothesis it can be made on the basis that the *p*-value falls in a particular range of values (fixed level testing). One way to measure reliability could be to look at how often you would make the correct decision for a given significance level. This is what power calculations try to do.

There are four possible outcomes to a fixed level hypothesis test.

1. You accept the null hypothesis incorrectly.
2. You accept the null hypothesis correctly.
3. You reject the null hypothesis incorrectly.
4. You reject the null hypothesis correctly.

Power calculations work out the chance of the last possibility occurring. Notice that this does not consider all of the times that you might be correct in accepting the null hypothesis (case no. 2). Power calculation does not consider all of the ways that you could make the correct choice, only one of them (see also Chap. 15).

There are ways in which power calculations are useful.

- They can be done before a study starts in order to predict how big a sample needs to be to give a result of a required power (sampling requires time so this is a way of figuring out how little you have to do in order to get a 'good' result). This is called an a priori calculation.
- They can be done after a sample has been performed to see the power the study had (maybe a much smaller number of questionnaires were returned than you wanted and you want to find out how worthwhile it is to use the limited number that you have). This is called a post hoc calculation.

Trying to collect new data when an original sample has proved insufficient can be problematic; it involves more time and there are potential problems with dependency (i.e., one answer affecting another, for example, if subjects in the group you are sampling have talked about the research before you sample the second time).

The best way to use power calculations is as an a priori tool to try to predict how effective certain sample sizes will be. When collecting data using questions with human subjects there will always be problems about compliance. When conducting studies over a period of time, subjects may also drop out (failing to complete a study). This can to some extent be corrected by trying to predict the rates at which data might be lost and combining this with information from power calculations to find out how big a sample should be to leave data that give a reliable result. For example, a patient satisfaction survey might generate one return for every two patients in which case the sample needs to be twice the size of the sample calculated by a power calculation in order to achieve the required reliability.

It is important to realise that if you have obtained ethical approval for a study, then there are ethical issues surrounding the failure to collect enough data as this may involve waste of resources and misuse of subjects (especially if a study is patient based).

The 'power' of a test is often given as a percentage. The higher the percentage means the better the chance that your findings will be reliable. For example, the power of a test may be calculated as 80%. This means that if the sampling was repeated many times, then for 80% of those occasions the null hypothesis would be correctly rejected. By now you will hopefully have spotted that this is a probability of making the correct choice (for 80%, $p = 0.8$—it is just two different ways of

writing the same thing). A very important point here is that because it is a probability it cannot tell you about the exact occasions when you make the right or wrong choice, only how likely it is.

In addition to a decision about what level of reliability you will accept you also need to have an idea of what constitutes a 'significant' change or difference for your study. For example, if you are trying to find out whether recovery rates following treatment are improving, at what point does a change of a significant size occur? Finally, you will need to have an idea of the amount of spread in your data. If it is an a priori calculation you will need some estimate (maybe from a pilot study). This needs to be expressed in terms of standard deviation in order to perform the calculation. While it is possible to manually perform power calculations, the theory is heavily reliant on mathematics. In practice these calculations are performed using statistics software packages. You can do the calculations yourself or you can ask a statistician to do them for you.

10.8 Reliability and Validity

Two other terms often used in statistics are reliability and validity. Reliability seeks to describe the consistency or repeatability of a measurement. Validity refers to the strength of the conclusions drawn. In order for a study to be 'good' it must be both valid and reliable: neither by themselves is enough.

To illustrate this many authors refer to a metaphor of darts, as described and shown in Fig. 10.8. Four people have each thrown three darts at the board. Each used

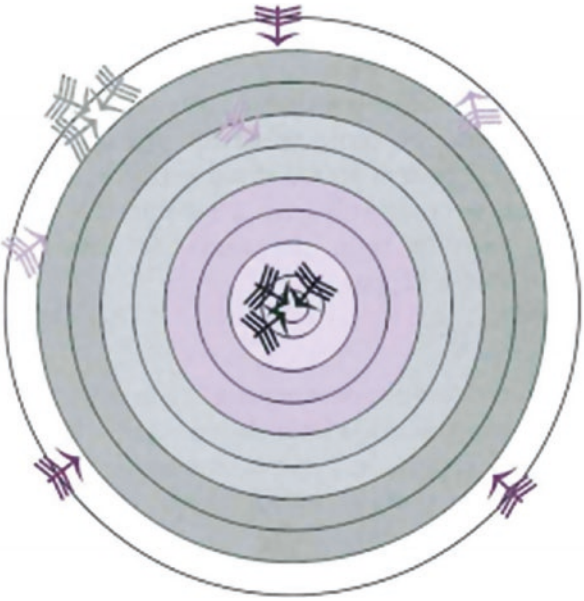


Fig. 10.8 Reliability and validity metaphor

a different colour dart. We are interested in how far on average they are from the bull. Two people (grey and black darts) managed to group the darts closely together, i.e., they were both consistent and reliable. Two people could also be described as being accurate (black and purple darts); the average of both colours is being the bull which equates to validity. Only one person (black darts) was consistent, accurate, reliable, and valid. This is what we should be aiming for in our research. The remaining person (mauve darts) was neither accurate nor consistent, so lacks both reliability and validity. But how do we know if we have a reliable and valid study?

10.8.1 Reliability

There are four common types of reliability: the inter-observer reliability, test–retest reliability, split halves reliability, and parallel forms reliability. Each is briefly described below. At the end of each description there is a test quoted. It is a common report of the type of reliability described.

- Inter-observer reliability

  In this instance a questionnaire is tested by a number of people or judges before being given to the test population. For example, if we wanted to know if a form was good at distinguishing between good and bad radiographs we could ask three experienced practitioners to each use the tool (questionnaire) to examine a number of different radiographs we have given them. As they will all be reporting on the same set of radiographs they should get the same or similar scores for each radiograph. If the raters do not agree with each other, all is not lost; we could retest the tool after training the testers to see if we get a better score. If we do, we may still have a reliable tool, but we would have to train the users to ensure this. Test: Cohen's kappa.
- Test–retest

  For this the test or questionnaire is administered to a sample and then re-administered after a time gap. The time gap is very important. Too short a time and the subjects remember what they said and try to emulate their first response, too long and there may have been a construct change, (things may have happened which means their response to the same question would be different). Test: Bland–Altman plot/test.
- Split halves test

  The test is only given once and then the test is divided into equivalent halves; a Pearson's correlation is then calculated between the scores from each half of the test. The closer the scores are between the two halves then the better the internal consistency. Test: Correlation.
- Parallel forms reliability

  This is used to assess the consistency of the results of two tests that were constructed in the same way. A large set of questions, which measure the same

construct, first needs to be produced. A major problem with this approach is that a lot of items that reflect the same construct have to be generated. The questions are then randomly divided into two sets and both instruments administered concurrently to the same sample. Test: Correlation.

10.8.2 Validity

As with reliability there are different types of validity: internal validity and external validity.

Internal validity is concerned with causality in the sample group studied. Internal validity asks the basic question, 'Did the experiment make a difference?' Another way of saying this: 'Was the experiment carried out in such a way that we are confident that the independent variable altered the dependent variable?' i.e., how confident are we about the cause and effect? Establishing internal validity can be threatened by a number of issues, such as confounding variables, outside influences/events, regression towards the mean and attrition from the study.

External validity is concerned with how results can be generalised to a population. The main threats to external validity are: the sample itself; sampling method; and time. For example, if a study was performed looking at radiography practitioners' perception of continuing professional development and was conducted predominantly using newly qualified radiographers, or the data were collected during a year when the HCPC audit to monitor radiographer registrants' compliance, these conditions would both affect external validity. In the first instance ensuring a random sample from all radiographers would help reduce the threat; in the latter instance a replication of the study would help eliminate the threat and demonstrate the generalisability of the results.

10.8.2.1 Construct Validity

Construct validity relates to a survey instruments, questionnaires, or tests and gauges how well we might expect the selected tool to perform at measuring what we think it is measuring. Do not get confused here. It does not refer to how well a questionnaire is constructed. A construct is the attribute, proficiency, ability, or skill that is being measured. Three variants of construct validity are briefly described below.

- Convergent validity

  This relates to the degree to which the test is similar to (converges on) other tests that it theoretically should be similar to. For instance, to show the convergent validity of a questionnaire that purports to measure fatigue in radiotherapy patients, we could compare the scores to a second fatigue test; high correlations would be evidence of convergent validity.

- Discriminant validity

  This is almost the opposite of the above. It is validity obtained when we measure two constructs that are thought to be dissimilar and the measures can discriminate between them. For instance, to show the discriminant validity of a spatial ability test, we might compare the scores with a test that looks at intelligence. Low correlations would be evidence of discriminant validity.

10.9 Conclusion

In this chapter the need for sampling was considered, as were the various forms of bias, random and systematic errors, and the concepts of reliability and validity. Statistics are created when describing/investigating samples. Sampling relies on being able to define a population precisely and to use an appropriate technique to avoid sampling error and to obtain an unbiased sample. There are many types of sampling, some are best in certain circumstances, but the best overall group of methods use probability sampling, which utilises some form of random selection. A random method of sampling gives each person an equal chance of being included in a study. Bias, a systematic error, and errors may be introduced into a study if it is designed incorrectly. Different types of bias occur depending on the study type. Deciding on the sample size of the study is very important: too small and it may not be representative of the population; too large and it is wasteful. It is also possible to calculate the power of a study after it has been conducted. In order for a study to be deemed 'good', the results must be both valid and reliable.

References

1. Gleason CE, Dowling NM, Friedman E, Wharton W, Asthana S. Using predictors of hormone therapy use to model the healthy user bias: how does healthy user status influence cognitive effects of hormone therapy? Menopause. 2012;19(5):524–33.
2. Roberts RS, Spitzer WO, Delmore T, Sackett DL. An empirical demonstration of Berkson's bias. J Chronic Dis. 1978;31:119–28.
3. Latkin CA, Mai VT, Ha TV, Sripaipan T, Zelaya C, Minh NL, Morales G, Go VF. Socially desirability response bias and other factors that may influence self-reports of substance use and HIV risk behaviors: a qualitative study of drug users in Vietnam. AIDS Educ Prev. 2016;28(5):417–25.
4. Culpepper RA, Zimmerman RA. Culture-based extreme response bias in surveys employing variable response items: an investigation of response tendency among Hispanic-Americans. J Int Bus Res Arden. 2006;5(2):75–83.
5. Van Herk H, Poortinga Y, Verhallen TMM. Response styles in rating scales. Evidence of method bias in data from six EU countries. J Cross-Cult Psychol. 2004;35(3):251–62.

Further Readings

Altman DG. Practical statistics for medical research. London: Chapman and Hall; 1991.
Bland M. An introduction to medical statistics. 4th ed. Oxford: Oxford University Press; 2015.
Petrie A, Sabin S. Medical statistics at a glance. 3rd ed. Oxford: Wiley Blackwell; 2009.

Web Resources

There are a number of web pages that deal with sampling and probability, including some with Java applications. Other useful resource on the web are random number generators.

Research Outcome Measures

11

Fiona Mellor and Karen Knapp

Chapter Points

- Outcome measures need to be selected at the research planning stage. Selecting an outcome measure after data have been collected increases the risk of bias and type I/type II errors.
- Primary and secondary outcome measures are used to report research findings. A primary outcome measure answers the most important question and, in quantitative pilot studies, may inform the power calculation that gives rise to a sample size.
- Outcome measures are used in all methodologies (quantitative, qualitative, and mixed methods). Examples of outcome measures include imaging tests (e.g. radiographs to assess rheumatoid arthritis), physiological processes (e.g. blood pressure), performance based (e.g. sit to stand tests), and questionnaires which may be administered by a researcher or self-completed by a participant.
- Some outcome measures have data from healthy populations allowing for comparison; others allow for baseline assessments.
- Outcome measures are assessed with a variety of instruments. It is important to select an established outcome measure instrument to facilitate consistency in reporting between/across studies, and to aid in the comparisons of findings and systematic reviews/meta-analyses.

F. Mellor (✉)
AECC University College, Bournemouth, UK
e-mail: fmellor@aecc.ac.uk

K. Knapp
South Cloisters, University of Exeter, Exeter, UK
e-mail: k.m.knapp@exeter.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_11

11.1 Introduction

Within research the acronym PICO (population, intervention, comparator, and outcome) lends weight to the importance of choosing an outcome measure at the initial stage of designing research. The acronym refers mainly to quantitative clinical research but in all research selecting the correct outcome is an important considerations in a design stage of research. Selecting outcome measures at a design stage is crucial in ensuring the correct order, quality and quantity of data collected. An outcome measure may also be known as a 'construct' or 'domain' [1] and they provide a common language for reporting, allowing comparisons to be made with previous research. Outcome measures are used for many reasons including the following.

- Discriminating between patients with differing disease severity at any one point in time
- Predicting patient outcome
- Measuring patient experience
- Evaluating change following an intervention
- Comparing results to normative populations
- Appraising patient safety
- Determining effectiveness of care and clinical outcomes

Different methodologies (quantitative, qualitative, and mixed methods) require different outcome measures and knowing the focus of the research is the first step in selecting an appropriate measure. Noting that 'not everything that can be counted counts, and not everything that counts can be counted' [2] can help target the selection of an outcome measure. For example, investigating 'personal stress'[1] can be undertaken with a quantitative or qualitative approach but because there is no direct way of measuring stress a quasi-measurement/outcome measure must be selected. Within a quantitative paradigm a suitable proxy outcome measure may be blood pressure or cortisol levels, which are widely associated to rise during periods of stress [3]. A qualitative paradigm meanwhile may utilise patient-reported outcome measures (PROMs) such as interviews, focus groups, symptom status, physical function, mental health, and well-being [4]. It is beyond the scope of this chapter to discuss levels of data (nominal, ordinal, interval, and ratio, also known as numerical or categorical) but it is important to understand the classification, or level, of data to select an appropriate outcome measure. In the example of stress, the research methodology, question, aims, objectives, and research design will determine whether blood pressure and cortisol or social function and well-being, are the most appropriate outcome measures.

[1] Defined as a state of mental or emotional strain or tension resulting from adverse or demanding circumstances.

The judicious selection of outcome measures at a research design stage reduces the risk of searching for an outcome, any outcome, in the data (a type 1 error) and also provides the basis for an estimation of the sample size necessary for an adequately powered future study [5, 6]. Thus primary and secondary outcome measures should align directly with a study's aims and objectives; a primary outcome measure is often incorporated in the hypothesis.

Researchers are often tempted to develop their own outcome measures pertinent to their research question; however this approach limits the usefulness of comparing the findings to other similar studies. Outcome measures are commonly selected for inclusion in systematic reviews and meta-analyses, and existing guidelines (COSMIN, consensus-based standards for the selection of health measurement instruments) [7] exist specifically for the selection of patient-reported outcome measures (PROMs) in systematic reviews (see PROMs in Sect. 11.3 below). Researchers are therefore cautioned against developing their own outcome measure, particularly as creating a new outcome measure requires extensive evaluation to determine validity, reliability, consistency, and ability to measure the change [8] in both a statistically and clinically meaningful way. Developing new outcome measures is a branch of research in itself and beyond the scope of this chapter.

11.2 Classifications of Outcome Measures

Outcome measures may be concerned with just one aspect (uni-dimensional) such as pain, or they may be multi-dimensional and look at many facets, such as physical, emotional, and social well-being. They may relate to one condition only (disease-specific) or be more generic and relate to overall health and well-being.

Various agencies have developed taxonomies for outcome measures; many of which overlap and complement each other. Hundreds of outcome measures exist for research, patient care, and service evaluation, and these can help focus a search for a suitable outcome measure. Outcome measures for research may focus on a range of scenarios which are not exclusive to research but are pertinent when considering the primary and secondary aims and objectives. The COMET (Core Outcome Measures in Effectiveness Trials) initiative provides one such taxonomy for outcome measures which states are 'intended for the classification of what, rather than how, outcomes are measured' [9].

Wilson and Cleary [10] categorised outcome measures into five types.

1. Biological/physiological variables (e.g., ranges of motion, radiographic changes, pulse rate)
2. Symptoms status (e.g., pain)
3. Functional status (e.g., return to work)
4. General health perceptions (e.g., various aspects of global health)
5. Quality of life (e.g., general well-being, patient satisfaction)

More recently, NHS England and the Better Care Taskforce [11] developed classifications for outcome measures for service evaluation and monitoring which are also pertinent to research. These include:

1. Developing a baseline from which to measure changes.
2. Mapping and demonstrating the needs of the population being studied.
3. Monitoring clinical practice.
4. Health and disease area.
5. Target population.
6. Methodology/methods.
7. Stakeholders.
8. Study type (e.g., longitudinal/cross-sectional design).
9. Cost effectiveness and financial analysis.
10. Modelling impact.
11. Measuring quality/experience and cost.

Similarly the University of Oxford categorised outcome measures into seven domains with examples and explanations of each category [12]. These focus on PROMs and include the following.

1. Disease-specific (e.g., asthma)
2. Population specific (e.g., child health)
3. Dimension specific (e.g., Beck Depression Inventory)
4. Generic (e.g., SF-36)
5. Individualised (e.g., patient generated index)
6. Summary items (e.g., general household survey questions about long standing illness)
7. Utility measures (e.g., EQ-5D)

In 2011 the British Dietetics Association (BDA) developed a model for dietetic outcomes after recognising that many current therapy outcome measures were not generally applicable, amenable, or transferable to the work of dietitians and where the emphasis needed to be more focused on nutrition [13]. There were six domains in their classification.

1. Symptom changes
2. Physical (e.g., anthropometry/body)
3. Biochemical
4. Psychological
5. Behaviour changes
6. Patient focused

The 2011 model was superseded by a model for dietetic practice which informs education and practice and demonstrates how outcome measures can also be profession specific [14]. There are however no profession specific outcome measures for radiography due to the vast and diverse nature of this profession.

There is clearly overlap within these categories. For instance, most outcome measures can be used as a baseline (which is a domain within the NHS Better Care Taskforce taxonomy). The general health perceptions categorisation from Wilson and Cleary [10] matches with the generic criteria from the Oxford PROM group [12] and the psychological classification [13] from the British Dietetics Association (health and disease) are defined categories in both the NHS Better Care Taskforce and Oxford PROM group classifications. Some disease sites have very well-developed outcome measures, for instance within rheumatology the OMERACT initiative (Outcome Measures in Rheumatology) developed a consensus of a core set of outcome measures for rheumatology drug trials. This has now developed into an international community of health professionals with an annual conference and large database of outcome measures [15].

Developing a taxonomy for diagnostic and therapy radiography would be difficult given the breadth and depth of the profession and research across many arenas. As noted by the OMERACT group [16] the seemingly simple questions of 'what' and 'how' to measure belie a complex structure which is difficult to untangle. Thus a radiography researcher needs a thorough knowledge of the literature and research within their area of study to select appropriate outcome measures. The primary outcome measure should align closely with the stated aim of the study, and secondary outcome measures should align to the objectives.

11.3 Patient Reported Outcome Measures (PROMs)

Patient-reported outcome measures (PROMs) are essentially any outcome directly reported by a patient. They are generally subjective in nature and range from simple visual analogue scales (VAS), used to report pain intensity on a scale of 1–10, to more in-depth questionnaires completed over a lifetime of a longitudinal study. They are typically used to collect patients' perceptions and views about their health, health status, quality of life, and care.

Health research used to be dominated by outcome measures selected by researchers without acknowledgement of their bias (a surgeon may be more likely to select a surgical outcome measure such as blood loss rather than an outcome measure of importance to a patient such as quality of life). However the inclusion and acceptance of patient and public involvement (PPI) in research, along with the phrase 'nothing about us without us' pushed PROMs to the fore. Patient reported outcome measures ensure that what matters to a population under investigation is included in the results and can be used to influence future decisions and policies and there is now a trend in healthcare research to use PROMs to focus more on patient-centred research [17, 18]. They can also act as a quality improvement strategy for patient care through feedback; it has however been noted that in some cases, such as palliative care and psychotherapy, clinicians viewed individualised PROMs as more useful to build rapport rather than substantially change communication practices [19].

Patient reported outcome measures became popular in the NHS in 2009 following Lord Darzi's 2008 report *High quality care for all* [20]. Since then the NHS has collected PROMs in four surgical procedures: hip replacement, knee replacement,

varicose vein surgery, and hernia surgery, and they publish an annual web based report [21] and an annual national conference on the development and use of PROMs [22]. The use of PROMs in research ensures that important knowledge about the impact of an intervention is not lost because the selected measure was unable to capture it or, even worse, distorted the true results [23].

It is essential to consider the use of PROMs in research to demonstrate impact. Such measures (including quality of life and symptoms), if collected, analysed, and reported appropriately, can be used to inform shared decision-making, clinical guidelines, and health policy. However one problem with interpreting PROMs is distinguishing how much movement on a scale equates to a clinically meaningful change. While scores on scales can be subject to rigorous statistical analysis, a statistically significant difference may not equate to a clinically meaningful change. A statistical test is used to determine whether an effect is likely to be due to chance or not and is a study is sufficiently powered then a small change can be statistically significant but this does not mean it is clinically significant. For many physiological measurements (such as temperature or blood pressure), experience and clinical judgement inform whether the results are clinically meaningful but with more subjective measurements, such as pain or stress, it becomes harder to define [24].

Patient reported outcome measures were recently sub-classified into patient satisfaction measures and patient-reported experience measures (PREMs) [25] and are discussed below in Sect. 11.3.2.

11.3.1 Patient Reported Experience Measures (PREMs)

Patient reported experience measures (PREMs) differ from PROMs in that the former capture participants' perception of their experience (with their healthcare or service) whilst PROMs focus on participants' perception of their health. Examples of PREMs include quality of communication, time spent waiting, and whether they would recommend the service to family and friends (the NHS Friends and Family test) [28].

In England, the Department of Health [20] defines the three domains of care as: patient safety, clinical effectiveness, and patient experience; commissioners and service providers are increasingly using PREMs in their assessments. A study in 2014 noted that there was a weak positive correlation between PROMs and PREMs in elective surgery in the UK; the authors [29] stated that PREMs may not be used as a proxy for a good outcome. Thus caution is advised in the use of PREMs as a primary outcome measure if the objectives of a study are not directly related to patient experience.

11.3.2 Patient Satisfaction

Patient satisfaction may be both a PROM or a PREM and are frequently used to enable researchers to focus on the impact of their research.

Measures of satisfaction differ from outcome measures such as quality of life (QoL); they address the process of treatment rather than its outcome [26]. However, patient satisfaction scales are sometimes referred to as 'happy scales' because they can mask negative experiences. Patients tend to score their care highly and there is little discrimination between items which can lead to a noted ceiling effect[2] [25]. A study of 21 EU countries [27] indicated that satisfaction is also linked to 'broader societal factors' such as the wealth and prosperity of a country.

11.4 Considerations for Choosing an Outcome Measure

Pertinent outcome measures are identified at the same time as the literature review, which is undertaken to justify a research question. However consulted research may utilise a number of different outcome measures and it may be difficult to choose the most useful way of measuring the outcome. The challenge is to ensure selected outcome measures truly reflect the change intended to be assessed. For example a scale measuring pain may be broad enough to include both acute and chronic pain, or focused enough to measure just one aspect of pain. If you are interested purely in comparing the efficacy of one intervention to manage immediate post-operative pain then a scale that measures both chronic and acute pain would likely to be misleading, as would a scale measuring chronic pain only.

With respect to PROMs both barriers and facilitators to their implementation have both been examined, and specifially in palliative care their routine use has been slow and difficult. An educational component (e.g. understanding how to complete and score PROMs) alongside an understanding of the emotional and cognitive processes of a patient, were deemed as crucial in this area of health [30].

A 2016 Delphi study [1] of 120 participants developed guidelines for selecting a 'core outcome set' (COS) of outcome measurement instruments (OMIs) to be reported in all clinical trials of a specific disease or trial population. This was in cooperation with the COMET initiative [9] which has an online searchable database of outcome measures and a wealth of advice on how to select the most appropriate outcome measure. Specific considerations for selecting outcome measures are presented below.

11.5 Licencing and Costs

Some well-known outcome measures are free to use; many others incur a charge. Outcome measures are copyrighted and may have a licence with associated terms and conditions which a researcher needs to be aware. Whilst it may be possible to

[2] A ceiling effect is when the top scale on the measurement instrument is consistently reached, thus reducing the ability of the scale to accurately capture data beyond the top of the scale.

develop a similar outcome measure with new wording it is important to bear in mind that reliability and validity are based on exact wording. Consequently changing the structure would mean the new outcome measure has no evidence to support its use and it would be difficult to prove that it consistently measures a result as expected. It is recommended that a developed and established outcome measure is used as intended and that licensing conditions are checked and adhered to before research begins.

11.6 Population and Stakeholders

In research a population is defined as a collection of individuals or objects with similar characteristics and they may be categorised by disease type, profession, or demographics. When selecting an outcome measure it is important to establish whether it has been used in a population similar to the one under consideration. Questions to ask include: Has the outcome measure been used in the same condition and disease severity as the sample being studied? Is it responsive to the differences you hope to detect? If level of pain is the primary outcome measure is acute or chronic pain being measured? Is a patient or clinician reporting the results?

Stakeholders by contrast are consumers of research outcomes and include patients, providers, payers, regulators, industry, academic, society, and policymakers [1]. The increasing acknowledgment of stakeholders in research is one driver for using PROMs. Orthopaedic research interest in PROMs dates back to the 1980s [31] but useage in this and other clinical areas has been slow. It is still usual to see clinical outcome measures that focus on the technical success of an operation such as mortality, morbidity, and complications, as opposed to patient satisfaction and quality of life.

A population and stakeholders may have different priorities in terms of what constitutes meaningful information from research and this needs to be accounted for in the selection of primary and secondary outcome measures. It is worth noting that PROMs need to be meaningful to many groups including patients, clinicians and researchers [32].

11.7 Administering and Scoring Results

It is essential to ensure a selected outcome measure is acceptable to both researcher and participant. It should be quick and simple to use, reliable, valid, specific to the question being investigated, and cost-effective. Longer outcome measures may collect full and comprehensive data but will demand a greater input from a respondent and a more thorough analysis beyond the scope of a research project. Conversely basic or simple outcome measures may not provide enough information to quantify results and report differences within or between groups.

The length of an outcome measure is important when considering a population under study. For instance, paediatric or brain injury research requires short and easy understandable outcome measures if thay are to be completed by a patient (PROMs). Equally clinical outcome measures need to be short and easy to complete if a large sample size in a short period of time is under investigation. Conversely, and with respect to the study design, a longitudinal study with a smaller sample size may require an established but complicated outcome measure, which could be more easily justified than in a cross-sectional case study design.

Whilst some outcome measures may be relatively easy and simple to administer, scoring and reporting may be more complicated and may require the use of complex statistical packages. An established outcome measure should include clear standardised instructions on how to implement and score the results and whether specialist training may be required, in which case this will need to be costed within the research.

11.8 Responsiveness

The responsiveness of an outcome measure relates to how well it can detect a change over time (longitudinal) or differences between groups. Responsiveness applies to the efficacy and effectiveness of an intervention [33]. For a measure to be responsive it also needs to be reliable and valid. The characteristics of an outcome measure affect its responsiveness, for example, ordinal data (data that place participants into categories such as ‘constant’, ‘frequent’, ‘occasional’, ‘rare’, or ‘never’) are likely to be less responsive than interval or ratio data (data on an established scale such as temperature) if large changes in status are required to change categories.

This consideration needs to be balanced against the amount of data collected and the time needed to analyse the data. Additionally outcome measures with ceiling or floor effects (that do not account for improvement or decline, or where the baseline includes a very high or very low score) may not be responsive. An example would be using an activity outcome measure which assesses level of assistance needed for certain tasks elderly post-surgical patients versus young amputees being trained for high level sports, would give a very different outcome in both groups [34].

11.9 Reliability and Validity

Reliability refers to how effectively an outcome measure can be repeated on different occasions with the same conditions and provide the same result. It is essentially the consistency of a measurement. Conversely validity refers to how well an outcome measure can assess a feature being measured. A simple explanation of reliability and validity would be an alarm clock set for 7.00 am every morning but which consistently rings at 6.30 am. IT would be a very reliable alarm clock, but it would not be valid [35].

There are different types of reliability and validity: both are divided into three sub-domains.

Reliability sub-domains include the following.

- Internal consistency is the extent that all items on a scale measure the same thing. For instance, does a pain scale only measure pain or does it also include questions which measure disability?
- Stability is the consistency of results in repeated testing in the same and different populations.
- Equivalence is the level of agreement between the interpretations of the scores; sometimes known as intra- and inter-rater agreement.

Validity sub-domains include the following.

- Content validity is how well a measure captures all the features of the domain. For instance, if both acute and chronic pain are being measured does the intended outcome measure cover both aspects or only acute pain?
- Construct validity is how well an intended outcome is measured. For instance, does a pain measurement only allow a score of 1–10 on a visual analogue scale (VAS) or does it also allow for responses measuring the use of painkillers?
- Criterion validity is how well a measure relates to other measures which examine the same outcome. For instance, does a VAS for disability relate to other disability outcome measures? [36].

Both reliability and validity of a chosen outcome measure should be noted when choosing the most appropriate way to record and interpret the results of a research. Although these are not always recorded it is imperative that they are investigated if a new outcome measure is being developed.

11.10 Subjective Versus Objective Outcome Measures

Outcome measures can be subjective or objective. Both have challenges in practice and there is no 'one size fits all' approach. An objective measure is more likely to consider medical data and be collected by professional equipment such as pulse oximeters, cardiac monitors, or biochemistry data. Such measures are precise and reduce bias thus they are widely used in research. They tend to collect short-term data which can change quickly.

Subjective measurements are defined as those which are open to interpretation such as questionnaires and visual analogue scales (VAS). They are generally quick and easy to administer and can be completed by a researcher or patient but they are more prone to bias and errors in both completing and reporting results. Pain as an outcome measure is subjective as demonstrated in a study of nursing students who were asked to identify pain on a VAS [37]. They reported a wide range of perceptions demonstrating the highly subjective nature and different terminology associated with pain. Radiographic scoring methods, often used in an assessment of onset

and progression of joint disease/degeneration, are also subjective. Numerous methods for this exist [38] and the results are often statistically reported because their scales and data distribution meet parametric properties and standards. To reduce bias in such cases it is preferable to have more than one independent assessor to enable intra- and inter-rater agreement to be presented [39].

11.11 Confounders

It is important to consider extraneous uncontrolled independent variables, or confounders, when designing a study and selecting an outcome measure. A confounder is a distortion of an association of an intervention and outcome. Unlike other kinds of research bias, such as selection or researcher bias, a confounder can be adjusted in an analysis providing the researcher knows the confounding variables in advance. Such variables often include patient demographics, such as height, weight, age, gender, race, but might also include a multitude of other variables such as stage of disease, co-morbidity, signs and symptoms, duration of disease, and types of imaging equipment used [40]. For example, bone density reduces with age in normal and osteoporotic subjects. If an outcome measure is the odds ratio and there is no adjustment for age then this will lead to an over-estimation of the discrimination of fracture cases from controls. Therefore age adjustment is required in the analysis to correct the problem [39].

11.12 Searching for an Outcome Measure

It is useful to search relevant literature and systematic reviews to identify an appropriate outcome measure. There are also a number of online searchable databases for outcome measures although these are subject to change with little or no notice. As with literature it is worth searching a number of databases to identify more popular outcome measures. When searching online databases take careful note of the owner of the URL and also the number of studies registered that have used, or are using, the outcome measure of interest.

The databases presented here are current at the time of writing.

11.12.1 The COMET Initiative

The COMET initiative focuses on effectiveness trials and brings together a core outcome set (COS) of minimum outcome measures which should be reported in clinical trials of a specific condition. It also explains problems and key issues in using outcome measures in research. This database (http://www.comet-initiative.org/) is endorsed by the European Commission, The Medical Research Council (MRC), The National Institute for Health Research (NIHR), and the Seventh Framework Programme.

11.12.2 Rehabilitation Measures Database

This is a privately held and free to search database of outcome measures associated with rehabilitation which are predominantly questionnaire based. This database (https://www.sralab.org/rehabilitation-measures) gives an overview of the listed outcome measures including the time expected to complete and costs of the licence. It is organised by assessment type (including PROMS and performance measures), area of assessment including activities of daily living (ADLs), patient satisfaction, and populations including allied health care professions, joint care, and fractures.

11.12.3 Patient Reported Outcome Measurement Group

This PROM group is based at the University of Oxford. The website (http://phi.uhce.ox.ac.uk/home.php) includes guidance on the selection of appropriate outcome measures for use in clinical trials, practice, and population surveys. It is no longer supported with updates but it is still a useful resource.

11.12.4 Proqolid

This database (https://eprovide.mapi-trust.org/about/about-proqolid) [41] provides information on over 2000 clinical outcome assessment tools. There are two levels of access; free and subscription. The free version provides basic information on outcome measures including the therapeutic area, indication, and bibliographic references for the description of the outcome measure.

11.12.5 OMERACT

Outcome Measures in Rheumatology (OMERACT) is an online database of outcome measures pertinent to rheumatoid diseases (https://omeract.org/) as well as clinical and radiographic outcome measures [42]. There is a heavy emphasis on patient involvement in endorsing outcome measures and advice on how to select an appropriate outcome measure.

11.13 Reporting Outcome Measures

Both quantitative and qualitative research follow a structured approach for writing and reporting results and the choice and use of outcome measures need to be clearly stated in research papers. A good example is presented by Hardy et al. [43].

The methods section acts as a recipe and provides details to allow full replication. Within this section it is important to clearly state the outcomes measured and the tools used for measurement. The primary outcome measure needs to align with

the overall aim of a study and secondary outcome measures align with its objectives. It is important to provide enough detail for a reader to know exactly what was measured and how.

The results section displays the outcomes of research and many outcome measures may have standard criteria for this. Following the stated criteria enables comparison with other research that used the same outcome measures.

The discussion section is where the results are compared to other research and the comparison is easier if other research has used the same outcome measures. It is also the section where the choice of outcome measures can be justified. This includes providing substantiated evidence of the outcome measures' reliability and validity, and relating these to the primary and secondary outcome measures with evidence of how they collectively support a study's aims. There are many instances where more than one outcome measure would have captured the relevant data but including too many outcome measures can lead to an unfocused research question and present problems with interpretation if the effect differs across the outcomes [44].

The conclusions of the study state the composite endpoints which are the amalgamation of the selected outcome measures that have been correlated to support or refute the hypothesis. The components of selected outcome measures should be complementary, with secondary outcome measures lending supporting evidence to the primary outcome measure.

11.14 Outcome Measures in Radiography Research

There are no recommendations for specific outcome measures for radiography research because the profession spans a range of conditions and diseases, and research within radiography is varied and far ranging.

However radiographic scoring methods are a well-established outcome measure in many diseases including rheumatology [38] and PROMs are frequently used in cancer research to establish the impact of cancer treatment [45]. A systematic review in 2006 found there were no PROMs developed specifically for radiology [46].

A focus of research within imaging and oncology is new techniques and technology development which are regularly introduced and require evidence to test their ability to detect pathologies or to predict clinical outcomes, cost effectiveness, and patient-centred outcomes. Such research often uses tests of reliability and accuracy as outcome measures, including sensitivity and specificity.

If there is already a technique for the diagnosis of a pathology of interest it is important to directly compare the techniques. The best method is to undertake this within the same patients so they would have two imaging or diagnostic techniques rather than one, however this is not always practicable or possible. There is also the potential that a combination of a new and old technique can improve the diagnostic accuracy above and beyond either technique when used individually, in which case this needs to be included in the study design and analysis.

Secondary outcome measures of new technology research often include the radiation dose associated with each of the techniques, the cost of the technique in terms of duration, cost of the equipment and consumables, how invasive the technique is, and consumers' perception with the treatment (PROMs). While some lower dose and cheaper techniques may be introduced it is important that they match or exceed the diagnostic accuracy of an existing gold standard [39].

In radiotherapy outcome measures are used extensively in quality assurance such as cancer wait times with statistics widely available from the NHS website [47]. This outcome measure does not alone capture the depth and breadth of radiotherapy and is rarely used in cancer research studies. Many radiographers may be involved in cancer research and may work within a research network ensuring the quality assurance of radiotherapy trials. Within the UK the Radiotherapy Trials Quality Assurance group (RTTQA) ensures all radiotherapy trials are conducted to the same standard. Whilst the RTTQA group is not concerned with outcome measures per se, it does hold the full trial protocols of radiotherapy studies and these protocols include the selection and justification of the outcome measures selected.

An example of such a trial called Fast Forward, which is a trial of radiotherapy in breast cancer, lists the primary outcome measure as ipsilateral local tumour control and secondary outcome measures as early and late adverse effects in normal tissues, quality of life, contralateral primary tumours, regional and distant metastases, and survival and publications regarding skin toxicity (a secondary outcome measure) have already begun to influence practice [48]. Quality of life is the only PROM in this study because its primary objective is to identify whether a five fraction schedule of curative radiotherapy is at least as effective and safe as the 15 fraction regime, but other breast cancer studies have centred on PROMs such as the START trial that examined patient reported breast, arm, shoulder symptoms, and body image [49].

Whilst there are no recommended outcome measures for radiography research, the Society and College of Radiographers (SCoR) have set research priorities which fall under four domains: accuracy and safety; effectiveness of technical approaches; the patient experience; and service delivery and organisation [50]. Aligning radiography research aims and objectives within these priorities would aid in selecting the most suitable primary and secondary outcome measures.

11.15 Conclusion

The final take home message is that outcome measures are a central aspect of research and need to be considered from the initial design stage. It is important to take time to select the most appropriate outcome measures, taking into account the considerations for choosing an outcome measure as covered in this chapter. Finally researchers need to justify their choice of outcome measure(s) and ensure they align with the aims and objectives of their research.

References

1. Prinsen CAC, Vohra S, Rose MR, Boers M, Tugwell P, Clarke M, et al. How to select outcome measurement instruments for outcomes included in a "Core Outcome Set" a practical guideline. BMC Trials. 2016;17(1):449.
2. Cameron WB. Informal sociology a casual introduction to sociological thinking. New York: Random House; 1963.
3. Rainforth MV, Schneider RH, Nidich SI, Gaylord-King C, Salerno JW, Anderson JW. Stress reduction programs in patients with elevated blood pressure: a systematic review and meta-analysis. Curr Hypertens Rep. 2007;9(6):520–8.
4. Nelson EC, Eftimovska E, Lind C, Hager A, Wasson JH, Lindblad S. Patient reported outcome measures in practice. BMJ (Clinical research ed). 2015;350.
5. Andrade C. The primary outcome measure and its importance in clinical trials. J Clin Psychiatry. 2015;76(10):1320–3.
6. Ferreira JC, Patino CM. Types of outcomes in clinical research. J Bras Pneumol. 2017;43(1):5.
7. Prinsen CAC, Mokkink LB, Bouter LM, Alonso J, Patrick DL, HCW DV, et al. COSMIN guideline for systematic reviews of patient-reported outcome measures. Qual Life Res. 2018;27(5):1147–57.
8. Mohtadi NG. Outcome measure development. Instr Course Lect. 2016;65:577–82.
9. Dodd SR, Clarke M, Becker L, Mavergames C, Fish R, Williamson PR. A taxonomy has been developed for outcomes in medical research to help improve knowledge discovery. J Clin Epidemiol. 2018;96:84–92.
10. Wilson IB, Cleary PD. Linking clinical variables with health-related quality of life. A conceptual model of patient outcomes. J Am Med Assoc. 1995;273(1):59–65.
11. NHS Better Care Fund Taskforce. 'How to guide'; the better care taskforce technical toolkit. Section 3. Outcomes and impact measurement. London; 2014.
12. University of Oxford. Patient reported outcome measures: instrument types. Oxford; 2010 [cited 2019 May 13]. http://phi.uhce.ox.ac.uk/inst_types.php.
13. The Association of UK Dietitians. Model for dietetic outcomes. Birmingham; 2011 [cited 2019 May 15]. https://www.bda.uk.com/publications/archive/bda_outcome_model__2011_archive.
14. The Association of UK Dietitians. Model and process for nutrition and dietetic practice. 2017 [cited 2019 May 14]. https://www.bda.uk.com/professional/practice/process.
15. OMERACT (Outcome Measures in Rheumatology) [cited 2019 May 20]. https://omeract.org/.
16. Boers M. How outcome measures in rheumatoid arthritis clinical trials works to develop outcome measures in rheumatology clinical trials. Choosing and developing core outcome measurement sets for clinical trials: outcome measures in rheumatoid arthritis clinical trials filter 2.0. Rheumatology. 2017;56:kex060.116.
17. Speight J, Barendse SM. FDA guidance on patient reported outcomes. BMJ. 2010;c2921:340.
18. Staniszewska S, Haywood KL, Brett J, Tutton L. Patient and public involvement in patient-reported outcome measures: evolution not revolution. Patient. 2012;5(2):79–87.
19. Greenhalgh J, Dalkin S, Gooding K, Gibbons E, Wright J, Meads D, Health Services and Delivery Research, et al. Functionality and feedback: a realist synthesis of the collation, interpretation and utilisation of patient-reported outcome measures data to improve patient care. Southampton: NIHR Journals Library; 2017.
20. Darzi A. High quality care for all: NHS next stage review final report. London: The Stationery Office; 2008.
21. NHS Digital. Patient reported outcome measures (PROMs). 2019 [cited 2019 May 15]. https://digital.nhs.uk/data-and-information/data-tools-and-services/data-services/patient-reported-outcome-measures-proms.
22. Gibbons E, Calvert M, Bostock J, Skryban M. Proceedings of patient reported outcome measures (PROMs) conference Birmingham 2018: Birmingham, UK. 20 June, 2018. J Patient Rep Outcomes. 2018;2(Suppl 2):58.

23. Coster WJ. Making the best match: selecting outcome measures for clinical trials and outcome studies. Am J Occup Ther. 2013;67(2):162–70.
24. Bolton JE. Sensitivity and specificity of outcome measures in patients with neck pain: detecting clinically significant improvement. Spine. 2004;29(21):2410–7; discussion 8.
25. Hodson M, Andrew S, Michael Roberts C. Towards an understanding of PREMS and PROMS in COPD. Breathe. 2013;9:358–64.
26. McKenna SP. Measuring patient-reported outcomes: moving beyond misplaced common sense to hard science. BMC Med. 2011;9(1):86.
27. Bleich SN, Ozaltin E, Murray CJL. How does satisfaction with the health-care system relate to patient experience. Bull World Health Organ. 2009;87:271–8.
28. National Health Service. Friends and family test. 2018 [cited 2019 May 20]. https://www.nhs.uk/using-the-nhs/about-the-nhs/friends-and-family-test-fft/.
29. Black N, Varaganum M, Hutchings A. Relationship between patient reported experience (PREMs) and patient reported outcomes (PROMs) in elective surgery. BMJ Qual Saf. 2014;23:534–42.
30. Antunes B, Harding R, Higginson IJ. Implementing patient-reported outcome measures in palliative care clinical practice: a systematic review of facilitators and barriers. Palliat Med. 2014;28(2):158–75.
31. Brinker M, O'Connor D. Stakeholders in outcome measures: review from a clinical perspective. Clin Orthop Relat Res. 2013;471(11):3426–36.
32. Duncan E, Murray J. The barriers and facilitators to routine outcome measurement by allied health professionals in practice: a systematic review. BMC Health Serv Res. 2012;12(1):96.
33. Tarrant C, Angell E, Baker R, Boulton M, Freeman G, Wilkie P, et al. Responsiveness of primary care services: development of a patient-report measure – qualitative study and initial quantitative pilot testing. Health Serv Deliv Res. 2014;2(46):1.
34. Roach KE. Measurement of health outcomes: reliability, validity and responsiveness. J Prosthet Orthot. 2006;18(6):P8–P12.
35. Heale R, Twycross A. Validity and reliability in quantitative studies. Evid Based Nurs. 2015;18(3):66.
36. Boonstra AM, Schiphorst Preuper HR, Reneman MF, Posthumus JB, Stewart RE. Reliability and validity of the visual analogue scale for disability in patients with chronic musculoskeletal pain. Int J Rehabil Res. 2008;31(2):165–9.
37. Bergh I, Sjostrom B. Quantification of the pain terms hurt, ache and pain among nursing students. Scand J Caring Sci. 2007;21(2):163–8.
38. Boini S, Guillemin F. Radiographic scoring methods as outcome measures in rheumatoid arthritis: properties and advantages. Ann Rheum Dis. 2001;60(9):817–27.
39. Mellor F, Knapp K. Research outcome measures. In: Ramlaul A, editor. Medical imaging and radiotherapy research: skills and strategies. London: Churchill Livingstone; 2010. p. 111–21.
40. Crewson PE, Applegate KE. Data collection in radiology research. Am J Roentgenol. 2001;177(4):755–61.
41. Pinotti R. PROQOLID. J Med Libr Assoc. 2016;104(1):91–2.
42. Kelly A, Tong A, Tymms K, March L, Craig JC, De Vera M, et al. Outcome measures in rheumatology - core domain set for trials of interventions for medication adherence in rheumatology: 5 phase study protocol. Trials. 2018;19:204.
43. Hardy M, Hutton J, Snaith B. Is a radiographer led immediate reporting service for emergency department referrals a cost effective initiative? Radiography. 2013;19(1):23–7.
44. Vetter TR, Mascha EJ. Defining the primary outcomes and justifying secondary outcomes of a study: usually, the fewer, the better. Anesth Analg. 2017;125(2):678–81.
45. Faithfull S, Lemanska A, Chen T. Patient-reported outcome measures in radiotherapy: clinical advances and research opportunities in measurement for survivorship. Clin Oncol. 2015;27(11):679–85.
46. Mathers SA, Chesson RA, Proctor JM, GA MK, Robertson E. The use of patient-centered outcome measures in radiology: a systematic review. Acad Radiol. 2006;13(11):1394–404.

47. National Health Service. Cancer waiting times. [cited 2019 May 15]. https://www.england.nhs.uk/statistics/statistical-work-areas/cancer-waiting-times/.
48. Brunt AM, Wheatley D, Yarnold J, Somaiah N, Kelly S, Harnett A, et al. Acute skin toxicity associated with a 1-week schedule of whole breast radiotherapy compared with a standard 3-week regimen delivered in the UK FAST-forward trial. Radiother Oncol. 2016;120(1):114–8.
49. Hopwood P, Haviland JS, Sumo G, Mills J, Bliss JM, Yarnold JR. Comparison of patient-reported breast, arm, and shoulder symptoms and body image after radiotherapy for early breast cancer: 5-year follow-up in the randomised standardisation of breast radiotherapy (START) trials. Lancet Oncol. 2010;11(3):231–40.
50. Society and College of Radiographers. Society and College of Radiographers Research Priorities. [cited 2019 May 20]. https://www.sor.org/system/files/article/201312/scor_research_priorities.pdf.

Part III

Health Technology Assessment

Health Technology Assessment

12

Heidi Probst and Aarthi Ramlaul

12.1 Introduction

HTA involves the assessment of all new technologies used in radiography and radiotherapy. Effectiveness refers to the extent to which benefits are brought to patients in routine circumstances, and efficiency refers to the extent to which acceptable effectiveness is achieved with the best use of resources.

In this chapter we will discuss the following research methods used to assess health technologies in medical imaging and radiotherapy. In order to give each of these areas their deserved attention, each is covered within a separate sub-chapter, as follows.

- Researching diagnostic tests
- Researching therapies using randomized controlled trials (RCTs)
- Health economic assessment and RCTs
- Systematic reviews and meta-analyses of RCTs

H. Probst (✉)
Radiotherapy and Oncology, College of Health, Wellbeing and Life Sciences, Sheffield Hallam University, Sheffield, UK
e-mail: h.probst@shu.ac.uk

A. Ramlaul
Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: a.ramlaul@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_12

12.2 Researching Diagnostic Tests

Diagnostic testing can be seen as the collection of information which will clarify a patient's clinical condition and help to determine prognosis. This information can include patient characteristics, signs and symptoms, clinical history, physical examination or clinical tests. Practitioners working in diagnostic imaging are particularly interested in providing high quality images which will permit an accurate medical diagnosis. Diagnostic imaging is a rapidly evolving specialty, and numerous imaging procedures such as Barium enemas, angiography and intravenous pyelography, to name a few, are being replaced by computed tomography (CT) and magnetic resonance imaging (MRI). New technologies, however, are complex and expensive and research is therefore required to evaluate them, in order to decide if and when they should be introduced into clinical practice.

The purpose of this section of the sub-chapter is to provide an overview of what is meant by evaluation of diagnostic technologies, focusing on research that measures the diagnostic performance (or accuracy) of an imaging modality and provides estimates of observer variability.

12.2.1 Evaluative Hierarchy of Diagnostic Technologies

Choices between alternative healthcare policies may be explored within healthcare evaluation, including the investigation of the efficacy and efficiency of available diagnostic technologies. It is not always apparent how the diagnostic technology itself brings about improvements in the prognosis or physical health of a patient. An imaging examination provides information from which a reporting radiographer or radiologist makes a report. This is then used by the clinician in combination with clinical findings and other tests to make or refine diagnosis and plan treatment which ultimately might affect patient outcomes. Therefore, to evaluate the effectiveness of imaging requires the measurement of a chain of events between the application of the technology and any potential influence on disease. With the development of CT in the 1970s, Fineberg and colleagues [1] suggested that hierarchy could be used to evaluate the effectiveness of diagnostic technologies. This has subsequently been extended to include whether the costs for a given examination are acceptable, providing an efficient use of resources [2]. Figure 12.1 presents the evaluative hierarchy as applied to the assessment of MRI [3].

Technical performance is the first level of the evaluative hierarchy and is concerned with whether, for example, MRI produces good quality images from which diagnostic and therapeutic decisions may be made [4].

The next level is diagnostic performance, which is concerned with whether imaging, such as MRI of the knee, correctly or incorrectly assesses the presence or absence of disease, such as meniscal or ligamentous injury, as corroborated by a 'gold standard' test (such as arthroscopy in this instance). Assessment of diagnostic performance is expressed using statistics such as sensitivity and specificity. Sensitivity is the percentage of correct abnormal diagnoses in patients with disease;

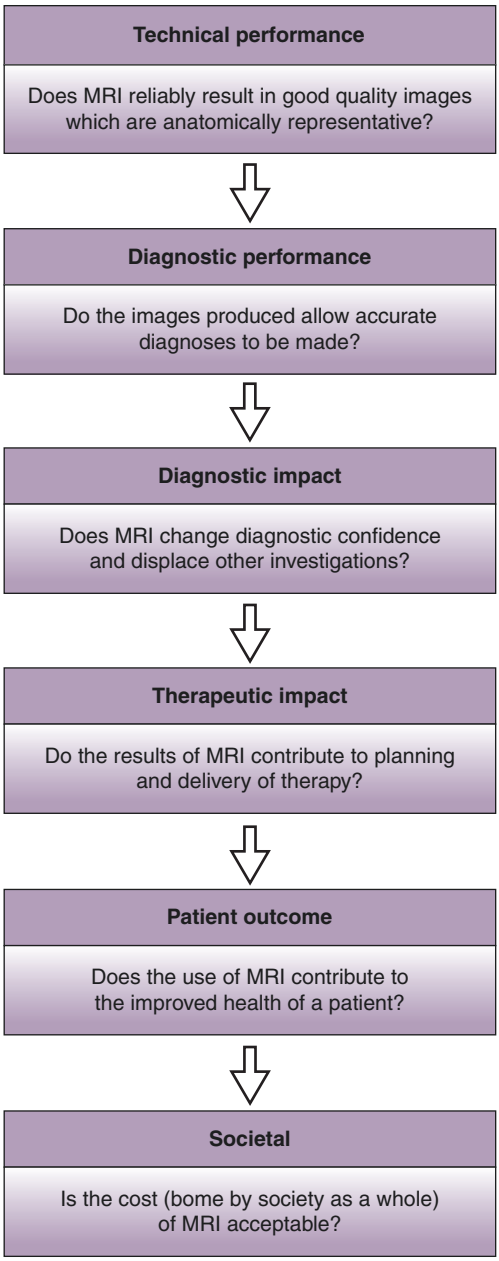


Fig. 12.1 The hierarchy used to evaluate MRI

and specificity is the percentage of correct normal diagnoses in patients without disease. Furthermore, observer variation in the interpretation of medical images is substantial and has been described as radiology's 'Achilles' heel' [5]. Thus, it is important to estimate observer variability, since the accuracy of the diagnostic test

can be a joint function of the images produced and the performance of the observers [2]. This level in the evaluation of a diagnostic technology is discussed further in the next section.

The following three levels of the evaluative hierarchy are concerned with:

- diagnostic impact, e.g., does MRI replace existing technologies?
- therapeutic impact, e.g., do MRI findings lead clinicians to make changes in treatments?
- patient outcome, e.g., does MRI improve patients' prognoses?

These levels of the hierarchy are often assessed using observational research designs. In these the technologies are simply observed and compared, without the experimental intervention that would take place in a randomized controlled trial. An example might be recording pre-imaging diagnosis and management plans and comparing this with post-imaging plans. Such studies assume that any change in diagnosis and management plan, or change in patient outcome, is attributable to MRI. The effectiveness of MRI, however, might be explained by the influence of other variables. One possibility is that there is a tendency for measured outcomes to 'average out' over time following the introduction of a new policy, due to random fluctuations in performance results, if enough results are taken. This is referred to statistically as 'regression towards the mean'. Another reason could be the Hawthorne or 'guinea pig' effect, which is the tendency for data to be biased because research subjects become aware they are being observed [6].

The best method for evaluating the effectiveness of technologies such as MRI is the randomized controlled trial (RCT), which will, in a controlled way, randomly allocate patients to receive either one diagnostic test or an alternative. Although there are logistical and financial implications to using RCTs, this method promotes study validity and provides a good basis for making statistical inferences [4]. The randomized controlled trial design is discussed later in the chapter.

The final level of the evaluative hierarchy moves beyond merely measuring the clinical effects of a technology to determining whether the cost of that technology is acceptable to society. For the policy maker entrusted with making resource allocations, it is necessary to assess the extent to which MRI is an efficient use of resources to provide benefits to society [2]. This could take the form of, for example, a cost-effectiveness study which involves computing a cost per unit of output for a medical technology such as cost per arthroscopy avoided by using MRI of the knee. The different methods of economic evaluation are discussed later in the chapter.

12.2.2 Studies of Diagnostic Test Accuracy

When diagnosing a patient, clinicians seldom have access to the gold standard or reference standard test for the disorders they suspect since these tests can be expensive, painful and/or invasive. There are many alternative tests that can be used for

patient diagnosis, such as taking a patient's history, physical examination, laboratory tests and diagnostic imaging. Diagnostic accuracy studies, which comprise the second level of the evaluative hierarchy, are vital to the assessment of imaging technologies, since they help to understand how they should be best used in clinical practice.

12.2.3 The Research Question

Sackett and Haynes [7, 8] identified four types of research questions that can be used to assess the real value of a diagnostic test such as an imaging modality.

- Do diagnostic test results in patients with the target disorder differ from those in normal people (a phase I question)?
- Are patients with certain diagnostic test results more likely to have the target disorder than patients with other test results (a phase II question)?
- Does the diagnostic test result distinguish patients with and without the target disorder among patients in whom it is clinically reasonable to suspect the disease is present (a phase III question)?
- Do patients who undergo this diagnostic test have better health outcomes than similar patients who are not tested?

Phase III questions are the most frequently asked in studies of diagnostic test performance and are concerned with the validity of the diagnostic test or rather whether it measures what it proposes to measure.

To evaluate whether a test can distinguish normal from abnormal patients during routine clinical practice requires the results of the test to be compared against the gold or reference standard that is acknowledged as being the best available test to accurately diagnose the patient's true disease status. To compare measurements, i.e., the diagnostic test and reference standard results, is to assess validity and this will be the main focus of this section. Studies of the diagnostic accuracy, or validity of a test, particularly for imaging modalities, should also consider whether the different observers responsible for interpreting medical images are doing this consistently; this provides an assessment of reliability. The design and analysis of reliability studies will also be briefly discussed.

12.2.4 Design of a Study of Validity

As described above, a diagnostic accuracy study involves the assessment of whether a diagnostic test can distinguish patients with and without the target disorder, as corroborated by gold or reference standard, among patients in whom it is clinically reason- able to suspect the presence of disease. If the study design is inadequate, there is experimental evidence that the performance of diagnostic tests might be exaggerated [9]. The STARD (Standards for the Reporting of Diagnostic accuracy

studies) statement, which is a checklist used to guide the reporting of studies of accuracy [9], and the QUADAS (Quality Assessment for Diagnostic Accuracy Studies), which is a generic tool used to appraise the quality of primary studies in systematic reviews of diagnostic accuracy [10], provide thorough descriptions of the relevant design issues when considering the validity of a diagnostic test. These design issues are also discussed in Chap. 10. In summary then, when designing a diagnostic accuracy study, it is important to consider the following areas as they pose an element of risk to the study validity [11].

- Patient selection—a consecutive series of patients suspected (but not known) to have the target disorder should be prospectively selected as a cohort of patients for inclusion in the study. There should be a clear description of the selection criteria and the setting, e.g., primary, secondary or tertiary care.
- Choice and application of the reference standard—the reference standard chosen should produce results close to the truth, or the performance of the diagnostic test will be poorly estimated.

  The reference standard should be applied within a clinically acceptable timeframe after the diagnostic test and preferably to the whole or at least a random sample of patients to avoid partial verification of patients. Nor should the index test form part of the reference standard.
- Measurement of results—a study should fully report indeterminate test results that occur due to factors such as technical faults or inferior image quality, and withdrawals that may occur due to patient death, move in residency or no longer wanting to cooperate. It is important to consider whether they are non-random exclusions and the effect on generalizability.
- Independence of interpretation—the reference standard should be interpreted blind, i.e., in total ignorance of the diagnostic test result and vice versa.

12.2.5 Analysis of a Study of Validity

Various measures can be used to assess how well a diagnostic test discriminates between patients with disease from those without disease. The diagnostic test will detect the presence of a disease, such as a lesion on a digital mammogram, and then be correctly classified as being present or absent by biopsy, as the reference standard. This 'binary' classification of results allows individuals to be classified either as true positives (TP) or true negatives (TN), which means that the test results are correct; or false positives (FP) and false negatives (FN), which means that the test results are incorrect (Fig. 12.2). Positive and negative refer to the presence or absence of the target disorder.

The number of individuals classified as TP, TN, FP and FN permits the calculation of sensitivity and specificity, predictive values and likelihood ratios to answer different questions as described below:

Sensitivity is the proportion of patients with disease who have a positive test result: i.e., how good is my diagnostic test in detecting patients with disease?

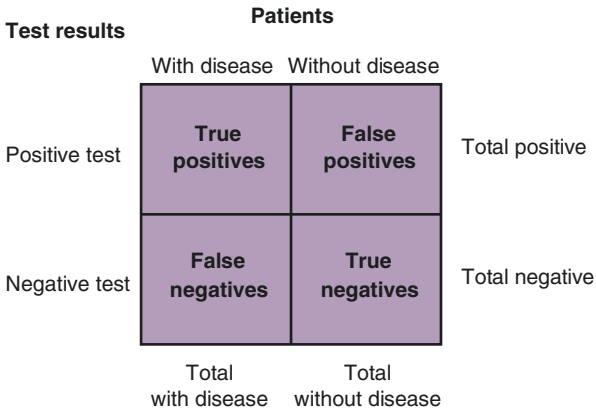


Fig. 12.2 Binary classification of results

$$\text{Sensitivity} \frac{1}{4} \text{TP} = \text{total with disease } \eth \text{TP} \text{ þ } \text{FN} \text{Þ}$$

Specificity is the proportion of patients without disease who have negative test results: i.e., how good is my diagnostic test in detecting patients without disease?

$$\text{Specificity} \frac{1}{4} \text{TN} = \text{total with disease } \eth \text{FP} \text{ þ } \text{TN} \text{Þ}$$

Positive predictive value is the proportion of patients with positive test results who have the disease: i.e., how well does a positive test result predict the presence of disease?

$$\text{Positive predictive value} \frac{1}{4} \text{TP} = \text{total positive } \eth \text{TP} \text{ þ } \text{FP} \text{Þ}$$

Negative predictive value is the proportion of patients with negative test results who do not have the disease: i.e., how well does a negative test result predict the absence of disease?

$$\text{Negative predictive value} \frac{1}{4} \text{TN} = \text{total negative } \eth \text{FN} \text{ þ } \text{TN} \text{Þ}$$

Positive likelihood ratio is the ratio of the true positive rate to the false positive rate: i.e., how much are the odds of the disease increased when a test is positive?

$$\text{LR} \text{ þ ve} \frac{1}{4} \text{sensitivity} = \eth 1 - \text{specificity} \text{Þ}$$

Negative likelihood ratio is the ratio of the false negative rate to the true negative rate: i.e., how much are the odds of the disease decreased when a test is negative?

LR-ve ¼ ð1 − sensitivityÞ=specificity Likelihood ratios can be applied to clinical practice to estimate the chances of disease in a patient according to their test result using Bayes' theorem [12]. In order to calculate the post-test odds of disease, you need to specify the pre-test odds: i.e., the likelihood that the patient would have a specific disease prior to testing. The pre-test odds are usually related to the prevalence of the disease, though you might adjust it depending on characteristics of the individual patient. Once you have specified the pre-test odds, you multiply them by the likelihood ratio. This gives you the post-test odds. Suppose a woman had a negative mammogram when screening for breast cancer and the local prevalence of

cancer among women is 5% and the negative likelihood ratio for a mammogram is 0.20. Using Bayes' theorem we can estimate that the woman's probability of breast cancer prior to screening will be reduced after a negative mammogram from 5% to 1%.

- pre-test odds ¼ prevalence/(1 − prevalence) ¼ 0.05/0.95 ¼ 0.05
- post-test odds ¼ pre-test odds ∗ LR-ve ¼ 0.05 ∗ 0.20 ¼ 0.01
- post-test probability ¼ post-test odds/(1 þ post-test odds) ¼ 0.01/1.01 ¼ 0.01 (or 1%)

Sometimes, however, the test under evaluation might yield results as a continuous measurement or ordered categories. The images from MRI of the knee, for example, might be used to describe some anatomical feature such as degenerative changes in the menisci as definitely, probably or possibly present, and probably or definitely absent, and then confirmed as present or absent by arthroscopy. Sensitivity and specificity could still be calculated by combining categories above and below a threshold, such as combining definitely, probably or possibly present compared to combining probably or definitely absent.

Changing the threshold will alter the estimates of sensitivity and specificity. A more useful method, however, of measuring the performance of MRI across a range of thresholds, or 'cut-offs', is the receiver operating characteristic (ROC) curve (see also Chap. 8). The ROC curve, as shown in Fig. 12.3, shows

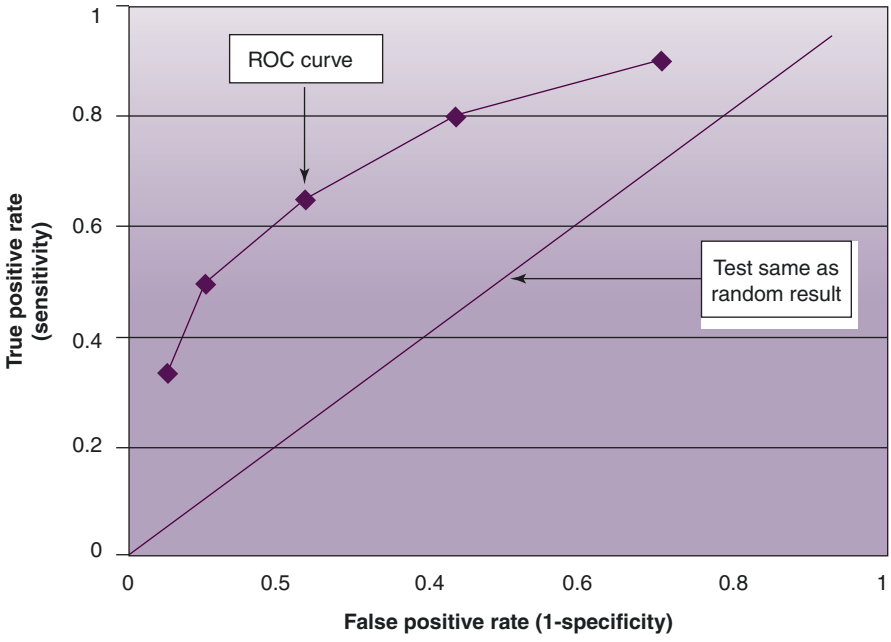


Fig. 12.3 Example ROC curve for an imaging procedure with ordinal categories

graphically the trade-offs at each cutoff for any diagnostic test that uses an ordinal or continuous variable. Ideally, the best cutoff value provides both the highest sensitivity and the highest specificity. This can be located on the ROC curve by finding the highest point on the vertical axis and the furthest to the left on the horizontal axis [13]. Alternatively, depending on the target disorder, it might be more important to exclude disease, so a higher sensitivity is chosen at the cost of lower specificity. Furthermore, it is possible to calculate the area under the ROC curve. When this is 0.5 (i.e., 50% sensitive and 50% specific) it represents a totally uninformative test, as shown in Fig. 12.3, by a straight diagonal line extending from the lower left corner to the upper right. A test that perfectly separates diseased from non-diseased patients would have an area under the curve of 1.0 (i.e., 100% sensitive and 100% specific). If the area under the curve of MRI of the menisci of the knee is 0.85, then the interpretation of the value is as follows. If two patients are drawn randomly from a sample of patients, in whom degeneration of menisci is present and absent, respectively, and, are both subjected to MRI to determine which patient had degeneration of the menisci, then MRI will be correct 85% of the time [14].

12.2.6 Assessment of Reliability

Diagnostic performance studies of imaging modalities require observers to interpret images and it is observer variability in this task that is considered to be the weakest aspect of clinical imaging [5]. It is important to estimate the variability of observers' performance, or the reproducibility with which an observer interprets an image, as this will influence the decisions made by clinicians and could ultimately affect patient outcome. The assessment of reliability involves different observers interpreting the same sample of images, known as an inter-observer test, or the same observers interpreting the same images on separate occasions, known as an intra-observer test [15]. We shall restrict our discussion of reliability to inter-observer variability as the principles of study design and analyses also apply to an assessment of intra-observer variability as well. In addition, inter-observer variability demonstrates observer consistencies within and between both sets of observers in the interpretation of images. As with studies of validity, similar principles apply to the design of a reliability study such as the need for a representative sample of patients and blinding in the interpretation of images. Selection bias is less likely when a consecutive or random sample of images is included, and blinding avoids the knowledge of one observer's interpretation influencing the interpretation of another observer. Availability of clinical data to observers should also be considered. It is important in a reliability study to carefully choose which observers are involved in the interpretation of images. For example, a study that includes highly specialist observers is likely to produce less generalizable results but in contrast could help to produce the best estimates of observer variability. Characteristics of observers that have been considered important in the assessment of reliability include the number of observers and their areas of training and expertise.

In studies of inter-observer variability, it is not assumed that one particular observer produces the correct report, but rather there is a genuine difference in interpretation of images between observers. The measure of performance used to analyse whether observers' reports agree is called the Kappa statistic [16]. It can be calculated when the classification of an image by an observer is binary, e.g., the presence or absence of a fracture on a plain radiograph, or ordinal, e.g., a normal mammogram, one which shows benign disease, the suspicion of cancer or the presence of cancer.

Kappa is defined as $K = (P_o - P_e)/(1 - P_e)$, where P_o is the observed proportion of agreement, and P_e is the proportion expected by chance. Kappa has a maximum of 1.0 when there is perfect agreement between observers and a value of zero indicates no better than chance. Kappa can be calculated for agreement between:

- a single observer interpreting the same image on two separate occasions
- two different observers on the same occasion
- comparisons of multiple observers [5].

When considering Kappa for ordinal categories, it might be preferable to use weighted Kappa which gives different weighting to disagreements in accordance with the extent of the discrepancy.

12.3 Researching Therapies Using Randomized Controlled Trials (RCTs)

12.3.1 Introduction

Diagnostic imaging and radiotherapy practitioners will be aware of the pace of technological change. However, the introduction of a new technology should be accompanied by a careful assessment of its value over existing methods. Meticulous assessment of any new technology should involve a controlled analysis of the new technology compared with the current approach [17]. The aim of this section is to provide an overview of randomized controlled trials (RCTs) and how they can be used within radiation therapy and imaging. By the end of this section practitioners should understand how to apply RCT designs for their own investigations as well as appraise RCTs published within the literature for applying evidence in practice. This section will start with a brief review of the benefits of RCTs and why they are considered a powerful research tool within HTA. Following this the specific characteristics and types of RCTs will be presented with examples of how the design characteristics could be used to investigate topics of relevance to clinical practitioners and those working in healthcare education.

The quality of a RCT, i.e., how stringent the design of the study is in limiting opportunities for bias, can influence potential outcomes by either overestimating or underestimating the benefit of the intervention. Such distortions have the potential to lead to ineffective treatments or interventions being employed and

effective treatments being discarded [18, 19]. Quality can be affected at many different stages of design and implementation and so throughout the following section attention will be paid to limitations of RCTs and the factors that may affect internal validity.

The final part of this section will focus on the use of economic evaluations alongside RCTs as part of HTA utilizing a case study from a radiotherapy trial as an example of how this can be of value.

12.3.1.1 Benefits of Randomized Controlled Trials

RCTs are a research design under the positivist research paradigm. For example, there is an emphasis on neutrality with an attempt to keep researcher and research participant's remote from each other to avoid any influence on the study results. Characteristically RCTs seek to explain the whole by a study of one aspect or parts. RCTs are based on a science model in which there is a belief in universal laws measuring and analysing relationships using numbers to quantify effects or behaviour. Objectivity is a primary aim and specific aspects of the approach are designed to provide neutrality and to avoid personal biases. Control over potential biases or confounding variables is integral to this approach. Owing to the strict controls and statistical strengths of RCTs, this design sits high within the hierarchy of evidence. Table 12.1 shows the Scottish Intercollegiate Guidelines Network (SIGN) hierarchy of evidence. It is clearly demonstrated in Table 12.1 that studies where there is a high risk of bias are given a lower ranking than similar designs where bias is deemed low [20].

Why Are RCTs So Useful?

Consider the following scenario.

Post-operative radiotherapy for breast cancer is the accepted treatment for the majority of women following surgery. Radiation treatment to the breast can lead to a mild skin reaction (erythema), reactions usually start in the second week of treatment and increase as the treatment course progresses. Traditionally skin care advice

Table 12.1 Levels of evidence from the Scottish Intercollegiate Guidelines Network

1++	High quality meta-analyses, systematic reviews of RCTs or RCTs with a very low risk of bias
1+	Well conducted meta-analyses, systematic reviews or RCTs with a low risk of bias
1-	Meta-analyses, systematic reviews or RCTs with a high risk of bias
2++	High quality systematic reviews of case control or cohort studies High quality case control or cohort studies with a very low risk of confounding or bias and a high probability that the relationship is causal
2+	Well conducted case control or cohort studies with a low risk of confounding or bias and a moderate probability that the relationship is causal
2-	Case control or cohort studies with a high risk of confounding or bias and a significant risk that the relationship is not causal
3	Non-analytic studies, e.g., case reports, case series
4	Expert opinion

has been to undertake a variety of practices with limited evidence base to support the advice given which may include the below three.

1. Washing with mild soap
2. Using mild creams in the treated area
3. Allowing air to get to the skin

Recently, barrier dressings have been developed to try to reduce the impact of friction from clothing exacerbating skin reactions and to reduce radiation induced erythema [21].

If investigators wanted to study the benefits of using a barrier film on radiotherapy patients to identify its usefulness in preventing erythema they might formulate a research question as follows.

> Does the use of a barrier film on the irradiated skin during breast cancer irradiation reduce the skin reactions experienced by patients?

Here it would be useful to take a few minutes to consider some of the different research approaches that can be used to answer this question. For the research question above, what would be the strengths and limitations of the research methods posed in Box 12.1

Box 12.1 Different Approaches That Can Be Used for Research on the Use of a Barrier Film on Irradiated Skin

Method 1

A prospective evaluation of skin reactions on all patients irradiated for breast cancer using a barrier film on the affected skin during treatment.

Method 2

A prospective evaluation of skin reactions on all patients irradiated for breast cancer using a barrier film on the affected skin during treatment compared with the results of a previous study to evaluate skin reactions in irradiated patients with breast cancer using conventional skin care instructions.

Method 3

A prospective evaluation of skin reactions on all patients irradiated for breast cancer using a barrier film on the affected skin during treatment, compared with a control group of irradiated patients who are given the conventional skin care instructions. Patients can opt for either the current skin care approach or the barrier film intervention.

Method 4

A prospective evaluation of skin reactions on all patients irradiated for breast cancer using a barrier film on one half of the affected skin during treatment, the other half of the breast or chest wall, patients use the conventional skin care instructions (no barrier film).

As method 1 has no comparison group it is not possible to place the results in any context, so we would still be unsure which skin care regimen was most effective.

In method 2 a comparison group is available to provide some way of assessing the performance of the new intervention. However, using a historical control group as the comparator has a number of problems and would mean any results obtained could be viewed as unsound. For example, if we assume the results identified a statistically significant reduction in erythema in patients using the barrier film, it is possible that this result may have occurred not because of the intervention but due to other extraneous factors including the following.

1. A difference in patient characteristics between the two study groups. If the historical control group contained a higher proportion of patients with larger breasts than in the barrier film group it is possible that this might account for the difference in skin reactions seen as it is known that breast size has an influence on subsequent adverse events [22].
2. Technological differences between the two periods of study. As time passes changes in technology may mean application of treatment is no longer the same. The introduction of a new planning technique or a change to the immobilization device between the two data collection periods could account for differences in skin reactions observed.

In method 3 the use of a comparison group treated in parallel with the intervention group eliminates potential confounding variables associated with a historical control group. However, patients choosing between treatments could mean that patient numbers might be unbalanced between the two skin care interventions and it is likely that patient characteristics would be unbalanced between the two study arms. Furthermore, where there is the option for choice it is possible that any patient reports of symptoms may be underplayed, especially where patients have read favourable information about a specific intervention, e.g., the benefits of using a barrier film, again limiting any confidence the researchers can have in the results obtained.

In method 4 the researchers would need to take care to make sure the radiation dose received to the skin underneath the barrier film was the same as that received by skin in the section of the breast not covered by the barrier film. They would also need to take care that the skin care used on the breast tissue not covered by the barrier film did not itself cause an increased irritation. For example, some topical creams can cause dryness or irritation that may exacerbate any radiation skin reaction.

Using the above scenario it is possible to see the need for strict control of possible confounding variables as well as the benefits of blinding participants to the intervention, and the use of methods to ensure a balance of patient characteristics between the intervention and control arms. RCTs allow rigorous evaluation of a single variable in a defined patient group. Within the RCT design it is possible to eradicate potential bias by comparing two or more groups with balance in patient characteristics. Where RCTs are used this also allows the opportunity for

meta-analysis comparing studies of the same investigation across different populations or geographical areas to provide a larger overall sample size and a potentially powerful analysis (see later in this section). In the next section the specific design characteristics of RCTs will be presented and some of the terminology associated with RCT design will be explained so practitioners can evaluate different RCTs presented in the literature.

12.3.1.2 Design Characteristics of RCTs

As the name indicates, RCTs involve random allocation of participants to treatment or control groups. Both groups are generally followed for a specific period and measurements taken at the same time points for both groups. The groups are analysed in terms of an outcome that is defined at the outset. For example, in the previous scenario a patient's skin reactions may be measured using a standard skin toxicity score such as the Radiation Induced Skin Reaction Assessment Scale (RISRAS) [23, 24] or the Radiation Therapy Oncology Group (RTOG) [25] scoring system at specific points throughout the treatment course. A pre-treatment (baseline) assessment of skin colouration should be undertaken to ensure patients do not have erythema, perhaps associated with sun exposure, prior to the start of radiotherapy that would alter any post-treatment results. This baseline measure would also be used to ensure parity between the two groups at the outset. Measurements may be taken weekly during the course of radiotherapy and also at 2 weeks post irradiation when skin reactions may be at their peak. The timing of outcome measurements is crucial to the accuracy of the study and thought needs to be given to this aspect of the study design.

Within the RCT design controlling bias is a main focus so researchers need to consider any potential confounding variables that may influence the outcome and control for these within the analysis. For example, using the skin study scenario we have already identified that patient size can influence the skin reactions experienced so it would be important to record patient size, either chest separation or breast volume, at the outset and test the two treatment arms for equality of this characteristic. Researchers would need to consider all possible confounding variables so other factors may include the level of homogeneity of the dose distribution [26] within the planning target volume (PTV). In the next few sections we will consider in a little more detail some of the specific design characteristics of RCTs.

Types of RCTs

RCTs are often defined by:

- the purpose of the study, i.e., explanatory, efficacy or pragmatic trials
- how participants are exposed to the intervention, i.e., parallel, cross-over or factorial designs
- number of participants
- how the intervention is assessed [27].

When assessing health technology, RCTs are usually pragmatic trials where the study is designed to reflect normal clinical activities. The aim of a pragmatic trial is to determine if the intervention works but also to describe any consequences of implementation of the technology. Pragmatic trials often have wider inclusion criteria to ensure the sample studied represents the normal group of patients that are likely to be seen in everyday practice. The comparison group in a pragmatic trial is often the current treatment or current imaging technique. Effectiveness trials aim to assess whether an intervention works in people who are offered the intervention. They tend to be pragmatic studies as the aim is to assess the effects under normal daily practice. They have simpler designs with less strict inclusion criteria than efficacy studies allowing participants to accept or reject the intervention offered. An example of an effectiveness study would be the early evaluation of breast cancer screening where RCTs were used to identify the impact of a screening intervention. Patients would be called for screening but may opt not to attend. Follow-up of this arm would include all patients offered screening irrespective of whether they attended the screen or not and compared with patients in a control group (who were not offered any intervention) [28, 29]. An efficacy study is where the aim is to identify if an intervention works in those that receive it. Figure 12.4. shows a pictorial presentation of two basic RCT designs [27].

In its simplest form an RCT has two arms, an intervention arm, that may be a new process or technology being tested, compared with either a control arm, that receives no intervention, or a second intervention arm, which in HTA is usually the current treatment or current imaging modality. Cross-over designs can be a powerful way to study the impact of a new technology (see design (**b**) in Fig. 12.4). Here patients or subjects are used as their own control and this avoids the need for matching

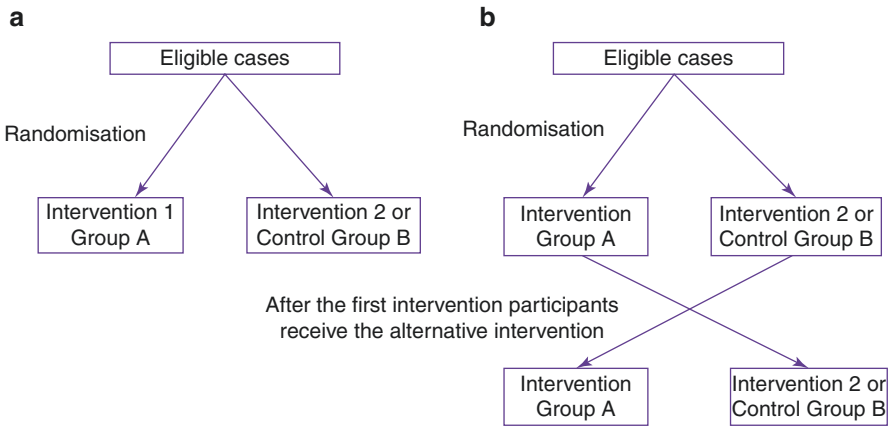


Fig. 12.4 Diagrammatic presentation of simple randomization and cross-over randomized controlled trial designs. (**a**) Simple randomization, (**b**) Cross-over or within-subjects design where participants receive both interventions in a different order (or the intervention and the control depending on the design)

characteristics across groups with different subjects that occur with the simple parallel design. However, cross-over or repeated measures designs can only be used in HTA where the first intervention has no lasting effect on the primary outcome measure. So, in the scenario used above it would not be appropriate to use a barrier film for the first two weeks of treatment and then apply traditional skin care for the remainder of the treatment course as the effect of the first skin care regimen would impact on subsequent skin reactions measured during application of the second regimen. In educational studies this design could not be used where subjects would learn through the first phase of the study. For example, if you wanted to test the effectiveness of two formats of patient information on a patient's ability to perform a breath hold technique during radiotherapy planning, you could use either of the below.

- A patient education video versus
- A traditional written information leaflet

The participant's ability to hold their breath following exposure to the video would influence subsequent performance so it would be difficult to distinguish if the video or the pamphlet had the impact on overall ability to perform the technique.

However, this cross-over design can be used successfully when used with consideration for potential learning effects. For example, it has been used to assess the impact of work speed on the accuracy of setting up a patient for a complex technique using a phantom [30]. In this study each pair of staff was asked to set up the phantom as they would for a normal treatment, twice, under two different conditions. In condition 1, participants were given a scenario whereby they could take as much time as they needed and a radiographic image was taken of the final setup position to assess positional accuracy. In condition 2, participants were given the same technique to apply but the scenario was that they were treating a child that was distressed and so it was important to work fast but accurately, in order to assess the impact that a time pressure might have on treatment accuracy. It was important that groups alternated in the order in which they undertook the test, i.e., condition 1, then condition 2, for one group and condition 2, then condition 1 for another group, to ensure if there were any learning effects these would not affect the overall results.

Factorial trials offer the opportunity to test individual interventions as well as studying the impact of two or more interventions applied together. In imaging the factorial design has been used in experimental conditions to test the factors that influence image quality and radiation dose [31].

In addition, trials can be described as being single-blind, double-blind or triple-blind. Blinding refers to either the participants being blind to the intervention, i.e., they are unaware of which intervention they have been allocated to, or the investigators, such as the statistician. The purpose of blinding is to minimize opportunities for bias as a direct result of knowledge of the intervention received either by the participants or the investigators applying the treatment or collecting the data.

For example, where the participants have knowledge of the intervention they are to receive, there is a possibility any patient self-reports may be influenced by this knowledge. Where possible patients should be blind to the intervention; this is not always possible as it may often be obvious which of the interventions participants

have received. For example, in the skin care example above patients that are randomized to receive the barrier film will know that is the group they have been allocated, it is impossible to blind the patients in this scenario. Further opportunities for bias can occur during the assessment of the study outcomes. Researcher knowledge of the intervention arm can influence interpretation of key outcome measures, especially where the researcher has a hypothesis to test. For example, in a study to evaluate the effectiveness of using tattoos to improve radiotherapy treatment accuracy during breast irradiation compared with gentian violet pen for marking the skin, it was necessary to blind the researcher undertaking the analysis to the intervention each subject was allocated to during the measurement of the treatment images that were used to establish treatment accuracy [32]. Knowledge of the intervention could have resulted in the favourable measurement of some images in order to prove the hypothesis being tested. To completely reduce the opportunity for any bias researchers, where possible, should aim to blind the patient, the researchers undertaking measurement of the outcomes, and the researchers undertaking the statistical analysis (i.e., single-, double- or triple-blinding) [27]. Treatment effects may be overestimated by approximately 17% where double-blinding is not employed as compared with studies where double-blinding is used [18]. The impact of not blinding patients has been shown to have a significant impact on patient reported outcomes, with non-blinded patients giving more optimistic reports of the intervention (exaggerated in the region of 0.56 SDs) [19]. Where studies are reporting true intervention effect sizes this could mean exaggeration of effect by over 100% [19].

Randomization

The rationale for using randomization is to prevent bias occurring as a result of inequalities between the treatment options or intervention arms. For example, when looking at the effectiveness of breast cancer screening, it would be important for researchers to ensure equity of characteristics between the screening group and the control arm, such as age at time of entry into the study, as incidence of breast cancer is known to increase with age [33].

There are a number of methods available to researchers for achieving random allocation. The simplest way is by tossing a coin, throwing a dice or use of a table of random numbers. For example, it can be agreed at the start of the study that heads on a coin will indicate treatment arm A and tails treatment arm B or the control group. However, simple randomization methods such as this may still result in unequal numbers or unbalanced characteristics between the groups [27], especially in small trials [34]. To overcome this one method is the use of block randomization. Generally, blocks of four are used for a simple RCT design with two intervention arms, A and B, as follows: AABB, ABBA, BBAA, BABA, BAAB, and ABAB. One of the six possible combinations is selected and participants allocated to an intervention arm based on the sequence of four; the process is repeated as required depending on the sample size.

Even with block randomization some inequalities may still arise simply by chance, hence researchers need to be aware of this possibility and test baseline characteristics between the groups for equality. Where differences occur, it may be necessary to control for imbalances in subsequent analyses of the outcome data.

Alternatively, stratifying randomization by an important characteristic may reduce the potential for inequality. When considering the study to look at the impact of a barrier dressing to reduce skin reactions during breast cancer radiotherapy discussed above, it may help to stratify on patient size, i.e., large or small patients, as this is a contributing factor for skin reactions during breast irradiation. In this case a separate list of block sequences would be produced for each stratum, although as you increase the number of strata the risk of errors in application also increases [34]. A further alternative is the use of a technique called minimization. This method is successful at obtaining equality between groups for a set of relevant characteristics even in trials with small samples [34]. Here for the characteristics that require balance, e.g., age, patient size, menopausal status, etc., a running total of how many participants have been allocated with each characteristic to each intervention arm is kept. Following random allocation of the first participant, subsequent participant randomizations are weighted to the intervention arm that would maximize balance, i.e., minimize inequalities, with totals for each arm updated after each participant is entered into the study.

A further option for researchers is the use of cluster randomization. In contrast to most randomized trials where the individual is randomized, with cluster randomization groups of participants are randomized [27, 35]; clusters can be either general practitioner practices or imaging/oncology departments. The benefit of cluster randomization is a possible reduction in contamination of the control arm. For example, if you wanted to investigate the impact of a new electronic information service for patients, it is possible that those in the experimental arm might pass on to patients in the control arm, simply by chatting while in the waiting room, useful information they have gleaned as a result of the intervention. Cluster randomization may not be necessary for the majority of trial designs and therefore individual randomization should be used where possible to avoid some of the limitations of cluster randomization (see Box 12.2 for details [35]).

Box 12.2 Limitations of Cluster Randomization

1. Selection bias—different types of participants maybe recruited into different arms of the study due to the geographical locations of the clusters which may result in differences, for example, in socio-economic status between arms.
2. Selection bias—in cluster trials participants are not asked to consent to the study but to consent to being included in the study analysis; if a substantial proportion of the cluster participants refuse, then an imbalance will occur between the trial arms.

Cluster trials need larger sample sizes than trials that use individual randomization to ensure sufficient statistical power. If there is not full uptake of the intervention within the cluster, then a dilution effect may further influence the power of the study.

Concealment of Randomization

Randomization is generally accepted as the best way of removing opportunities for selection bias by removing any predictability in the assignment process. Yet the process of randomization itself can be fraught with opportunities for bias that may invalidate or reduce the quality of the subsequent results. A common approach adopted by novice researchers to the issue of randomization is to alternate participants to interventions as they are referred to the clinic or department, as they consider referral to be in itself, a random process (Table 12.2).

Looking at the process in Table 12.2, can you foresee any problems with this approach? Primarily there is an identifiable pattern that may introduce bias. For example, where the pattern is known, there is the opportunity for researchers to selectively change the detail of the information given to potential participants. This is done to discourage entry into the trial where that patient has co-morbid disease or any potential characteristic that the researcher considers may influence or skew the results in an unfavourable direction. Inadequate concealment of this nature can result in overestimation of the potential effect of the intervention in the order of 40% when compared with trials with adequate concealment of randomization [18].

One method used to reduce the opportunity for bias during randomization is to use sealed opaque envelopes containing random allocations. However, this system may be prone to interference. Clinicians can open envelopes in advance, or view allocations by holding the envelope up to a bright light. Block randomization of four is used, if three of the previous participant allocations are known, the fourth can be predicted allowing the clinician to reserve entering patients into a trial until specific participants present with desired characteristics. Subversion of allocation concealment has also been shown in one study to have a significant impact on the age of patients enrolled in the experimental arm compared with the control arm [36]; the median age in the experimental arm was 59 years compared with a median age of 63 years in the control arm when a sealed envelope system was used. For lone researchers undertaking a simple RCT as part of perhaps an undergraduate or postgraduate course of study the use of sealed opaque envelopes may be the only practical solution on offer; in these circumstances researchers should be aware of the potential for interference and subsequent effects on the study quality. In most cases attempts should be made to use a system that removes the randomization process from the researchers, such as a central randomization service available through local trials units [37].

Table 12.2 One method sometimes used by students or novice researchers to randomize participants

Participant number	Allocation
1	Intervention A
2	Intervention B
3	Intervention A
4	Intervention B
5	Intervention A

Sample Size Requirements

As well as randomization of patients into the control or intervention arms, RCTs rely on statistical analysis of the primary outcome to demonstrate effectiveness of the intervention. In order to demonstrate a statistically significant difference in treatments between the study groups it is important that an adequate sample is studied to demonstrate an effect. In HTA, improvements in outcomes may be small and therefore where studies have small sample sizes it may not be possible to demonstrate a difference even where a difference exists [38]. For this reason, researchers undertaking RCTs must consider at the outset what improvement in the primary outcome would be appropriate for a clinically significant improvement or benefit and then calculate the sample size required to establish this statistically. This calculation is referred to as a power calculation. For example, in a study to establish the effectiveness of a radiotherapy protocol to reduce lung morbidity for patients undergoing breast or chest wall irradiation following surgery for breast cancer, it was calculated that a sample of 200 patients in each group would be required to detect a difference of 0.3 (in the primary outcome measure) with 5% significance and 80% power [39] (see Chap. 10 for more information on power calculations).

Recruitment of Subjects

Recruitment of patients into clinical trials is often problematic. In the study of the effectiveness of a radiotherapy protocol to reduce patient reports of lung morbidity mentioned above [39], recruitment of subjects to the study was slow despite a feasibility study indicating sufficient eligible patients were available in the host centre. Recruitment was hampered by:

- clinicians forgetting to mention the study to eligible patients,
- patients refusing to participate partly due to poor information about the possible side effects of treatment at the early referral stage. Within the patient information sheet for the study details of lung morbidity were highlighted, and patients unknowing of this aspect of their treatment feared that inclusion in the study would cause unwanted respiratory side effects, even though this was a possible corollary of treatment regardless of inclusion in the study,
- limited patient awareness of clinical trials during the early stages of the study,
- a strong preference for one of the intervention arms with patients not wishing to take a chance of receiving the alternative option through randomization.

Of 452 patients assessed as eligible for inclusion in the study, 92 (20%) refused to participate [39], which is similar to reports from other cancer trials [40]. As well as an effect on the overall sample size, this loss of potential participants can have an effect on the generalizability of the results as the sample recruited may not fully represent the population of patients as intended. Where studies include a placebo arm it is possible that a reduction in acceptance to randomization may also occur [41]. Generally factors reported as influential in a patient's decision to join a study include the belief that they may help future patients, or that they may

benefit from inclusion [41]; hence researchers should ensure potential participants are aware of the benefits of the study during the recruitment stage. In many cancer trials a lack of participants can be reflective of strict inclusion criteria excluding a substantial proportion of patients, perhaps in the region of 30% [40, 42]. Hence more pragmatic trials with less strict inclusion criteria may enhance the proportion of patients eligible for study and thus increase the potential for recruitment and generalizability [40, 43].

A comparison study of two community-based RCTs undertaking similar palliative care interventions identified a number of positive recruitment strategies. The more successful of the two trials, studied in terms of reaching an adequate sample size, employed the following strategies to maximize recruitment [43]:

- use of an inflated sample size to account for expected high attrition from early withdrawal or death
- maximal inclusion criteria and minimal exclusion criteria
- dedicated recruitment nurse
- triage process to screen for eligible patients
- recruitment interview included key messages
- patients approached for consent before GP consent was requested
- extensive marketing to raise the profile of the study topic
- effort was placed on ensuring clinician input to the study to encourage feelings of inclusion and reduce concerns
- realistic timeframe to recruit sufficient sample size
- adequate funding to support an extensive recruitment strategy.

Other strategies that have been shown to have a beneficial effect is telephone reminders to non-responders [44]. Recruitment may be hampered where potential participants or referring clinicians have strong preferences for one of the intervention arms that leads to a refusal to be randomized. Again, where these patients refuse consent to randomization, a reduction in generalizability of the results may be a consequence. Furthermore, where patients with a strong preference accept being randomized, subsequent results may be biased by strong beliefs about the treatment received where blinding of the patient is not possible [45]. A solution to this dilemma is the use of patient preference trials and there are a number of different designs currently being used (Fig. 12.5) [27, 45, 46].

While patient preference designs may allow a greater proportion of patients to be included in a study, the disadvantage of such designs is the resultant unknown or uncontrolled confounding variables in the preference arms [45]. It is suggested that the analysis for these studies includes comparison of the two randomized arms alone and perhaps an analysis using randomization status as a co-variate [45]. A concern of using the Zelen design, where participants are randomized before giving consent, and where those randomized to the standard treatment only consent to treatment and not to participation in a study, is a possible ethical implication in therapeutic scenarios [46]. However, it has been suggested that this design is specifically helpful for population-based screening studies [46].

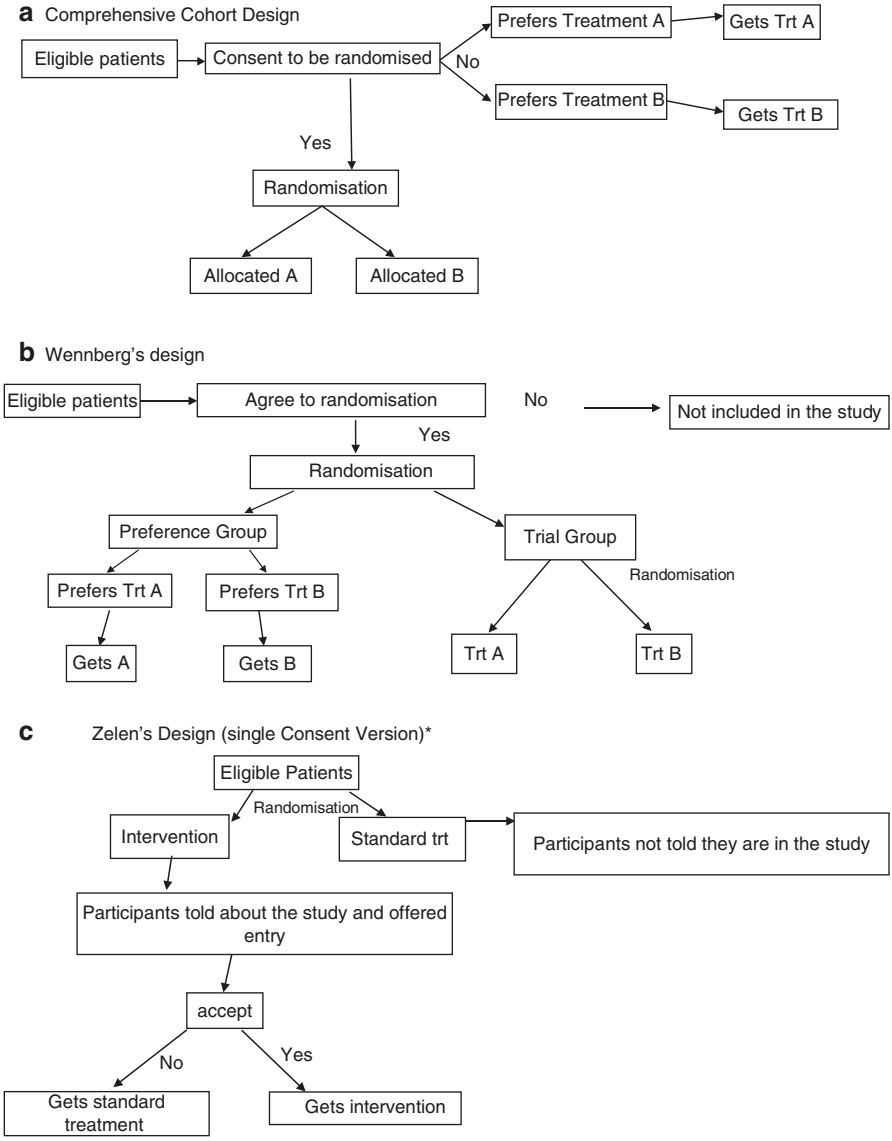


Fig. 12.5 Patient preference study designs [37, 46]. (**a**) Comprehensive cohort design. (**b**) Wennberg's design. (**c**) Zelen's design (single Consent Version)∗. ∗ In the double consent version participants are told which intervention they have been randomized to and offered the opportunity to switch to the alternative treatment [46]

Attrition

Even when researchers manage to recruit sufficient numbers to their trials problems with attrition can lead to a reduction in the strength of the reported findings. It is common for participants to fail to completely finish the allocated treatment or

intervention for a number of reasons: the patient may move to a different geographical area, the intervention may cause adverse side effects and the participant opts to withdraw leaving an incomplete data set. In addition, there may be missing data as a result of incomplete collection, perhaps due to staff absence at the time of collection, participants not adhering to the protocol or staff failing to record the information on the correct forms.

Attrition through a loss of patients to follow-up or incomplete data sets can bias results when the characteristics of those with missing data differ between the randomized groups [47]. It is therefore advised that missing data be presented by researchers in publications by providing normal baseline characteristics data for the whole study sample, and also separate data on those lost to follow-up from those remaining in the analysis so that readers can judge any imbalances between the intervention arms as a result of the missing data [47]. Using strategies to minimize attrition is beneficial and these may include minimizing patient burden by attention to the data collection methods [43]. For example, reducing the need for patients to attend clinics by visiting them at home may increase cooperation and reduce missing data; although more costly than other approaches this method was successful in the comparison of two community-based palliative care trials [43]. Ensuring all staff involved in data collection are fully informed and included in the trial process may ensure adequate data recording and protocol compliance. In addition, regular assessment of data accrual may highlight the need for a change in strategy where rising missing data becomes apparent.

12.3.1.3 Protocol Deviations

In circumstances where there have been protocol deviations it is appropriate to use an 'intention to treat' (ITT) analysis where participants are analysed as part of the group they were assigned to irrespective of whether they competed their allocated treatment/intervention or not [27, 48]. The ITT analysis should be applied to a full data set [49] but frequently protocol deviations are accompanied by missing data. Failure to include participants with missing data can result in an overestimation of the benefit of the intervention [27]. Consider the barrier film example used previously, if patients stopped using the barrier film because of exacerbation of the skin reaction, or left the study due to adverse reactions, this would result in missing data for some patients. If these data are excluded from the analysis, the assumed benefit of this product may be exaggerated [27].

When considering missing data it may be appropriate to use a 'sensitivity analysis' [27] or imputation [50]. Here it is proposed that either a worst case scenario is used, or a value is chosen that is credible given the rest of the patient's data set [50]. Sometimes it may be appropriate to use the last recorded response or to assume that responses remained constant [49]. For example, in the study of the effectiveness of a radiotherapy protocol to reduce patient reports of lung morbidity mentioned previously, 'no symptoms' were used for missing data in both groups as this was a plausible outcome given the rest of the data set [39]. However, imputations of this nature are provided to give some estimation of treatment effect and should be considered carefully; producing a range of potential outcomes for readers using different imputation methods may be the most beneficial policy [48].

It is suggested that the use of the ITT analysis is the most cautious approach to take when handling protocol deviations [51]. However, it is proposed that using an ITT approach can lead to type II errors and there may be justified circumstances when patients with specific criteria could be excluded from the analysis [51]. These would include participants that were randomized for inclusion in a study but who were in fact ineligible, i.e., they did not meet the eligibility criteria for the study [51]. Even in these cases it is prudent to consider individual exclusions with care. In addition, the ITT approach is appropriate for effectiveness or pragmatic studies where the aim of the study is to evaluate the impact of an intervention under normal clinical circumstances where it is likely that some deviations from protocol would also occur [49]. Hollis and Campbell suggest a strategy for the full implementation of the ITT approach that researchers designing RCTs may find helpful [49].

When evaluating RCTs presented in the literature it is important to consider how protocol deviations were handled by the researchers. In a study of RCTs published in a number of high impact journals, Hollis and Campbell discovered only 50% of RCTs published in 1 year stated explicitly that results were analysed on an ITT basis [49]. Of those stating they used an ITT analysis 13% did not actually analyse patients as randomized (which is the criteria for the ITT approach). Furthermore, the handling of missing data was variable across the studies, emphasizing the need for practitioners to undertake rigorous appraisals of published RCT study results before considering applying the evidence to practice.

Drug Trials

When a new drug is developed the development process can be time-consuming. Initially, safety and efficacy of the drug will be tested through animal studies. The first human studies of cancer drugs (Phase I trials) are usually tested on volunteers. There is no randomization and incremental doses of the drug are administered so that side effects can be monitored. Once the safety of the drug has been established in humans the drug can then be administered to a small group of patients (approximately 20) with the condition to establish efficacy, i.e., where the aim is to establish if the drug works in people who receive it, with different doses and frequencies. There are very strict inclusion criteria to exclude patients with coexisting disease. These Phase II studies may involve randomization if the outcome measure is appropriate, i.e., pain. Where the end-point is reduction in number of deaths there may not be randomization. Phase III trials are conducted once the drug has been shown to be effective and safe in Phase II studies. Usually Phase III trials are randomized effectiveness studies.

In the cancer field a prominent and well publicized cancer drug trial was an evaluation of the effectiveness of trastuzumab (Herceptin®) for the adjuvant treatment of early breast cancer in HER2-positive cases. This monoclonal antibody against HER2 had proven efficacy in advanced breast cancer and the first interim analysis to be published in early breast cancer trials showed such promising results [52] that there was a desire for clinicians to consider its use in HER2-positive patients with early stages of the disease. The primary outcome in this study was disease free survival with early results showing 92.5% of patients in the

trastuzumab arm free from disease at year 1; as compared with 87.1% in the control arm [53]. These results led to acceptance of the drug for treatment in early stage cancer by the National Institute for Clinical Excellence (NICE) in the UK despite a relatively short follow-up period (median follow-up 2 years) [54].

12.4 Health Economic Assessment and RCTs

Economic assessments in conjunction with RCTs have become increasingly important due to the need to allocate scarce health resources in the most efficient and beneficial way. Economic evaluations deal with both costs and outcomes of activities and the basic purpose of an economic evaluation is to 'identify, measure, value and compare the costs and consequences of the alternatives being considered' [55]. Economic evaluations are comparable in the way they measure costs but differ in the way outcomes or consequences are derived. Essentially evaluations can be divided into three main types [56].

- cost-benefit analysis
- cost-utility analysis
- cost-effectiveness analysis.

Cost-benefit analysis involves the measurement of costs and benefits in comparable monetary terms. An example of the use of a cost-benefit analysis is the evaluation of an intensive follow-up regimen for patients diagnosed with breast cancer. This involves oral history, physical examination, blood tests including biological markers, annual hepatic echography, chest X-ray and a bone scan as compared with a standard clinical follow-up in breast cancer patients to identify early signs of relapse [57]. In this study the authors undertook a simple RCT comparing the two follow-up methods for number of relapses identified during scheduled follow-up appointments. The results identified no difference in the early detection of relapse between the two methods, so no benefit cost but a substantial increase in costs for the intensive follow-up schedule that was three times the cost of the less intensive follow-up regimen [57].

Cost-utility analysis involves the use of a utility based measure such as quality adjusted life years (QALYs). By using a single measure of benefit (QALYs) across RCTs, it is possible to compare the effectiveness of different interventions and hence this type of analysis allows the assessment of the benefit of employing a particular treatment or intervention in one area against the loss in benefit caused by redirecting resources from other programmes, i.e., productive efficiency and allocative efficiency, and is considered as a variation of cost-effectiveness analysis [56].

Cost-effectiveness analysis measures outcomes or benefits in units such as quality of life or improvements in function; in radiotherapy this may be measured as improvements in accuracy of treatment. To illustrate how an economic evaluation can be undertaken, consider the work by Shah et al. [58]. This work compares the cost-effectiveness of:

- standard whole breast irradiation (where patients are treated in 15 treatments with or without a boost at the end of treatment) and
- accelerated partial breast irradiation (APBI, where patients receive 5 treatments in total over 10 days).

The aim of this study was to identify the cost and cost-effectiveness of APBI compared with the current standard whole breast irradiation treatment protocol. The cost-effectiveness evaluation was undertaken from both the health care system perspective and also the societal perspective. A healthcare perspective includes direct costs for staffing, and equipment. Individual staffing costs for each patient attendance can be calculated based on the procedure time and the pay rate for the highest staff grade performing the procedure. In the study by Shah et al. [58] a breakdown of the direct costs is presented in a supplementary file on the journal web site, this is helpful to understand the breakdown of costs and where these differ between the two different approaches. A societal perspective takes in to account the impact on the patient of the treatment regime. In this study the authors calculated lost work time and parking costs for attendance at the hospital for treatments and appointments. Effectiveness in this study was determined by QALYs. Table 12.3 shows the final cost analysis. It can be seen from the table that the effectiveness of the two approaches is similar but the costs (both direct and with indirect costs considered) favour the APBI technique. It is the individual cost per treatment that influences the overall cost-effectiveness outcome in this case; the cost per treatment fraction is lower for whole breast irradiation but as there are 15 treatments compared with only

Table 12.3 Direct and indirect costs from a cost-effectiveness analysis assessing whole breast irradiation versus APBI (reproduced from Shah et al. [58] with permission)

Treatment	Cost	Incremental cost	Effectiveness (QALYs)	Incremental effectiveness (QALYs)	Incremental cost-effectiveness ratio
Direct cost only (in US dollars)					
APBI	2966	–	0.2300	–	
Whole breast irradiation (without boost)	3666	700	0.2289	−0.0011	Dominated
Whole breast irradiation (with boost)	4551	1585	0.2289	−0.001	Dominated
Direct and indirect costs (in US dollars)					
APBI	3569	–	0.2300	–	
Whole breast irradiation (without boost)	4940	1371	0.2289	−0.0011	Dominated
Whole breast irradiation (with boost)	6160	2591	0.2289	−0.0011	Dominated

APBI accelerated partial breast irradiation, *QALYs* quality adjusted life years

5 treatments in the APBI approach the overall cost is higher for the whole breast irradiation technique.

12.5 Systematic Reviews and Meta-Analyses of RCTs

During clinical activities practitioners may come across questions about practice that they do not know the answer to. They may choose to ask an expert who may or may not know the answer; or they may turn to the published literature for an answer. In Chaps. 3 and 4, literature reviews were discussed and the method for searching for literature was presented as an important aspect of the research process. This section focuses on the method for undertaking a systematic review of published literature in relation to HTA: it starts with a discussion of the differences between discussion papers (or narratives), systematic reviews and meta-analyses. Following clarification of the different types of reviews the discussion concentrates on the method for undertaking systematic reviews with particular attention paid to the review process and aspects of the search strategy including the assessment of study quality. The final subsections describes the common principles of meta-analyses and standards required for the presentation of systematic reviews.

12.5.1 Types of Reviews

Literature reviews or discussion papers found in journals are an informal collection of literature on a specific topic and are often invited papers from experts in the field. They are common in journals as they are easy to read and synthesize by practitioners and are often quick to produce. One of the main disadvantages is the variability in the level of detail that is presented about the search strategy employed, making replication of the review difficult. In addition, they may lack rigour and objectivity, with conclusions and recommendations based on a narrow examination of the available data. However, they can provide an opportunity for debate and allow the authors to provide an interesting perspective on a topic of current interest.

A systematic review is a formal review of the evidence on a particular topic with a specific research question that is to be addressed and a detailed search strategy that would allow replication. The search strategy includes details about inclusion and exclusion criteria, databases used and the method used to assess the quality of the studies identified by the search, the process for selecting research and the method used for data extraction and synthesis. There is also an attempt to reduce potential bias by using standardized tools for the assessment of study quality as well as using more than one assessor to evaluate selected studies and blinding of reviewers to the authors and journal names of selected studies.

A meta-analysis is a review where the results of RCTs undertaken independently are combined and a statistical analysis produced, usually graphically, to provide an estimate of the effect of an intervention. By combining a number of individual studies it is possible to essentially increase the overall sample size and hence increase

the strengths of the conclusions that can be drawn about an intervention, making meta-analyses a major asset for practitioners needing to make decisions about clinical interventions. However, meta-analyses do have some limitations and these are covered in more detail below.

12.5.1.1 Systematic Reviews

Planning a systematic review is crucial to its success and subsequent quality. Figure 12.6 provides a schematic presentation of the process required to plan and execute a systematic review.

Planning the Review

Before embarking on a systematic review, it is important to be clear about the clinical question that needs to be answered. The research question will be used to define facets of the search strategy and any lack of clarity may reduce the effectiveness of the search. In addition, before any detailed work is undertaken in preparation of the review it is important to identify if:

- a systematic review already exists on the topic area
- sufficient data are available to undertake a systematic review.

Therefore, being clear about the question and the topic of interest is important. Once this has been clarified it is beneficial to undertake a scoping exercise to identify how much literature exists in the field. This takes the form of a small search using the main electronic databases relevant to the topic area; for example, this might include MEDLINE, CINAHL and the Cochrane databases, using key terms. A simple search should allow the opportunity to identify whether any up-to-date systematic reviews already exist and indicate the amount of literature available to answer the proposed question [59]. Once the need for a systematic review in the field has been established a research proposal should be prepared. Box 12.3 highlights the key subheadings that practitioners may find useful to incorporate in a proposal for a systematic review [60].

Once the proposal has been written it may be helpful to gain an independent scientific review (ISR) of the proposal prior to a protocol being implemented, replicating the process undertaken for a primary study. Whereas in a primary study there is a need to gain the relevant research ethics and governance approvals, for systematic reviews there may not be such stringent requirements. However, gaining some peer review of the proposed work prior to the project being initiated is helpful for a number of reasons.

- Reviewers may identify additions to the search strategy that could improve the overall quality of the study.
- Poorly designed reviews will be ineffective and may produce results that are biased or inaccurate leading potentially to an inappropriate technology or treatment being implemented. ISR can identify potentially poor quality reviews and prevent resources being wasted on projects that may not be effective.

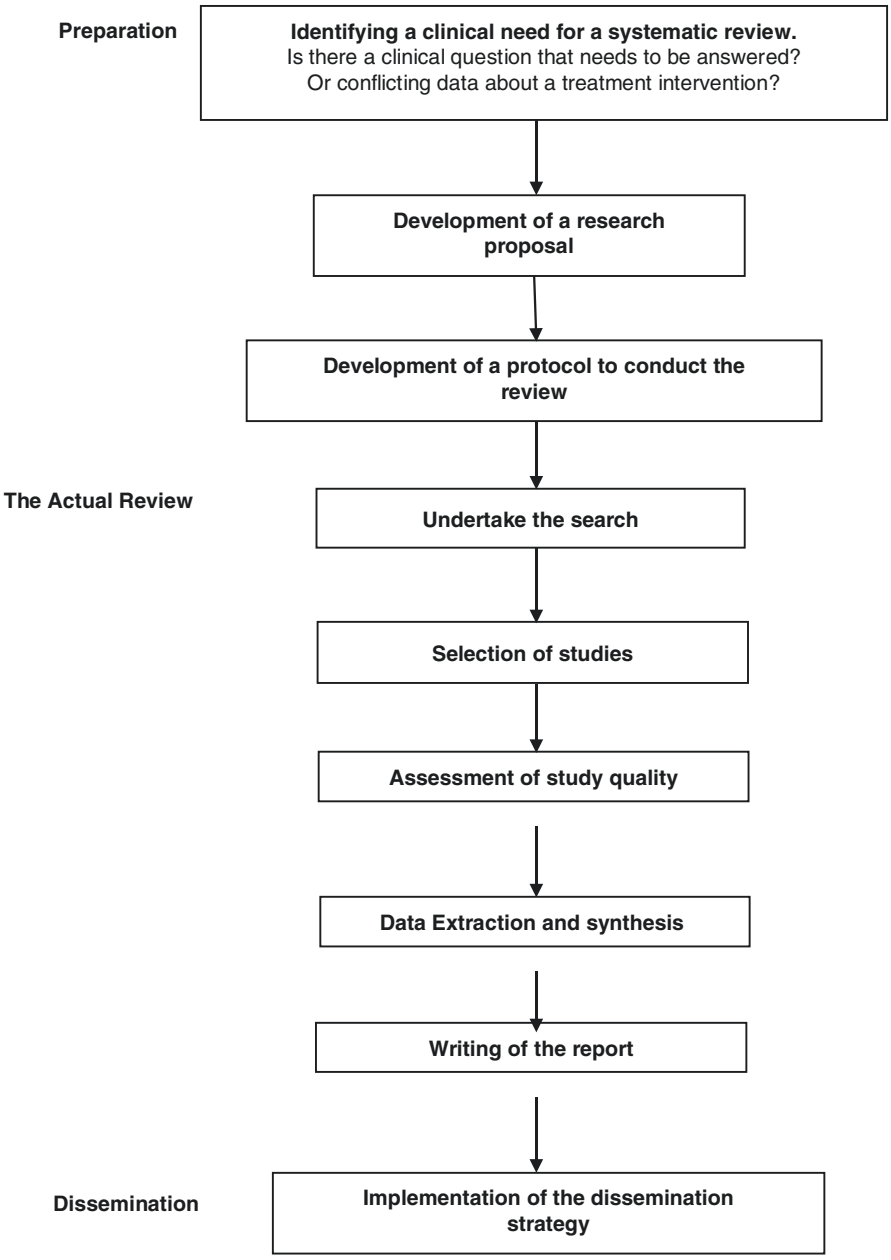


Fig. 12.6 The process for undertaking a systematic review

Box 12.3 Headings (and Content) for a Proposal for a Systematic Review

- Title.
- Summary—a brief synopsis of the aims of the review and the significance of the work will give readers an instant understanding of the importance of the proposed project.
- Aims—detail of the study aims and the research questions that the review is aiming to answer as well as the end-points for the study. End-points may include development of key research questions that remain unanswered and need further primary study or identification of a specific intervention to apply in practice.
- Background—this section should include a brief review of the literature to place the proposed work in some context, it might focus on the political, economic or social drivers for the project or give a historical perspective to current treatment or imaging rationales. Evidence identified from the scoping exercise may be beneficial in this section.
- Method—this section should include the search strategy, databases to be used, key terms, inclusion/ exclusion criteria, search limits, data extraction method, approach to be taken to quality assessment of the individual studies, how data will be synthesized (including information on any quantitative analysis), how reliability of the review will be determined and how bias will be minimized.
- Timeline—a detailed breakdown of the key milestones for the study.
- Project management—how the study will be managed and the key roles of members of the project team.
- Dissemination strategy—details of how the results will be disseminated should be multifaceted and practitioners may find it useful to consider how to measure impact, this work by Cruz Rivera et al. [71] helpful.
- Costs—identified costs including time for researchers undertaking the review, costs of searching databases (there may be a cost for access to some databases), costs of retrieving articles as well as costs to disseminate the results.

Funding bodies provide ISR during the application approval process but practitioners may wish to seek peer review prior to a funding application and this may be available locally through a university or via the local research and development department of the employing organization. Undergraduate and postgraduate students can use the experience of their supervisors to review the quality of their proposal.

The Search Strategy

Developing a multifaceted search strategy should ensure the review identifies as much of the available evidence as possible. The search strategy should detail the databases to be searched, the key terms for the search, inclusion and exclusion criteria and any limits placed on the search. Box 12.4 provides an example of a search

Box 12.4 A Sample Search Strategy

1. *Databases*
   a. MEDLINE
   b. CINHAL
   c. EMBASE
   d. Cochrane Reviews database,
   e. National Research Register including the ongoing reviews database (CRD Register of reviews)
   f. LILACS Latin American and Caribbean Literature in Health Sciences
   g. ISI Web of Knowledge to search Science Citation Index to follow citations from key papers.
   h. ScienceDirect to search for articles from journals not listed on MEDLINE
2. *Websites- to identify professional reports*
   a. National Institute for Clinical Excellence (NICE)
   b. National library for Health
   c. TRIP (Turning Research into Practice)
   d. Intute: Health and Life Sciences Medicine http://www.intute.ac.uk/healthandlifesciences/medicine/
   e. UK Society and College of Radiographers www.sor.org
   f. UK College of Radiologist www.rcr.ac.uk
3. *Key Journal hand Searches*: these will vary according to the topic area but common journals of relevance may include:
   a. *Radiation Therapy*
      i. Radiotherapy and Oncology
      ii. International Journal of Radiation Oncology Biology Physics
      iii. Journal of Radiotherapy in Practice
      iv. European Journal of Cancer
      v. Clinical Oncology
   b. *Imaging*
      i. British Journal of Radiology
      ii. Clinical Radiology
      iii. Radiology
   c. *Imaging and Radiation Therapy*
      i. Radiography
      ii. Journal of Medical Imaging and Radiation Sciences
4. *Author Searching*: searching databases by author may be beneficial where an author is known to publish or is a known expert in the topic area, this may be identified from literature retrieved in the original scoping exercise.
5. *Grey Literature*
   a. Index to Theses
   b. Index to Scientific and Technical Proceedings (via ISI web of knowledge)

c. Conference Papers Index
d. British Library Integrated Catalogue
e. COPAC–merged online catalogue of major university and national libraries in the UK and Ireland
f. Clinical trials databases here are the UK and US web addresses https://bepartofresearch.nihr.ac.uk https://clinicaltrials.gov

6. *Key Words*: for each facet of the research question key words and MEDLINE subject headings should be identified, for example, if a facet of the question included 'patients with cancer', then keywords might include:
   a. Carcinoma, tumour, tumour, cancer, invasive carcinoma
   b. MEDLINE subject heading—neoplasms
7. *Inclusion/Exclusion Criteria*: these may be specific to the topic area, for example, factors in a review to identify the effectiveness of partial breast irradiation inclusion criteria may be studies that consider external beam methods as well as brachytherapy (including balloon catheter methods). Alternatively, the focus for inclusion may be on the types of studies to be included. For example, in effectiveness reviews it may be relevant to include RCTs or quasi-experimental studies (trials without randomization).
8. *Search Limits*: For studies in HTA it is sensible to limit the review to data produced once the technology under question was implemented. For practical reasons undergraduate and postgraduate students often choose to limit studies to those published in the English language but possible bias needs to be considered where this is adopted.

strategy with common databases, websites and other strategies that may be useful for those working in imaging or radiation therapy. The Cochrane Collaboration provides a useful starting point for a search strategy along with the other major electronic databases. The Cochrane Library contains the Cochrane Database of Systematic Reviews (CDSR) and the Cochrane Central Register of Controlled Trials (CENTRAL). It may be beneficial to also search the Cochrane Methodology Register (CMR), the National Institute for Health Research (NIHR) Health Technology Assessment Database (HTA) and the NHS Economic Evaluation Database (NHS EED).

A detailed protocol using the main databases listed in Box 12.4 will go some way to helping retrieve as many of the relevant research studies as possible. However, in complex reviews it is possible that the protocol itself may only identify a proportion of the available data, and researchers should try to broaden their approach to include a range of strategies that develop as the review progresses.

For example, use of snowballing, which is using the reference lists of retrieved articles and forward tracking from a selected article to identify articles that have subsequently cited this paper—citation tracking, can increase the yield of relevant articles, and has been shown to account for approximately 53% of articles used in a

complex systematic review [61]. Other strategies to consider include using personal networks to contact individuals who may know of relevant research. This type of informal approach has been found to increase the proportion of relevant articles for a review by approximately 60% [61].

Searching the grey literature is also of importance as this may limit the effect of publication bias [62]. In a systematic review of studies including grey literature as well as published trials, it was identified that published trials tended to show a greater treatment effect than grey literature. This may be due to differences between published and unpublished trials such as sample size differences, and grey literature studies finding the intervention has no effect, which is a less interesting result and less likely to be published [62]. Grey literature refers to studies not yet formally published and may be found in conference proceedings, indexes to theses or on trial registers.

A common problem with using electronic databases as the primary search strategy is their lack of sensitivity in some cases to identify all the relevant RCTs that have been published. The Cochrane Collaboration has developed a sensitive search strategy that should allow greater search precision and using this database to identify effectiveness trials should be a fundamental part of the search strategy. Other strategies to maximize retrieval of all relevant trials is the use of electronic databases searches that contain journals not registered with Medline. Some new journals may not be registered with electronic databases such as MEDLINE or CINAHL so individual hand searching of these journals and other key journals known to publish research in the field of interest should be considered. Hand searching has been shown to identify between 92% and 100% of the total number of trials identified from both hand searching and electronic searching [63], with MEDLINE identifying 55% of the total trials identified [63]. While hand searching is a useful additional strategy it is time-consuming, involving review of each article, review and letter published in each issue of the chosen journal to identify relevant work.

Another aspect of the search strategy that practitioners need to consider is the restriction of the search to English language journals. This is often undertaken for simplicity in undergraduate and postgraduate studies and where funding is not available for translation. There is a possibility that limiting the search in this way may bias the outcome of the review, but evidence about the impact of such a strategy is unclear. It has been identified that the quality of English language versus non-English language articles is the same [64, 65], but it is possible that research published in non-English journals is less likely to demonstrate a significant result [64], so by their exclusion may alter the outcome of any meta-analysis. However, in a review of language-restricted and language-inclusive meta-analyses, no difference in estimates of benefit was identified [66, 67]. It is therefore difficult to predict the overall impact of excluding non-English language studies.

Quality Assessment

Chapter 4 highlighted the importance of critical appraisal of the published literature and identified a range of tools that can be used to help in the appraisal process. A number of tools are reported in the literature and these include checklists, as well

as scales, with many different quality assessment tools available. Quality is a difficult construct to define for the range of research that a practitioner is likely to come across and no one tool may be appropriate for a range of topic areas. The QUADAS-2 tool for the assessment of the quality of studies of diagnostic accuracy is a validated tool that has built on the original QUADAS tool based on a consensus Delphi study [68, 69]. All quality assessment tools should be developed using formalized methods of development with assessments of face, content and construct validity and tested for reliability across different raters [27]. A tool developed initially to assess the quality of RCTs in pain research (the Jadad scale) was also based on a Delphi consensus method of agreement of experts and has been proposed for use across a range of clinical trials [70]. This tool uses a scale from 0 to 5 with reviewers scoring the answers to three questions as either yes (scores 1 point) or no (scores no points), with additional points awarded where blinding and randomization were appropriate [70].

In contrast, the Cochrane Collaboration recommends a domain-based approach to quality assessment of RCTs including assessment of the following [71].

- sequence generation
- allocation concealment
- blinding of participants, personnel and assessors
- incomplete outcome data
- selective outcome reporting
- other sources of bias.

Assessment tools often consider the internal validity of the study as reported but published trials judged by assessment tools to be low quality may actually reflect poor reporting rather than poor design quality, resulting from a lack of understanding on how to report a clinical trial, a problem of under-reporting. To overcome problems associated with poor reporting, it is now possible to publish trial protocols in peer reviewed journals; allowing readers to see greater detail of study designs and allowing better assessment of study quality of the final published results manuscript.

A quality assessment threshold should be identified to exclude weak studies from the review and can be achieved by applying a cutoff level for study selection. This may be based on quality assessment criteria identified above, as well as using a hierarchy of study designs. For example, in effectiveness studies the primary research question is based on an assessment of one intervention over another, which is best studied using a RCT with concealment of allocation. Where these are not available the next best design should be chosen, i.e., quasi-experimental studies where there is no randomization or cohort studies [59].

The ROBINS-I (Risk Of Bias In Non-Randomized Studies of Interventions) tool has been designed to facilitate researchers in assessing the quality of research involving non-randomized cohorts. Details of the ROBINS-I can be found here https://sites.google.com/site/riskofbiastool/welcome/home, and a useful guide to using the tool was published by Sterne et al. [72]. For reviews considering test

accuracy the hierarchy of study designs differs and the method at the top of the hierarchy is a blind comparison where there is a reference standard and a broadly defined sample of consecutive patients. Similarly, where these do not exist or are limited for the test under review it may be necessary to include studies where there is a narrow population or differential use of a reference standard [59].

When attempting to assess trial quality it is helpful to use a data collection/extraction form that includes details of the bibliographic reference, description of study characteristics and the quality assessment. This can then be used to develop a table of evidence comprising all the included trials. Examples of such forms can be found on the Scottish Intercollegiate Guidelines Network (SIGN) website (https://www.sign.ac.uk/assets/sign50_2015.pdf).

Regardless of the chosen assessment tool or threshold level chosen it is important that the quality assessment is not only integrated into the selection of studies for inclusion in the review but also incorporated within the results that are presented. However, in many published systematic reviews, while quality assessment is apparent in the selection of included trials, the quality of the selected studies is not always transparent in the final reporting of the results [73]. Quality assessment should be incorporated into the systematic review process at the selection of studies phase, in the interpretation of conflicting trial results, in the weight apportioned to trials within a meta-analysis and in the conclusions and recommendations of the review [59]. This can be achieved in its simplest form by a description of the results with a review of any risks of bias within the individual studies included. It can also be achieved by listing the quality score, or using the method adopted by SIGN, where ++ refers to high quality, + refers to acceptable quality and '-' refers to low quality (see the SIGN checklists https://www.sign.ac.uk/checklists-and-notes.html) against the tabulation of the individual trial characteristics so that readers can instantly see how the study quality may influence the overall outcomes of the review. See Table 12.4 for an example of where the quality assessment has been included in an evidence table in a published systematic review [74]. In this systematic review only research that scored '+' or '++' was used to draw conclusions.

The Cochrane handbook [71] provides good guidance on how to report the risk of bias in studies included in a systematic review and it is worth taking a look at the Cochrane handbook available here https://training.cochrane.org/handbook.

Meta-Analysis

Where individual studies allow, a formal quantitative analysis of the results may be undertaken in the form of a meta-analysis. This quantitative analysis provides a precise estimation of intervention effects and can indicate heterogeneity between studies where this exists. Including inappropriate studies in the meta-analysis can lead to misleading results, hence care needs to be taken in the execution of the analysis. For systematic reviews that include meta-analysis inclusion criteria need to prescribe the characteristics of studies that allow them to be combined in the meta-analysis; this may be trials studying the same intervention with the same outcome measures, undertaken on patients with similar characteristics (such as age or disease type).

Table 12.4 Example of evidence table-immobilization literature (supine position) from Probst et al. [74]

Author+ year	Description	Accuracy	*n*	Materials used on the breast	Skin reactions	Advs/disad	QA
Latimer et al. (2005) [13]	A micro-shell vs two other breast rings	Not measured	8	Polyacrylic micro-shell shaped into a horse-shoe	Micro-shell increased surface dose by 9%, other devices increased by 22%	• Shaped to reduce skin dosage • Reusable • Expandable capacity	+
Carter et al. (1997) [26]	Retrospective review	CLD variability average = −1.2 mm	20	Alpha Foam cradle	Not applicable	• No patient demographic available so unable to assess impact of patient size on reproducibility • No control group for comparison	-
Thilmann et al. (1998) [18]	Comparison between a positioning support cushion and no immobilization	Mean error without support 8.4 mm vs 6.1 mm	55	Foam	Not observed	Accuracy significantly improved with support (72% more comfortable)	+
Graham et al. (2000) [19]	Randomization to armrest or vacuum bag immobilization	Lung exposure (mean SD): Vac-bag 0.21 cm (95% CI 0.17–0.26) Arm-rest 0.21 cm (95% CI 0.17–0.24)	30	None thorax stabilization	Less skin folds present in armrest	Armrest more comfortable, vacuum bag allowed less lung exposure, no difference in stability and setup time	+
Nalder et al. (2001) [20]	Comparison of standard breast board and vacuum bag attached to a breast board	mean and SD of the systematic errors (mm): *With VB* AP −1.8 (2.9) *No VB* AP −1.7 (2.8) SD of the random errors: *With VB* AP 2.6 *No VB* AP 2.2	17	Not stated	n/a	• Minimal improvements found using the VB • Majority found the VB more comfortable	+

Bentel et al. (1999) [21]	Patients with large and/or pendulous breasts underwent radiotherapy using a breast ring; comprised of a hollow tube and fitted around the breast in contact with the skin	n/a	56	PVC tube (other material of tube tested was nylon)	Moist desquamation in 60.7% Surface dose under the ring approximately 85% of D_{max} dose. Without ring surface dose 35%	• Reduce skin folds and lateral movement in supine position no quantitative data • Good cosmetic outcome reported	-
Strydhorst et al. (2011) [22]	Assessment of the effect of a thermoplastic immobilization device on minimizing breast/chest wall movement during chest wall/breast irradiation	Inter-fraction motion: average random error Left/rt = 4 mm Sup/inf = 12 mm and AP = 4.5 mm Intra-fraction motion: av = 1 mm	$N = 8$	Thermoplastic shell	Not measured	Inter-fraction motion appears large which would indicate this method of immobilization does not work well	-
Cross et al. (1989) [23]	Feasibility study to assess the usefulness of the lateral decubitus position for women with very large breasts	Not measured	$N = 4$	Styrofoam block plus alpha cradle	all developed moist desquamation inferiorly due to contact with styroform foam, surface dose increased from 40 to 80%	Conclude lateral decubitus position feasible for women (cup size EE). technique does not allow matching of an scf	-

(continued)

Table 12.4 (continued)

Author+ year	Description	Accuracy	*n*	Materials used on the breast	Skin reactions	Advs/disad	QA
Goldsworthy et al. (2010) [24]	RCT comparing positioning on a breast board with either both arms abducted (intervention group) or single arm abducted. (control group)	*CLD* systematic error mean = −1.7 mm vs −1.9 mm $p = 0.06$, population systematic error 4 mm vs 2.3 mm $p = 0.005$ in favour of intervention. Population random error 2.1 mm vs 1.6 mm p=0.055	50	Traditional breast board with armpole device	not measured	The use of bi-lateral arm abduction resulted in smaller set up errors than the single arm positioning, although differences small	+
Zierhut et al. (1994) [25]	A repeated measures design to test the usefulness of a thermo plastic immobilization device. Patients were treated in the thermoplastic but simulation data available with and without the device	AP mean deviation = 3 mm with the device. sup-inf 4.1 mm	7	Thermoplastic	Surface dose increased from 47% to 64% on patients, on the phantom the surface dose was increased from 51 to 64% (of the maximum dose). The increase in skin dose was 17%	The increase in skin dose was 17%	-
Chopra et al. (2006) [29]	A case series	Displacements: Sup-inf = 1.3 mm Med-lat = 1.3 mm Ant-post = 4.4 mm	5	Vacuum bag immobilization	Not measured/Not applicable	Patient demographics not reported, no control group for comparison	-

Creutzberg et al. (1993) [27]	Non-randomized trial 1. patients lying flat with plastic mask (n = 17) 2. patients no mask (n = 14) 9 on inclined wedge, 5 lying flat	Ventral-dorsal displacement: With mask = 3.2mm Without mask = 4.6 mm	31	Plastic mask vs no mask And flat vs inclined on a wedge	Not measured	Not clear the criteria for allocation (except for those with additional nodal fields), no patient demographic data	-
Valdagni and Italia (1991) [28]	Case series	Ventral-dorsal shift = 2.7 mm (± 2.2 mm) Craniocaudal shift = 1.9 mm (± 1.8 mm)	20	Plastic mask immobilization		No control group for comparison Patient demographic data, no information on observer reliability	-
Keller et al. (2013) [30]	A commercially available bra/bustier compared with no bra	Not measured	N = 246	Commercial bra using thin plastic stays	Bra: 90% of cases grade 2 dermatitis No bra: 70% (p = 0.003)	Baseline characteristics were uneven across control and intervention (i.e., more cases with larger breast cup size in the intervention group), no randomization between control and intervention	-

++ = high quality low risk of bias, + = acceptable quality with some risk of bias, - = high risk of bias

The meta-analysis itself involves combining the results of all included studies that are combinable, i.e., have the same outcome measure. The individual trial results are weighted according to trial size although weighting based on trial quality has been proposed [75]. The methods used to combine the data are defined by two models, 'fixed effects' and 'random effects'. The choice of model depends on the presence of heterogeneity or variability between studies. Variability across studies, i.e., between-studies heterogeneity, can be assessed using either a Q statistic or an I^2 index [76]. The Q statistic produces a binary outcome identifying whether heterogeneity is present or absent. The I^2 index has been proposed as it gives a better indication of the level of heterogeneity that is present. Studies with an I^2 index of 25%, 50% and 75% would be classified as having low, medium or high variability, respectively [76].

The 'fixed effects' model combines the results of studies assuming that the effect of the intervention is constant across studies so only within-study variation is included in the analysis. In contrast, the 'random effects' model is based on the premise that the true treatment effect is different across individual studies [77] and this method is preferred when variability across studies is high [75, 76].

The results of combining data are often presented in graphical form; traditionally, this has been using a forest plot like the one in Fig. 12.7.

Figure 12.7 is a forest plot from the independent review of breast cancer screening trials [78].

Each trial is described by one line. Squares indicate the relative risk of death from breast cancer from screening versus the non-screened population for each trial. The horizontal line on forest plots usually defines the 95% confidence intervals. The solid vertical line indicates a ratio of 1.0 (i.e., 1.0 indicates no difference between screened and non-screened populations) trials that fall on the solid line would indicate no benefit from screening. For each category of trial the total ratio (relative risk) are shown as a diamond. The overall results of this meta-analysis identified a benefit from screening; the reduction in breast cancer mortality in those invited for screening was estimated to be 20% (95% CI 11–27). There was some heterogeneity between the trials but this was not statistically significant. However, the confidence interval around the RR of 0.8 is reasonably large (i.e., 0.73–0.89).

This meta-analysis serves to highlight an important dilemma in HTA primarily that when mature data are available for analysis, the technology and treatments for the condition may have moved on substantially, making the outcomes difficult to interpret within a new context. In some of the breast cancer screening trials included in this analysis there has also been discussion about the internal validity and therefore the accuracy of the predicted benefits of screening programmes.

Meta-analyses do have limitations which may be ascribed to the quality of the original RCTs available for analysis. As described in the previous section, inadequate sample sizes or opportunities for bias, such as inadequate concealment, may reduce the quality of the research which may then lead to inaccuracies in subsequent meta-analysis. Furthermore, research with a positive result is more likely to be published than a study showing no treatment or intervention benefit. Therefore,

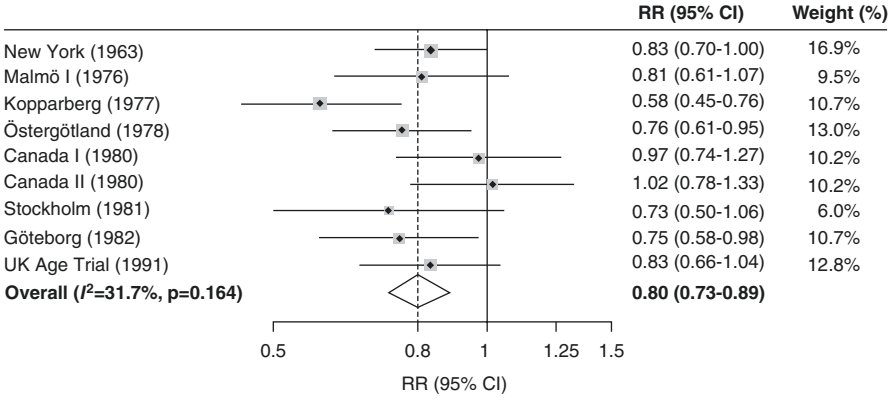


Fig. 12.7 Forest plot example—Meta-analysis of breast cancer mortality after 13 years of follow-up in breast cancer screening trials. Adapted from the Cochrane Review [78]. *RR* relative risk. Malmö II is excluded because follow-up of about 13 years was not available; the Swedish Two County (Kopparberg and Östergötland) and Canada I and II trials are split into their component parts; the Edinburgh trial is excluded because of severe imbalances between randomized groups. Weights are from random effects analysis.

meta-analyses may suffer the effects of publication bias if search strategies to identify eligible studies exclude the grey literature. In addition, meta-analyses suffer the risk of bias that may occur from the process of undertaking a systematic review including bias in the selection of studies, the assessment of study quality by the reviewers and problems with poor reporting of study results or errors in the data of the published reports [77]. A method proposed to identify publication bias in meta-analyses is the use of a simple graphical presentation of the individual trials estimate of treatment effect plotted against the trial sample size (funnel plot). If there is no bias, the plot should be symmetrical, depicting an inverted funnel with greatest dispersion of effects among trials of small sample sizes and a less marked dispersion in trials with larger sample sizes [75, 77], with meta-analyses that contain bias demonstrating asymmetrical funnel plots [77].

Reporting the Results of a Systematic Review

In a review of the methods of reporting of systematic reviews of diagnostic tests in cancer, Mallett et al. [79] identified significant variability in reporting of critical criteria such as defining the target condition where 51% failed to report if tumours were primary, recurrent or metastatic, with equal failings when it came to reporting tumour stage. To improve the quality of reporting of systematic reviews a consensus report (by the QUOROM group) proposed a checklist of items and a flow diagram that should be included in systematic reviews and meta-analyses [80]. The Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) checklist and flow diagram can be downloaded from the author section of most leading journal websites or from

here http://www.prisma-statement.org, and consists of 27 headings and subheadings to guide authors in the reporting and quality assessment of this type of research [81]. The PRISMA guidance covers the detail that is needed in the reporting of the search strategy, selection of studies for inclusion in the review, quality assessment of the selected trials, method of data extraction, details of the study characteristics, the quantitative data analysis (if there is any) the discussion of the results and the reporting of funding of the review. PRISMA also suggests the use of a flow diagram to indicate the number of trials identified, those included, and information about trials that were excluded. Figure 12.8 is an example of a PRISMA diagram from a published systematic review [74].

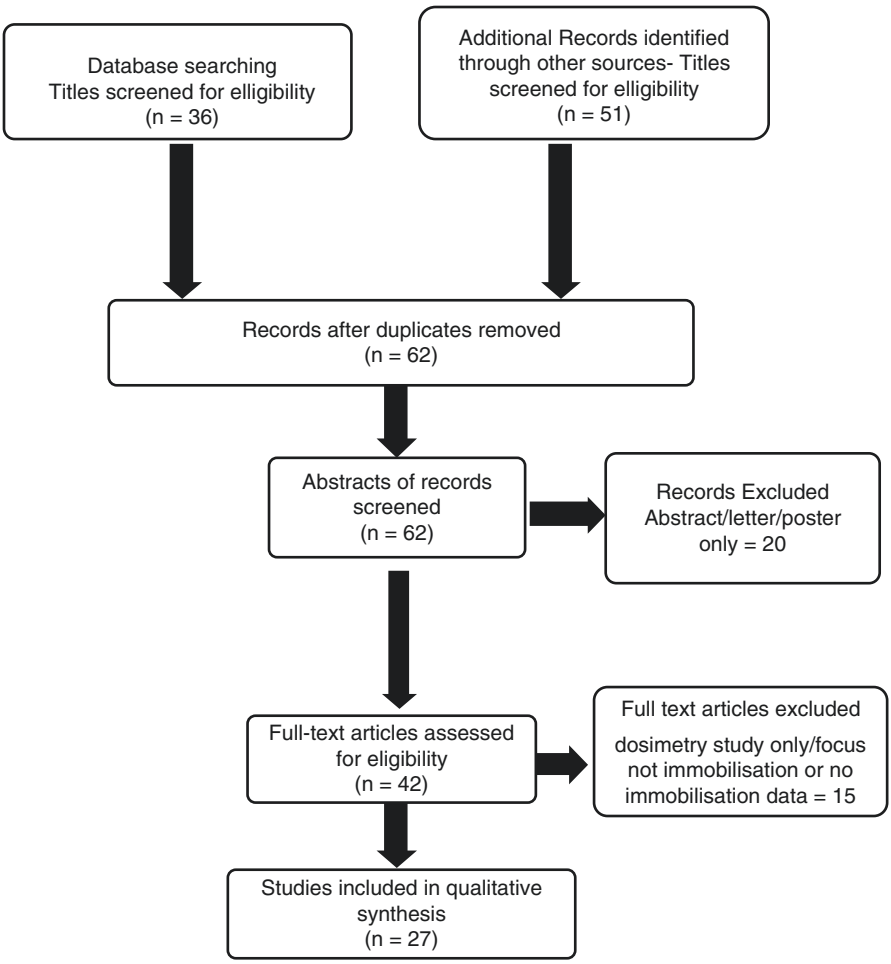


Fig. 12.8 An example flowchart demonstrating how articles were included and excluded from a systematic review by Probst et al. [74]

12.6 Conclusion

This chapter has provided an insight into the research design approaches that are useful for assessing new diagnostic imaging technologies and radiotherapy interventions. The use of healthcare technology has medical, social, ethical and economic considerations. Although in diagnostic imaging and radiotherapy this technology tends to be complex, many healthcare approaches can rely on quite simple devices. The full evaluation of a diagnostic test requires assessment at every level of the evaluative framework to demonstrate how good quality images contribute to accurate diagnoses, beneficial changes to diagnoses and management plans and improved patient outcomes, at acceptable costs.

Randomization is important for ensuring a balance in the characteristics of patients between groups and should be performed remote from clinical practice to help ensure adequate concealment in treatment allocation. The sample size calculations should be conducted based on clinically significant improvements in the primary outcome measure. The recruitment of patients is a major challenge in clinical trials. The methods to facilitate recruitment must include careful consideration of the participant consent process, inclusion of important members of the multidisciplinary team to encourage recruiting participants and a realistic timeframe to recruit a sufficient sample size and adequate funding. The attrition in the follow-up of patients should be limited by considering methods to reduce participant burden such as questionnaire length and minimizing the collection of missing data. With the increasing emphasis on resource allocation it is important to consider the economic implications of any new technology or new process. In HTA a cost-effectiveness analysis maybe appropriate and can be considered alongside the design of the RCT. Systematic reviews differ from the conventional type of review in that they adhere to strict scientific design to make them more comprehensive, to minimize bias and errors thus providing more reliable results to support evidence-based decision-making in policy and practice. It is important to examine variation or heterogeneity across studies to inform the choice of statistical model (‘fixed effects’ or ‘random effects’) for pooling the results of studies. ‘Healthcare technology’ is a broad term and encompasses a variety of instruments and techniques which promote health, prevent and treat disease, and enhance rehabilitation.

References

1. Fineberg HV, Bauman R, Sosman M. Computerised cranial tomography: effect on diagnostic and therapeutic plans. JAMA. 1977;238:224–7.
2. Fryback DG, Thornbury JR. The efficacy of diagnostic imaging. Med Decis Mak. 1991;11:88–94.
3. Mackenzie R, Dixon AK. Measuring the effects of imaging: an evaluative framework. Clin Radiol. 1995;50:513–8.
4. Bothwell LE, Greene JA, Podolsky SH, Jones DS. Assessing the gold standard-lessons from the history of RCTs. N Engl J Med. 2016;374:2175–81.
5. Robinson PJA. Radiology’s Achilles’ heel: error and variation in the interpretation of the Röntgen image. Br J Radiol. 1997;70:1085–98.

6. Brealey S, Scally AJ. Methodological approaches to evaluating the practice of radiographers' interpretation of images: a review. Radiography. 2008;14(1):e46–54.
7. Sackett DL, Haynes RB. The architecture of diagnostic research. In: Knottnerus JA, editor. The evidence base of clinical diagnosis. London: BMJ Books; 2002. p. 19–38.
8. Sackett DL, Haynes RB. Evidence base of clinical diagnosis: the architecture of diagnostic research. BMJ. 2002;324:539–41.
9. Cohen JF, Korevaar DA, Altman DG, Bruns DE, Gatsonis CA, Hooft L, Irwig L, Levine D, Reitsma JB, de Vet HCW, Bossuyt PMM. STARD 2015 guidelines for reporting diagnostic accuracy studies: explanation and elaboration. BMJ Open. 2016;6(11):e012799.
10. Whiting P, Westwood M, Rutjes AWS, et al. Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol. 2006;6:9.
11. Kelly S, Berry E, Roderick P, et al. The identification of bias in studies of the diagnostic performance of imaging modalities. Br J Radiol. 1997;70:1028–35.
12. Deeks J. Systematic reviews of evaluations of diagnostic and screening tests. In: Egger M, Smith GD, Altman G, editors. Systematic reviews in health care: meta-analysis in context. London: BMJ Publishing Group; 2001. p. 248–82.
13. Hajjan-Tilaki K. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med. 2013;4(2):627–35.
14. Habbema JDF, Eijkemans R, Krijnen P, et al. Analysis of data on the accuracy of diagnostic tests. In: Knottnerus JA, editor. The evidence base of clinical diagnosis. London: BMJ Books; 2002. p. 117–44.
15. Brealey S, Scally AJ. Bias in plain film reading performance studies. Br J Radiol. 2001;74:307–16.
16. Cohen J. A coefficient of agreement for nominal scales. Educ Psychol Meas. 1960;20:37–46.
17. Lee W. Technology assessment: vigilance required. Int J Radiat Oncol Biol Phys. 2008;70(3):652–3.
18. Kunz R, Oxman AD. The unpredictability paradox: review of empirical comparisons of randomised and non-randomised clinical trials. Br Med J. 1998;317:1185–90.
19. Hróbjartsson A, Emanuelsson F, Skou Thomsen AS, Hilden J, Brorson S. Bias due to lack of patient blinding in clinical trials. A systematic review of trials randomizing patients to blind and nonblind sub-studies. Int J Epidemiol. 2014;43(4):1272–83.
20. Scottish Intercollegiate Guidelines Network Health Improvement Scotland. SIGN 50: a guideline developers handbook. Quick Reference Guide. 2015.
21. Herst PM, Bennett NC, Sutherland AE, Peszynski RI, Paterson DB, Jasperse ML. Prophylactic use of Mepitel Film prevents radiation-induced moist desquamation in an intra-patient randomised controlled clinical trial of 78 breast cancer patients. Radiother Oncol. 2014;110(1):137–43.
22. Goldsmith C, Haviland J, Tsang Y, Sydenham M, Yarnold J. Large breast size as a risk factor for late adverse effects of breast radiotherapy: is residual dose inhomogeneity, despite 3D treatment planning and delivery, the main explanation? Radiother Oncol. 2011;100(2):236–40.
23. Noble-Adams R. Radiation induced reactions 2: development of a measurement tool. Br J Nurs. 1996;8(18):1208–11.
24. Noble-Adams R. Radiation induced reactions 3: evaluating the RISRAS. Br J Nurs. 1999;8(19):1305–12.
25. Radiation Therapy Oncology Group. Cooperative Group Common Toxicity Criteria. Minimize. 2019. https://www.rtog.org/ResearchAssociates/AdverseEventReporting/CooperativeGroupCommonToxicityCriteria.aspx.
26. Neal A, Torr M, Helyer S, et al. Correlation of breast dose heterogeneity with breast size using 3D CT planning and dose volume histograms. Radiother Oncol. 1995;34(3):210–8.
27. Jadad A. Randomised controlled trials: a user's guide. London: BMJ Books/Wiley; 2004.
28. Moss S, Thomas I, Evans A, Thomas B, Johns L. Randomised controlled trial of mammographic screening in women from age 40: results of screening in the first 10 years. Br J Cancer. 2005;92(5):949–54.

29. Hendrick RE, Smith RA, Rutledge JH, et al. Benefit of screening mammography in women aged 40–49: a new meta-analysis of randomised controlled trials. J Natl Cancer Inst Monogr. 1997;1997(22):87–92.
30. Probst H, Griffiths S. Increasing the work speed of radiographers: the effect on the accuracy of a setup of a complex shaped cranial field, part of a matched cranio spinal junction. Radiother Oncol. 1996;38(3):241–5.
31. Norrman E, Persliden J. A factorial experiment on image quality and radiation dose. Radiat Prot Dosim. 2005;114(1–3):246–52.
32. Probst H, Dodwell D, Gray JC, et al. An evaluation of the accuracy of semi-permanent skin marks for breast cancer irradiation. Radiography. 2006;12(3):186–8.
33. Cancer Research UK. Breast cancer incidence 2018. Accessed May 2019. https://www.cancerresearchuk.org/health-professional/cancer-statistics/statistics-by-cancer-type/breast-cancer/incidence-invasive#heading-One.
34. Roberts C, Torgerson D. Understanding controlled trials randomisation methods in controlled trials. Br Med J. 1998;317:1301.
35. Torgerson DJ. Contamination in trials: is cluster randomisation the answer? Br Med J. 2001;322:355–7.
36. Kennedy ADM, Torgerson DJ, Campbell MK, Grant AM. Subversion of allocation concealment in a randomised controlled trial: a historical case study. Trials. 2017;18(1):204.
37. Torgerson DJ, Roberts C. Understanding controlled trials randomisation methods: concealment. BMJ. 1999;319(7206):375–6.
38. Pocock SJ. The size of a clinical trial. Clinical trials: a practical approach. Chichester: Wiley; 2008. p. 123–41.
39. Probst H, Dodwell D, Gray J, Holmes M. Radiotherapy for breast carcinoma: an evaluation of the relationship between the central lung depth and respiratory symptoms. Radiography. 2005;11(1):3–9.
40. Corrie P, Shaw J, Harris R. Rate limiting factors in recruitment of patients to clinical trials in cancer research: descriptive study. Br Med J. 2003;327:320–1.
41. Welton A, Vickers M, Cooper J, et al. Is recruitment more difficult with a placebo arm in randomised controlled trials? A quasi-randomised, interview based study. Br Med J. 1999;318:1114–7.
42. Hancock BW, Aitken M, Radstone C, et al. Why don't cancer patients get entered into clinical trials? Experience of the Sheffield Lymphoma Group's collaboration in British National Lymphoma Investigation studies. BMJ. 1997;314(7073):36.
43. Mitchell G, Abernethy AP, Investigators of the Queensland Case Conferences Trial, Palliative Care Trial. A comparison of methodologies from two longitudinal community-based randomized controlled trials of similar interventions in palliative care: what worked and what did not? J Palliat Med. 2005;8(6):1226–37.
44. Treweek S, Lockhart P, Pitkethly M, Cook JA, Kjeldstrøm M, Johansen M, et al. Methods to improve recruitment to randomised controlled trials: cochrane systematic review and meta-analysis. BMJ Open. 2013;3(2):e002360.
45. Torgerson DJ, Sibbald B. Understanding controlled trials. What is a patient preference trial. BMJ. 1998;316(7128):360.
46. Torgerson DJ, Roland M. What is Zelen's design? BMJ. 1998;316(7131):606.
47. Dumville JC, Torgerson DJ, Hewitt CE. Reporting attrition in randomised controlled trials. BMJ. 2006;332(7547):969–71.
48. Pocock SJ. Protocol deviations. Clinical Trials: a practical approach. Chichester: Wiley; 2008. p. 176–86.
49. Hollis S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ. 1999;319(7211):670–4.
50. Altman DG, Bland JM. Missing data. BMJ. 2007;334(7590):424.
51. Fergusson D, Aaron SD, Guyatt G, et al. Post-randomisation exclusions: the intention to treat principle and excluding patients from analysis. BMJ. 2002;325(7365):652–4.

52. Piccart-Gebhart MJ, Procter M, Leyland-Jones B, et al. Trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer. N Engl J Med. 2005;353(16):1659–72.
53. Smith I, Procter M, Gelber RD, et al. 2-year follow-up of trastuzumab after adjuvant chemotherapy in HER2-positive breast cancer: a randomised controlled trial. Lancet. 2007;369(9555):29–36.
54. National Institute for Clinical Excellence. Trastuzumab for the adjuvant treatment of early-stage HER2-positive breast cancer. London: NHS Department of Health; 2006 NICE Technology Appraisal Guidance 107.
55. Drummond MF, O'Brien B, Stoddart GL, et al. Basic types of economic evaluation: methods for the economic evaluation of health care programmes. 2nd ed. Oxford: Oxford University Press; 1997. p. 6–26.
56. Palmer S, Byford S, Raftery J. Economics notes: types of economic evaluation. BMJ. 1999;318(7194):1349.
57. Amparo O, Santaballa A, Munarriz B, et al. Cost-benefit analysis of a follow-up program in patients with breast cancer: a randomized prospective study. Breast J. 2007;13(6):571–4.
58. Shah C, Ward MC, Tendulkar RD, Cherian S, Vicini F, Singer ME. Cost and cost-effectiveness of image guided partial breast irradiation in comparison to hypofractionated whole breast irradiation. Int J Radiat Oncol Biol Phys. 2019;103(2):397–402.
59. Centre for Reviews and Dissemination. Undertaking systematic reviews of research on effectiveness: CRD's guidance for those carrying out or commissioning reviews. Centre for Reviews and Dissemination, University of York; 2001. Centre for Reviews and Dissemination Report 4.
60. Cruz Rivera S, Kyte DG, Aiyegbusi OL, Keeley TJ, Calvert MJ. Assessing the impact of healthcare research: a systematic review of methodological frameworks. PLoS Med. 2017;14(8):e1002370. https://doi.org/10.1371/journal.pmed.1002370.
61. Greenhalgh T, Peacock R. Effectiveness and efficiency of search methods in systematic reviews of complex evidence: audit of primary sources. BMJ. 2005;331(7524):1064–5.
62. Hopewell S, McDonald S, Clarke M, et al. Grey literature in meta-analyses of randomised trials of healthcare interventions. Cochrane Database Syst Rev. 2007;(2):MR000010.
63. Hopewell S, Clarke M, Lefebvre C, et al. Handsearching versus electronic searching to identify reports of randomised trials. Cochrane Database Syst Rev. 2007;(2):MR000001.
64. Egger M, Zellweger-Zahner T. Language bias in randomised controlled trials published in English and German. Lancet. 1997;350(9074):326.
65. Moher D, Fortin P. Completeness of reporting of trials published in languages other than English: implications for conduct and reporting of systematic reviews. Lancet. 1996;347(8998):363.
66. Moher D, Pham B, Klassen TP, et al. What contributions do languages other than English make on the results of meta-analyses. J Clin Epidemiol. 2000;53(9):964–72.
67. Juni P, Holenstein F, Sterne J, Bartlett C, et al. Direction and impact of language bias in meta-analyses of controlled trials: empirical study. Int J Epidemiol. 2002;31(1):115–23.
68. Whiting P, Rutjes A, Reitsma J, et al. The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. BMC Med Res Methodol. 2003;3(1):25.
69. Whiting PF, Rutjes AWS, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, Leeflang MM, Sterne JAC, Bossuyt PMM. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529–36.
70. Jadad RA, Moore D, Carroll C, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? Control Clin Trials. 1996;17:1–12.
71. Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions version 5.1.0 [updated March 2011]: The Cochrane Collaboration; 2011. http://handbook.cochrane.org.
72. Sterne JA, Hernán MA, Reeves BC, Savović J, Berkman ND, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ. 2016;355:i4919.

73. Moja LP, Telaro E, D'Amico R, et al. Assessment of methodological quality of primary studies by systematic reviews: results of the meta-quality cross sectional study. BMJ. 2005;330(7499):1053.
74. Probst H, Bragg C, Dodwell D, Green D, Hart J. A systematic review of methods to immobilise breast tissue during adjuvant breast irradiation. Radiography. 2014;20(1):70–81.
75. Egger M, Smith GD, Phillips AN. Meta-analysis: principles and procedures. BMJ. 1997;315(7121):1533–7.
76. Huedo-Medina TB, Sanchez-Mecca J, Bottela J, et al. Assessing heterogeneity in meta-analysis: Q statistic or I2 index? Psychol Methods. 2006;11(2):193–206.
77. Egger M, Davey SG, Schneider M, et al. Bias in meta-analysis detected by a simple, graphical test. BMJ. 1997;315(7109):629–34.
78. The benefits and harms of breast cancer screening: an independent review. Lancet. 2012;380(9855):1778–86.
79. Mallett S, Deeks JJ, Halligan S, et al. Systematic reviews of diagnostic tests in cancer: review of methods and reporting. BMJ. 2006;333(7565):413.
80. Felson DT. Bias in meta-analytic research. J Clin Epidemiol. 1992;45(8):885–92.
81. David M, Cook D, Eastwood S, et al. Improving the quality of reports of meta-analyses of randomised controlled trials: the QUOROM statement. Lancet. 1999;354:1896–900.

Part IV

Dosimetry and Service Evaluation

Dosimetry

13

Martin Vosper

13.1 Introduction

Although dosimetry research may appear challenging at first sight, it can provide a relatively straightforward route to completing a diagnostic imaging or radiotherapy research project. The steady pace of change in diagnostic imaging and radiotherapy means that new dosimetry research is always valuable. The subject of radiation dosimetry could easily fill a volume in its own right. In this chapter an overview of research opportunities and issues, rather than a comprehensive text, is provided.

13.2 Why Should We Undertake Research in Radiation Dosimetry?

Dosimetry refers to the recording of doses of ionising radiation received by patients, staff and members of the public during medical imaging and radiotherapy procedures. Many practitioners feel inclined to avoid research in radiation dosimetry, perhaps because they believe it is a very 'dry' subject or that they have to be high-flying experts in physics (and maths) to manage it. But neither of these worries is really justified. In reality dosimetry can be a useful introduction to medical imaging and radiotherapy research.

- It is easy to argue that dosimetry research is useful and relevant, as all hospital departments of imaging or radiotherapy must keep radiation doses to patients and staff as low as reasonably achievable (ALARA) and take steps to monitor this.

M. Vosper (✉)
Department of Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hertfordshire, UK
e-mail: m.r.1.vosper@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_13

- Dosimetry is a form of experimental research, giving an opportunity to control experimental variables and reduce sources of bias. This provides a structured scientific project that is objective and clearly defined.
- The methods used are already well described and can be followed quite easily.
- If the project is equipment-based and does not involve human subjects (such as patients or staff), there is usually no need for formal research ethics approval.
- Since every piece of clinical equipment is to some extent unique and every hospital department has individual practices, dosimetry research can easily provide original information.
- The rapid pace of technical change, for example, in digital imaging, multi-slice CT and intensity modulated radiotherapy means that new dosimetry studies are always valuable and welcome.
- There is a range of 'phantoms' available that can simulate the human body for dosimetry purposes.
- It might be possible to undertake some dosimetry research in a convenient laboratory, perhaps at your university or clinical site.
- Dosimetry usually avoids the need for time-consuming questionnaires or interviews.

This all sounds promising, but are there any disadvantages to dosimetry research? Well it does require access to the necessary equipment, such as radiation detectors, dose readers and possibly phantoms (if not involving human subjects). Some of these things are expensive or might not always be available. It is best to check and book them in advance, since it is not good to find out at the last minute that the apparatus has broken down or is already on loan to another student. There is also a need for training in equipment use. More will be said about equipment and techniques later in this chapter. Anyone undertaking dosimetry research needs to be comfortable with presenting numerical data using tables and graphs, as well as doing some fairly straightforward calculations. People who are interested in human viewpoints and feelings will probably steer away from dosimetry research, but how about doing some qualitative research on the way that patients and staff feel about radiation risks?

As in any other sort of research, it is important to ask a question. This question should be clinically relevant and of interest, but it would not normally be expected that the dosimetry research method should be original. It is completely acceptable, and in fact necessary, that tried and tested methods are used, in order to permit comparisons between study data. This is a big help to any new researcher. Every hospital radiography environment is different, and so fresh dosimetry readings will always be of value, both for audit purposes and for addition to existing data collected by national surveys of medical radiation protection.

The constant pace of technical and procedural change in radiography provides plenty of opportunities for original dosimetry research. There should not be any excuse for trying to ask research questions which were answered many decades ago and published in standard textbooks. An example of unoriginal and unnecessary research would be a study of whether a posteroanterior (PA) or anteroposterior (AP) technique affects patient dose in diagnostic lumbar spine radiography. The answer

is ‘yes: of course it does and the PA technique reduces dose’. Studies to demonstrate this were published over 30 years ago. It is not likely that the situation will be any different today. But what could perhaps be usefully asked in a research project is: Since the PA lumbar spine projection reduces patient dose, why is this not the standard radiographic technique today? This might identify a lot of interesting factors such as image quality issues, personal preferences, resistance to change.

13.3 Topics for Radiation Dosimetry Research

In medical imaging, which involves ionising radiations (such as X-rays and gamma rays), there is a clear need to obtain a high quality diagnostic image whilst giving the lowest practicable radiation dose to patients and staff. However, the very best possible quality images tend to require larger doses of radiation; this has always been the case, ever since the discovery of X-rays in 1895, and continues to be so, despite recent improvements in technology. Given the trade-off between the two factors of image quality and dose, it is a good idea to also consider image quality when studying the doses of radiation given during medical imaging procedures. It must be remembered that a very low dose X-ray procedure might not be of much benefit to a patient if the image is too indistinct to enable a proper diagnosis of their disease.

In medical imaging the following suggestions might provide good topics for dosimetry research.

- The effects on dose of manipulating the technique for a given procedure
- New and improved techniques for dose reduction
- The effectiveness of radiation protection measures, such as shielding and filtration
- Magnitudes of doses to the ‘whole body’, individual organs or to the foetus
- ‘Environmental doses’, in areas close to radiation sources such as X-ray tubes or radioisotopes
- Dose variations between imaging departments, different pieces of imaging equipment, different practices or even different operators
- Dose variations over time
- Doses received by staff, e.g., imaging practitioners and other groups
- The relationships between dose and image quality
- The effects of new technology, such as digital imaging or multi-slice CT
- The effects of staff training and radiation awareness
- The effects of systems of work, including administration, image archiving and patient referral.

In radiotherapy, dosimetry is very much an integral part of the work process and is an essential element of treatment planning. ‘In vivo’ dosimetry plays an increasing part in providing rapid verification data during the external beam treatment itself. While in medical imaging the damaging effects of radiation are always an unwanted side-effect, in radiotherapy the damaging effects of radiation on tumours

are deliberate and planned. Here the aim is to promote an adverse effect, but only on cancer cells. There is still a need to restrict the unwanted dose to normal and vulnerable tissues, both within and without the radiation field. Advances in radiotherapy have often focused on the need to deliver optimum doses to cancers, including the tumour margins, while limiting the doses to non-cancerous tissues and minimising the risks to neighbouring vulnerable organs. Radiotherapy dosimetry involves a greater variety of ionising radiations; not only X-rays from linear accelerators (LINACs) and gamma rays from cobalt sources, but also fast-moving particles such as electrons, protons or neutrons.

In radiotherapy the following suggestions might provide good topics for dosimetry research.

- The effects of altering the technique for a given procedure, such as when using high energy beams, electrons, multiple treatment fields
- New and improved techniques for dose reduction in at-risk tissues
- The effectiveness of radiation protection measures, such as shielding, filtration, distance
- In vivo dosimetry during radiation treatments, for example, using electronic portal imaging
- Doses in total body irradiation prior to bone marrow transplantation
- 'Environmental doses', measuring leakage radiation in areas close to radiation sources such as LINACs or radioisotopes
- Dose variations between different pieces of radiotherapy equipment, different practices or even different operators
- Doses in the simulator suite and in the CT-simulator
- Doses in brachytherapy
- Doses in superficial radiotherapy
- Fractionation regimes
- New developments in areas such as 3-D conformal radiotherapy, intensity modulated radiotherapy
- The effects of staff training and radiation awareness
- The effects of systems of work, including administration

13.4 Effects of Ionising Radiation

The damaging effects of ionising radiations are well known and fall into two major categories: stochastic and deterministic effects.

13.4.1 Stochastic Effects

Stochastic effects can occur at all magnitudes of dose and are a product of chance. The end result can be the induction of cancer, or the conveyance of a genetic abnormality to future offspring. The likelihood of the effect increases with radiation dose,

but is neither impossible at low doses nor inevitable at high doses. Dosimetry in medical imaging is mostly concerned with measuring small radiation doses which, although relatively low-risk, might still cause stochastic effects. In radiotherapy, although doses are much higher and stochastic effects therefore more likely, it must be remembered that the patient already has a cancer and that the risk from 'under-dosing' a malignant tumour is likely to be greater than the risk from inducing cancers in surrounding normal tissues. However, the risks of genetic damage to reproductive cells such as ova or sperm, or radiation-induced secondary cancers, are always relevant considerations. Stochastic effects tend to be 'late effects', which means that they only appear after a time period, which is usually measured in years, and not immediately. Thus, it is normally hard to link a stochastic effect to a single episode of previous radiation exposure with certainty.

13.4.2 Deterministic Effects

Deterministic (or tissue reaction) effects only occur at relatively high doses and are of increasing severity as dose increases. The result is extensive cell death and tissue damage, which may manifest as erythema (skin reddening), fibrosis, cataract, reduction in blood cell count, infertility, bowel disturbances and other serious changes. Deterministic effects are largely predictable and are common side-effects of radiotherapy. They can be minimised by delivering a radiation dose over an extended period of time rather than in a single large 'burst'. This is one of the benefits of dose fractionation in radiotherapy. Dosimetry in radiotherapy is mostly involved with measuring large radiation doses, which are intended to kill tumour cells, but might have unintended deterministic or stochastic effects on normal tissues.

In medical imaging, deterministic effects are rare but can result from extended fluoroscopy. Some types of detriment to the developing embryo or foetus in utero can also be regarded as deterministic in nature and might result from over-use of relatively high dose diagnostic procedures such as pelvic CT in an unsuspected pregnancy. Deterministic effects tend to be 'early effects' that appear within a period of days or weeks.

13.5 Some General Principles Regarding the Use of Ionising Radiations

Some general principles regarding the use of ionising radiations, which apply both in medical imaging and radiotherapy, are as follows.

- Increasing the number of ionising rays (or particles) will always increase the radiation dose.
- Increasing the size of the beam (or field) of radiation will always increase the radiation dose.

- Increasing the energy, or penetrative power, of the radiation beam will tend to reduce the skin dose where the beam enters the human body and tend to increase the dose at depth. It will also tend to produce more 'forward scatter' of radiation and less 'back scatter'.
- Increasing the distance between the source of radiation and the human body will reduce the radiation dose. This is particularly so if considering electromagnetic radiations such as X-rays and gamma rays and is covered by the 'inverse square law'. It should be noted that the dose from particle radiations such as electrons and protons may actually increase in some regards as the radiations pass deeper into the human body, due to the Bragg effect.
- Increasing the number of radiation exposures will always increase the dose, assuming that the size of the individual exposures is unchanged.
- Shifting sensitive tissues and organs away from an incoming radiation beam, for example, by turning a patient to face away, will tend to reduce the dose.
- Leaving sensitive tissues and organs outside the radiation beam, either by angling, redirecting or limiting the beam, will reduce the dose.
- A rapid radiation procedure is not necessarily a low dose one, since improvements in technology mean that large doses can be given very quickly.
- No dose of radiation is ever totally without risk, no matter how small the exposure, although the risk from very small exposures is negligible.
- The size or weight of a patient will affect the dose they receive, since radiation absorption will be influenced by body composition and dimensions.
- Young patients, especially children and babies, are more likely than elderly patients to suffer harm from a radiation exposure.

It is important to consider physics and radiobiology principles like these, as they will help the researcher to explain unexpected findings and interpret results.

To give an example of the appliance of science principles to real X-ray doses, let us consider the plain and simple PA chest projection in medical imaging. Here a high kVp (kilovoltage peak), and hence high beam energy, technique should give a lower dose to a patient's chest than a low kVp technique. This seems quite reasonable, as the more penetrating rays should pass through a patient's chest without being absorbed. But increasing the kVp of the exposure increases not just the energy of the X-rays, but also the number of rays produced. To maintain the image density and keep patient dose at an acceptable level, a practitioner will decrease the mAs (milliampere-second) setting to compensate for the increased kVp. A general rule is that a 15% increase in kVp can be compensated for in terms of image density by halving the mAs. This enables a high beam energy technique to provide dose reductions.

The next practical issue is that higher energy beams tend to produce scatter within the human body that is more energetic and more forwards in direction. Thus, it is more likely to reach the imaging plate. To avoid this scatter degrading the image, a practitioner may be obliged to use a secondary radiation 'grid' in order to maintain image contrast and quality. Unfortunately, such a grid typically requires a three to four factor increase in the number of X-rays produced by the X-ray tube, to

compensate for the removal of X-rays by a grid. These increased numbers of rays will still strike the patient, although they will not reach the image. The increase in dose resulting from the use of a grid might outweigh any dose benefit arising from the use of the high kVp technique.

Another issue is that some modern digital imaging plates may be less efficient at high beam energies, requiring a further increase in the number of X-rays produced by the tube. Finally, tissues lying outside the beam during PA chest radiography (such as the thyroid and eye) may receive a slightly higher dose at high beam energy, since the scattered radiation will be more penetrating and more likely to reach them. This example illustrates that patient dose is affected by a number of practical considerations in 'real world' radiography. It also shows that a dosimetry study may provide incomplete information if image quality is not also considered.

13.6 Devices, Quantities, and Units of Measurement in Radiation Dosimetry

To become familiar with radiation dosimetry we need to know the meaning of several terms used to describe 'dose', as these can be a source of misunderstanding. We need to know what we are measuring and why. In radiation dosimetry it is much easier to take physical measurements of the amount of energy deposited by ionisation in non-living materials, rather than to take biological measurements of the actual 'injury' inflicted by that ionisation in living tissues. Although it is the tissue injury that most interests us, usually this can only be extrapolated or calculated indirectly from the more direct physical measurements of ionisation in non-living materials. Microscopic measurement of cell injury, for example, through examination of chromosomes, is not available for most student projects, although macroscopic skin damage from large radiation doses (only likely during radiotherapy) is easily visible and can appear quite quickly. Some of the non-living materials used in radiation dosimetry to record doses are summarised in Table 13.1.

None of the materials listed in Table 13.1 will perfectly mimic the absorption of radiation by the human body, although some of them, such as lithium fluoride thermoluminescent dosimeters (TLDs) may be fairly similar in density and atomic number to soft tissue. But all of them will be able to give a good idea of the relative sizes of different radiation exposures. Often in radiation dosimetry it is these relative values that are used to compile local and national surveys.

Bulky ionisation chambers might be useful for providing reference dose standards and very sensitive measurements. They are not convenient for estimating a dose within a small patient volume. Large handheld Geiger counters are valuable for detecting radiation and for environmental monitoring, but likewise not very practical for monitoring doses to individuals. Personal monitoring of staff is usually based on compact devices such as pocket dosimeters, TLDs or film badges, none of which are perfect tools but do give an idea of relative doses.

Table 13.1 A selection of non-living materials used to absorb radiation in dosimetry

Non-living material used to absorb radiation	Applications
Air in a graphite or other container	Ionisation chambers such as—thimble chambers and pencil chambers
	Dose area product (DAP) meters
	Pocket dosimeters ('pen' meters or 'bleepers'), which are modified Geiger devices
Argon or helium gas in a glass/graphite container	Geiger counters
Phosphors such as lithium fluoride with other additives	Thermoluminescent dosimeters (TLDs)
Phosphors such as sodium iodide doped with thallium (scintillators)	Scintillation detectors
Silver halides	Film dosimeters
Silicon, germanium	'diode' type solid state semiconductor devices
Metal oxide and silicon	Metal oxide field effect transistors (MOSFETs), a type of solid state semiconductor device
Amorphous silicon	Flat panel detectors used in electronic portal imaging devices (EPIDs), for portal dosimetry
Ferrous sulphate solution	Fricke dosimeters

The most accurate way to monitor doses to patients' tissues would be the direct method of placing a small dosimeter within the body. This is not normally practicable of course, although it may be possible to place a dosimeter in the body in some radiotherapy situations. As an alternative, tissue-equivalent whole body phantoms, of about 75 kg weight for an adult male or about 50 kg weight for an adult female, such as the Alderson ART phantoms, may be used to accurately simulate typical organ doses for diagnostic imaging or radiotherapy procedures. These phantoms consist of a stack of body slices, which can be dismantled, with holes at regular intervals to accept TLDs. Of course, these standard phantoms cannot simulate the wide range of body sizes and weights found in real patients. An entire Alderson body phantom costs over £15,000, but access to one is very useful and it may remove the need to involve actual patients in the research. In radiotherapy, a simple water-equivalent epoxy resin phantom with holes to accept ionisation chambers can be very useful to measure absorbed doses arising from treatment fields for audit purposes and for comparisons between radiotherapy treatment devices or hospital centres.

In practice, it is often possible to attach small dosimeters, such as TLDs, MOSFETs, diode type semiconductors or thimble ionisation chambers, to a patient's skin surface. This is an indirect method for judging doses to internal organs. It does give a direct measurement of entrance surface dose (ESD) and exit dose (where the radiation beam enters and exits the body). TLDs are the most convenient devices for a patient as they do not come attached to wiring. They are also better than semiconductor devices for measuring skin dose or lens of eye dose, since they are not surrounded by a radiation-attenuating cap.

The most commonly used measurement for radiation dosimetry is the gray (Gy) which refers to energy absorbed in joules (J) per unit mass of matter in kilograms (kg). It is described as a unit of absorbed dose. The energy deposition in the matter takes place by ionisation and the irradiated matter can consist of air, water or solids such as body tissue. The gray is a physics-based unit and does not tell us much about the biological effects of radiation on living tissues.

Some publications mention the term kerma, which, like the absorbed dose, is measured in joules per kilogram of matter. Kerma refers to 'kinetic energy released (per unit) mass'. The process of absorption of ionizing electromagnetic radiations, namely photons such as X-ray and gamma ray, or particles such as neutrons, in matter leads to the release of 'secondary' charged particles (electrons) from the atoms in that volume of matter. These electrons may come to a halt within the volume or pass outside it. Put simply, while the kerma only records energy deposited by electrons, which arise from ionisations within the volume, the absorbed dose also records energy deposited by electrons which arise from ionisations just outside the volume, but which come to a halt within it.

13.7 How Can We Get an Idea of the Biological Harm Caused by Ionising Radiations?

To do this we need to consider not just the absorbed dose in body tissue but also how damaging different types of radiation are to tissue. This consideration gives us a unit called the equivalent dose, measured in sieverts (Sv). The damaging ability of the radiation is termed the relative biological effectiveness (RBE), which is quantified using a radiation weighting factor. In some older books you might find the term 'quality factor'. It means the same thing as the more recent term radiation weighting factor. In order to simplify calculations, X-rays, gamma rays and electrons of all energies are assumed to have a radiation weighting factor of 1. For these radiations, the equivalent dose in sieverts is equal to the absorbed dose in gray multiplied by 1. In reality of course we know that relatively low energy X-rays, for example, in mammography or superficial radiotherapy, are absorbed more than high energy X-rays and will thus be a bit more damaging, but these slight differences are ignored for the purposes of clinical dose limitation, control and routine assessment.

A different radiation weighting factor has to be used if we are measuring the damaging effects of alpha particles, protons or neutrons. These heavier particles will of course do more harm, rather like an elephant crashing around inside a china shop. Proton and neutron beam treatments may be encountered in radiotherapy. Current recommendations are that protons have a radiation weighting factor of two, while the radiation weighting factor for neutrons varies from about 2 to 20, depending on energy. Alpha particle emitters are not nowadays much encountered in medical imaging or radiotherapy but may arise from inhaled or ingested heavy elements like radon, radium or thorium. Alpha particles do not travel far. They are very ionising and are given an estimated weighting factor of 20.

The equivalent dose concept is used in the personal dose monitoring of staff via film badges or TLD tablets. The personal dose equivalent is that dose in tissue at a point just below the monitoring badge. It is used as a guide to likely doses received by the whole body. It is also possible to measure an ambient dose equivalent when monitoring doses received in an area such as a location in a workplace.

To get a more meaningful measurement of the real damaging effects of radiation on a human body, we need to know the damaging power of the radiation and also: (a) the individual doses received by all the organs and tissues exposed to that radiation and (b) the individual sensitivities of those organs and tissues. This is not easy to measure. The term effective dose takes account of the damaging power of radiation and its effects on all of the body's vulnerable tissues. The tissue weighting factor describes the relative contribution from various organs and tissues to the likely total harm that a person will receive from ionising radiation. More vulnerable organs have a higher weighting factor and the factors are given in Table 13.2.

The effective dose to the whole body provides a dose value in sieverts and is used to calculate the risk of stochastic radiation-induced events, chiefly cancer. Thus, it is widely used measure for radiation protection of staff, patients and the public. To calculate an effective dose, we need to know the equivalent doses to each organ or tissue and multiply these by their appropriate tissue weighting factors. The total effective dose is then the sum of the individual tissue-weighted equivalent doses. In most clinical situations the radiation beam will not cover the whole body and thus many organs will only receive a small dose from scattered radiation. It is those organs which lie within the primary beam that will receive the greatest equivalent doses. Stochastic effects (both whole body and to individual organs) are also of interest in radiotherapy, as dose to normal tissues may arise from beam passage through the body on the way to the target volume, from scatter and from radiation leakage via the LINAC treatment head. Newer intensity modulated radiotherapy techniques, although they reduce the dose to at-risk structures near the tumour, may actually increase the volume of normal tissue irradiated.

Table 13.2 Tissue weighting factors for radiation dosimetry (ICRP Publication 103, [1])

Organ or tissue	Individual weighting factor for each organ or tissue	Total contribution
Lung, stomach, colon, bone marrow, breast, remainder[a]	0.12	0.72
Gonads (mean of ♂ and ♀)	0.08	0.08
Thyroid, oesophagus, bladder, liver	0.04	0.16
Skin, brain, bone, surface, salivary glands	0.01	0.04
Whole body		1.00

[a]The 'remainder', a combined contribution of 0.12, is due in equal parts to the following fourteen tissues: adrenals, gall bladder, heart wall, kidneys, lymph nodes, muscle, oral mucosa, pancreas, prostate, small intestine, spleen, thymus, uterus/cervix and extra-thoracic tissue

13.8 How Can Effective Doses Be Calculated in Practice?

One solution would be to insert TLDs in appropriate positions to record organ doses within the slices of an anthropomorphic tissue-equivalent whole body phantom, such as the Alderson ART phantom. However, this might require many measurements, looking at the long list of organs in Table 13.2. Although published in 2007, they are still currently applicable. An alternative is to calculate the equivalent doses to the organs using mathematical calculations based on standardised mathematical models of the volumes, shapes, densities, and atomic numbers of the structures in the human body. These are often called Monte Carlo calculations; as exotic as that seaside resort, but based on the mathematical laws of chance just like the gambling in Monte Carlo's casinos. In this way, equivalent doses to internal organs can be calculated from equivalent doses measured at the surface of a real human body. This is made possible by the fact that entrance surface doses (where the radiation beam enters the human body) will tend to be the highest equivalent doses for many types of radiation beam and thus depth doses can be estimated from these values without too much inaccuracy. Computer programmes to provide Monte Carlo calculations are now more powerful than previously and a number of software packages are available. A full discussion of Monte Carlo techniques is beyond the scope of this chapter and advice should be sought from a local medical physicist.

Effective doses can be estimated from easily measurable quantities: the dose–area product and the entrance surface dose. These quantities are described as follows. They have the advantage that they can be readily reproduced using standard equipment and are a good basis for comparisons of doses between hospitals and between X-ray equipment in national surveys. Conversion coefficients have been published to give estimates of effective dose from quantities such as dose–area product and entrance surface dose. These coefficients will vary with the radiation procedure being undertaken.

The dose–area product or DAP reading is obtained from an ionisation chamber attached directly to the source of radiation. The measurement is a dose in air, recorded in grays per square centimetre. The dose value depends not only on the radiation output of the radiation source, but also on the size of the radiation field or beam. The value will increase as the tube output and radiation field increase. It is a mistake to describe DAP readings as patient doses: this error is often made in project write-ups. A DAP reading does provide us with an indication of the relative sizes of doses given to different patients. It is however not a direct measurement of patient dose. In order to estimate effective doses for given procedures in a given X-ray room calibration factors can be used. One pitfall is that the recorded DAP value increases greatly with a larger beam size, whereas the beam size will not greatly affect the dose to an organ that is in the centre of the beam. DAP values are more useful if field size and radiation output factors are also recorded. One big advantage is that a DAP meter gives an immediate digital readout.

In radiotherapy, the electronic portal imaging device (EPID), rather like the DAP meter, acts 'in beam' during actual patient procedures. Portal imaging devices, originally used for verifying patient position, may now be used to provide 2-D dose

readings during LINAC treatments. EPIDs are usually of an older liquid-filled ionisation chamber or a newer flat panel solid state amorphous silicon design. These devices permit calculation of either dose rates or absorbed doses, by applying software to the image data. Like the DAP meter, there is some dependency on field size. There is also some over-sensitivity to relatively low energy photons, which can be reduced using a copper filter. Solid state EPIDs, due to their fine matrix of individual detectors, are likely to be useful in connection with intensity modulated radiotherapy (IMRT).

The entrance surface dose (ESD) is perhaps a rather more reliable measurement than the DAP as it produces equivalent dose values at the skin surface, either using TLD tablets of near tissue density or the air of a small ionisation chamber, either being attached at the point where the radiation beam enters the body. The reading is only affected by beam size to a relatively small degree, assuming that the beam is large enough to cover the TLD tablet or ionisation chamber and surrounding area. Entrance surface dose readings can be affected by back-scatter of radiation from the patient or phantom and thus a correction needs to be made for this. The equivalent doses of organs lying quite close to the skin surface, such as the thyroid or breast, can be easily estimated from the ESD after making a reduction for attenuation. Monte Carlo techniques, as mentioned above, can derive effective doses from ESD values. TLD tablets are useful in that they do not interfere with the diagnostic image or radiotherapy treatment field. However, their sensitivity varies, both between batches and with beam energy. Also, they cannot be read directly; they need to be processed to produce a visible light output, which is proportional to the radiation dose absorbed. This delay brings about the possibility of fade and experimental errors, due to the need for ancillary equipment to read the tablets.

In radiotherapy, where doses to the tumour target volume must be precise and preferably subject to less than 3% variation from the planned dose, solid state semiconductor devices (termed diodes) attached to the skin surface in vivo have the advantage that they provide rapid ESD readouts, permitting alteration of radiotherapy technique during the treatment if necessary to deliver the planned dose. Strict calibration of the devices is required, as there is some variation in sensitivity in response to field size, focus to skin distance (FSD) and use of a wedge. A semiconductor device is surrounded by a ‘cap’ whose thickness may be needed to be adjusted for different beam MV values or electron treatment. In vivo dosimetry is preferable to data transferred from the planning system, due to possible errors with the latter. ESD measurements help to check radiation output, patient positioning and the calculation of number of monitor units (MU) needed.

The exit dose, measured at the skin surface where the beam leaves the patient, is a very useful additional value, especially in radiotherapy, as it allows better estimation of the actual dose delivered inside the patient, when used in conjunction with the ESD. It also permits verification of the dose delivery calculation and checking the effect of the patient’s size and body composition on dose. Exit dose, like entrance surface dose, can be measured using semiconductor devices or TLDs. If semiconductors are used, the entrance and exit devices should not be positioned directly above each other as the former may interfere with the beam reaching the latter. The

increased use of 3-D and conformal radiotherapy has brought about the need for a high precision in dose measurement. In intensity modulated radiotherapy (IMRT) there is a need for rapid and accurate dose readings from small volume detectors, especially when dynamic multi-leaf collimation (that moves during the actual radiation delivery) is used.

13.9 Dosimetry in Computed Tomography (CT)

This is a distinct topic in radiation dosimetry with its own terms and measurements, due to the nature of the tomographic process. The technique gives some of the highest effective doses to patients in medical imaging, due in part to the speed and ease with which large volumes can be scanned. Indeed, many doses from spiral and multi-slice scanners have risen in recent years, making CT a very good area for dosimetry research. The dose received in CT depends on a large number of factors: kVp, mAs (tube current), volume coverage, pitch (CT table distance in millimetres moved per tube rotation), slice thickness, slice spacing or overlap (or 'bed indexing') and beam collimation. Many of these parameters are within the control of an operator. Image noise (which reduces image quality) can be suppressed if higher dose parameters such as high mAs, low pitch and narrow slice spacing are used. This presents a familiar conflict between image quality and dose and there has been a tendency to over-dose, including in CT of children. A common term in CT dosimetry is the computed tomography dose index or CTDI. It represents the area under the CT slice profile curve in a graph of absorbed dose (Gy) versus horizontal position across the slice (mm). CTDI is expressed in Gy divided by mAs and slice thickness. The weighed CTDI or CTDIw takes account of the fact that dose in a phantom (or a real human body) decreases from the edge to the centre. Doses in CT are influenced by the total width of the number of slices used (or the detector array width for a multi-slice scanner), as well as the mAs and kVp. Values of CTDI can be obtained 'free in air' simply by using a line of TLD tablets or a thin pencil ionisation chamber placed along the long axis of the scan profile. This may provide useful comparisons between scanners. Readings of CTDIw can be obtained by using a pencil ionisation chamber inserted into holes both at the centre and four points around the periphery of a Perspex CT phantom. Another use of a line of TLDs is to record the shape of the dose profile across a single slice, which ideally should be rectangular but is usually a curve-sided peak in practice. A further term you might encounter is the volume weighted CTDI or CTDIvol. This is the CTDIw divided by the pitch of the scanner.

As can be seen from the complexity of the topic, most people would be well advised to seek the advice of a medical physicist when recording and using CTDI data. As an alternative, some straightforward readings of equivalent doses to organs could be obtained by placing TLDs in suitable holes within the slices of an anthropomorphic phantom. This is another useful way to compare doses between different scanners or between different scan parameters. Remember that the traditional concept of entrance surface dose does not really apply in CT, due to the rotation of the

X-ray beam around the body. Some modern CT scanners provide digital readouts of CTDIvol values, although these do not always correlate totally with independent measurements. Effective doses can be obtained from CTDIvol values, by multiplying by the total length of the scanning volume and applying correction coefficients based on patient age and the body areas being scanned.

13.10 Some Practical Issues in Radiation Dosimetry

As in other forms of research, when undertaking radiation dosimetry a researcher need to remind him or herself of the following questions.

- What am I aiming to achieve with this research?
- How am I going to apply my results to clinical practice?
- Are there any errors in my data?
- What is feasible?

Complex calculations of reliable effective doses to patients are necessary if it is intended to calculate the small but real increased risk of stochastic effects such as radiation-induced cancer. But is this information vital to a department's clinical practice? Will procedures be altered as a result? Possibly yes, but realistically we are probably more interested in the relative magnitudes of equivalent doses due to different practices, decisions and equipment set-ups. For this purpose, the simple application of TLDs to patients' skin surfaces will suffice. In radiotherapy, an accurate and reliable determination of equivalent doses to at-risk normal tissues and malignant cancers is vital, owing to the large sizes of the doses and their immediate impact on patients' well-being. Choosing appropriate techniques to fit the research question is a clear message here.

When dealing with small doses, perhaps arising from scattered radiation outside the primary beam, we need to ask whether the values are actually measurable with the equipment we have available. It might be possible to make multiple exposures in order to give a detectable reading and then calculate the dose per individual exposure, but this will not be feasible if a patient is involved. Similarly, some changes in radiographic technique might not produce a measurable difference in dose. It is wise to do a pilot study to check if this is the case.

Many of the methods used to estimate doses have inherent margins of error, due to assumptions and approximations made in calculations. Thus, it is not always justified to report results to a very high accuracy; error ranges can be usefully quoted. Also, since dosimetry is based on experimental methods, there are a number of biases that can affect our results. Let us consider the variables that can affect our experimental readings in dosimetry research, namely:

- faults and fluctuations in the radiation detector(s)
- alterations in detector efficiency according to beam energy
- variations in the output of the radiation source

- errors in the equipment used to record dose values
- differences in size and composition between human bodies, and between phantoms and human bodies
- inconsistency in the use of parameters such as source-to-skin distance, collimated field size, detector positioning, beam energy and intensity, beam filtration
- variations in the intensity of scattered radiation, due to lead shielding, presence of nearby solid objects, human body or phantom size and composition.

A researcher should be aware of these variables, try to minimise them where possible and always consider them in the discussion of results.

Data errors are a feature of dosimetry, since all of the measurement devices available have their own particular shortcomings, as can be illustrated by the following brief discussion of TLDs. A LD is a widely used device and has some advantages, including small size, good sensitivity, near tissue density equivalence and a response that is not much influenced by dose rate. But it is important to be aware of the variations that can occur when using TLDs. They need a complicated 'annealing' cycle to remove past readings and may show alterations in sensitivity due to various causes. Also, they suffer from fade and their response to radiation might not be linear at large doses. The process of annealing involves heating TLDs to temperatures of up to 400 °C and then allowing controlled cooling. This process removes any residual readings from absorbed radiation and affects sensitivity. The temperature in the annealing cycle must reach equilibrium and avoid variations. It is possible that TLDs can lose sensitivity after many anneal cycles and sensitivity is also affected by the cooling rate after removal from a high temperature oven.

It is important that TLDs are handled with tweezers during a radiation dosimetry process or kept sealed in containers, as dirt contamination can affect the readings. It is always a good idea to calibrate TLDs against a reference ionisation chamber placed in the same radiation beam and using a beam energy that will be used in the subsequent dosimetry experiment. This should take place no more than 2 h or so after annealing. If there are wide variations between TLD values, 'rogue' TLDs can be removed before the actual dosimetry measurements.

It is known that TLD responses can vary by 30%, due to manufacturing variations. Another pitfall of using TLDs is that steps must be taken to identify them individually, especially when they are applied to different locations on a human body or phantom. This is not always easy when dealing with these tiny tablets. After irradiation, a TLD reader, such as those manufactured by Harshaw pre-heats the TLDs in an inert nitrogen environment in order to eliminate some low energies (which are liable to fade) absorbed from the radiation exposure and improve accuracy at low dose levels. Heating in the reader during the 'acquire' cycle emits light which is recorded by a photomultiplier and converted to an electrical signal. Following the acquisition of signal, the TLDs are annealed in the reader device as described above.

Practical issues when using air-filled ionisation chambers include the fact that sensitivity to radiation increases according to the size of the chamber. Thus, large chambers are more accurate for recording small doses, including those arising from scatter or background radiation. However, a large chamber will be cumbersome and difficult to attach with much anatomical precision to a human body or phantom. Ionisation chambers tend to be less effective at high beam energies, due to reduced absorption and the greater range of the secondary electrons produced by ionisation. Air ionisation chambers are sensitive to changes in atmospheric pressure and temperature, and are also subject to 'drift', requiring regular accurate calibration.

13.11 Summary of Effective Doses and Radiation Dosimetry

Although effective doses are widely published and give an idea of the relative risks of stochastic effects, it must be remembered that derivation of effective doses from DAP or entrance surface dose values is prone to large uncertainties. The mathematical con version values employed tend to assume an average body composition, such as the 70 kg adult male, and do not take much account of the reductions in radio-sensitivity that occur in real people from childhood to old age. While considering artificial models, we can note that the Alderson range of anthropomorphic phantoms, such as the Rando and ART, is very useful for dosimetry research but was originally designed for radiotherapy beam energies and may give overestimates of internal organ doses when used with diagnostic exposures of less than 70 kVp.

Radiation dosimetry provides much opportunity for interesting and valuable research and is expanding due to continuous development in medical imaging and radiotherapy. Original work can still be done with relatively modest equipment. Although the science appears to provide a firm foundation, a researcher should not take this apparent certainty for granted and should be aware of the pitfalls and possible uncertainties highlighted in this chapter.

13.12 Conclusion

In diagnostic imaging radiation doses are relatively small. The main role of dosimetry is to measure the stochastic risks of cancer induction. There is interest in effective doses to the whole body, measured in sieverts. These are obtained indirectly via calculations and estimates. In radiotherapy the radiation doses are relatively high and the main role of dosimetry is to ensure that radiation treatments to tumours are optimised. Unwanted side-effects of radiotherapy include deterministic damage such as skin erythema and stochastic risks such as cancer induction. There is interest in measuring absorbed doses to target tissues, which can sometimes be recorded directly. The measuring devices and techniques used in dosimetry are not infallible and all have limitations, which can provide sources of error.

Reference

1. ICRP. The 2007 Recommendations of the International Commission on radiological protection. ICRP publication 103. Ann ICRP. 2007;37(2–4):1–332.

Further Reading

Bomford CK, Kunkler IH. Walter and Miller's textbook of radiotherapy. 8th ed. Edinburgh: Churchill Livingstone; 2019.

Graham DT, Cloke P, Vosper M. Principles and applications of radiological physics. 6th ed. Edinburgh: Churchill Livingstone; 2012.

Intensity Modulated Radiation Therapy Collaborative Working Group. Intensity-modulated radiotherapy: current status and issues of interest. Int J Radiat Oncol Biol Phys. 2001;51(4):880–914.

Stabin MG. Radiation protection and dosimetry: an introduction to health physics. New York: Springer; 2007.

Clinical Audit

14

Kirti Thakor, Vicki Major, and Aarthi Ramlaul

14.1 Introduction

Clinical audits are a good way for new researchers to participate in studies, using some research methods without the complexity of ethical approval, participant information sheets or gaining consent from participants; so if you are new to research this is a good way to start. Projects involving clinical audit are normally easier to perform compared to research, if you have a limited amount of time or resources, as they usually have a set methodology, and data analysis is normally via descriptive statistics making it easier to implement and conduct whilst adding value to the improvement of services. If you are an early researcher, you can gain valuable experience from participating in these investigations.

14.2 What Is Audit?

Audit is the assessment of an activity which is measured against a national or local standard (known as the gold standard) in order to check for compliance [1, 2]. This may be in the form of, for example, guidance notes, protocol, procedure, trial specification. The aim of an audit is to achieve and maintain a high quality of care and services through the process of setting standards, observing practice, evaluating results, communication of results and when necessary implementing changes. The

K. Thakor (✉) · V. Major
Paul Strickland Scanner Centre at Mount Vernon Cancer Centre, Northwood, UK
e-mail: kirti.thakor@stricklandscanner.org.uk; v.major@nhs.net

A. Ramlaul
Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: a.ramlaul@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_14

audit is a cycle and should therefore be repeated at regular intervals. Documentation must be in place in clinical departments to assist the workforce with their audit activity and to ensure standardised practice. All such documents require ownership and recognition of their accuracy and worth. Audit is an umbrella term. It is an effective method of measurement and analysis of processes that are already in place and also provides a means of introducing improvements that can be assessed over a specific timeframe. All aspects of a quality management system can be tested for compliance through audit. Audit is an essential component of clinical governance. Under the clinical governance framework all clinicians are required to be involved in audit activity. According to the Health and Care Professions Council (HCPC), you should be participating in audit as part of your training and as a qualified radiographer [3].

14.3 What Is Clinical Audit?

Clinical audit focuses on the clinical practice of healthcare professionals to set high standards of practice. Clinical audit is defined as a quality improvement cycle that involves measurement of the effectiveness of healthcare against agreed and proven standards for high quality and taking action to bring practice in line with these standards so as to improve the quality of care and health outcomes [1]. The term 'audit' is widely used in the healthcare sector and there are numerous terms associated with audit activity. In health care most audit activity is categorised under the umbrella term of 'clinical audit'.

In the healthcare sector audits aim to provide:

- a systematic review of the clinical practice of the whole multidisciplinary team involved in the patient's care,
- a process of measuring current practice against specified standards aimed at improving patient care,
- a tool to enable healthcare professionals to disseminate good practice,
- a tool to demonstrate evidence-based practice (EBP),
- an organisational/management tool to assess activity,
- a quality assurance (QA) tool so that action can be taken to remedy discrepancies.

In England and Wales, there is a National Clinical Audit and Patient Outcome Programme (NCAPOP) managed by Healthcare Quality Improvement Partnership (HQIP) [4]. Local clinical audits should include a range of staff groups to provide a depth of understanding, so whether you are a student or a qualified radiographer you should be participating in audit in some way. Being a systematic process, a clinical audit allows us as radiographers to assess whether what we are doing is what we should be doing in relation to the service we provide to our patients and service

users [4]. By performing audits, we are ensuring that patients are receiving an effective service or treatment. Audits are a useful way of evaluating our services and their impact on patient care, outcomes and service delivery.

Clinical audits are very important in both medical imaging and radiotherapy practice. Without audits it is difficult to understand the current standards of work within a department and what requires improvement. For example, waiting times in an Accident and Emergency (A&E) X-ray department. How long does it take from a patient's arrival time in the department to the time their images are available on the PACS? Another important area which is commonly audited is the rate of rejected radiographs in imaging departments. This involves considering the number of times an examination is repeated before the final images are sent to the PACS, and the reason for the repeated radiographs. In radiotherapy, you could audit the time taken from a patient's referral for radiotherapy treatment to their first actual radiotherapy session and assess the consistency in referral times, or, if there are delays, explore the reasons for the delays. Another example could involve the assessment of the number of patients who develop skin reactions within a set timeframe following radiotherapy treatment for breast cancer. Without carrying out these audits it is difficult to highlight the problems which are likely to occur when providing these services or making recommendations to improve such services. Getting involved in clinical audits as a Band 5 radiographer helps you to understand the demands of the services you work for and gets you started in feeling involved as a part of the department. This involvement should inspire you to make a difference. If you have a chance you should get involved with audits as a student. This will help you decide on the type of audits you would like to conduct when you qualify. The experience will also help you to learn the stages of the audit cycle as well as how to analyse audit data and report on the results. Audits make feasible undergraduate projects because you can undertake them within a set period of time, as a small-scale project.

14.4 Forms of Audit

There are many ways to conduct an audit and they can be categorised depending on the method in which the data is to be collected.

1. Compliance audit: this involves ensuring compliance with a set standard and could be in the form of professional guidelines, national protocols or local policies and procedures.
2. Audit trail or process audits: this involves assessing the results of an audit from input to output to ensure that the information it yields are being effectively and efficiently managed and explained. For example, a patient from a radiotherapy treatment localisation or diagnostic scan to provision of follow-up with a view to improve the efficiency or effectiveness of the process.

3. Improvement audit: this involves using either of the audit types above to review a process or procedure to identity improvements. This audit process is used primarily in an area where an issue has already been identified and a systematic approach is required to implement change.
4. Documentation audit: this involves a review of a specific document to ensure that the content is appropriate and relevant to current practice.

14.5 How to Plan and Conduct a Clinical Audit

Before conducting an audit ensure that you know which type of audit you would like to undertake and that you have a clear strategy in place. Consider whether you will do the audit retrospectively or prospectively? Retrospective audits involve auditing existing information sources and are less complicated as there is no direct patient involvement. The disadvantage of retrospective audits is that some of the data required could be absent or incomplete.

Prospective audits gather more complete and accurate data sets but can take longer since a predetermined amount of data is collected over a period of time. They provide a more structured audit giving greater depth. Often the audit cycle starts with a retrospective audit showing areas for improvement followed by a prospective audit after recommended changes have been made. Prospective data is more accurate. It allows for real time data which reflects current, rather than historic, practice. Case notes are easily assessable and the pro forma can be designed to ensure that all relevant data are collected. Again this is an activity in which time must have a bearing on sample size. If you are reliant on others to record the data for you, the data collection again may be inaccurate or incomplete. Colleague cooperation will be essential. Audit should not inhibit normal clinical activity and data collection by colleagues should not be laborious.

Before performing prospective audits, you must check that research is not being undertaken, which is especially important if there is any direct patient involvement or change to the service a patient would receive. This can affect the nature of the patient involvement including how much time is being asked of the patient to participate in both the research and audit activities. Chapter 6 gives detailed guidance on ethical considerations.

Each department should ideally have their own audit plan template which must be used before starting a new audit. The audit plan template allows you to identify the question to be answered from your audit, the methodology that you will use including the timeframe to carry out the audit.

In addition to mandatory audit topics, the choice of topic should be based on the standard criteria of areas with high volume, high cost and high risk. Audit is an effective tool for change in specific areas where compliance is weak and an improvement in practice or assessment of process is required. Audit activity needs to be appropriate for an individual's level of influence, ensuring results will impact on changing activity. Traditionally persons independent of the task being audited undertake audit activity, for example, finance audits. This is to ensure that activity is

transparent and without bias. This method would be effective if a review of services was taking place or several work areas are being compared for efficiency. However, within the healthcare sector it is common for an individual to assess an activity that relates specifically to his or her own level of work, or an activity that is having a fundamental effect on how they work. In this way the audit is more likely to influence and change current practice. Care must be taken not to conduct an audit where criticism or blame is directed at another staff group or department without their knowledge or involvement. Joint audit is far more effective where the goal is to improve the quality of care provided and not to pass on blame.

14.6 The Audit Timeframe

The timeframe in which the audit is to be conducted has a major impact on the audit itself. If the data collection period is too long, interest will be lost and data may no longer be accurate. Any enforced time constraints should be respected; hence, the scope of the data collection and type of analysis required need to be considered at original audit design. There is little point in generating 6 months' worth of data requiring analysis of thousands of samples unless you have the means by which the statistical analysis can be conducted. Do not design separate data collection tools if the information is already recorded elsewhere. The sample size should be small enough to allow for rapid data collection but large enough to be representative. When designing an audit, potential seasonal fluctuations should be considered, for example, patient throughput may increase following the summer vacation period, hence a larger sample of data may be required to ensure data collection accurately reflects the current situation. A multidisciplinary approach will ensure that such patterns are recognised and the audit designed appropriately.

The aim of the audit must be defined from the outset. An awareness of what is to be achieved will keep continued focus on the audit and ensure that activity is worthwhile. An accurate definition of the audit question and the defined scope are essential to evaluate results within the context of the actual data collated. The audit questions need to be specific and unbiased. The data collection should be transparent to present facts that accurately reflect conclusions being made. The data collected should answer the audit question.

14.7 Who Should Be Involved in Clinical Audit?

A clinical audit should involve several staff groups within the healthcare environment. Patients, service users, carers, relatives, commissioners, managers and board trustees could also contribute. Having different groups of people involved could enrich the interpretation of the data since it will be viewed from a different perspective. Furthermore, for audit activity to be effective and meaningful it requires a coordinated multidisciplinary approach. There will be a named clinical governance lead for most departments performing a clinical activity. Departments with quality

systems in place based on the International Organisation for Standardisation (ISO) requirements will also have a named quality manager to manage the approach towards audit activity. Audit is an effective method of measurement and analysis of processes that are already in place and also provides a means of introducing improvements that can be assessed over a specific timeframe. This will ensure that the requirements of the Trust, directorate, local users and patients are continuously monitored to ensure a quality service is being delivered.

However, undergraduate student radiographers are increasingly conducting audits that involve collecting retrospective data. These audits typically involve patient or radiographer data only and not the patient, service user or radiographer directly. Students wishing to conduct an audit using retrospective data must write to the department manager to seek permission to access the data required. The permission must be granted before data collection can begin. The department manager would normally recommend that the PACS manager be available to help you retrieve the data. If undertaking an audit as a student radiographer, some hospitals may require you to complete a clinical audit proposal form and submit it to the appropriate person prior to the audit being approved. The department manager would be able to advise if this is the case.

14.8 Management of Audit Activity

Many Trusts will have a department dedicated to clinical governance/audit activity.

These departments will play a central role in project design, project management, data collection, data analysis and report production and often hold a list of specific audit activity that has to be conducted. This is often related to the current political agenda, for example, infection control audits. Diagnostic imaging or radiotherapy departments will have a named clinical governance lead. Departments with a quality management system (QMS) will have an audit plan for the year that ensures the system documentation is audited on a regular basis. This will be managed by the quality manager. Other departments may favour a team approach with multidisciplinary team meetings to coordinate audit activity. Whichever method is used, the audit reports and recommendation will need to be managed by named individuals to ensure the audit process is effective and efficient at implementing a quality service. Under the clinical governance framework all clinicians are required to be involved in audit activity. For all other clinical staff, although it is not a compulsory activity, audit activity is often assessed during the performance appraisal process with managers.

14.9 The Audit Cycle

Clinical audit is not a one-off procedure but is a continuous cycle to ensure a high standard of service is delivered. Attention needs to be paid to all aspects of the cycle in order to have a successful/safe and compliant service. Successful audits follow a cycle with defined stages of completion. Figure 14.1 shows a typical audit cycle.

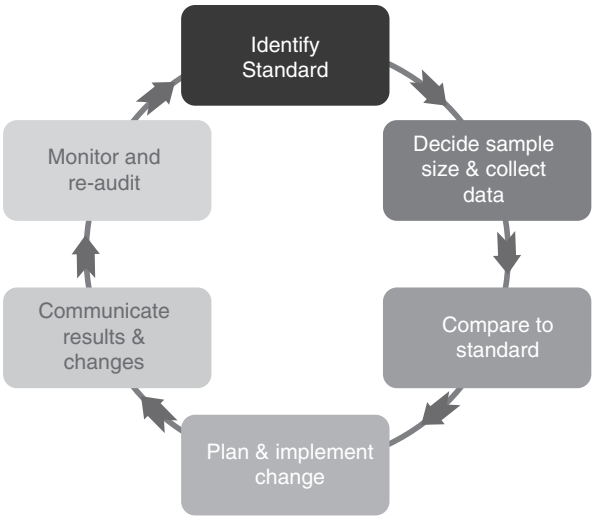


Fig. 14.1 The audit cycle

14.10 Identifying Standards

When you have decided on your audit you will need to refer that to a standard. The standard for an audit, as previously mentioned, may be derived from national guidance, local protocols, regulatory compliance, best practice, local expert opinion or previous audits. All audits require standards to measure the accuracy of the audit. Standards can be set nationally or locally depending on the improvement or clinical outcome being measured. You should ask members of staff in the department if the audit has been conducted before and if a local standard exists. High standards may only be achieved by a few, but they could encourage improvement. Where patient safety is a factor, for example, an audit to assess whether the World Health Organisation (WHO) checklist has being completed for all sedated, analgesia and anaesthesia patients in radiology, the target must be 100% [5]. The Royal College of Radiologists (RCR) in conjunction with HQIP has a list of completed projects with standards which can be used for local audit. You should have a look at these to find out if your audit is on the RCR list.

14.11 Indicator to Be Measured

The indicators are measurable variables that should be identified during the planning stage. If you are undertaking a local audit which has been performed before, you should use the same indicators. The indicator can be expressed as an absolute number, percentage, average or rate. Audits may need several indicators to assess a complex process. For example, you want to highlight to management that an extra X-ray room for the A&E X-ray department is required. To justify this you can do an audit to determine the work flow in the A&E X-ray department

at different times of the day. You would have to consider the average number of patients coming in, including the minimum and maximum number at various time points. You may also need to look at the transportation of these patients such as, are they ambulant or on a wheelchair or trolley, and how long each one occupies the X-ray room for. Or in the case of radiotherapy, your department may have introduced a new type of scan prior to radiotherapy treatment, such as a new MRI sequence to demonstrate the position of the brachytherapy rods. In this case each scan will be assessed for the accurate position of the brachytherapy rods. You will have to record the number of patients having an MRI scan prior to radiotherapy treatment and the percentage of patients with accurate positioning of the brachytherapy rods.

14.12 Data Collection

The sample size is very important for the accurate results of any audit, along with the timescale used as this will have direct bearing on the sample size. If the data collected in the audit is to be representative of a wider population, then it is important to have a representative sample size. During any audit you want to ensure that you have high confidence of accuracy levels in your results to give an overall picture of the entire service. There are calculators available online that can be used to calculate the sample size required in order for the results of the audit to yield a high confidence level, i.e., a confidence level of 95%. A larger sample size is required to identify a small margin of error which is usually considered in the range of ±2.5%. For example, in a population of 500 people you may require a sample size of 380 to give a 95% confidence level with a ±2.5% margin of error, whereas a sample size of 80 would give a ±10% margin of error. The results will be more powerful if the margin of error is smaller.

14.13 Analysis

If the standard is attained, it is an assurance of the quality of the service provided and gives positive reassurance that no change is necessary. The audit can be repeated at a relevant interval though, such as 1 year or 6 months, to check for consistent compliance. When results do not meet the standards, all possible reasons for not meeting the standards should be examined such as target level, system, process, technical reasons and so on. The possibility of sampling bias must also be considered before recommending any changes. Only after this analysis should a change in practice be implemented, if necessary. It is always best to have a list of suggested changes put together by the team in order to improve outcomes of a re-audit. Following any change in practice, a follow-up or repeat audit is required at an appropriate time interval. An improvement in service can only be proved following the repeat audit and only if it shows an improvement in results.

14.14 Audit Report

The audit must be written up at the end of the process in the form of a report with aim/s, methodology, results, analysis and conclusion along with recommendations for change in practice, policy or protocol. Accurate recording of the audit procedure allows the same audit to be repeated by any individual at a later date. The report should be discussed with the relevant staff members and changes can be planned and then implemented in the next cycle. The structure and presentation of the report is similar to that of the dissertation but shorter in length and more concisely written. If you are conducting the audit as part of your undergraduate training course, follow the layout that is required by your educational institution but typically the audit report will cover the following areas.

- *Contents page*: This page lists the sections that are included within the audit report.
- *Executive summary*: This is the summary of your whole report. It is usually in the region of about 300 words. It allows the reader to identify the key information. It should include a summary of background information, key findings and recommendations. It should also include aims and objectives and keywords. Much like the abstract in a dissertation, the executive summary is written last.
- *Introduction including background and rationale*: This section should include essential background information which puts the audit in context; it should set the scene and describe the reasons for undertaking the audit. Recent audits or related publications should be critically examined in order to justify the need for the audit.

  Perceived benefits to practice should be included as well as the standards and guidelines you are comparing practice against.

  Aims and objectives should be clearly laid out in this section. Objectives should be a statement of what you are trying to achieve by undertaking the audit and should reflect a commitment to improve practice.
- *Methodology*: This section should include all the details of the data that is to be collected during the audit. It must be written in a clear and logical manner so that if the audit is to be repeated by another person, it would enable them to conduct it in the same way. The following points are not indicated as subheadings to be used but rather as a guide as to what should be addressed within this section:
  - *Criteria*: Here you should consider what is to be measured? You can refer to your aim and objectives to help you with this information.
  - *Ethical considerations*: Here you should consider confidentiality and sensitivity of data. Acknowledge that ethical approval is not required in clinical audits; however, ethical considerations such as confidentiality, anonymity and data protection need to be explicitly made.
  - *Data collection*: Here you must consider who is/ are to be involved in the supervision of the collection of the data. In the case of students conducting the audit, most hospitals will nominate a qualified member of staff to guide the student during the data collection process.

  - *The design of the audit*: Here you should consider whether the audit is prospective or retrospective and the type of audit design. You should also state justification for why you think so.

    Data sources: Here you must specify where the data was obtained from, for example, PACS, RIS or DAP meters in retrospective audits or from radiographers, patients or service users if the audit is prospective. You can also provide details on who collected the data and whether the data were validated by the Trust staff involved in the audit. In addition, you should specify the time period that the data was collected and why.

    Sample: Here you can consider questions such as, 'what is the sample size?', 'how was this determined?'

    Procedure: Here you are required to give a clear outline of the procedure you followed in collecting your data by providing a step by step explanation of the method that was used. The procedure is important for replicating the study in case of a re-audit. Information on data analysis methods used should be included, as well as details of computer packages if used.

- *Results*: This section should report on the compliance against each of the standards/audit measures that you are comparing your practice against. This should include the number of cases that were compliant and the percentage compliance. There should be further investigation/explanation of where any non-compliance has been identified.

  Present findings in a logical, sensible order. Be selective with the use of charts, remembering to use the most appropriate method to present the data, e.g. pie charts to show proportions and bar charts for easy comparison between different areas/time periods. It is important to be consistent in presenting results; do not mix bar charts, pie charts and line graphs for similar data. You may refer to HQIP (2018) [6] for help with analysing and presenting your results.
- *Analysis*: This is the part of the audit which most clearly demonstrates your ability to discuss, evaluate, analyse and interpret the results in relation to the original reasons for undertaking the audit and findings of previous audits.

  Take care to critically analyse your results and not simply describe them. Refer back to the results and your initial rationale for conducting the audit. Where your results differ from previous published evidence, you should explain and justify why this might be. Include any real or perceived weaknesses of your audit design.
- *Conclusion and recommendations*: The conclusions should summarise the key findings from your results. This should identify the areas of good practice where standards are being met and identify the areas for improvement where there is a gap in compliance. This is the only section where you can express an opinion based on the application of the results. This section should summarise the extent to which the aims were achieved. It should end with recommendations for implementation or a re-audit. In the case of a re-audit being recommended, a timeframe for this to take place should be specified.
- *References*: This should be a list of the resources, including publications, that were used within the report.
- *Appendices*: These should be presented in numerical order as cited within the body of the report.

14.15 Communication of Results and Any Changes

Once an audit has been completed it is important that the outcomes of the results are disseminated to the relevant groups along with any actions identified. This will affect the success or failure of the audit if it is to be repeated. This can be done in the form of a simple discussion in a staff meeting, email or a presentation depending on the nature of the audit. If the current audit has not met compliance, then further action is required, and staff should be made aware of what is expected from them. If the current audit has met compliance, staff also need to be made aware in order for them to maintain their standards of practice in running a successful department.

At this point you should announce whether the audit would be repeated or closed down.

14.16 A Step by Step Guide to Help You Design an Audit

Figure 14.2 below gives an outline of a ten-step guide to designing and conducting a clinical audit. Figure 14.3 gives an example of how to implement the steps within the guide.

14.17 Data Protection and Information Governance

When performing a clinical audit it is important that you gain relevant permissions prior to starting. An audit will not require ethical approval or consent from patients because it is a review of the care provided but permissions may be required to access patient information. Wherever possible, data collected for audit purposes should be anonymised. Where data cannot be anonymised, it must be stored appropriately and not removed from the hospital/clinic where it was accessed and collated. Any reports produced must be anonymous and any non-anonymous data must be destroyed once analysis of the data has been completed. Patients must not be individually identified. See also Chap. 6.

Prior to the collection of data, you may need to consult the data protection officer and the data may need to be collected by a responsible person. When setting up the clinical audit the need for a data protection impact assessment should be considered. This is a process which helps you to identity and minimise any potential risks of your audit, including considerations regarding how the data will be collected and stored.

14.18 Statutory and Mandatory Requirements for Clinical Audit

Healthcare professionals are expected to take part in regular local and national clinical audits [3]. When clinical audits are carried out in accordance to best practice it:

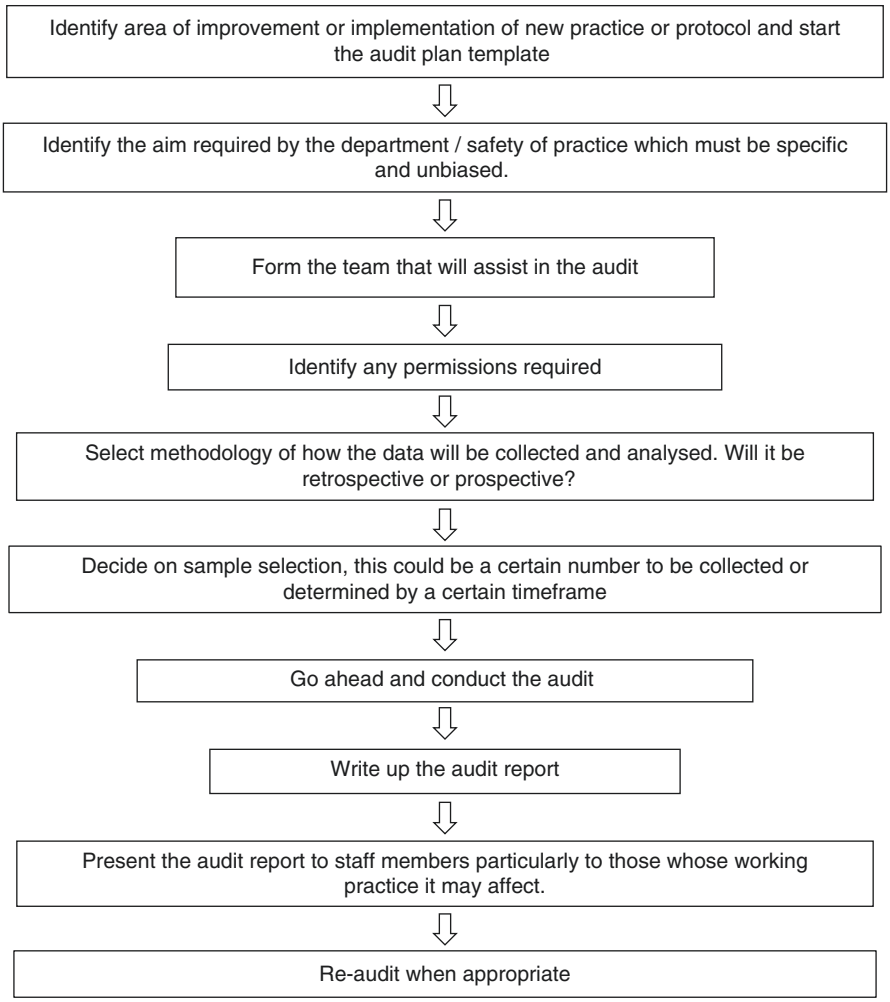


Fig. 14.2 Ten-step guide to conducting a clinical audit

- Improves the quality of care and patient outcomes [2, 4, 7]
- Provides assurance of compliance with clinical standards [2, 4, 7]
- Identifies and minimises risk, waste and inefficiencies [2, 4, 7]
- Complies with the ionising radiation (Medical Exposures) Regulations 2017 (IR(ME)R17, 2017) regulations [8] which states that employers' procedures must include provisions for clinical audit as appropriate.

The National Health Service (NHS) standard contract forms the agreement between commissioners and providers of NHS funded services who must do the following.

- Participate in national clinical audits within the NCAPOP relevant to their services.

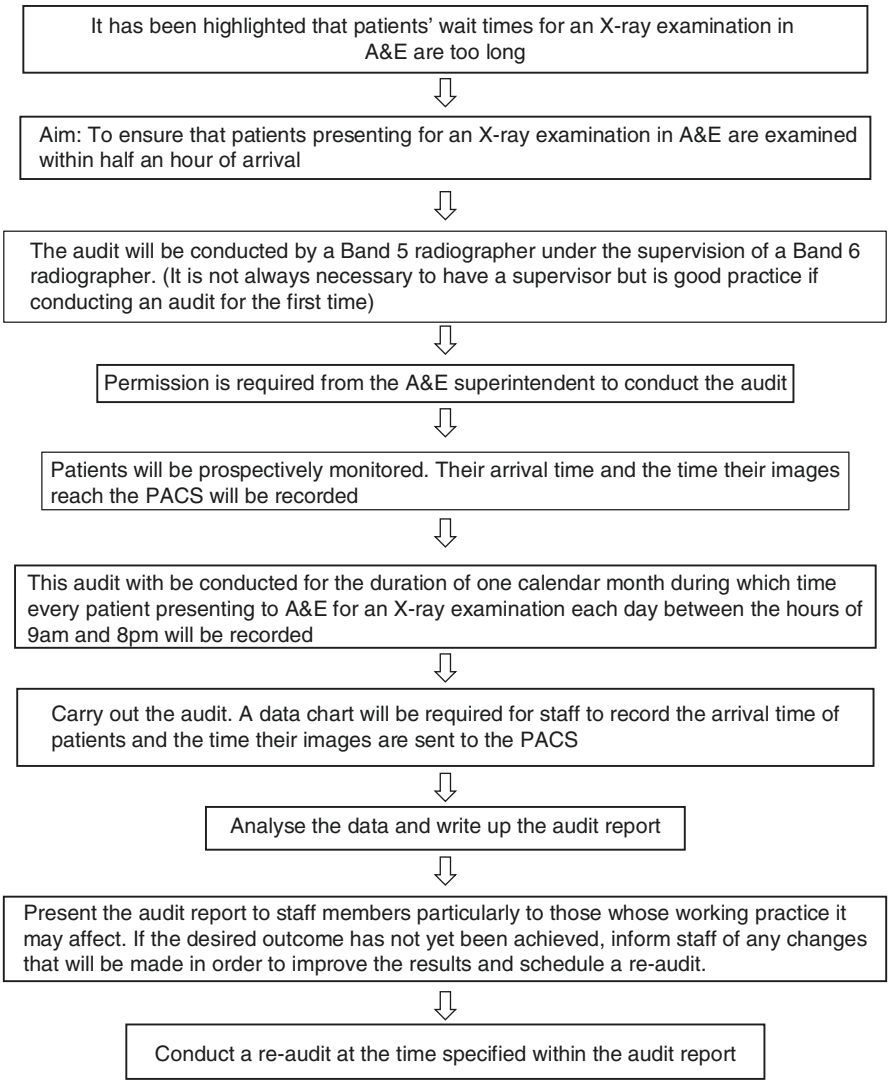


Fig. 14.3 An example of implementing the ten-step guide

- Make national clinical audit data available to support publication of consultant level activity and outcome statistics.
- Implement and/or respond to any outcome measures.
- Implement ongoing programmes in accordance with good practice.
- On request, provide the coordinating commissioner with the findings of any audits carried out especially in relation to locally agreed requirements such as Commissioning for Quality and Innovation (CQUIN) audits.

NHS England and NHS improvement, and National Institute for Health and Clinical Excellence (NICE) [7] provide the clinical governance frameworks which

feature strongly within NHS Trust clinical audit programmes. In addition, the Care Quality Commission (CQC) requires healthcare providers to constantly monitor the quality of their services. Access to the guidance from these organisations is available freely within the public domain.

14.19 Succeeding in Clinical Audits

In both diagnostic imaging and radiotherapy, someone embarking on clinical audit needs to consider the following principles.

- Do ensure that you have the permission of clinical managers, including radiologists or oncologists if necessary, to undertake the audit. It is best to obtain a written permission letter.
- Do consider the timeframe needed to undertake the audit. Although a 'snapshot' of clinical activity or performance at a certain point of time may be useful in many situations, such as in ongoing quality assurance, it may be necessary to collect data over a period of several months. Is this feasible for you?
- Do remember that you may not always be present at the clinical site to collect data. In your absence, are there people who have the time and ability to undertake data collection for you? And if so, will the data be collected in the same way that you would do yourself?
- Do check that there are measuring instruments or procedures in place which are capable of gathering the data you need. Also, are you capable of using them, or will you need training?
- Do not assume that retrospective data will always be available, or accurate.

14.20 Conclusion

Clinical audits act as a starting point for novice researchers and play an important role in measuring compliance and quality. To ensure a successful clinical audit it is important that the right persons are involved and that efforts are focused towards the question that needs to be answered. You must pay careful attention to the timeframe of data collection and ensure that your dataset is large enough to give you a high confidence level in your audit results. The audit report should clearly communicate the conduct and findings of the audit. The results and recommendations must always be disseminated within the department team and re-audits should be carried out as much as it is required in order to prove that standards have been met for the maintenance of a quality service.

References

1. Burgess R. New principles of best practice in clinical audit. Oxford: Radcliffe Publishing; 2011.
2. Healthcare Quality Improvement Partnership. Clinical audit. A manual for lay members of the clinical audit team. 2012. https://www.hqip.org.uk/wp-content/uploads/2018/02/developing-clinical-audit-patient-panels.pdf. Accessed 17 Oct 2019.
3. Health and Care Professions Council. Standards of proficiency-radiographers. 2018. https://www.hcpc-uk.org/resources/standards/standards-of-proficiency-radiographers/. Accessed 20 Oct 2019.
4. Healthcare Quality Improvement Partnership. Accessing national clinical audit and patient outcomes programme data. 2017. https://www.hqip.org.uk/national-programmes/accessing-ncapop-data/#.Xam972dwaM8. Accessed 17 Oct 2019.
5. Royal College of Radiologists. Sedation, analgesia and anaesthesia in the radiology department. 2nd ed. 2018. https://www.rcr.ac.uk/system/files/publication/field_publication_files/bfcr182_safe_sedation.pdf. Accessed 29 May 2019.
6. Healthcare Quality Improvement Partnership. An introduction to analysing quality improvement and assurance data. 2018. https://www.hqip.org.uk/wp-content/uploads/2018/10/final-an-introduction-to-data-analysis-october-2018.pdf. Accessed 17 Oct 2019.
7. National Institute for Health and Care Excellence. How to use quality standards. 2019. https://www.nice.org.uk/standards-and-indicators/how-to-use-quality-standards. Accessed 17 Oct 2019.
8. Ionising Radiation (Medical Exposure) Regulations 2017. Legislation.gov.uk. 2019. http://www.legislation.gov.uk/uksi/2017/1322/pdfs/uksi_20171322_en.pdf. Accessed 31 May 2019.

Part V

Data Collection Methods and Analysis

Quantitative Methods and Analysis

15

David M. Flinton and Christina Malamateniou

15.1 Introduction

Students often struggle with choosing the correct data analysis test tool. It is thus assumed that readers may have limited prior knowledge in this area. This chapter, therefore, is written and delivered in a basic, explanatory manner. We recommended that you consult it as a basic toolkit. Although it has all the information you may require for an undergraduate or master's level research project, you could also consult additional texts for more sophisticated data analysis.

The chapter is divided into various sections, each with information you will need to understand in order to carry out and interpret the results of quantitative data analysis. A summary of key points is provided at the end of each section. To help with your understanding of some of the more common tests, worked examples of analyses using SPSS are provided as an appendix at the end of this chapter.

15.2 Types of Data and Analysis Techniques

Every different type of collected data lends itself to a specific type of analysis. Data can exist in a variety of forms: qualitative research tends to work with words (spoken or written) and is analysed using pattern recognition (thematic analysis) techniques, discussed further in Sect. 10.2 in Chap. 10; quantitative research uses numbers (See also Chaps. 8 and 16). But even this statement is a simplification because words can be turned into numbers, i.e., we could count how often the word radiographer was used in an interview.

D. M. Flinton (✉) · C. Malamateniou
School of Health Sciences, City, University of London, London, UK
e-mail: d.m.flinton@city.ac.uk; christina.malamateniou@city.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_15

When looking at quantitative data, which this chapter focuses on, we can split the data into two groups: continuous and discrete data. Each contains two sub-classifications of data as shown in Fig. 15.1.

The type of data dictates the method of description and analysis. Therefore we need to be able to distinguish between these forms of data.

- Nominal data

This can best be thought of as labels that should be mutually exclusive and have no numerical significance. The radiography profession in the UK consists of two different Health and Care Professions Council (HCPC) registered professions: diagnostic radiographers and therapeutic radiographers. This is an example of nominal data. You are either on the diagnostic or the therapeutic register. We would code the data using numbers, e.g., diagnostic radiographers = 1 and therapeutic radiographers = 2, but the number has no numerical significance. It would be possible to count the number in each category to represent the sample. This example is not a particularly good one as some radiographers are on both registers. To overcome this we could create a third category: both. Lots of other examples of this type of data exist such as gender, religion, diagnosis and area being treated/X-rayed. If data only consist of two categories it may be referred to as 'dichotomous'.

- Ordinal data

With ordinal data the order of the values is known, but the differences between each one are not known. An example of this is seen in data from the TNM cancer staging system, the higher the T number the bigger the cancer. Consider two patients with breast cancer; one with a T1a and the other T1c. T1a means the

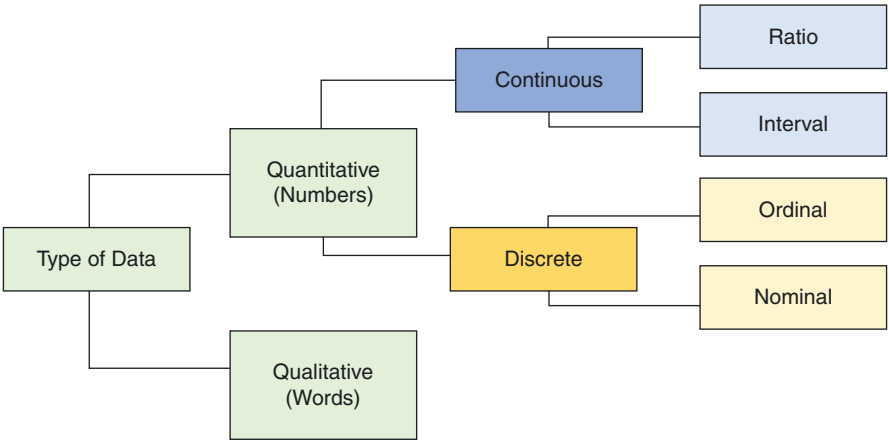


Fig. 15.1 Types of data

tumour is between 0.1 and 0.5 cm. T1c means the tumour is more than 1 cm but less than 2 cm in diameter. In this instance we can say a patient with the T1c tumour has the bigger tumour. This means that we can order the data (biggest to smallest). We cannot comment on the exact difference between the data because that information is missing.

In research terms this form of data tends to come from Likert scales. These tend to measure concepts such as satisfaction, pain or discomfort. A Likert scale has a closed-ended question followed by a rating scale. An example of a Likert scale is shown below.

	Very limited	Limited	OK	Good	Excellent
How would you rate your understanding of cone beam CT?	①	②	③	④	⑤

Most Likert scales have an odd number, typically 5 or 7 to allow a neutral response; in some instances you may see them with even numbers of options so as not to allow for a neutral response.

- Interval data

With interval scales we will know both the order and the exact differences between any values. An example of an interval scale is temperature using the Celsius scale. If the temperature last night was 2 °C and the day before it was 4 °C then we know which night was colder, and we know the exact difference in temperature between the two nights was 2 °C. From this form of measurement we get a lot more information than with either nominal or ordinal data. Negative values make sense, but there can be a problem with interpretation in some instances. If we look at the temperatures again you might be tempted to say that last night was twice as cold as the night before as the value appears to have halved. Interval scales do not have a true zero. In this instance 0 was arbitrarily set as the melting point of water and because of this we cannot use ratios as they make no sense.

- Ratio data

This form of data is like interval data in that we know both the order and exact differences between any values but, because the scale has a true zero, we can use ratios. Going back to the temperature example, temperature can be measured in Kelvin; it has an absolute zero. 2 °C is approximately 275 K and 4 °C is 277 K, so it is obvious looking at these figures that one temperature is not twice that of the other. Other examples of ratio scales are weight, blood pressure, length and time; if one patient awaited 5 min for an X-ray and another 10 min, the second patient waited twice as long as the first. A summary of the differences between the four different types of data is seen in Fig. 15.2.

Type of data	Possible to count	Possible to order	Know the exact difference between each value	Use ratios (multiply and divide)
Nominal	Yes	No	No	No
Ordinal	Yes	Yes	No	No
Interval	Yes	Yes	Yes	No
Ratio	Yes	Yes	Yes	Yes

Fig. 15.2 Summary of types of data

15.3 Using Summary (Descriptive) Statistics

While graphical methods give a good visual impression of what is going on, it is also useful to use numerical characteristics of data: summary or descriptive statistics. Some of these summary statistics are commonly encountered in everyday practice, such as averages, whilst others are a little less intuitive, for example, standard deviation. All of them relate to aspects of what the data look like when plotted in the form of a graph. Once you understand the way that they work it is possible to see how they can then be used to analyse data in more sophisticated ways. Sometimes symbols such as Σ or μ are used to represent these values. Be aware that they can vary between textbooks and websites. The important thing is to make sure that you know what the summary describes and not just which symbol is used to represent it.

There are two things that summary statistics are concerned with. The first is trying to describe where there are concentrations of measurements. If a lot of measurements crop up in the same position, then perhaps we can describe this with a single number that is close to the other numbers in the data. The second is to try to describe how spread out a collection of numbers is. If lots of measurements are very closely grouped, then this is a different pattern from a very loose grouping. Figure 15.3 shows a histogram, which has data spread across a range of values from 1 to 9. The measurements are grouped about the value in the middle, 5, and the spread of the data tails off reasonably evenly in both directions. Figure 15.3 shows the final dissertation marks of year three radiography students in one university on a given year. In Fig. 15.4 the marks of the same students, in their clinical knowledge vivas, are presented, where a novel marking scheme was trialled, to give students extra marks for coherence of their response as well as for content.

Compare Fig. 15.3 with Fig. 15.4 and consider the differences in the shape of each graph. They both have the same number of observations recorded on them as they refer to the same numbers of students. Note that the scale on the vertical axis is different. In Fig. 15.4 the data are not as spread out; they are grouped around a value of 7, compared to a value of 5 for Fig. 15.3. The two histograms are different, and it

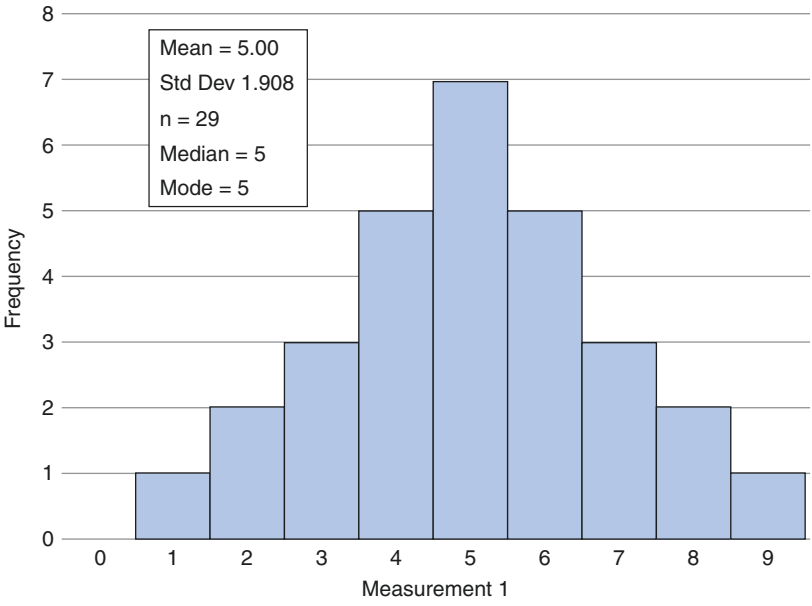


Fig. 15.3 Histogram A

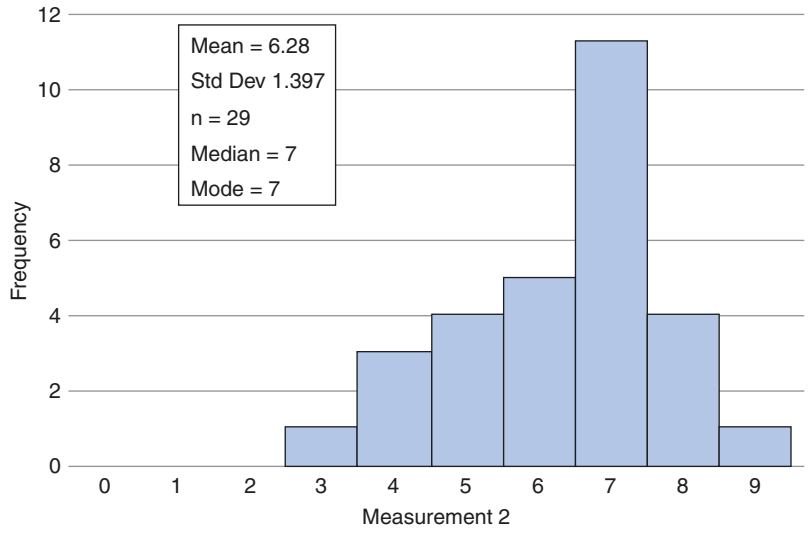


Fig. 15.4 Histogram B

is the job of summary statistics to show these differences in the form of numbers. These histograms are used to introduce some basic summary (descriptive) statistics.

15.3.1 Measurement of Central Tendency

The mean is commonly known as the average. It can be calculated for any set of data (interval or ratio) if you add the values up and divide by the number of data entries. If you look at the two histograms, you will see the mean values for the numbers used to generate each histogram. There are some other values, but you need not worry about these at this stage. The mean changes as the position of the peak of the histogram changes; note that it does not track it exactly. Move the peak to the right, as for histogram B compared to histogram A, and the mean increases, moving the mean to the right on the horizontal scale. Be aware that using the method for calculating the mean given here gives the 'arithmetic' mean (there is also a variation called the geometric mean which does not concern us here). For symmetrical data the mean, mode and median are all the same as shown in Fig. 15.3.

The median refers to the number in the middle of a list if all of the bits of data are placed in ascending order. If the middle of a list lies between two numbers, we take the mean of the two numbers, i.e., go for the mid-point between the numbers. A clinical example of this could be measurement of aortic diameters in CT angiography by two observers. The data in Fig. 15.4 are skewed; there is more data to the right than the left. The median lends itself to use with ordinal data and skewed interval/ratio data.

The mode is the number that appears more often in a set of data. For both histograms there is one mode; 5 for histogram A, and 7 for histogram B. The mode corresponds to the peak of the graph. It is not always the case that there is one peak (unimodal), so the mode needs to be used with caution as a second smaller peak would not be obvious by using the mode value on its own.

If there is more than one mode (peaks) in a graph, then the data sample is described as multi-modal: each obvious peak value would be given for a better description of the shape of the plot. A good example of this in medical imaging might be when looking at the energy of events detected from a radioactive source that emits gamma rays at different energies. A source such as gallium-67 emits three distinctly different photon energies; a histogram of energies detected would show a mode corresponding to each of the energies. The mode lends itself for use with nominal data although can be used for skewed interval/ratio data.

15.3.2 Measures of Spread

As discussed, there are different types of measures of spread, to how spread out a collection of numbers is. You can see from the two histograms that the two sets of

data are spread over a range of values, 1–9 for histogram A and 3–9 in histogram B. For histogram B the measurements are more tightly clumped around the peak than for histogram A. The tightness of this clumping is part of the pattern that the data make. This feature needs a different summary statistic; one that measures the spread of the data.

Standard deviation and variance are two common methods of describing how spread out the data are. It is important to notice that it is the spread around the mean and not any of the other measurements of central tendency (median or mode). Look at the histograms again. Just below the values for the mean are values for standard deviation and the shape of the graph is again reflected by the values. For histogram A the data are more spread out and the standard deviation is higher than for histogram B. The greater the degree of spread in the data, the higher the standard deviation will be.

There is no mention of variance in the diagrams. It is very easy to find the variance: simply square the standard deviation. You might wonder why we have two different ways of expressing the degree of spread when one is based on the other. You will need to get a little more mathematical to see the reasons for this, but at this level it is enough to say that there are good reasons and that the standard deviation is usually the better value to quote as it is expressed in the same units as the mean.

15.3.3 Symmetry

There is something else we can say about data concerning the shape of a graph. This relates to the way that data are distributed either side of a mode. If the data are spread equally on either side of a mode then the pattern is symmetric, if not the pattern is asymmetric. If there is asymmetry in the pattern, then the data can be called skewed as in Fig. 15.4.

It is important to identify skewness as it indicates whether it is valid to use certain statistical tests. When looking at Fig. 15.5 notice what effect skew has on the position of the measures of central tendency relative to the peak of the graph. If there is no skew, then all the measures of central tendency and the peak are in the same place. Skew, when present, can occur in two directions; the greater the degree of skewness, the less reliable the mean is as a way of describing where most of the

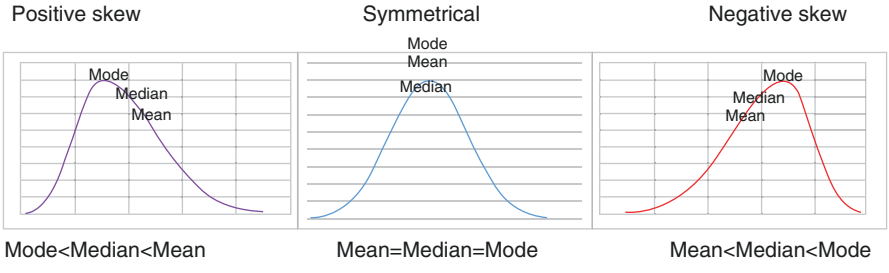


Fig. 15.5 Symmetry and asymmetry of the distribution

data sit. Consider the example below showing the salaries for six radiographers in a small radiography department with skewed data. As a measure of central tendency, which is most accurate?

Band 5	Band 5	Band 5	Band 5	Band 5	Band 8
£23,000	£23,000	£23,000	£23,000	£24,000	£72,000

The average salary of the radiographers is £31,333.

The median salary of the radiographers is £23,000.

Most radiographers (mode) earn £23,000.

In general the median does not move from the peak as much as the mean does when the data are asymmetric, and so in these cases can be thought of as a better measure of central tendency. Remember that if one or more extra modes exist then this complicates things and may make the mean and the median redundant for measurement of central tendency. Using the example of a gamma ray emitting source such as gallium-67, the mean or median gamma ray energies are meaningless because neither of these measures will correspond to any of the particular emitted gamma rays; they occur in the gaps between modal peaks. In this case these measures of central tendency do not relate to any real aspect of the overall pattern of distribution.

Another useful graph to describe data is a box and whisker plot. A box and whisker plot is made up of three elements: the whiskers that describe the range of the data (maximum and minimum value); the box that describes the data's interquartile range; the middle 50% of the data and a bar which represents the median value. In Fig. 15.6 we see two box plots. The box plot on the left has data that is approximately symmetrical as the median is approximately in the middle of the box and the box is in the middle of the whiskers. On the box plot on the right this is not the case and we can see we have a bigger range of data above the median than below, it is positively skewed. Although not always present you can see a further element on the box plot on the right, a circle representing a data point a long way away from any other point. This is called an outlier and the common definition for an outlier is a

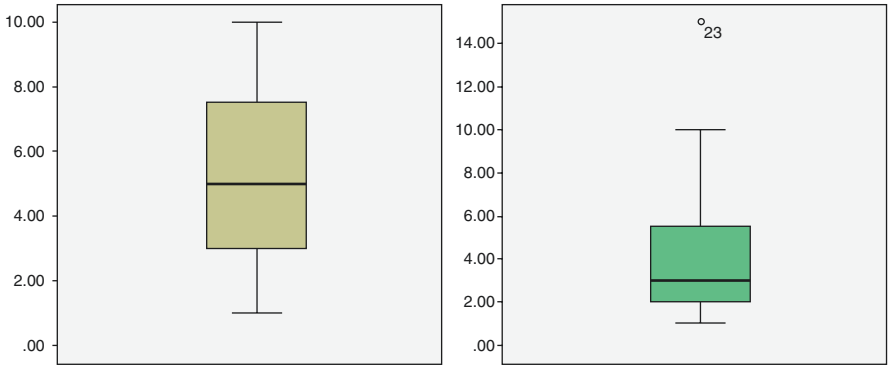


Fig. 15.6 Box and whisker plots

number that is greater or less than the interquartile range by more than 1.5 times the interquartile range. Although not shown here SPSS also identifies extreme values, which are numbers that are greater or less than the interquartile range by more than three times the interquartile range and are usually shown by an asterisk. It is important to look at these points as they can considerably influence the outcome of certain tests or may simply be an error such as typing 333 in a cell rather than 33 that can be corrected before the analysis is undertaken.

Key Points

- Descriptive statistics can be based on graphical representations and on numbers that represent key features of the data collected.
- These techniques attempt to summarise the data in some way so that any patterns can be interpreted more easily.
- They cannot reveal anything beyond the properties of the data collected and therefore have no predictive value.
- In order to start to make predictions (whether right or wrong) a different branch of statistics is used; this is called inferential statistics.

Both descriptive and inferential statistics are of equal importance. Descriptive techniques can inform a reader about the sample being reported on and also they frequently suggest ways in which data can be examined in order to perform inferential testing. As an example descriptive statistics can access whether data follow a mathematically predictable pattern. Orderly patterns suggest ways in which we can make predictions. A commonly encountered application might be the suggestion that data are normally distributed; in which case a *t*-test would be applicable (this will become clearer when the normal distribution and the *t*-test are described).

15.4 Distributions

In the summary of statistics in the previous section you saw one of the points of interest about a set of data is when it follows a particular pattern. Some patterns are quite complex and difficult to describe, or may even be non-existent in a set of data, while others are rather more orderly. A good example of a non-orderly pattern would be produced if you recorded the occurrence of the UK national lottery numbers. Great pains are taken to make sure that they are random and to prevent any pattern from emerging that would make it possible to predict the next set of numbers. Orderly patterns suggest ways in which we can make predictions. For this reason data in orderly patterns are of particular interest. They are of even greater interest if the patterns observed follow a mathematically predictable pattern. An example of an orderly pattern would be way in which the intensity of a monoenergetic X-ray beam changes as it is attenuated by thicknesses of a uniform (homogeneous) material. The mathematical pattern which is observed in this case is an exponential decrease and it can be used to ensure that the thickness of lead used in particular circumstances is enough to give a predictable level of protection.

15.4.1 Mathematical Distributions

Mathematical relationships can be represented graphically as patterns. These patterns are not real; they are mathematically generated patterns. Some patterns fit what we observe in the real world quite closely; we know this because when data are collected, they fit mathematical predictions very well. You may have had met some orderly observations of graphs of exponential relationships for attenuation of X-rays and radioactive decay. In these cases the data that are collected are distributed in such a way that they fit a mathematically predicted pattern. Other sets of data are not so orderly, as in the example of the UK national lottery draw used above.

Because predicted patterns allow us to make future predictions they are of particular interest when looking at data. Not surprisingly the predictions made by statistical testing rely heavily on predictable patterns. There are many such patterns. However, this section focuses on one particular and very special pattern called the normal (or Gaussian) distribution.

15.4.2 The Normal Distribution

This is also known as the Gaussian distribution after Karl Gauss, a German mathematician who did early work in statistics. It is special because large numbers of natural observations follow this particular pattern. Variations in the lengths of particular bones, errors in blood tests, variations in height and weight of patients all follow this pattern. (The first written description of the pattern followed a series of measurements of the chest sizes of soldiers in Scottish army regiments.) Another interesting point is that if people are given a series of related questions where they have to rate their responses on a Likert scale (see Sect. 15.2 above), then the total scores of their answers are usually normally distributed. There are some key points about the pattern of a normal distribution.

- It is symmetric and unimodal: the mean, median and the mode have exactly the same value.
- The spread of the data follows a predictable pattern and this is best demonstrated on a graph.

The graph in Fig. 15.7 contains some algebra that needs explaining. The symbol μ (Greek letter 'mu') represents the mean of the data. Notice it sits at the peak and the distribution is symmetrical about it. The symbol σ (Greek letter 'sigma') represents the standard deviation of the data. The diagram indicates regions that are one standard deviation, two standard deviations and three standard deviations either side of the mean. The percentage of observations that falls in certain regions (when real world data follow a normal distribution) is predictable (subject to some random variation) and the numbers on the graph give some values for these percentages. It is possible to calculate a percentage value for any distance away from the mean; the values for σ, 2σ and 3σ are convenient benchmarks and it is worth committing these to memory if you can.

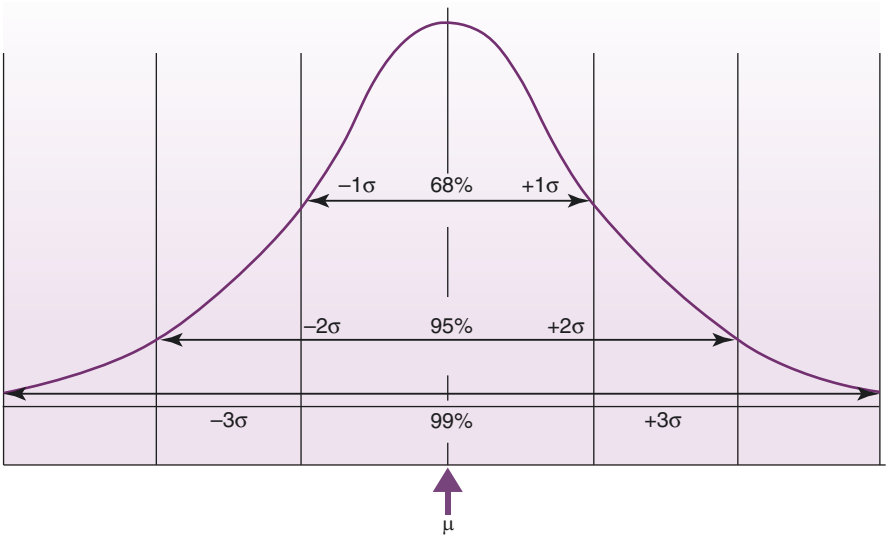


Fig. 15.7 The normal distribution

Looking at this graph it is possible to determine how likely a measurement is to fall within a certain distance from the mean before taking it. In other words, it gives a prediction of how likely an observation is. Also, the further an observation occurs away from the mean, the less likely it is that it belongs to the main group. It is important to realise that there is no definite prediction here; it is rather a prediction based on how likely something is. There is always room in the real world for unlikely events to happen; individual UK national lottery winners show us this on a regular basis.

15.4.3 Other Distributions

There are some other distributions that are useful and each one is useful in certain circumstances.

Poisson distributions tell us about the likelihood of random infrequent events such as major incidents in casualty departments. Uniform distributions tell us about occasions when it is just as likely for one thing to occur as any other, such as observing heads or tails when flipping a coin. Exponential distributions tell us about the decay of radioactive tracers. There are others but the important point to make here is to ensure that you understand what a distribution is, and not necessarily to be able to recall all of the different kinds.

Key Points

- Patterns can be observed when observations are made. Patterns can also be generated using abstract mathematics. For example, the Poisson distribution is generated by a mathematical equation, which relies on the rate at which events occur. It could be used to predict the numbers of patients arriving, in an X-ray department, as casualty referrals, within a given period of time.

- When it is possible to match real world observations to theoretical mathematical patterns we can develop ways of predicting how likely it is that some event will be observed. Using the example above the number of staff required to deal with a given rate of arrivals in a department could be estimated and provide a basis for staffing levels.
- This does not give the power of absolute prediction, but when likelihoods (or unlikelihoods) get very extreme then it is possible to get reasonably close to absolute prediction.
- The real world never allows us to give definite predictions from statistics and is full of unlikely events.

15.5 Probability and *p*-Values

Probability expresses the chance that a particular result may occur when there is a choice of possible outcomes. In the context of quantitative analysis, probability is commonly expressed in the form of a probability value, called '*p*-value'. It is important to understand what a *p*-value tells you as it allows you to make judgements about your own research findings when you use statistical testing. It also allows you to understand what other people's research discovered. You should always report the *p*-value so that your readers can draw their own conclusions.

It is common to find that *p*-values are either reported as an exact figure or as being 'less than a certain figure' (e.g., $p < 0.05$). The rationale behind statistical testing is dealt with in hypothesis testing in Sect. 15.6. At this point it may be useful to briefly explore what *p*-values mean in this context.

The *p*-value reflects the probability that the outcome observed by the test happened as a product of random chance as opposed to being an effect or real association. The logic used here is such that if the *p*-value is high then there is a very real chance that what was observed could just happen accidentally without any underlying reason. Alternatively, if the associated *p*-value is very small then the opposite is likely to be the case, that is to say that there is some underlying reason for the observation made. This gives rise to the idea of 'significance'. If the *p*-value for an observation is found to be below a certain level, decided by a researcher based on study design, then it is commonly referred to as a statistically significant finding. This means that something has happened and it does not seem to be an accidental occurrence and therefore there must be an underlying reason for it. This is why when researchers report low *p*-values in papers, this should act as a flag in a reader's mind saying that there is something interesting going on. Equally when high *p*-values are encountered, it is good practice to check these findings against the claims of researchers who report that they have discovered something happening. For example, a researcher claims that a significant improvement in patient recovery was observed following a change in treatment regime. This is supported with evidence from an inferential test with a resulting *p*-value of 0.11 in favour of their claim. *p* values above 0.05 provide weak evidence to support the claim. Clearly, in this case

the evidence is weaker, and many would consider it too weak to be relied on for support. It is important to say here that this 'hunting for statistical significance' in research can create a lot of distress, particularly to inexperienced researchers, when results proved statistically non-significant. From an academic point of view discovering both significant and non-significant statistical relationships is equally vital for the progression of the evidence base.

Key Points

- 'Likelihoods' are expressed formally as *p*-values.
- P-values use an assumption that in the event that there are many possible outcomes for an observation, all possibilities add up to 1.
- A commonly encountered *p*-value in research papers is $p \leq 0.05$. This reflects a probability of equal to or less than 1 in 20.
- Many researchers consider that if their observations have this probability of occurring by chance then they have observed something significant happening, that is to say that it was not down to chance and the observation signifies that there is some real effect or association involved. This choice of what is significant is purely arbitrary. A randomised controlled trial in clinical and preclinical research would most certainly require the use of a lower *p*-value ($p < 0.01$) as an indicator of significance. A pilot study though would work well with a higher *p*-value ($p < 0.05$) as an indicator of significance.

For other types of data such as nominal (categorical data) the possibilities for quantitative analysis become a little more restricted. In Sect.15.8 different types of data are presented and tabulated. Some examples and suggested analysis techniques are covered later in this chapter. The various types of research information/data are discussed in more detail Chap. 8.

15.6 Hypothesis Testing

In previous chapters and sections in this book a number of terms and ideas are discussed: hypotheses, summary statistics, mathematical and real world distributions and *p*-values, for example. These form a toolkit from which it is possible to develop a way to find out what information may be contained within a collection or collections of data.

Hypotheses form the backbone of a research process by giving us a question to answer and to test. Summary statistics allow us to condense key features about data into a manageable form. It is not always necessary to use summaries; the data can be analysed in their raw form. Distributions allow us to try and determine the probability of obtaining observations; and p-values give us a formal way of expressing them.

This section looks at how you can combine them into an analytical package, which can then be used to address the purpose of your research.

15.6.1 Hypotheses: Asking the Right Kind of Question

As mentioned in Chap. 2, a hypothesis is a question or idea that your research sets out to find out about. For qualitative research a hypothesis is not required as a robust research question would suffice. For quantitative research designs, hypothesis testing is vital and is part of the golden thread of a research design. It is important at this stage to ensure that you distinguish between two different kinds of hypotheses: experimental and statistical. An experimental or research hypothesis sets the context of your research, giving a reader an idea of what your research aims to do. This is not quite the same as the way in which the word hypothesis is used in statistics. Because statistics involve mathematical methods, a statistical hypothesis has a very different and somewhat strict meaning. Whereas in the real world the way in which we describe things can offer many different options; statistical hypotheses only consider two options. The observations either fit a hypothesis (alternative hypothesis H_1), so a researcher aims to prove the hypothesis as correct, or the observations do not fit the hypothesis (null Hypothesis H_0), so a researcher aims to disprove the hypothesis. The choice often is defined by previous studies in the field and the context and it relates to what is the expectation that a research will find out, without excluding some often surprising findings.

It is a very simplistic, rather black and white version of the world. Many people new to research experience problems because they have too much confidence in the power of statistical testing to provide answers to complex questions. The findings of your test will only be as good as the quality of the questions that you ask. Vague questions are usually very complicated when you look at them in depth. A good example would be a study where a researcher wants to 'know how to optimise a particular type of MRI examination in relation to a given clinical condition or diagnosis'. A noble endeavour no doubt, but how do you do it in practice? The following questions may arise.

- Is that optimal?
- Is it dose related?
- Is it image quality related?
- Does it relate to the observer's perception?
- Is it all of these factors together?

It is critical to formulate clear hypotheses otherwise you will find yourself in the situation of setting out to investigate a problem, collecting data and then, when it comes to drawing conclusions, there will be a mass of confusion. Clearly, vague questions are to be avoided.

That is not to say that vague questions cannot be tackled, but that they are best broken down into a number of simple related questions. Normally, unless you are a full-time researcher, where time to tackle complex problems is available, it is best to keep it simple.

The alternative or experimental hypothesis refers to the idea that there is a difference. The alternative hypothesis is commonly given the symbol H_1.

Null hypothesis refers to the idea that there is no difference between some particular quality concerning sets of data, perhaps a summary statistic that you are testing, and contradicts the alternative hypothesis. This leads to the inference that the groups of data are the same or are similar in some way, e.g., similarities between the means of the groups or patterns in the data sets, etc. The null hypothesis is commonly given the symbol H_0.

Notice at this point that for the purposes of conducting a statistical test, there are only two possibilities. Questions, which potentially allow more than two possible answers, cannot be tackled using these techniques. The pair of hypotheses (null and alternative) that you use are also described as one-tailed or two-tailed. If an alternative hypothesis is that there is a difference, but it does not matter what the difference is, then it is two-tailed. If an alternative hypothesis is that one group of data has a quality that is greater than or less than another, then it is one-tailed. A good example would be a drug trial where an investigator is looking to see whether a new kind of drug is better than the existing one. This is a one-tailed situation. If an investigator just wanted to show that they were different to each other, then this would be two-tailed.

15.6.1.1 Matching Statistical Hypotheses to the Point of your Research

Having developed the idea of a statistical hypothesis the next step is to see how you might use it to address the experimental hypotheses of your research. The process of testing relies on trying to figure out how likely it is that observations happen by chance. If there is a high probability that what you have measured could happen by chance, then it will be difficult to make predictions based on these observations. Think of this in terms of winning a lottery: the process is accidental. Because of this it is impossible to make predictions about subsequent winners based on observations of previous winners. If, on the other hand, there is a very small probability of your observations being due to chance, then the opposite will be the case and your observation will allow you to make good predictions about future events. This suggests that the most useful tests will look for circumstances where chance is unlikely to have played a big part in what you have observed. So how does this work in practice?

A statistical test looks at the probability that one of the statistical hypotheses is unlikely to have happened by chance. Let us think about what this means by considering the following points.

1. If the chosen statistical hypothesis (remember there are two, null and alternative) is unlikely to have been observed by chance then the opposite must be true; something must really be going on to give the observation.
2. If the chosen hypothesis is likely to have been observed by chance (the opposite of point 1 above) then it means nothing, as it could be just a bit of random luck, so we choose to ignore it.
3. If we are going to ignore one hypothesis and there are only two possibilities, null or alternative, then the other hypothesis must be the case.

In practice statistical tests always look at the likelihood of the null hypothesis having occurred by chance. To see how this works it is better to construct a flow chart (see the flow chart in Sect. 15.7) from the points above.

Perhaps it is worth repeating that the logic used here is simplistic and therefore only accommodates simple questions. If you try to do anything more sophisticated with a statistical test on its own, then it does not work. An example might be a patient survey that investigated degrees of satisfaction with their X-ray examinations. A single test, used on its own, only distinguishes between one source of satisfaction between two groups of patients. Any other number of possibilities cannot be directly explored as the test only permits two possibilities of a researcher's choice. In this case a more complex approach would be necessary to investigate a patient's feelings.

After the next section an example is provided to show how statistical testing might be used to investigate a real problem.

Key Points

- The process of hypothesis testing requires questions to be formulated in terms of hypotheses.
- The null hypothesis considers the idea that there is nothing special about what has been observed. It could be random chance that produced it.
- The alternative hypothesis usually represents what the researcher is looking for.
- The observation has to be shown to be improbable in order for the alternative hypothesis to be accepted.
- We accept the alternative hypothesis by gathering evidence that allows the rejection of the null hypothesis.
- Remember that 'absence of proof' is not 'proof of absence'.

15.7 Choosing the Correct Data Analysis Test Tool

Choosing the correct statistical tests depends on a number of conditions about the data.

- Matched or related (these words are used interchangeably).
- The number of tails required to tackle an experimental hypothesis (see previous section).
- The size or anticipated size of the groups of data and how many groups there are. Many statistical tests do not produce accurate results if the samples of data are too small. Each test has its own limits.
- Whether the data are from a normal distribution (therefore it can be tested using parametric methods) or not (so non-parametric methods can be used).

All parametric tests assume that the data are normally distributed and so for interval and ratio data, where this type of distributions are possible, we first have to ascertain this fact in order to use the correct test. This can be ascertained by using the Shapiro–Wilk or Kolmogorov–Smirnov tests.

Having figured out how you will use the process of statistical testing, the next step is to choose an appropriate test. There are many tests available and it can be a bewildering choice. The choice of a test is relatively logical as each test is a tool, which has its own particular application. The way to figure out which one suits your particular job is by considering a number of key points about your research.

1. Regardless of the chosen test the data used should be random. The less random the data can claim to be, the weaker the results of a test are.
2. The choice of test is influenced by dependency in the data. Some tests do not work at all if one measurement could influence another. For other tests it is necessary that the data are dependent, could fit a mathematical distribution or not. It is important to stress the word 'could' here because unless you can collect impossibly large amounts of data it is very difficult to say that that they definitely fit a theoretical distribution. Random chance always plays a part in stopping you from going this far. The idea of a mathematical distribution was covered earlier but it is worth looking at an example to make sure that it is quite clear.

Figure 15.8 shows a set of data from a fictitious sample. The sample is made up of time observations. These have been grouped together to make a histogram with intervals of 1 minute. The histogram appears symmetrical and unimodal.

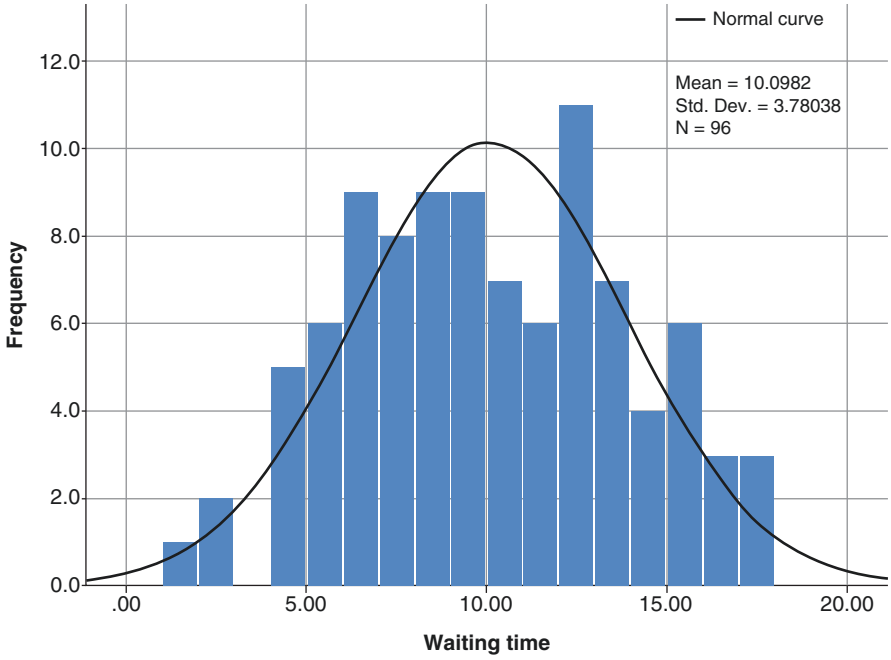


Fig. 15.8 Histogram and normal curve fit for department A

Superimposed over the histogram is a curve with a mean of 10.0 and a standard deviation of about 3.8 (these figures are at the right-hand side of the diagram and are true for both the histogram and the curve). The curve is a theoretical mathematical shape placed on the same plot as the histogram. It was drawn using two 'parameters' that determine the shape of normal curves: the mean and the standard deviation. To create the curve the mean and standard deviation of the histogram data were used, which is why they are the same for both. It looks like the shapes of the histogram and the 'fitted normal curve' are very similar so in this case the data seem to be 'parametric' in nature and the parameters relate to a normal distribution. If you consider Fig. 15.9 the fitted curve does not look normal and the data appear skewed.

The distribution does not have to be 'normal'; it can be any shape (see the section on other distributions). As long as the data seem to fit a theoretical pattern, then a parametric test can be used. Particular tests relate to particular distributions, a common one being the normal distribution. An important point to remember is that theoretical mathematical distributions are made from variables on a scale, so the underlying parameters will be on an 'interval' scale. This type of test will generally only apply to interval data. A notable exception is where some types of ordinal questionnaire-based data from Likert scales can be combined to give data that fit a normal curve.

Having small sample sizes can have a great deal of random variation and it is difficult to identify a pattern convincingly and ideally we need a reasonable sample size in order to check if the data is normally distributed. If you cannot give a convincing case for the data being parametric, then treat them as if they are not.

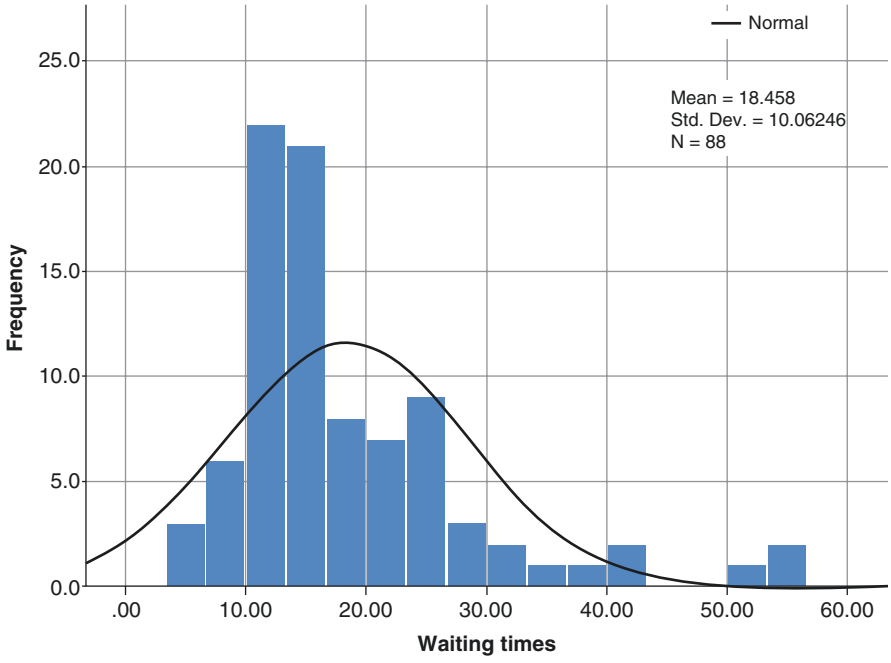


Fig. 15.9 Histogram and normal curve fit for department B

		Kolmogorov-Smirnov[a]			Shapiro-Wilk		
		Satistic	df	Sig.	Satistic	df	Sig.
Waiting Department A	time.	.061	87	.200*	.982	87	.269
Waiting Department B	time.	.181	87	.000	.830	87	.000

*. This is a lower bound of the true significance.

a. Lilliefors Significance Correction

Fig. 15.10 Normality test results for departments A & B from SPSS

There are other ways of showing that data may fit a pattern besides plotting the data, but these often rely on software or the use of a statistical test with a null hypothesis that there is no difference between the data observed and the required mathematical distribution. Figure 15.10 shows the SPSS printout for two tests of normality, by default two statistical tests, the Shapiro–Wilk and Kolmogorov–Smirnov test are both done each time. The null hypothesis of this test is that the population is normally distributed and if the significance value is below 0.05, we reject the null hypothesis and accept the alternate hypothesis that the data is not normally distributed. If we look at the significance values, we can see that the significance value for department A (Fig. 15.8) is 0.200 and 0.269 for the tests; both tests are in agreement hence we do not reject the null hypothesis so we can accept that the data are normally distributed and we can use parametric tests. For department B however the significance level for both tests is <0.001 which means that we reject the null hypothesis and accept the alternate hypothesis, the data are not normally distributed and we would use a non-parametric test. Other options might be available to us such as transformations to change our data into a normal distribution, but this is beyond the scope of this book.

Finally, in consideration of data being normally distributed is the quantile–quantile plot in Figs. 15.11 and 15.12. The purpose of the plot is to compare two sets of data to see if they come from the same distribution. The solid line represents a theoretically normal distribution and the circles the data being tested. If the circles fall on the line, then they must be normally distributed and although there is some small variations we can see that the data from department A is very close so we can say that the data approximates a normal distribution. When looking at department B however we can see that the data points do not cluster around the line and therefore the data is not normally distributed.

Having discussed interval/ratio data, we now consider other sorts of data. Any investigation involving other nominal or ordinal data is called non-parametric. Non-parametric tests are also called distribution free tests; they make no assumption about the distribution and this is why we can also use them for interval/ratio data that is not normally distributed. It is critical to identify whether to use parametric or non-parametric techniques. In general when you have interval/ratio data you should attempt to use parametric tests whenever possible as they are more powerful than their non-parametric counterparts, providing stronger evidence. On the other hand,

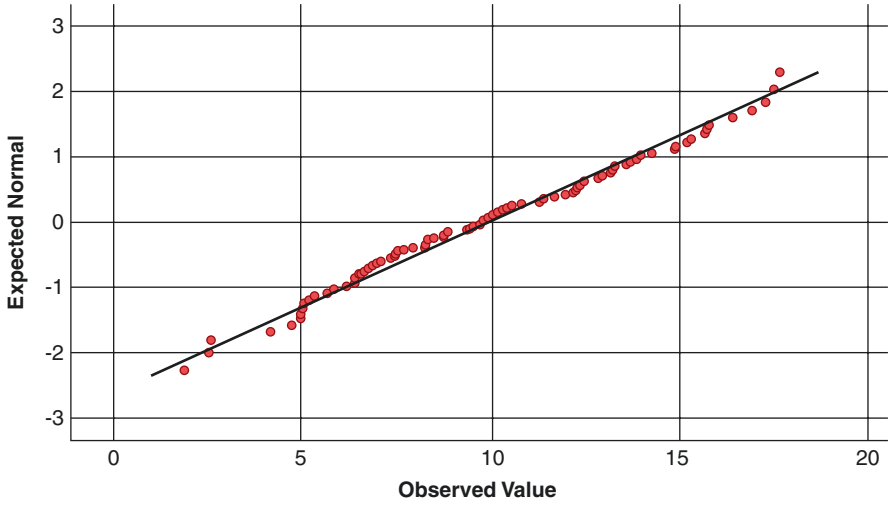


Fig. 15.11 QQ plot for department A

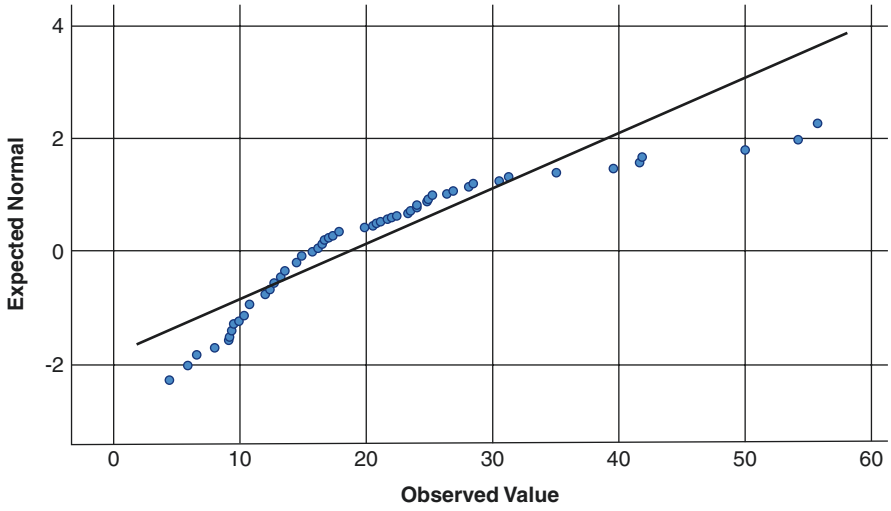


Fig. 15.12 QQ plot for department B

non-parametric tests are much more robust and there are fewer factors that can upset the validity of the findings. The choice of test depends on how much evidence you have to support your choice. If sufficient evidence exists in favour of a parametric test then this is the one you should choose.

A flow chart that summarises the most commonly statistical tests used and how to choose these is shown in Fig. 15.13.

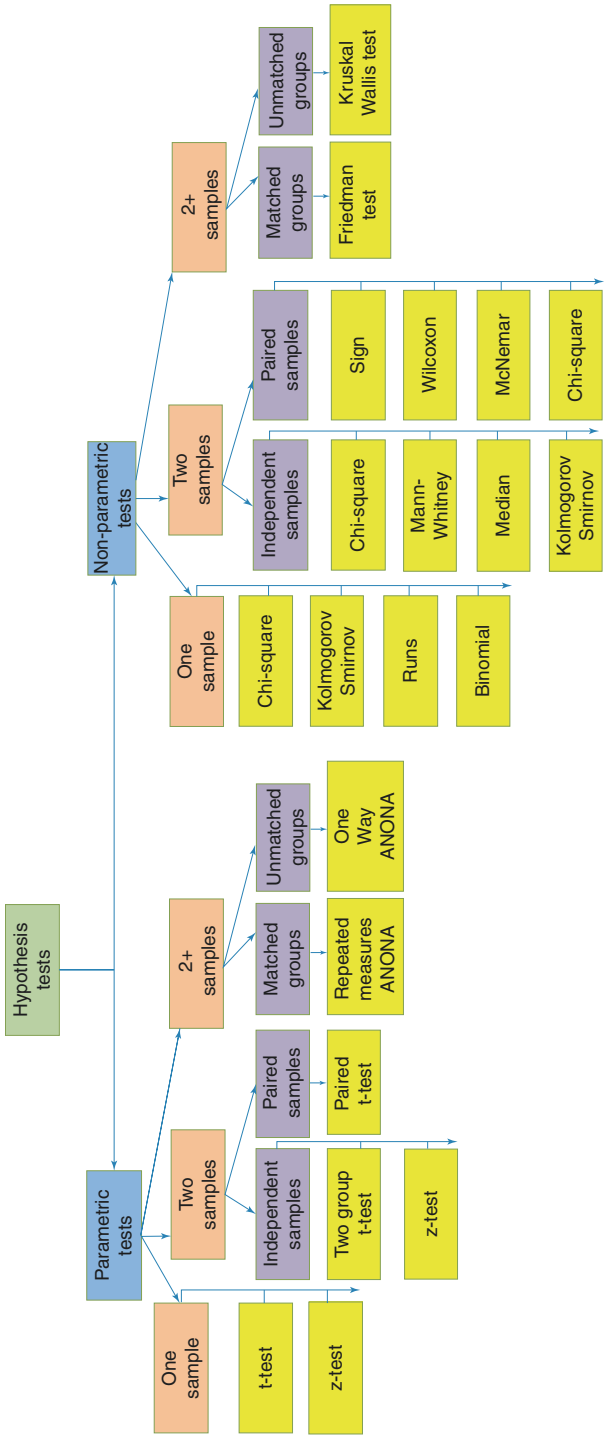


Fig. 15.13 Flowchart for tests selection

15.8 Interpreting Results

By applying a statistical test to data a 'test statistic' is produced. There are many different kinds of test statistic each relating to a particular test, or tests. This result in itself is meaningless; it tells you nothing about the probabilities of 'accidentally' observing the data; it is a means to an end. The test statistic is used to generate a probability, or *p*-value, which is of much more interest. The *p*-value generated gives evidence in favour of rejecting the null hypothesis. The lower the *p*-value, the less likely it is that the data observed occurred by chance and the more likely it is that you have observed something 'really happening'. In mathematical language, a low *p*-value provides evidence to reject the null hypothesis in favour of the alternative hypothesis. If the *p*-value is high, we cannot get rid of the null hypothesis so we favour it instead.

There are two schools of thought about how you should consider the *p*-value. One says that a pre-determined limit should be applied to the *p*-value and if the value equals or falls below a fixed limit then the alternative hypothesis is 'true'. If the *p*-value is above the pre-determined value, then the null hypothesis is 'true'. This is called fixed level testing. The other way of using the result is to present it to the reader and let them make their own mind up, pointing out how significant the findings are. Not surprisingly this is called significance testing. A fairly standard way of interpreting the significance of *p*-values is on the verbal scale shown in Table 15.1.

15.9 The Process for a Typical Statistical Test

Now we can draw all of the points above together and an example of the resulting statistical package is examined. As you read through this section, keep in mind the sections that come before; it is at this point that the relevance of each section becomes apparent.

The experimental hypothesis will be that for men born in the last 70 years, adults of later generations are taller than those of earlier ones. The data will be height measurements on an interval/ratio scale. The data will be taken by randomly selecting unrelated adult male individuals from the population and placing them into one of two groups: below 40 years of age and over 40 years age. We will not concern ourselves with how this is achieved but assume that it is.

The initial null hypothesis is that there is no difference between the two groups (an older and a younger group), giving an alternative that there is a difference in heights between the two groups. Since you are looking to see if there is an increase in height the process will involve one-tail.

Table 15.1 Interpretation of *p*-values for significance testing

p-value	Interpretation
≤0.01	Strong evidence to reject the null hypothesis
>0.01 and ≤0.05	Moderate evidence to reject the null hypothesis
>0.05 and ≤0.10	Weak evidence to reject the null hypothesis
>0.10	Little evidence to reject the null hypothesis

It was mentioned earlier, when discussing the normal distribution, that this kind of measurement would be expected to be normally distributed and can be tested before the analysis to see if we are going to undertake a parametric or non-parametric test.

For this example you can use a commonly encountered family of tests called *t*-tests. More particularly an unrelated or unmatched *t*-test can be used; see Fig. 15.2 in Sect. 15.2. This test is sensitive to a number of assumptions; namely,

- the data are randomly collected
- the data are independent
- the data sets are normally distributed with a similar spread about their mean values.

Notice that the way that the experimental hypothesis has been set up and the way that the collection of the data has been specified makes this *t*-test the ideal choice. It is always worth thinking about how you may test your data before deciding on a final experimental hypothesis.

The first step would be to calculate the test statistic or '*t*' statistic. How exactly this is arrived at is not important at this stage. If you are interested a more advanced book on statistics will tell you this. What is of interest is that the calculation involves the use of the means and variances of the two groups. The process is based on summary statistics, which reflect properties of each group of measurements. Also, the particular summary statistics involved are closely related to the theoretical distribution that the groups fit (see Sect. 15.14) The test statistic links properties of the groups to the chosen theoretical model (the normal distribution). If software is used to do the test, then the test statistic will probably be included in the output.

If you are doing the test manually, this statistic is now taken to a set of tables and converted into a *p*-value. If you are using software, then the *p*-value will be given in the output with the test statistic. At this point it would be wise to check that the *p*-value has been generated for a one-tailed test; do not assume that software can read your mind. Always check that the requirements of the test have been satisfied.

The *p*-value now needs to be interpreted. If you chose a fixed level approach, then you now need to compare the *p*-value obtained against your chosen value. A widely accepted *p*-value is 0.05, representing a 1 in 20 chance that the results could have occurred accidentally. At this point, if a fixed level approach is used and the *p*-value obtained is equal to or falls below 0.05, then the result leads to the rejection of the null hypothesis and you conclude that younger males are (on average) taller than older ones. If significance level testing is being used and the *p*-value is just below 0.05, then you have 'moderate evidence' to reject the null hypothesis (see Table 15.1) and 'moderate evidence' in favour of the experimental hypothesis. The *p*-value chosen here is not the only possible choice. If you wanted to be really tough about making conclusions with a fixed value test, then you could always lower the *p*-value at which the result becomes significant. A glance at Table 15.1 gives you an idea of how much this affects the credibility of your findings.

The steps taken for a non-parametric test are no different, but because there is no direct link between the data and a mathematical distribution the process is not quite as obvious. Non-parametric tests work because the test statistic calculated in the process relates to a theoretical distribution as opposed to the data themselves.

Key Points

- A correct choice of a statistical test is determined by the kind of information or data that has been collected.
- Tests can broadly be separated into two classes: parametric and non-parametric.
- A low *p*-value gives evidence to support the acceptance of the alternative hypothesis that observations are not occurring by random chance.
- *p*-values can be interpreted using a rigid 'fixed level' approach or a more flexible 'significance' approach.
- The general process is similar for all tests.

15.10 Common Statistical Tests

This section covers some of the commonly used tests. The hypothesis for each test is included. It is very important that you have a clear idea of what each test actually does and how it allows you to draw conclusions.

Different tests are required depending upon the number of sets of data to be compared. The tests are considered to involve one, two or three or more sets of data. This is indicated for each test and examples are given to clarify how they are applied.

15.10.1 The *t*-Test (or Student's *t*-Test): Compares Two Groups of Data

The different names for this test reflect its origin. It was created by a nineteenth century statistician working for Guinness breweries in Ireland, but the statistician's name had to be kept secret. The owner of the brewery did not want his competitors to know that such 'underhand' industrial methods were being employed to control the quality of his product, and at the time it was quite a radical change of method. Because of this the test was published under the pseudonym 'Student'. It is called the *t*-test because it uses a mathematically determined distribution called the '*t*' distribution in order to generate its results.

The *t*-test can be used to compare two groups of measurements where the particular observations are on a continuous interval scale. An example of this would be an ionising radiation dose. The doses recorded are on a continuous scale, down to the level orientations on the dose to the thyroid gland during chest radiography. It is important to realise that it could only compare two orientations at a time and no measurement can affect any of the others for this test to work.

The null hypothesis of this test is that the means of both groups are the same. A low *p*-value indicates evidence that this is not the case and that the means (and hence the two groups) are different.

Other versions of the *t*-test exist, such as the dependent, matched or single sample test. In this case the test is used to examine the possibility that a single group comes from a larger population which has a particular mean associated with it or to examine matched measurements of a group.

- Example

This test could be used to examine if the waiting times in a satelite hospital were different to those in the main hospital. The waiting times of patients would be recorded in each hospital. This would give a dependent variable of time which is continuous and we would be comparing 2 groups of data (the data from each hospital). Looking at WCC score for each patient before and after the administration of the antibiotic would provide a single data set comprising the difference in counts for each patient.

The null hypothesis would be that there was no difference between the mean waiting time between the two hospitals. A low *p*-value would give evidence to suggest that there was a difference.

In both cases the choice of tails for the test is important. For the waiting time study it might be that the waiting times are shorter in the satellite centre or longer, therefore both possibilities exist and we need to do a two-tailed test.

Generally, it is also important that the spread of data in both sets is not too dissimilar. There are versions of the *t*-test that are specifically designed to cope with large difference in variance or standard deviation. Unless your test states this to be the case a useful rule of thumb is to make sure that the variances of the two data sets are within a factor of three of each other (i.e., neither one is three times bigger or smaller than the other).

15.10.2 Mann–Whitney 'U' Test: Compares Two Groups of Data

This test draws its name from its originators and is much less interesting from a historical perspective.

The Mann–Whitney U test uses a calculation to produce a result called the 'U' test statistic, which is then used to generate associated *p*-values. There is no directly underlying mathematical distribution involved therefore this is a non-parametric test. Indirectly the distribution of 'U' is related to a normal distribution whose mean and variance are related to the size of the samples used and a mathematical pattern known as the uniform distribution. It is a highly versatile test that can be used to compare two independent sets of observations. The observations can be continuous or discrete and so it can be used on laboratory style measurements and whole numbers such as the ones that come from surveys or counting exercises. When it is not possible to perform an independent samples *t*-test this test is frequently chosen.

- Example

If the recovery time following radiotherapy treatment was to be compared for different treatments, interval data could be collected for each treatment. If examination of one or both sets of data revealed a skew in the data, then the Mann–Whitney U test could be used to compare the treatments.

The null hypothesis would be that the two groups of recovery times belong to a single underlying set (population) of recovery times. A low *p*-value would give evidence to suggest that the two data sets do not belong with each other and that they are therefore different.

The null hypothesis would be that the responses of the two groups are the same as the responses of a single larger group. A low *p*-value would give evidence to suggest that the satisfaction levels of patients at the two sites are different.

Notice that this test looks for differences, which is fine when addressing two-tailed hypotheses. If it was important to determine if one department was better, in terms of satisfaction delivered, than the other, then some other evidence would be required. In this case comparison of a simple bar chart would indicate if one site scored higher than the other. There is nothing wrong with addressing one-tailed hypotheses in this way.

15.10.3 Chi-Square (χ^2) Test

The chi-square test, or Pearson's chi-square test to give it its full name, can be used to look at contingency tables. A contingency table is constructed of a number of boxes each of which is mutually independent, i.e., when an observation is made it can only belong to one box. An example of this would be the sorts of tables that are used to record data in cohort studies (see Chap. 9). In this case data are in a table that looks at exposure to a condition and whether or not a particular outcome is observed. Exposure to habitual smoking, for example, with an outcome of lung cancer could be represented on a contingency table. Each person observed can only belong in a distinct category or 'contingency' constructed from the options exposure/non-exposure and lung cancer/no lung cancer.

Although there is a mathematical distribution involved (the chi-squared distribution), the data used in this are discrete, and categorical or ordinal. This means that the test is considered to be non-parametric. It is not only useful for cohort studies. It can also be applied to any contingency table. There is, however, one condition to bear in mind here. When looking at a table, each box or cell belongs on a particular row or column. Each cell has an 'expectation value' associated with it. This is calculated by multiplying the total number of observations for its row by the total number of observations for its column and dividing this by the total number of all observations in the table. If the expectation value for any cell is lower than five then the test cannot be reliably used. The reason for this is mathematical and requires an understanding of conditional probability, which goes beyond the purpose of this book and is not elaborated on here. In these circumstances a similar test called Fisher's exact test should be used. This test deals with all expected values for a cell. A particularly useful aspect of this test is that it can deal with ordinal or categorical data from questionnaires.

- Example

As part of an audit in an X-ray department, patients are asked to comment on their satisfaction with the courtesy of the staff. The question asks for a yes/no response. The answers for males and females are counted separately and placed in a contingency table (Table 15.2) with a view to investigating whether there is gender

Table 15.2 Contingency table for patient satisfaction

Satisfaction	Males	Females	Totals
Yes	Satisfied males	Satisfied females	Sub-total satisfied patients
No	Dissatisfied males	Dissatisfied females	Sub-total dissatisfied patients
Total	Sub-total males	Sub-total females	Grand total

equality regarding the way the staff behaves. The auditors check that all of the cells have an expectation value of five or more using the sub-totals and grand total and find that all cells have at least this value.

The null hypothesis is that there is no association between gender and satisfaction. A low *p*-value suggests evidence that this is not the case and that the males and females are being treated differently.

Note that this test only comes in a two-tailed form hence we cannot determine whether males or females are getting preferential treatment. Other evidence would be needed to confirm this.

This test can be used to investigate the response from Likert scales, but the data need modifying before they will fit into a suitable table. The problem is that attitudinal scales are not discrete enough to prevent overlaps between cells. Using the example of patient satisfaction, if a Likert scale were to be used, there would be degrees of satisfaction or dissatisfaction and a patient's response at any one time would be determined by their mood. A recent bad experience in another department could affect the response, e.g., a long wait in clinic before being sent for an X-ray examination. Because of these possible factors the responses are not reliable in this case. The data from the scale would possibly need to be simplified into three options in this case: for example, satisfied, no preference or dissatisfied. This reduces the subtlety of the data and there is still a problem that extreme recent experiences could adversely affect the data. It would be unacceptable to choose a simplification that would ignore some of the responses.

A final point that needs to be made is that the results of this test depend on the number of degrees of freedom involved. Again, the basis of this is beyond the remit of this book, but the number of degrees of freedom is very easy to determine.

15.10.4 Goodness of Fit (Chi-Square Goodness of Fit): Can Be Used to Compare Two Groups of Data or One Group with a Theoretical Distribution

The chi-square distribution can also be used to look at what is called 'goodness of fit'. This is a process of investigating how similar or dissimilar two distribution patterns are to each other. It can be used to investigate patterns in two sets of collected data and also to compare collected data with a theoretical pattern. The examples below indicate how each works.

- Example 1

A researcher wants to compare the work patterns of two radiotherapy departments and collects data about the kinds of therapies that they conduct. The patterns can be plotted in a bar chart for visual assessment, but to find if the work patterns are significantly different (in a statistical sense), a researcher could conduct a goodness of fit test comparing the categories of therapy against themselves, across the sites.

The null hypothesis will be that the patterns of work are not different. A low *p*-value indicates evidence that they may be different from each other.

In this case the patterns are not necessarily mathematical; they are generated by the frequency that observations fall into certain categories. The test could also be used for data that could be displayed in frequency histograms. Any patterns of matched categories or intervals can be used. The important point is that the data must be in a discrete form and the categories or intervals in the data must correspond to each other and be independent.

- Example 2

In order to compare the waiting times for two diagnostic X-ray rooms data are collected and placed into groups or 'bins'; e.g., waiting times are binned into 5 min intervals: 0–5 min; 5 min. 1 s to 10 min; 10 min 1 s to 15 min. A goodness of fit test could be performed to determine any differences between waiting times.

The null hypothesis is that there is no difference between the two rooms. A low *p*-value provides evidence that the waiting time for each room is different.

As with the Pearson chi-square test this is two-tailed and further evidence would be required to discover which room had the shorter waiting times.

From the examples above it can be seen that this test can be used for all types of discrete numerical data (i.e., involving counting), and this makes it extremely flexible.

The same process is followed in order to establish whether a collection of data follows a mathematical pattern. In this case the data must be interval data as all mathematical distributions are based on numerical scales. This test should only be used when the mathematical model is also discrete. Predictions can be made from discrete mathematical distributions and identifying this property in a set of observations can be extremely useful.

- Example 3

A researcher wants to discover whether a set of observations follows a uniform distribution. This would mean the responses in all of the categories are the same. A clinical department might, for example, wish to determine whether the rate at which patients arrive is uniform throughout a working day. Graphical representation might indicate that this is not the case and a goodness of fit test could then be performed to determine whether the pattern seen graphically was statistically significant.

The null hypothesis would be that patients are arriving at a uniform rate. A low *p*-value would provide evidence that there was a lack of uniformity and indicate if there was an underlying significance in the pattern. This could then be used to inform staffing levels based on the pattern.

15.10.5 Analysis of Variance (ANOVA): Compares Three or More Groups of Interval/Ratio Data

Analysis of variance is a technique that uses the spread (or variance) of each set of data. If the spread of the data sets covers similar ranges of values, then it is likely that there is no difference between the data sets. If the spread of any of the data sets are different, then ANOVA will identify this.

With the notable exception that ANOVA is for three or more groups, this test is similar to the *t*-test. The test is used for interval data (e.g., dose measurements on a continuous scale) and there is an assumption that the data sets are normally distributed. This means that ANOVA is considered to be a parametric test.

A point of particular interest to notice here is that the test uses variance, not the mean of the samples as a *t*-test does. A potential problem here is that variance in a group of measurements comes from two sources. It can occur as a result of the process being measured (which is what the ANOVA test is looking for). It can also occur because of errors of measurement (due to the measurement process and random variation). If measurement error provides too much contribution to the overall variance, then the ANOVA test can give unreliable results. In other words the major contributor to a group's variance should be members of a group being measured, not the way that they are measured. A second related point is that variance in a group of measurements is also related to the number of measurements for independent random variables. This means that sample sizes are an important issue for ANOVA.

- Example

A researcher is investigating the radiation doses given to different patient groups undergoing different radiotherapy treatment regimens. There are three different regimens and separate (independent) groups are used. Provided that the sets of data for each group could be shown to be normally distributed then an ANOVA test could help to identify whether the different techniques deliver the same doses. If the variances in the groups are similar to the variances in measurement due to the dosimetry used, then there are potential problems. Also, it is important to note the number of patients in each group.

For this case the null hypothesis would be that all groups are receiving the same radiation dose. A low *p*-value would provide evidence that there was a significant difference across the groups.

There are some important points to notice about this example.

- The case above is a relatively straightforward one that involves only one variable, ionising radiation dose.
- The subjects in each group are independent of each other.
- The result of the test only allows a difference across all of the groups to be identified. It does not allow a researcher to identify which of the groups are different from each other.

The number of variables, which the test can handle, is referred to by the number of 'ways' that the test runs. A one-way test looks at one variable (as is this case if measuring radiation dose), a two-way test looks at two variables, a three-way test looks at three variables, and so on. This book only considers one-way analysis, the simplest case. A further way of simplifying matters is to keep the size of each group of patients the same.

The results of the test are based on multiple null hypotheses that there are no significant differences between pairs of groups in the study. A low *p*-value for any pairing of groups provides evidence that there is a significant difference between the two groups. A simple comparison for the two mean scores then provides evidence of the sense of relationship.

We continue with the above example. ANOVA provides evidence that there is a significant difference between treatment regimens A, B and C. A post-hoc test is conducted using the ANOVA results and the means of the measurements on groups A, B and C. The post-hoc test reveals that there is a significant difference between group A and the others with no significant difference between groups B and C. At this stage the process is still two-tailed, providing evidence for difference but not saying whether group A's dose is higher or lower than the others. Direct comparison of the group means it gives an indication of this sense allowing a one-tailed result to be obtained.

15.10.6 Kruskal–Wallis Test: Compares Three or More Groups of Data

This test can be applied to three or more groups of data and uses an extension of the logic, which lies behind the Mann–Whitney 'U' test. As with the Mann–Whitney 'U' test, the method is useful for ordinal data, or interval data that are not normally distributed (or where insufficient evidence exists to assume normality). The test is considered to be non-parametric and is used to investigate differences between groups.

- Example

A researcher investigating patient dose as a result of different radiotherapy regimens has reason to believe that the data sets do not follow a normal distribution. In this case a Kruskal–Wallis test will enable a comparison of the groups as a whole.

As with ANOVA the process is two-tailed and does not allow comparison of individual pairs of groups. In order to compare pairs of groups a post-hoc test called Dunn's test would need to be applied (in an analogous way to the Scheffé test for ANOVA for independent groups, which can be used to look for relative differences between groups). It is called a post-hoc test because it is carried out after ANOVA. This test uses the mean scores of the groups as well as the results of the ANOVA. To use it two conditions must be met.

- The results from ANOVA must be significant.
- It cannot be performed on its own, but only following ANOVA.

The process would still be two-tailed. In this case the comparison between median scores would provide further information. Remember that the median is a better summary statistic than the mean for asymmetrically distributed data.

The use of post hoc testing: a general note. If post hoc testing is used, it is considered dishonest to then retrospectively pretend that direct comparisons between pairs of data sets were performed and to 'forget' to mention that this route was suggested by a post hoc test. It could be particularly tempting to modify the point of a piece of research, based on the findings of this kind of testing, and make it appear as if something is happening when in fact it was not the original intention. The hypothesis that generates the research should come first. You should not start off with a hypothesis and then modify or re-engineer it as a result of your findings. Retrospective alteration means that the data are used to find the hypothesis and this is not the basis of scientific investigation.

Key Points

- Inferential testing involves drawing evidence from observations, which can then be used to infer ideas that go beyond the observations themselves.
- Inferential statistical tests perform the same function. They do vary depending on what kinds of observations have been made.
- Fixed level testing requires the assumption of a fixed *p*-value, which is then used as a basis for accepting or rejecting a null hypothesis. Rejection of the null hypothesis leads to acceptance of the alternative hypothesis.
- Significance testing is more flexible and seeks to find the probability of the observations occurring and then using this as evidence to support the experimental (research hypothesis).
- Parametric tests provide stronger evidence.
- Non-parametric tests are more robust and generally require fewer assumptions.

15.11 Reality vs Idealistic Conditions for Data Collection

The examples used in this section were based on the assumption that the data collection was flawless; however, in the real world it is very easy to introduce errors while collecting data. The errors of interest to this chapter are those introduced at the interpretation stage of a statistical test and there are two ways to get it wrong.

1) It is possible to reject the null hypothesis (i.e., to say the alternative is likely), when in fact it should not be. For fixed level testing this is less likely to happen if the *p*-value chosen for the level is higher rather than lower. The problem with raising the level is that the validity of your conclusion will be lowered, so some sort of compromise is necessary, which for most researchers occurs at $p < 0.05$. This type of error is referred to as a type I error.

Consider reclassifying responses into yes/no/don't know responses.

2) It is also possible to fail to reject the null hypothesis (i.e., say that the alternative is unlikely), when in fact it is not the case. The effects of this kind of error can be overcome by using large samples repeating the research exercise in a variety of circumstances (a luxury which you may not have).

This is referred to as a type II error. The ability of a test to resist this is called its power. Power calculations are mentioned in Sect. 10.7 in Chap. 10 thus it is not necessary to revisit them in-depth; however, we need to say that the power of a test is a measure of the probability that a type II error will be avoided. The power of a test can be calculated manually or more probably arrived at using software. With many software packages it is possible to select an appropriate power and then work backwards to decide how big a sample should be to make your findings significant.

Remember that all testing of data will be prone to these errors; there is no way to know when they happen or predict their occurrence.

Note that not all of the conditions for the use of the tests are covered here. This discussion should be used as a general guide. Small sample sizes could be problematic and this should be discussed with a statistician.

Note that agree/disagree responses can be treated in the same way as yes/no responses. The stages where the data are reclassified are a way of making dependent classes into independent classes. Remember that attitudinal scales have to be treated with caution. The combination of several responses into a score can be tested with the Mann–Whitney 'U' test.

15.12 Getting it Wrong

By now it should (hopefully) be apparent that the whole process is driven by probability. By being careful you will probably be able to gain some insight into the circumstances leading to the data you have collected and into the workings of the world at large, except its power. Power calculations are mentioned in Chap. 10. Suffice to say that the power of a test is a measure of the probability that a type II error will be avoided. The power of a test can be calculated manually or more probably arrived at using software. With many software packages it is possible to select an appropriate power and then work backwards to decide how big a sample should be to make your findings significant.

Remember that all testing of data will be prone to these errors. There is no way to know when they happen or predict their occurrence.

Summary

- Conducting a hypothesis test follows a well-defined procedure, which is common to all tests; data are used to generate a test statistic. The test statistic is used to produce a *p*-value, the *p*-value is then interpreted to accept or reject the null hypothesis and this is then interpreted in light of the original research hypothesis.

- The selection of hypotheses has a great influence on the success of the process. The exact wording of an experimental hypothesis can be influenced by the anticipated use of a particular test.
- Parametric tests should be favoured over non-parametric tests when it is practical to do so.
- The process of testing is subject to errors, not only from the original data used, but also from faulty interpretation of the results due to the influence of random chance. These errors are inherent in the process of testing because a fundamental condition for their use is the random acquisition of data.

So far we have been concerned with describing key aspects of sets of data and using these to answer questions related to research topics. These have related to the same quantity or variable being measured for different groups, e.g., the heights of adult males in different age groups. But what can we do when we have the opposite arrangement, which is when we have one group and different variables? This is discussed in the following section.

15.13 Correlation and Regression

Both correlation and linear regression are alternative methods of examining a relationship between variables. The purpose of the two tests is distinct; which test you choose depends on the aim of the research, whether you are looking to describe a relationship (correlation) or you are using the data for modeling or prediction (regression). This section is only concerned with linear relationships between two variables, known as bivariate.

15.13.1 Correlation

A correlation is used to measure both the direction and magnitude of a relationship between two variables. The variables are usually from the same individuals or matched cases. The test is quite descriptive. If a relationship exists between variables, this does not infer any causality.

Correlation can only be used for data where the numbers can be ordered, i.e., ordinal or interval data. If both the variables are measured on an interval scale, the summary statistic used, assuming that each variable is normally distributed, is the Pearson's correlation coefficient (r). The Spearman's correlation coefficient (rho, r), which is the non-parametric equivalent, is used when one of the variables is measured on an ordinal scale or the data do not approximate a normal distribution.

The tests produce a correlation coefficient that ranges from -1 to $+1$. The closer the correlation is to either -1 or $+1$, the stronger the correlation. Zero represents no relationship between the two variables. The direction of the correlation, either positive or negative, informs us how the two variables are related. A positive correlation exists when the two variables move in the same direction; a negative correlation is

one where the two variables move in opposite directions. If we were to increase the repetition time in an MRI scan for a given number of slices, and this was plotted together with the overall scan time, we would see a negative correlation. As repetition time increases overall scan time decreases.

The best way to visualise the direction and magnitude of a correlation is to produce a scatter plot. Figures 15.14, 15.15 and 15.16 illustrate three different scatter plots demonstrating a strong positive relationship, a weak negative relationship and no relationship between the two variables in question. The plots give you an idea of the strength of relationship but no exact figure. A strong positive correlation is shown in Fig. 15.14.

15.13.1.1 Obtaining a Correlation Coefficient

The first step is to look at the data and decide whether a Pearson's or Spearman's correlation is appropriate. To perform a Pearson's correlation the variables then both have to be interval data; they both have to approximate a normal distribution. A weak correlation is shown in Fig. 15.15. Figure 15.6 shows no correlation.

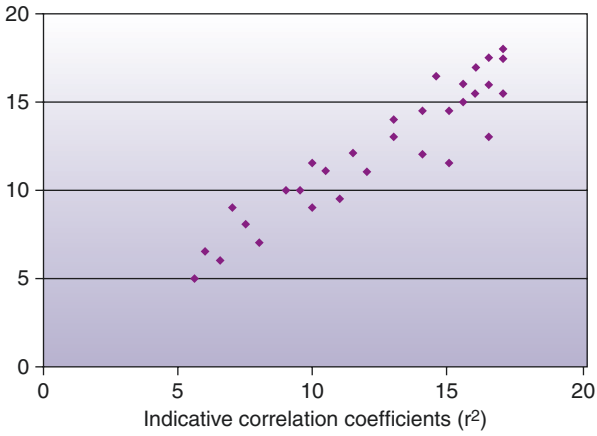


Fig. 15.14 Strong positive correlation

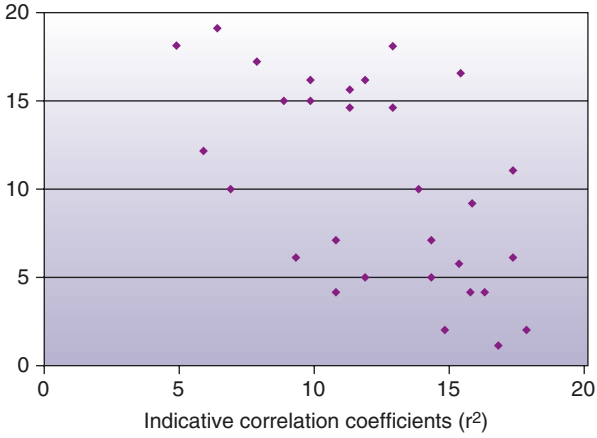


Fig. 15.15 Weak correlation—inverse relationship

The output from a correlation analysis should provide you with a correlation coefficient and a probability value. If the probability value is ≤0.05 (or pre-specified value), you can reject the null hypothesis that there is no correlation between the two variables.

The strength of the association between variables can be classified using the following guidelines provided by Cohen [1].

- Large: $r^2 = 0.5–1.0$
- Medium: $r^2 = 0.3–0.5$
- Small: $r^2 = 0.1–0.3$

In some instances, researchers take a good correlation to mean good agreement. As shown in Fig. 15.17 the correlation is good, but there is no agreement between the scores as there is no data on the line, which represents equal scores from both variables.

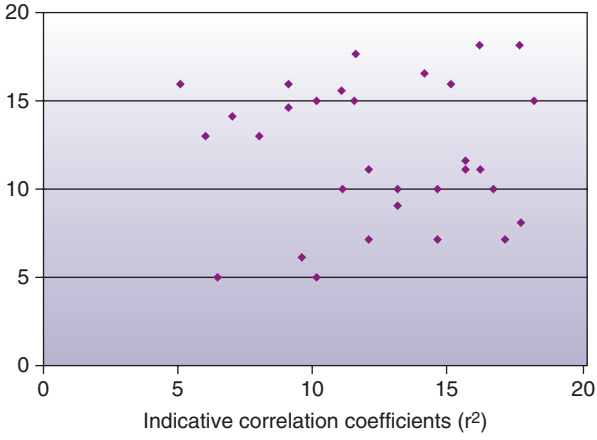


Fig. 15.16 No correlation

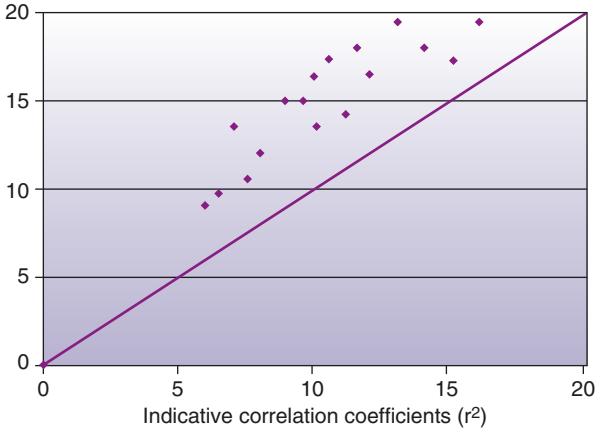


Fig. 15.17 Difference between correlation and agreement

Correlations: points to consider.

- Do not assume that a smaller *p*-value implies a greater strength of association; this is what the r value explains. When sample sizes are large it is relatively easy to get significant correlations, but they may be weak associations that are of little if any practical significance. This is the most common mistake people do in correlation.
- A good correlation does not necessarily mean good agreement. Figure 15.12 illustrates a scatter plot with a correlation of 0.963, $p < 0.001$, a significant, very large association. A line has been added showing where the x and y values are equal, but of the 31 data points on the graph there is only one point on the line, so is there good agreement?
- Using a correlation on a rating scale is controversial. The numbers in the scales often do not have a true meaning and so are more ordinal in nature and in this case a correlation should not be used. However, if you are clear in the write-up that the correlation only provides a general indication rather than a specific value, then it may be used.
- Correlation does not mean causation. It may suggest causation and so may need further studies of a different design to investigate to see if causation really exists. A classic example of this problem is given by the significant and strong correlation between ice cream sales and the number of shark attacks on swimmers in Australia. From this do we conclude that eating more ice cream causes an increase in shark attacks on swimmers? Or do we consider another variable. The hotter it gets the more ice cream is eaten and equally it is the more likely that people enter the water to cool off.

15.13.2 Regression

Regression is used to examine a relationship between variables. Unlike correlation, regression can explore non-linear relationships although only an agreement. Linear regression is discussed here. Unlike correlation, regression implies a direction of influence of one variable on the other. Another difference is that both the dependent (observed or outcome) and independent variable (predictor) must be identified. When drawing a scatter plot of data for a regression the dependent variable is always placed on the ordinate (y axis) and the independent variable on the abscissa (x axis). In addition a line is fitted to the data. There are a number of methods of fitting a straight line, but a common method is the use of the least square regression (LSR). This is a straight line that minimises the sum of the squares of the distances between the line and the data points (Fig. 15.18). The fit of the line to the points is explained by the value r^2, which is a value between 0.0 and 1.0, and it has no units. An r^2 value of 0 indicates that knowing x does not help you predict y. There is no linear relationship between x and y, and the regression line is a horizontal line going through the mean of all y values. When r^2 equals 1.0, all points lie exactly on a straight line. Knowing x allows you predict y perfectly.

The resulting line, called the regression line, is characterised by the formula $y = a + bx$ where y and x relate to the dependent and independent variables, a is the

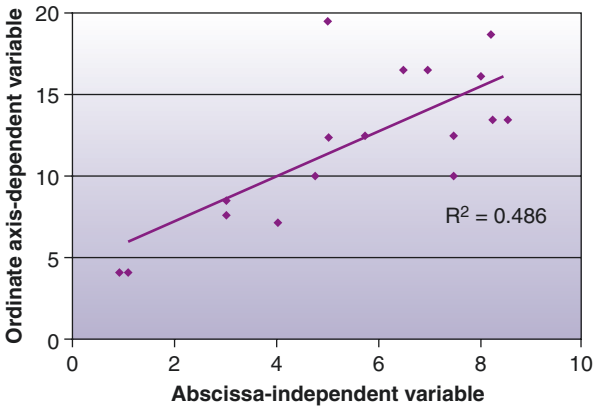
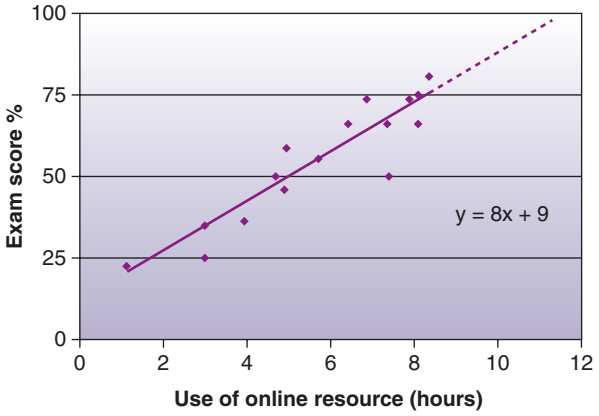


Fig. 15.18 Regression plot

Fig. 15.19 Regression line—formula and extrapolation



intercept and b the slope of the line (see Fig. 15.19). This is explained further in Sect. 15.14 when looking at the regression output for SPSS.

As with other tests it is possible to give a confidence interval (CI) for the regression line. This can be plotted on the scatter plot and it gives curved lines above and below the regression line. The area contained within these lines is the area that has a 95% chance of containing the true regression line. A second set of curved lines (95% prediction interval) can be added to the graph that gives an area in which you expect 95% of all data points to fall. The area within these lines is larger than the 95% confidence interval and would represent our confidence in predicting a single value on the regression line.

Aspects to consider about regression

- Regression uses interpolation (the construction of new data points within the range of known data points) to make predictions of values. Do not use it for

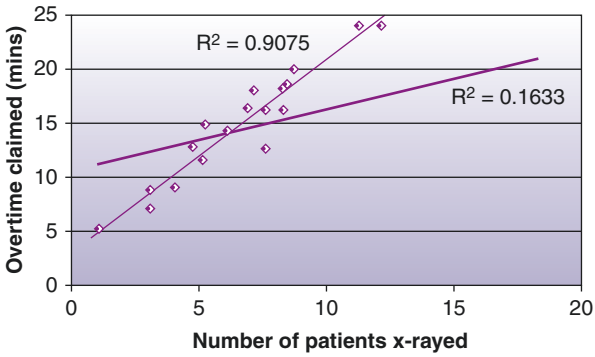


Fig. 15.20 Effect of outliers on regression

extrapolation (the construction new data points outside of known data points), as the predictions may be less reliable, as indicated by the dotted line in Fig. 15.18.

- Regression is sensitive to outliers, particularly in certain locations, as is correlation; the model's ability to predict what is happening is seriously affected by outliers. The data should be checked very carefully for outliers. If we look at Fig. 15.20, we can see some data points on the left and a regression line (light purple). If an outlier is added (purple circle), the regression line changes (solid purple line) and extends. The information between the outlier and the data point closest to it is based almost entirely on the outlier and even though we would be interpolating the data the confidence we have in the regression line being accurate must be suspect. Also, the outlier has a large impact on the slope of the regression line.

Key Points

- When considering proof it is necessary to rely on repeatability and inference from other related research to suggest that proof has been obtained, e.g., ionising radiation is taken to cause cellular damage because no experiment has ever found otherwise, and slightly different styles of experiments will confirm the same result.
- One study on its own does not give enough evidence.
- Any bias in your data means that your evidence becomes progressively weaker as more bias is introduced.
- It is not only important to understand how fragile an argument for causation is when you are presenting one, but it is also very important to recognise inappropriate claims of causation made by others.
- The safest course of action is to present the evidence and comment on its strengths and weaknesses.
- Allow the readers to form their own opinions guided by your observations rather than telling them what to conclude.

15.14 Conclusion

Choosing the correct statistical test is important because the answer it provides is the culmination of all the work that has gone before it. A brief overview of the different types of data was covered in this chapter. Also covered was how it might be best to summarise data using measures of central tendency. Methods of reporting dispersion, which should be used, were presented together with measures of central tendency to describe data. We looked at the basic inferential procedures that allow us to make inferences from the data. The assumptions behind the parametric tests were discussed. When each test might be used was considered and some worked examples were provided of more common tests performed using SPSS.

We also discussed the null and alternate hypotheses, probability, significance levels and type I and type II errors: all of which are key to understanding the process of getting and reporting statistical results.

15.15 Appendix

Independent Samples *t*-Test

The test below relates to emotional intelligence scores of radiography students, referred to in the printout as test scores.

The independent samples *t*-test printout consists of two tables. The first table describes the data. In this case the dependent variable was called test score. The next two columns contain the names of the two groups being compared and the number of subjects in each group. The mean score is then seen. In this case we can see that the diagnostic radiographers had a higher test score (36.91) than therapeutic radiographers (32.29). We then can see the standard deviation and standard error of the mean (SEM), which is the standard deviation divided by the square root of the sample size ($SEM = \sigma/\sqrt{n}$). This table tells us the difference between the groups, but not whether this difference is significant. For that we need to look at the second table.

Group statistics

	D/T	*N*	Mean	Std. deviation	Std. error mean
Test_Score	Diagnostic	80	36.91	4.450	0.498
	Therapeutic	21	32.29	5.781	1.261

The second table should be read as two tables in one and to help with this colour has been added splitting the two tables up. The first table coloured blue looks at the result of the Levene's test, a test that tells you if we have equal variances or not between the two groups. The null hypothesis of the Levene's test is that the variance of the two groups is equal. If the test reaches significance ($p \leq 0.05$), then we reject the null hypothesis that they are equal and accept that they must be different. In the

example below the significance of the Levene's test is 0.449 and therefore we do not reject the null hypothesis and we can assume equal variances. With that decided we now know to read the top row of the *t*-test result which is coloured yellow. If the Levene's test had been ≤0.05, we would have read from the bottom row of the yellow *t*-test as equal variances could not be assumed.

Reading the test we can see the test statistic is 3.974 and the significance value is being reported as 0.000. This would be written up for publication as $t = 3.974$, $p < 0.001$. The significance value is only reported to three decimal places so we cannot report an exact value in this instance, just that it is less than 0.001.

Independent samples test

		Levene's test for equality of variances		*t*-test for equality of means						
									95% confidence interval of the difference	
		F	Sig.	*t*	df	Sig. (2-tailed)	Mean difference	Std. error difference	Lower	Upper
Test_ Score	Equal variances assumed	0.578	0.449	3.974	99	0.000	4.627	1.164	2.316	6.937
	Equal variances not assumed			3.412	26.544	0.002	4.627	1.356	1.842	7.411

The difference between the two groups is 4.627. This can also be calculated by subtracting the means in the group statistics table.

This finding might be reported as follows: The tests scores for diagnostic radiographers were significantly higher (36.91) than therapeutic that of radiographers (32.29), $t = 3.974$, $p < 0.001$.

ANOVA

When using SPSS it is possible to test more than one assumption at a time; the table below contains two analyses. The analyses are all part of one data set and are looking at the diagnostic and therapeutic scores for a mental rotations test looking at comparing the 3 years of students. The first table as with the *t*-test looks at homogeneity of variance and as the significance values are above 0.05 we can assume equal variance. The ANOVA is quite robust to deviations from homogeneity of variance but it is an assumption of the test. Descriptive of the variables is not done automatically with this test as with the t-test but can be toggled on if required.

Test of homogeneity of variances

		Levene statistic	df1	df2	Sig.
Diagnostic students' scores	Based on mean	0.160	2	94	0.852
	Based on median	0.176	2	94	0.839
	Based on median and with adjusted df	0.176	2	93.675	0.839
	Based on trimmed mean	0.183	2	94	0.833
Therapeutic students' scores	Based on mean	0.431	2	92	0.651
	Based on median	0.273	2	92	0.762
	Based on median and with adjusted df	0.273	2	91.977	0.762
	Based on trimmed mean	0.461	2	92	0.632

The second table is the ANOVA analysis. The important columns are the final two, which give us the test statistic for each test and the significance value. The first test looking at the 3 diagnostic years was not significant ($F = 0.349$, $p = 0.707$). The second test looking at therapeutic students is significant ($F = 6.014$, $p = 0.004$). What we do not know at the moment is what year group is different from which other year group. The test simply states that there is a difference, it could be that all year groups are different from each other or just 1 year might be different from another, any variation is possible and the more groups you have, the bigger the possible number of possibilities.

ANOVA

		Sum of squares	df	Mean square	*F*	Sig.
Diagnostic students' scores	Between groups	15.927	2	7.963	0.349	0.707
	Within groups	2147.558	94	22.846		
	Total	2163.485	96			
Therapeutic students' scores	Between groups	258.663	2	129.331	6.014	0.004
	Within groups	1978.537	92	21.506		
	Total	2237.200	94			

In order to find out where the difference is we have to undertake a post hoc test. As there was no significant difference between the diagnostic student year groups, just the therapeutic students we only need to do a post hoc test for the therapeutic students. There are a variety of post hoc tests that can be done and in this instance a Tukey's HSD (honestly significant difference) test was used. In the table below I have highlighted the significant differences. The table repeats itself, but we can see that the first year's score was significantly different to the second year's score ($p = 0.016$), but not the third year's score ($p = 0.833$). The second year's score was also different to the third year score ($p = 0.007$).

Multiple comparisons

Tukey's HSD

Dependent variable	(I) year	(J) year	Mean difference (I–J)	Std. error	Sig.	95% confidence interval	
						Lower bound	Upper bound
Therapeutic students' scores	1	2	3.138*	1.110	0.016	0.49	5.78
		3	–0.692	1.201	0.833	–3.55	2.17
	2	1	–3.138*	1.110	0.016	–5.78	–0.49
		3	–3.830*	1.230	0.007	–6.76	–0.90
	3	1	0.692	1.201	0.833	–2.17	3.55
		2	3.830*	1.230	0.007	0.90	6.76

*The mean difference is significant at the 0.05 level

Mann–Whitney *U* Test

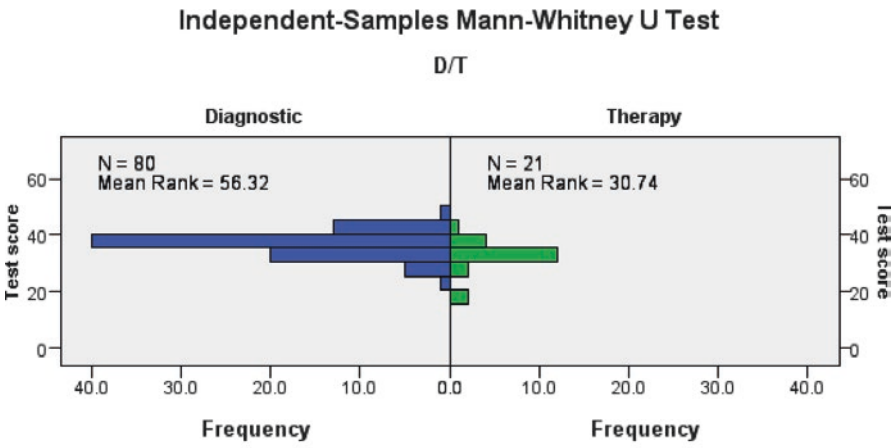
The SPSS printout for non-parametric tests is far simpler than for parametric tests. The table below is the printout for the same data as the ANOVA test. Again, two tests were performed and each row reports one test. The first column tells you what is being tested. The second column what statistical test was performed and then the significance level of the test. The final column tells you whether or not to reject the null hypothesis.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Test score is the same across categories of D/T.	Independent-Samples Mann-Whitney U Test	.000	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

Double clicking on this box brings up further information seen below. First, we see a bar chart that shows us the frequency of the test scores for each of the two groups. Second, we see a further box that tells us the test statistic for the test we carried out. In this case $U = 414.5$, $p < 0.001$.



Total N	101
Mann-Whitney U	414.500
Wilcoxon W	645.500
Test Statistic	414.500
Standard Error	119.133
Standardized Test Statistic	-3.572
Asymptotic Sig. (2-sided test)	.000

Kruskal–Wallis Test

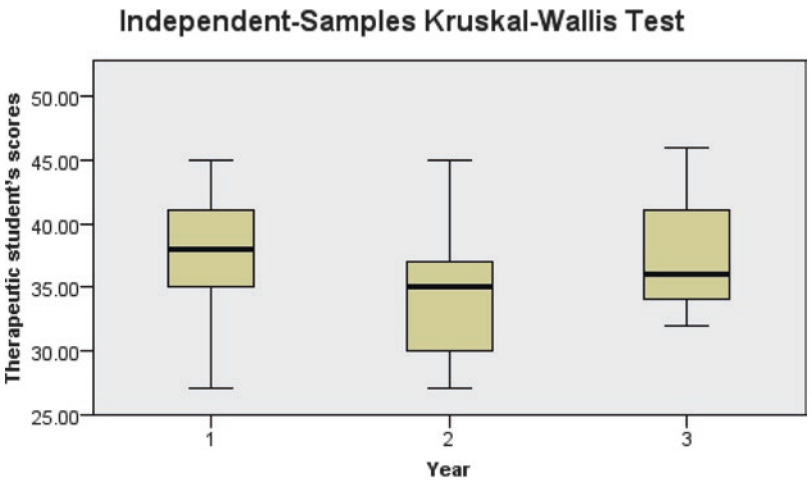
The printout for the Kruskal–Wallis test is very similar to the Mann–Whitney test above consisting of the same columns. As with the ANOVA test in this case two tests were performed and each row reports one test. The first column tells you what is being tested. The second column what statistical test was performed and then the significance level of the test. The final column tells you whether to reject the null hypothesis.

Hypothesis Test Summary

	Null Hypothesis	Test	Sig.	Decision
1	The distribution of Diagnostic student's scores is the same across categories of Year.	Independent-Samples Kruskal-Wallis Test	.808	Retain the null hypothesis.
2	The distribution of Therapeutic student's scores is the same across categories of Year.	Independent-Samples Kruskal-Wallis Test	.009	Reject the null hypothesis.

Asymptotic significances are displayed. The significance level is .05.

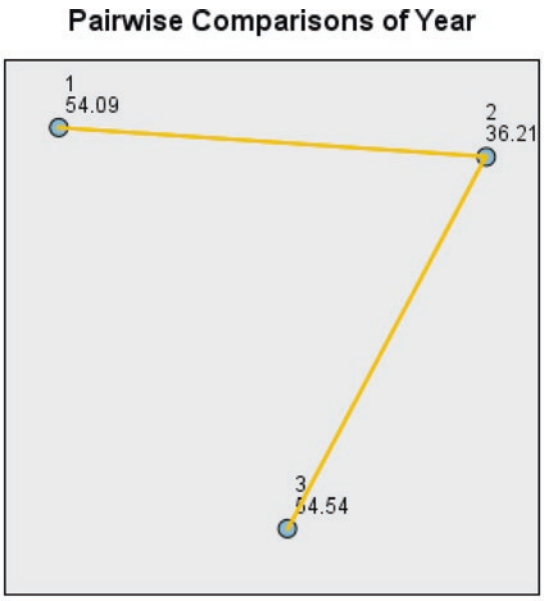
Again, double clicking the box in SPSS brings up a pop-out with more information. The pop-out shown below shows you each year's score for the therapeutic students in the form of a box and whisker plot. The second box gives you the test statistic and repeats the significance level of the test. We can now report the finding of the test, $H = 9.321$, $p = 0.009$.



Total N	95
Test Statistic	9.321
Degrees of Freedom	2
Asymptotic Sig. (2-sided test)	.009

1. The test statistic is adjusted for ties.

Clicking on pairwise comparisons brings up further information seen below. This is the post hoc test which takes into consideration that we are doing multiple tests. Just as with the ANOVA test the post hoc test is telling us which groups are different to which other groups. In this instance group 1 and 2 are different to each other as are groups 2 and 3.



Sample 1-Sam...	Test Statistic	Std. Error	Std. Test Statistic	Sig.	Adj.Sig.
2-1	17.882	6.575	2.720	.007	.020
2-3	-18.328	7.281	-2.517	.012	.035
1-3	-.445	7.110	-.063	.950	1.000

Correlation

The following data looks at patient information leaflets and how readable they are to the public. The analysis compared readability score with average sentence length and the use of passive voice in the text. As with other tests you can undertake more than one test at a time with correlations. Colour has been added to the table to show that each test is undertaken twice in the table and each variable is correlated with itself, which of course will give a perfect correlation of 1. Reading horizontally across the first row we can see that the readability score is correlated with sentence length ($r = 0.532$), a moderate positive correlation and that this correlation is significant ($p < 0.001$). It can also be seen that readability score and passive voice have a correlation coefficient of 0.098 which is not significant, $p = 0.379$. Looking at the second row the first box looks at sentence length and readability score and is repeat

information, the next box is the correlation coefficient with itself, but the third box gives us new information in that sentence length and passive voice are related to each other and have a correlation coefficient of 0.280, $p = 0.10$. The third row of boxes is all repeat information.

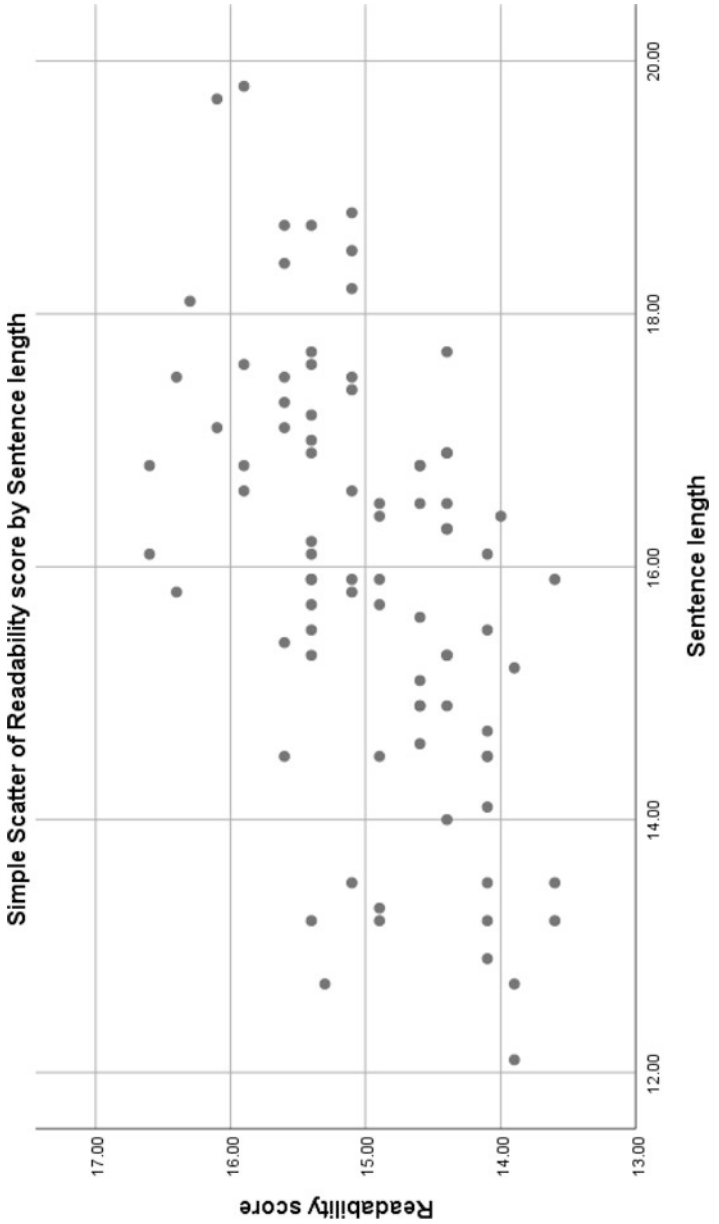
Correlations

		Readability score	Sentence length	Passive voice
Readability score	Pearson Correlation	1	.532**	.098
	Sig. (2-tailed)		.000	.379
	N	83	83	83
Sentence length	Pearson Correlation	.532**	1	.280*
	Sig. (2-tailed)	.000		.010
	N	83	83	83
Passive voice	Pearson Correlation	.098	.280*	1
	Sig. (2-tailed)	.379	.010	
	N	83	83	83

**. Correlation is significant at the 0.01 level (2-tailed).

*. Correlation is significant at the 0.05 level (2-tailed).

When doing a correlation it is very useful to view the result in the form of a scatterplot. This will show the relationship between the two variables. It is not usual to put a line of best fit on as this is technically a regression line and implies that you have done a regression.


Simple Scatter of Readability score by Sentence length
Readability score
17.00
16.00
15.00
14.00
13.00
12.00
14.00
16.00
18.00
20.00
Sentence length

Regression

Regression analysis is widely used for prediction and is used to infer a causal relationship between the independent and dependent variables. In the above example it might be possible to suggest that if we change the average sentence length of a document, we can change its readability score. If we undertake a regression analysis on the above data the first column tells us the correlation coefficient followed by the coefficient of determination (r^2 value). The r^2 value tells us that approximately 28% of the variance in readability score can be explained by the sentence length.

Model summary				
Model	*R*	*R* square	Adjusted *R* square	Std. error of the estimate
1	0.532[a]	0.283	0.275	0.62398

[a]Predictors: (Constant), Sentence length

The ANOVA table informs us whether our regression model explains a statistically significant proportion of the variance. The *F*-ratio in the ANOVA table (see below) tests whether the regression model is a good fit for the data. Remember we could add more variables to the model and build up a picture of what affects readability score. The table shows that the independent variable (sentence length) is statistically significant when predicting the dependent variable (Readability score), $F = 32.035$, $p < 0.0001$.

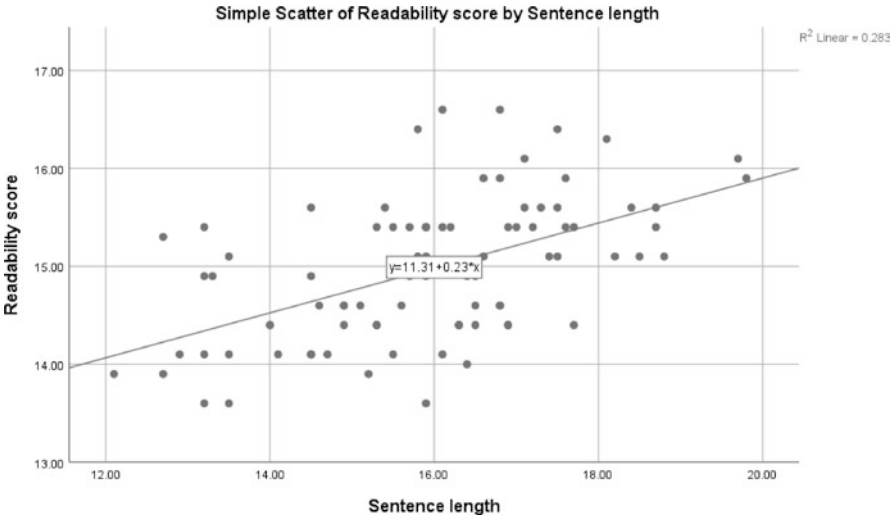
ANOVA[a]						
Model		Sum of squares	df	Mean square	*F*	Sig.
1	Regression	12.473	1	12.473	32.035	0.000[b]
	Residual	31.538	81	0.389		
	Total	44.011	82			

[a]Dependent variable: readability score
[b]Predictors: (constant), sentence length

The coefficients table informs us about the values of the regression line. The column marked B tells us where the line intercepts the axis (11.312) and the slope of the line (0.229), the model predicts an increase of 0.229 in readability score for every increase of 1 word in sentence length. The graph following the coefficients table shows this data and also has the data from the above table present in the form of a formula. The zero intercept is not shown on the graph but let us put some figures into the formula. The intercept is 11.3 and if we want to know the value for a sentence length of 12 the formula becomes $11.3 + (0.23 \times 12) = 14.07$, which looking at the graph seems to be about right. The final column is the significance value for the regression coefficient.

Coefficients[a]						
		Unstandardised coefficients		Standardised coefficients		
Model		*B*	Std. error	Beta	*t*	Sig.
1	(Constant)	11.312	0.650		17.403	0.000
	Sentence length	0.229	0.041	0.532	5.660	0.000

[a]Dependent variable: readability score



Reference

1. Cohen J. Statistical power analysis for the behavioral sciences. 2nd ed. Hillsdale, NJ: Lawrence Erlbaum; 1988.

Further Reading

Altman DG. Practical statistics for medical research. London: Chapman & Hall; 2018.
Campbell MJ. Statistics at square two. Understanding modern statistical applications in medicine. 2nd ed. Oxford: Wiley-Blackwell; 2006.
Campbell MJ, Swinscow TDV. Statistics at square one. 11th ed. Oxford: Wiley-Blackwell; 2009.
Claude CS, Longo G. The Deluge of spurious correlations in big data. Found Sci. 2017;22(3):595–612.
Miles J, Shevlin M. Applying regression and correlation: a guide for students and researchers. London: Sage; 2001.

Qualitative Methods and Analysis

16

Peter Williams and Susan Cutler

16.1 Introduction

First in this chapter we examine what is meant by the term 'qualitative research'. It is an approach that seeks to understand human behaviour (i.e., what its essential qualities are) and, in our context, health behaviour and understanding of human health. Over recent years qualitative methods of data gathering and analysis have gained popularity among healthcare professionals as we seek to question traditional approaches to the delivery of healthcare. Qualitative research uses a number of different methods to collect the data that generate narrative or non-numerical information; it tends to use 'language data': written or oral. In contrast, quantitative data collection focuses on collecting numerical data and then employs statistical analysis to test hypotheses. It should be remembered that similar ways of collecting data can be employed in both qualitative and quantitative methodologies. Qualitative studies can employ frequency counts. Language data can be used in quantitative studies. The overall aims of a study determine the methodical approach taken. A qualitative approach is utilised when you are asking 'how and why' questions rather than 'how often' or 'how many': a quantitative approach may be more appropriate for the latter questions. A qualitative approach is used to try and gain insight into an individual's view of their own world. It is important that a researcher does not make any value judgements about the data collected. The focus is on the meaning and experiences of individuals or groups, to analyse how and why people form associations with other people, things, and their immediate environment.

P. Williams (✉)
Department of Information Studies, University College London, London, UK
e-mail: peter.williams@ucl.ac.uk

S. Cutler
School of Health and Social Care, Teesside University, Middlesbrough, UK
e-mail: s.cutler@tees.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_16

There are two key aspects of a qualitative approach.

- We are not trying to quantify or count things by gathering interview data (except where closed 'survey-like' questions are administered to a large sample of people).
- We are interested in understanding people and how they behave or think. We need to explore the ideas they hold in their minds as ideas are one of those things that we cannot observe.

A qualitative approach can be beneficial in the drive to expedite the service improvement agenda for the benefit of patients, as it seeks to explore their and practitioners' experiences of contemporary imaging and healthcare delivery. The Society and College of Radiographers maintains that research will support change within diagnostic imaging, radiotherapy, and oncology departments so ensuring practice and patient-centred care becomes fully evidence based [1]. In order to change and develop our practice for the benefit of a patient and client, we need to understand the environment in which we live and work and a patient's experiences and expectations. Any research we undertake must have a purpose; it would be unethical to undertake a study that did not attempt to explore phenomena relevant to the development of contemporary practice.

16.2 Interviews

Any project that seeks to examine the attitudes, experiences or behaviour of people, in other words social research, will almost inevitably involve data collection by some form of interview or survey. An interview might appear to be the easiest form of data collection. After all, who has not sat down and had a good chat with someone? And you only have to turn on any form of media, be it YouTube, TV, podcasts, or whatever, to hear an almost endless stream of interviews/conversations/dialogues.

The first thing to say, however, is that an interview is not like a 'chat', despite the fact that the best interviewers make it appear so, and even though a research interview has been likened to a conversation. This distinction becomes clearer later, but suffice to say here is that unlike a 'chat' or a 'normal conversation', in which you may expect a roughly equal contribution from the participants, a research interview should aim to be about 90:10 in favour of an interviewee. It is their views we are interested in. By contrast, even the biggest pub bore would be hard-pushed to keep up that ratio, and if he did, his audience would soon be tiptoeing out of the back door.

What is an 'interview', in particular a 'qualitative' one, and why do we conduct them? Let us look at a couple of definitions.

- Qualitative interviews develop a rapport between interviewer and interviewee allowing the researcher to probe and explore complex issues [2].

- Qualitative interviews tend to ask about the details of [a phenomenon], how it happened, and how the respondent felt throughout. [This] can allow researchers to capture certain social phenomena in ways that other techniques simply cannot [3].

Also, importantly, we are interested in individual experiences: stories, insights, things that you would not capture in questionnaire tick boxes. Patients and medical staff were interviewed in a study [4] at various general practitioner (GP) practices where a touch-screen information kiosk had been installed. Interviewees not only had a wide variety of views, but ones which were totally unforeseen, and almost unforeseeable. One interviewee thought that the kiosk merely showed a plan of the forthcoming surgery renovation; receptionists at a large hospital said that as their kiosk was in a booth, it had been mistaken for a toilet cubicle. Given just these two responses, it is clear that a questionnaire could not have captured these perceptions, unless it approximated an interview format by using questions allowing for free-text answers.

One final point regarding the data is accrued from interviews: people often worry about getting enough information to generalise or to make the research 'meaningful'. Here it is useful to quote Henry Mintzberg, academic and author on business and management, who said, in a very famous quote: 'What is wrong with samples of one? Why should researchers have to apologise for them? Should Piaget apologise for studying his own children, a physicist for splitting only one atom?' [5]. Similarly, television documentaries often contain individuals' descriptions of a particular event and quite often the recounted experience of four or five people gives a very rich picture, offering an insight into what was going on beneath the surface.

As practitioners we deal with diverse and discrete client groups and their experiences of imaging or oncology may be considered for inclusion in this type of study. Interviewing of patient groups may be impractical for undergraduates owing to the complexities of the ethical approval process. But it is an avenue of exploration for qualified practitioners. Nevertheless, this approach can be utilised in a university setting, with appropriate ethical approval, accessing the student body, either undergraduate or postgraduate, employees of the organisation or the general public. For example, you could explore the experiences of practitioners or students of working with hearing impaired patients or patients with learning disabilities. This has the potential to be extended to compare experiences with other health professional groups within a university setting. Examples might be comparisons with students at different stages of their educational programme, different academic pathways, for example, pre-registration undergraduate and postgraduate routes.

Using an open interviewing technique, possible questions could include the following.

- What is your definition of hearing impairment/ learning disability?
- Could you describe any experience you have had with a hearing impaired/learning disability patient?
- How did this make you feel?

- Do you think deafness awareness training would have helped?
- From your own experiences, in what way do you think a hearing-impaired patient's visit to the medical imaging department could be improved?
- Are you in contact with or do you have experience of learning disabilities/hearing impairment outside the placement? (and follow-up with 'could you describe your experiences?')
- How confident do you feel about dealing with people with learning disabilities/hearing impairment?

These questions are by no means exhaustive, but hopefully give you a feel of how questions could be phrased to enable participants in the study to express their opinions and feelings about their experiences.

16.3 Focus Groups

Focus groups are unstructured interviews in which several participants discuss or explore together a specific set of issues. A focus group does not set out to change opinion or test knowledge, but like an interview, it seeks to explore participants' experiences or attitudes. The role of a researcher is to facilitate and guide the debate and discussion by posing a series of questions that the participants explore. This approach can be used in its own right to explore issues. It however is often used to complement questionnaire data or even as the preliminary tool for data collection from which a subsequent questionnaire is developed. The scope of an undergraduate dissertation possibly excludes utilising this multi-method approach due to the time constraints, but can be valuable in a clinical and post-registration context to explore in more depth issues relevant to practice.

The number of participants in a focus group varies; the general rule is to include 6–12 people. Below this number there is the potential for little stimulus or for dominance of one of the participants. With too few participants, you also run the risk of the discussion and commentary from the participants being limited. Above this number you run the risk of having too many people talking at once and the essence of their debate is lost in the overall chaos of conversation. Also, there is an opportunity for some of the participants to hide within the larger group and therefore not contribute. Group dynamics can be significant in the data collection as some do not interview well in groups and others may dominate. Sometimes the public setting of a focus group may inhibit the free flow of ideas or thoughts, which might be captured in a more intimate individual interview. Conversely the comment of one participant may provoke profound or animated debate among the others.

This type of interview usually takes between an hour and 90 min to conduct and can be quite demanding on a facilitator, who needs good social skills and interview technique. As a new researcher this can be quite a daunting aspect. It is worth trying to observe a skilled facilitator before commencing your data collection. Some of the key attributes required of the facilitator include the following.

- Alertness: this refers to the attentiveness of the participants within the group; are there individuals who are dominating or others who are keeping quiet?
- Assertiveness: this refers to making 'demands' of the participants to contribute to the discussion or allow others to speak, without being overbearing.
- Confidence: this refers to having enough self-assurance to ensure the focus group is appropriately conducted, and relevant data are collected.
- Diplomacy: this refers to the ability to diffuse any potential confrontational situations.
- Empathy: this refers to the ability to understand and imagine the feelings of others.
- Encouragement: this refers to gently coaxing participants to share their ideas and thoughts.
- Interpersonal skills: this refers to the ability to establish a rapport with the participants.
- Listening skills: this refers to listening attentively and being alert so that comments can be reflected back to the group for clarification.

Before commencing a focus group, you need to take into consideration the following three factors: the location of the interview, the physical environment, and the group composition.

1. The location of a focus group should be such as to prevent the participants incurring any unnecessary expense.
2. The environment in which an interview takes place should be an adequate size so as not to inhibit any interaction and enable the participants to see each other, but not so large as to detract from the group dynamics or to prevent recording and observing effectively. It is advisable that the location is free from distracting elements such as noise, wall furniture, and busy windows.
3. The participants' location. Are they going to come from a disparate population or are they going to know each other or come from the same background, such as students in a cohort, practitioners from the same clinical department or practice area? Within the context of radiographic practice, whether therapeutic or diagnostic, the gender of focus groups is unlikely to be representative of the population as a whole as the profession is still predominantly female. It is more important that focus group participants do reflect the knowledge base or experiences required for the study.

One of the limitations of collecting data using focus groups is that only a few questions can be addressed and are not usually explored in any detail. On the other hand, the experiences of a larger number of participants can be captured and it is less time-consuming than individual interviews.

It is suggested that focus group participants should be debriefed following their interview; this may include follow-up leaflets or contact details of relevant support groups if applicable. It is unusual for the transcription of focus group data to be sent

to the participants, unlike interviews. Data produced from a focus group can be analysed in the same way as data generated from individual interviews, and is discussed later in this chapter.

Focus groups can be used to explore patients' experiences of healthcare. Particular reference, for example, could be made to different client groups such as older patients, parent perceptions of their children's experiences and those with special needs. Conversely practitioners, students, assistant practitioners, and administrative staff perceptions could be expedited to give a contrasting perspective or to explore the context of the patients' or practitioners' experiences. For example, you could explore students' experiences of using the virtual learning environment while on placement. You could use a questionnaire to capture the diverse experience of as many students as possible as there would have been too many students to interview individually. You could then analyse the responses to the questions on the questionnaire and extract themes to explore within a focus group setting. The focus group could comprise of two groups of about six students each and this exploration will generate rich data which will account for the responses on the questionnaire.

16.4 Recording Your Interview or Focus Group

Recording the data can be done in three ways: voice recording using a digital recorder, video recording, or transcription of key points by a facilitator. The fear of doing this used to be whether it would work; the most infamous problem being tape jams or breakages. Nowadays the problem is more about the battery life of your device. So frustrating when you have undertaken a fantastic interview and obtained excellent data, then discover that your mobile gave up the ghost after 10 min because you forgot to charge it before going out. Apart from making sure this does not happen, you could be extra sure and download a battery charge alarm app; choose carefully as most of these apps alert you when the battery is full, not dying. They give a minority sound when it is low. Take a portable charger with you.

Instead of audio-recording, you could video record an interview or focus group. Participants may however find video recording inhibiting and become self-conscious. It is useful to take notes during the session; total reliance on this method does however mean your focus is on scribing and not on the dynamics of the focus group. As a consequence some comments may be missed and you may misinterpret meaning. It is important to ensure the participants do not all talk at once because data could be lost in the melee. When reviewing recorded data it can be difficult to determine which participant was speaking on each topic. It should be remembered that this type of data can be 'messy' in comparison to that of an interview. In addition, it can be time-consuming to transcribe and analyse the data.

16.5 Advantages and Disadvantages of Interviews and Focus Groups

Many of the considerations for conducting an individual interview apply equally in the context of a focus group. The advantages and disadvantages of interviews and focus groups are highlighted below.

16.5.1 Advantages

- Participants' own words: an interview/focus group offers an unrivalled opportunity for participants (respondents) to explain, reflect, and pontificate, using their own words and hence not being shoehorned into the words and the agenda of an interviewer.
- Interviews/focus groups develop a rapport between interviewer and participants allowing a researcher to probe and explore complex issues.
- Full and complete responses: participants have the opportunity to expand on their answers, and often provide details not even considered by an interviewer.
- Observational opportunities: these manifest when interviews take place in a subject's workplace or other location of interest to a researcher. It is amazing what important data can be gleaned from observing someone in action, fielding phone calls, having to break an interview to go and resolve some crisis, etc. Focus groups are unique in that participants build on the answers of others in the group adding to the production of rich data through social interaction. They can produce new thoughts a researcher may not have thought of, and participants encourage discussion of the topic collectively as a group.
- Non-verbal communication: this can also be taken into account.
- Clarification/follow-up questions: these are possible, given the synchronous and dynamic nature of interviews. Captive subjects: as mentioned below, interviewees should be given the right to terminate an interview at any time. This happens very rarely; usually because of work commitments or some unforeseen circumstance. Generally, the interviewee is more or less captive. Hopefully, interviewees will be interested enough in the topic to not think in those terms themselves, but even if they do, they are often prisoners of their own politeness.
- Known respondent: unlike in a postal or online survey, you know who is responding as they are sitting in front of you (unless they are not who they claim to be, which is pretty unlikely).

16.5.2 Disadvantages

It all seems so fine up to here, but of course, as with every other research method, there are disadvantages. We provide some examples; no doubt you will be able to find others if you think hard or cruise the internet for long enough.

- Conformity to expectations: as you are right there with them, participants sometimes feel they have to say things that they assume you want them to say: people like to be nice, and to be cooperative. It is important, therefore, to underscore that you are not judging them and also that you are not promoting anything. In our kiosk study we had to make it very clear we were not part of the company producing them.
- Interviews/focus groups can be time-consuming to conduct and interviewer bias can compromise the quality of data collected.
- The success of the interview/focus group depends upon the skill of an interviewer/moderator and the articulacy of a respondent.
- Attempts to be rational: in another manifestation of the desire to conform, people may try to be more logical than they really are. Having postulated that health information is very important in managing chronic illness, they may feel they cannot then say that, actually, they prefer to not think about it at all. A good approach, instead of confronting them with questions about how much information they need for their health, would be to simply ask how they cope, and look out for where information crops up in their answer.
- Reticence/shyness: you may come across people who find the whole interview process intimidating, especially a 'research' interview. Unless you are good at putting them at their ease, you may find the time and expense yielding very little return. Of course, you as an interviewer may also be somewhat introverted. Before you decide on your project, to the extent to which you have a choice, consider carefully if it is one in which interviews are necessary.
- In focus groups power dynamics can be problematic with some participants domineering and others not speaking up. Some participants may withhold true feelings so as not to create disagreement within their group. These issues can be overcome by having a supportive environment with established ground rules.
- Time/cost: clearly, interviews are time-consuming, and costly if you have to travel to undertake them. As noted above, the prize is usually rich, fascinating, and very worthwhile data. However, you do need to weigh up whether you have the resources and, if so, whether the benefits that accrue are worth all the effort.
- Data analysis is difficult: it is not easy to extract meaningful information from, sometimes, reams of transcripts or interview notes. More on this can be found below.
- Unrepresentative samples: sometimes, in carrying out qualitative research, there is a tendency towards finding and relying on a few key people or, alternatively, people who happen to have the most time to speak to you. This can be an issue with undergraduate research when students often reciprocate participation in studies.
- Poor articulation: depending on the group(s) you decide to interview (individually or in a focus group), there may be a problem with participants being able to articulate their thoughts and opinions. In addition to possibly limiting the amount of data you accrue, there is a potential problem here of an articulate individual's views being over-represented, simply by their ability to express them.

- Researcher bias: you need to be aware that your own social background assumptions, attitudes, beliefs and behaviour can affect a research process and should be acknowledged when writing up your project. It is important therefore to reflect on your own stance with regard to the research topic and acknowledge that your own personal experiences have the potential to lead to bias in the phrasing of your questions and interpretation of the data. This process of self-reflection is termed 'reflexivity'. Reducing investigator bias through admission of one's beliefs and assumptions in this manner increases a study's credibility.
- Participant overload: one of the problems associated with undergraduate research is the potential sample size being limited to the number of students studying diagnostic or therapeutic radiography. There is a real risk that students become exhausted by participating in a number of studies. This also means that results are contextualised to the population and may not be extrapolated to the wider population. As explored earlier with regard to researcher bias, there is the potential for participant bias in this context.

16.6 Other Techniques

To obviate some of the above problems, it is recommended to at least incorporate some of the following techniques to make your interview research more robust.

- Purposive sampling: this focuses on a specific population, so the participants have, as close as possible, the same or similar experiences of the phenomenon you are studying. This only works, of course, if you are not interested in a wide exploration.
- Choose deviant case: the exception proves the rule. If you can find someone who does not seem to conform, the data gathered will put the other research into perspective and will give you an overall richer picture.
- Member check: this refers to going back to the participants (if possible) or their peers and checking with them that this is what was said, and that the way you are interpreting it is correct.
- Researcher 'reflexivity': this is simply acknowledging your own views and where you fit in. Reflecting on your analysis helps you to decide whether your own biases and pre-suppositions have 'contaminated' your data.
- Triangulation: this uses another data gathering method and combines the results. This approach is often used in the study of some aspects of human behaviour. A good idea is to combine interviews or focus groups with a questionnaire. Much survey research actually starts with a qualitative phase in order to tease out the main issues, around which a survey questionnaire can be constructed. However, each approach should stand on its own merit.

Using just one method of data collection provides only a limited view of the complexities of human behaviour, so one method alone might introduce some bias. A multi-method approach that yields the same results can increase the confidence in those results.

16.7 Aims of Your Interview, Types of Interviews, and Preparation Needed

You will need, of course, ethical approval for your study, as described in Chap. 6, so this section assumes you have done this, and the study has been approved. In terms specifically of interview preparation, the first thing you need to do is to go back to your original aims and objectives and determine the extent to which your interview will address and inform these. The aims of your interview will also determine the style with which it is undertaken. Briefly, there are three main factors involved: how specific a topic may be; the number and type of questions, and the order of questions. At one extreme is an interview in which there are very few, if any, predetermined questions and the interview is led to a great extent by the interviewee. This is known as an 'unstructured' interview. This kind of interview is useful if the topic is vague or if you as a researcher have little prior knowledge of the field.

More common is the 'semi-structured' interview, where a researcher has a number of issues to cover and has a loose set of questions. Here the order of questions is not important, but a researcher tries to cover all the question areas, albeit not necessarily in a given order.

There is also a 'structured' interview, where the questions are more specific, and asked in a predefined order. A structured interview is often undertaken on the telephone.

If possible, try to choose the interview location yourself. You really want to talk to your participants in the environment of whatever it is you are studying. For example, if you were doing a study on some aspect of the work of radiographer or practitioner, it would be better to interview them actually at work so that you could see aspects of their work that they may (or may not) mention, and the contextual factors that may inform your study.

A very important part of your interview preparation is to look at individual questions/areas/and themes. Questions can include those seeking the following.

- Facts (e.g., age, gender, education, behaviour, experience).
- Opinion/preference/attitude/feelings.
- Motivation or intention (e.g., likeliness, willingness).

You also need to decide whether or not to record your interview. Here you need to consider the following.

- The possible effect on interviewees. Will they be self-conscious, or as frank as they would be if it were not recorded?

- Listening/transcribing time afterwards, which may be prohibitive.
- The reliability of machine/recording. A range of high-quality digital recording devices should be used to collect data and record verbal and non-verbal communication during the interview/focus group. This ensures data are recorded on multiple devices should a technical failure occur and data is not lost.
- Take notes. Think what a disaster it would be if the recording devices did not record an interview. It is always best to take notes as well, partly as a fallback and partly as orientation when playing the recording back or seeking a particular point.

16.8 Conducting an Interview

The guidance discussed below applies equally to individual and focus group interviews. We present a few basic rules for when you conduct your interview.

- Thank your participants. Needless to say, this is the first thing you will need to do. Tell them how much you appreciate their time and effort in speaking to you, whoever they are, and make them feel important. This is especially advisable with people who may not ordinarily feel too important; frequent GP surgery attendees, for example. Second, set the scene. Explain the why/how/where, and for whom of the study.
- Give an idea of the question areas, and more or less in what order they will come up (if there is a logical progression).
- Explain the ground rules. You need to lay these out anyway in the information sheet they will have about the interview, but normally these will be that they can decline to answer any question, as they wish; they can terminate interview at any time; they can choose to remain anonymous if you decide to quote them in your dissertation or report, etc.
- Ask them if they have any questions before you start.

As you can see, the resemblance of an interview to a chat is already receding. Just to emphasis this let us compare an interview with a conversation in a pub.

> Imagine, once you had got the 'hellos' and 'how are yours?' out of the way, then saying "first, during this chat, I am going to ask you about how your day has been, with particular reference to that bloke in your office whom you wish to strangle; then we will briefly discuss the weather, and finally I am going to lament the fact that my wife does not understand me and my teenage son is off the rails. I have a car problem I would like to discuss also, but that will probably have to go on the agenda for next time".

No doubt by this stage your friend would exercise his or her right to terminate the meeting at any time: probably doing so immediately!

Having considered a pub conversation we now return to an interview. It is common to start with basic demographic questions, which will be easy for your participants to answer; this should then make them feel comfortable. It is also pretty standard to then continue with a general question and then funnel to the specific.

16.9 Types of Questions

Types of interview questions are open, closed, probing, and leading.

- Open-ended question.
  - opening stages in a line of questioning,
  - invites opinion, general knowledge,
  - can cover areas where the interviewer's own knowledge is lacking,
  - makes no presumption about the response.
- Closed-ended question.
  - elicits hard facts,
  - controls pace/direction of interview.
- Probing question.
  - extracts more depth,
  - maintains a line of enquiry.
- Leading question.
  - confirms an interviewee's answer,
  - helps an interviewee, by rephrasing answer,
  - brings a line of questioning to an end (summarising).

16.10 Analysing Interview Data

This is the part that worries people the most. How do you get meaningful data from the mass of interview transcripts or notes that you have? Although all of your interviewees tell you individual and unique stories, each one is valuable in its own right, your task will be to look for commonalities, themes, and contrasts. Most important, of course, is to consider again your aims and objectives and to see how the interview data inform these.

There are several methods for sorting the raw notes and transcripts into meaningful research data. An approach which we favour is called 'framework analysis' [6]. This goes through a logical sequence which is relatively easy to follow. It consists of five stages: familiarisation, identifying a thematic framework, indexing, charting, and mapping and interpretation.

These are outlined below.

- Familiarisation

  This is the immersion in the raw data undertaken by listening to recordings, reading transcripts, notes, and so on, so you can just get a feel for all the different ideas and themes that emerge.
- Identifying a thematic framework

  Here you try to identify the key issues, concepts, and themes within the data. The best way to do this is to go through each interview transcript with the aims and objectives of your study in mind. Issues raised by the participants (respondents) themselves, which may not be central to your study aims, might nevertheless inform the overall research. Write a word or phrase beside each issue/concept

elicited. By the end of this stage you will have a series of keywords and phrases, and possibly questions, next to your main interview notes or transcripts.

- Indexing/encoding

  This stage involves taking the comments, etc., grouping them into the themes identified, and coding them, possibly adding a few notes to the codes. By indexing the data like this, you are categorising the original notes you made, and you may find you see the data in a new light.

- Charting

  Here you take each index entry, lift it from its original place, and paste it into a new document, which relates to one of the specific themes that have been itemised. Thus, all the text relating to a particular theme, such as 'changes in lifestyle', will be in a single document. Before you do this, it is a good idea to use some form of participant identification, so that you can trace back each comment to an individual. This is a good idea because it might be that all the participants who shared something in common, either demographically or in terms of job status or whatever, had similar views. This will only come out where each comment is attributable. There are a number of good suggestions about how to undertake the 'charting' process. One such suggestion is an approach that involves numbering each line of each transcript and printing transcripts relating to different kinds of interviewees on different coloured paper, for example, green for nurses, blue for doctors, and yellow for patients, etc. You can then write 'each question to be analysed', or each research aim, on a large sheet of paper (preferably 'flip-chart' size) and you then have to go through each page of coded transcript and relate the indexed entries to the research questions or aims, so that the interviewees' comments that related to each point are all together. There may well be other data that do not fall naturally into the predefined categories suggested by the research aims or questions. This does not mean that they do not have any value. They provide extra, perhaps contextual, data and may well suggest further areas for exploration in subsequent studies.

- Mapping and interpretation

  Once you have 'charted' the data, you are able to really get something from what you have collected. The next, and final, stage uses the charts you have created to map the range and nature of phenomena, create typologies, and find associations between themes with a view to providing explanations for the findings. The process of mapping and interpretation, just as the 'charting' of the previous stage, is influenced by the original research objectives as well as by the themes that emerged from the data. During this stage a researcher reviews the notes, draws comparisons, and matches similarities in the perceptions and experiences of participants, and resolves to explain these.

By the end of this stage you will have completed analysing the data and will be ready to either take your study forward onto a quantitative (e.g., survey) stage or make conclusions from your findings and write up your results, about which there is more below. An analysis of an interview using this method is described below. See also Parkinson et al., whose article describes a worked-through analysis [7].

A computer package, such as N-Vivo could be used to analyse the data. It is a systematic process that removes subjectivity can distance the researcher from the analysis process.

16.11 An Example of an Interview with Analysis

Below is an interview that has been transcribed. Pauses, silences, giggles, etc. have been removed, but in some instances verbatim transcription is essential as the pauses and hesitations enhance the data collected.

The five-stage process discussed above has been used as a guide for analysis of the following interview.

- Familiarisation: once you have read, listened, and transcribed the interview you start to analyse the data.
- Identifying a thematic framework: we suggest that you attempt to do this yourself first, but you need to consider the overall aims and specific questions while reading.
- Aim: to ascertain practitioners' perceptions of their role as a student mentor.

The specific questions were as follows.

- What do practitioners perceive as challenges to the students' learning in the practice setting?
- Do practitioners perceive there are any benefits of undertaking this role?
- How do practitioners perceive the students learn in the practice setting?
  - Indexing,
  - Charting,
  - Mapping and interpretation.

Here is an interview that you could practice on. Try this out yourself first. We include some of our framework analysis later for you to look at and compare. Depending on who is interpreting the data, some different themes can emerge. This does not mean that you are wrong, but rather that your analysis is different.

16.11.1 Transcription of an Interview with a Practitioner About Her Role as a Mentor

It is usual practice to use italics for verbatim transcriptions of collected qualitative data (responses of participants). We present 18 questions posed and the replies to each question.

- Tell me about your experience of supporting learning in the clinical setting?

Well I have been a mentor for 3 or 4 years now. I started by coming to the university for an induction procedure. I've gradually felt my way through it. I have a lot of experience so throughout the years I have had a lot to do with students training though not always on an official level. I've been a senior practitioner so I've always had some input into training but it is now more formalized with appraisal. I've just gradually got into it and do reports on the student's progress and I've done a few sessions where I have done some practical talks about basic views and the problems that they might encounter.

- What do you understand about the way adults learn?

I think they learn quite a lot from their co-workers or other people, in actual fact I think the students learn from other students who are further on in the process but I think learn from observation really and from just being there whilst it's going on.

- Do you think that the learning needs of individuals can be met in the practice setting?

Yes, yes I think so, it's just a normal thing that all students are different so that you match your teaching to what they are capable of taking in at the time. I think you have to play it by ear as some are capable of taking in things quicker than others. Some people are more confident and are prepared to go ahead and because they (the students) work with all of us (practitioners) and they get a bit of something different from every person that they work with. And here they get a good mixture of people to work with, so I think we can support individual needs. As mentors we often discuss students together so we can highlight strengths and weakness and put some support mechanisms in place to help students. The site coordinator is very good and helps us with ideas of how we can help the students.

- How do you use reflective practice to support the students' learning?

Practice can be hectic, so quite often we can't take 'time out' to reflect immediately on what has happened. It's often when we are relaxing, having a coffee when we actually have chance to discuss issues, but its more informal that way and we probably don't really think of it as reflection. But we do ask questions as we are going about practice. For example, what about that patient who came from such and such a clinic? What did you make of the image? Did you think it needed a lateral? Why did you think that? So, you do it amongst yourselves. I've never actually sent anyone off to reflect.

- Why do you think reflection is a key aspect of professional practice?

Practitioners are not always good at documenting events but when I think about it I've always reflected on my work, but not always by writing it down. We discuss things about practice, such as CT examinations with (names a colleague), you might not put a label on it as reflection, but it can be seen as reflection, so I do reflect all the time.

- What is important about clinical education?

It's basically a practical occupation, you can theorise all you like, but until you've actually been there and done it I don't think you are ever quite as capable as you think you are until you have encountered the challenges of practice. Every patient is different and they all have different needs, so they (the students) might be taught about ageing but they don't really understand about the implications of that for taking an X-ray until they meet older people.

- What about professional socialisation? How does the practice experience affect that?

I'm sure that is learning from example, the students need to find a good role model. I can say they get some good examples to follow and perhaps see some examples of things that aren't as they should be. When I was a student, I wanted to be a professional and so tried to exemplify that and would take my lead from whoever I thought was worthy of that niche. But I think it takes a while to learn to work and fit in with a department, but by the time they qualify, they know what is acceptable and what is unacceptable, we tell them if we don't think they are acting properly.

- What do you think are the key goals of clinical education?

Well basically it's to have the confidence to make your own decisions and that comes from having a lot of experience of undertaking examinations. So, they have a good base so that they know they can cope with any event that they encounter, well they can't do everything, but have the practical and thinking skills to work out what they need to do. It's about building up their knowledge so they have the confidence to make decisions about referrals and how to examine the patient. You have to lead them through this process so they have the confidence in the end to do this on their own.

- How do you help to develop their clinical reasoning?

I think we have to start with the basics and ask them to think about what they are going to do before they do it. They then go and do it and afterwards think about what they have done, what they did correctly and what they did wrong. Why did they do what they did? What might they do next? Unfortunately, it's time pressure really that means we don't always have the opportunity to discuss each case in depth as would be ideal, but over time we do iron out any issues over a period of time.

- What do you think your role is as a mentor in the management of clinical education?

I think it's about giving the students the chance to try things out but ensuring that they are doing things right the first time if possible. So sometimes it can be about saying that's not an appropriate examination for you to do at this stage of your training or saying it's about time you were doing these types of examination now. It's about pushing them sometimes, but also about holding them back, maybe that's not the right term, ensuring they aren't attempting things they aren't currently capable of.

- How do you maximise their learning opportunities?

We do try and let them get on with what they are capable of. We do rotate them around the department so they get a chance to see and do lots of different things. Some do make more of their opportunities than others. In some ways I think all we can do is give them the chance to learn, to some extent they have to take some responsibility. Some of them will stand back and need to be encouraged all the time; others will ask 'can I do that?'

- What about the ethical issues relating to clinical education?

Well sometimes it can be difficult to manage the learning and the patient, the students often want to ask questions, which I might not want to answer in front of the patient. They some-

times might say or think that all old people are demented, but that's not right. I try to get them to think about what is right and wrong.

- How do you get the students to link theory with practice?

I get them to look at anatomy and pathology for example. But it can be difficult to get them to realise that we apply theory all the time, but we don't talk in that way every day. But when I take an X-ray I do think about radiation protection all the time, but I might not say that I'm doing this or that as it helps to reduce the dose.

- How do you use assessment to promote learning?

I think it's good, as it makes them learn things, but it does scare them. They have done things lots of times, but when you suddenly say, well this time I'm going to assess you, they can lose the plot. When I was a student people were watching me all the time and they could see I was progressing, I think it can be harder now as there is so much technology. I think it's good that they can make mistakes and can learn from them. I think assessment is good it helps them realize what they need to know to do the job properly.

- What about feedback?

I do it informally all the time; I fill in progress forms for the students, but (names colleague) she seems to do the more formal feedback, but I think they could do with more. But sometimes it's hard, you don't want to make them feel bad, but sometimes you have to be cruel to be kind as they can be hopeless sometimes.

- How important is CPD in learning and teaching to you?

Well it's important as you want the students to have the best experience. So, it's important that you are actually doing best practice yourself. I've always done it but haven't always recorded it but need to now. I need to keep up to date with my practice, so that I can help and advise the students about what is now best practice, so in that way I'm using CPD to inform my teaching. But I haven't really done much about teaching as such; I've been on the training courses and update days, but not much else.

- Why did you decide to become a mentor?

Well I think I have a responsibility to do it, but it's a big responsibility. I am going to work with these people in the future. the future, so I want to make sure they are moulded into the right sort of person. I think you volunteer for this role, which is better than 'pressed men'. But that doesn't mean you will be good at it.

- Is there anything else you would like to say?

No I don't think so.

Thank you for taking the time to participate in the interview.

The data (responses) then had to be categorised. An example of this can be found in Fig. 16.1.

Question and answer. Issues and concepts are highlighted	Words and phrases relating to each issue/concepts
Tell me about your experience of supporting learning in the clinical setting?	
1.Well I have been a mentor for 3 or 4 years now. I started	
2.by coming to the University for an induction procedure.	
3.I've gradually felt my way though it. I have a lot of	3.Gradual
4.experience so throughout the years I have had a lot to do	4.Experience
5.with students training though not always on an official	5.Unofficial
6.level. I've been a senior and superintendent radiographer	
7.so I've always had some input into training but it is now	
8.more formalised with appraisal. I've just gradually got into	8.Gradual process
9.it and do reports on the student's progress and I've done	
10.a few sessions where I have done some practical talks	10.Teaching
11.about basic views and the problems that they might	
12.encounter.	
What do you understand about the way adults learn?	
13.I think they learn quite a lot from their co-workers or	13.Co-workers
14.other people, in actual fact I think the students learn from	14.Each other
15.other students who are further on in the process but I	15.Observation
16.think learn from observation really and from just being	
17.there whilst it's going on.	17.Being there
Do you think that the learning needs of individuals can be met in the practice setting?	
18.Yes, yes I think so, it's just a normal thing that all	
19.students are different so that you match your teaching to	19.Different/Individualised
20.what they are capable of taking in at the time. I think you	
21.have to play it by ear as some are capable of taking in	21.Pace of learning
22.things quicker than others. Some people are more	
23.confident and are prepared to go ahead and because	
24.they (the students) work with all of us (radiographers)	
25.and they get a bit of something different from every	
26.person that they work with. And here they get a good	
27.mixture of people to work with, so I think we can support	
28.individual needs. As mentors we often discuss students	28. Support individuals
29.together so we can highlight strengths and weakness	
30.and put some support mechanisms in place to help	
31.students. The site co-ordinator is very good and helps	
32.us with ideas of how we can help the students.	
How do you use reflective practice to support the students learning?	
33.Practice can be hectic, so quite often we can't take "time	33.Hectic
34.out" to reflect immediately on what has happened. Its	
35.often when we are relaxing, having a coffee when we	35.Informal discussion
36.actually have chance to discuss issues, but its more	
37.informal that way and we probably don't really think of it	
38.as reflection. But we do ask questions as we are going	38.Question practice
39.about practice, what about that patient who came from	
40.such and such a clinic, what did you make of the image,	
41.did you think it needed a lateral, why did you think that,	
42.for example. So you do it amongst yourselves, I've never	
43.actually sent anyone off to reflect.	
Do you think reflection is an important aspect of practice?	
44.I do, but I think I'm not very good at documenting all	44.Not formally recorded
45.these things, but when I think about it I've always	
46.reflected on my work. We discuss things about practice,	
47.such as barium studies with (names a colleague) you	
48.might not put a label on it as reflection, but I do reflect all	48.Active reflection
49.the time.	

Fig. 16.1 (**a**) An example of how interview data are categorised.

What is important about clinical education? 50.Its basically a practical occupation, you can theorise all 51.you like, but until you've actually been there and done it I 52.don't think you are ever quite as capable as you think 53.you are until you have encountered the challenges of 54.practice. Every patient is different and they all have 55.different needs, so they (the students) might be taught 56.about ageing but they don't really understand about the 57.implications of that for taking an x-ray until they meet 58.older people.	51.Practical 52.Challenges 54.Individual needs of patients
What about professional socialisation? How does the practice experience affect that? 59.I'm sure that is learning from example, the students need 60.to find a good role model. I can say they get some good 61.examples to follow and perhaps see some examples of 62.things that aren't as they should be. When I was a 63.student I wanted to be a professional and so tried to 64.exemplify that and would take my lead from who ever I 65.though was worthy of that niche. But I think it takes a 66.while to learn to work and fit in with a department, but by 67.the time they qualify, they know what is acceptable and 68.what is unacceptable, we tell them if we don't think they 69.are acting properly.	60.Role model 61.Good examples 63.Professional 64.Exemplify 66.Learning to work 67.Behaviours
What do you think are the key goals of clinical education? 70.Well basically it's to have the confidence to make your 71.own decisions and that comes from having a lot of 72.experience of undertaking examinations. So they have a 73.good base so that they know they can cope with any 74.event that they encounter, well they can't do everything, 75.but have the practical and thinking skills to work out what 76.they need to do. It's about building up their knowledge so 77.they have the confidence to make decisions about 78.referrals and how to examine the patient. You have to 79.lead them through this process so they have the 80.confidence in the end to do this on their own.	70.Autonomy 73.Capability 74.Competence 77.Confidence
How do you help to develop their clinical reasoning? 81.I think we have to start with the basics and ask them to 82.think about what they are going to do before they do it. 83.They then go and do it and afterwards think about what 84.they have done, what they did correctly and what they 85.did wrong. Why did they do what they did? What might 86.they do next? Unfortunately its time pressure really that 87.means we don't always have the opportunity to discuss 88.each case in depth as would be ideal, but over time we 89.do iron out any issues over a period of time.	81.Thinking 83.Reflection 86.Time barriers
What do you think your role is as a mentor in the management of clinical education? 90.I think it's about giving the students the chance to try 91.things out but ensuring that they are doing things right 92.the first time of possible. So sometimes it can be about 93.saying that's not an appropriate examination for you to 94.do at this stage of your training, or saying it's about time 95.you were doing these types of examination now. It's 96.about pushing them sometimes, but also about holding 97.them back, maybe that's not the right term, ensuring they 98.aren't attempting things they aren't currently capable of.	90. Providing opportunities 92.Encouraging them
How do you maximise their learning opportunities? 99.We do try and let them get on with what they are capable 100.of. We do rotate them around the department so they 101.get a chance to see and do lots of different things. 102.Some do make more of their opportunities than others, 103.in some ways I think all we can do is give them the 104.change to learn, to some extent they have to take some 105.responsibility. Some of them will stand back and need 106.to be encouraged all the time; other will ask "can I do that?"	99.Encourage 100.Provide opportunities 104.Learner responsibility

Fig. 16.1 (continued)

What about the ethical issues relating to clinical education? 107.Well sometimes it can be difficult to manage the 108.learning and the patient, the students often want to ask 109.questions, which I might not want to answer in front of 110.the patient. They sometimes might say or think that all 111.old people are demented, but that's not right. I try to get 112.them to think about what is right and wrong.	107.Difficult to manage 109.Appropriate responses 112.Reduce prejudice Right and wrong
How do you get the students to link theory with practice? 113.I get them to look at anatomy and pathology for 114.example. But it can be difficult to get them to realise 115.that we apply theory all the time, but we don't talk in 116.that way every day. But when I take an x-ray I do think 117.about radiation protection all the time, but I might not 118.say that I'm doing this or that as it helps to reduce the dose.	115.Constant use of theory 117.Not explicit
How do you use assessment to promote learning? 119.I think its good, as its makes them learn things, but it 120.does scare them. They have done things lots of times, 121.but when you suddenly say, well this time I'm going to 122.assess you, they can loose the plot. When I was a 123.student people were watching me all the time and they 124.could see I was progressing, I think it can be harder 125.now as there is so much technology. I think it's good 126.that they can make mistakes, and can learn from them. 127.I think assessment is good it helps them realise what 128.they need to know to do the job properly.	119.Active tool for learning 120.Intimidating 125.Reflect on errors
What about feedback? 129.I do it informally all the time; I fill in progress forms for 130.the students, but (names colleague) she seems to do 131.the more formal feedback, but I think they could do with 132.more. But sometimes it's hard, you don't want to make 133.them feel bad, but sometimes you have to be cruel to 134.be kind as they can be hopeless sometimes.	129.Informally/Continuous 131.Written – formal 132.Insufficient feedback 133.Consideration
How important is CPD in leaning and teaching to you? 135.Well it's important as you want the students to have the 136.best experience. So it's important to you yourself are 137.actually doing best practice yourself. I've always done it 138.but haven't always recorded it, but need to now. I need 139.to keep up to date with my practice, so that I can help 140.and advise the students about what is now best 141.practice, so in that way I'm using CPD to inform my 142.teaching. But I haven't really done much about teaching 143.as such; I've been on the training courses and up date 144.days, but not much else.	135.Enhance learning experience 138.Best practice 139.Responsibility 142.Not teaching
Why did you decide to become a mentor? 145.Well I think I have a responsibility to do it, but it's a big 146.responsibility. I'm going to work with these people in the 147.future, so I want to make sure they are moulded into the 148.right sort of person. I think you to volunteer for this role, 149.which is better than "pressed men". But that doesn't 150.mean you will be good at it.	145.Responsibility 146.Future practitioners 148.Volunteers 149.Quality of mentor
Is there anything else you would like to say? 151.No I don't think so 152.Thank you for taking the time to participate in the interview	

Fig. 16.1 (continued)

Thematic framework and indexing/encoding We read the interviews and immersed ourselves in the data. We looked at the issues and concepts that emerged from the indexing/encoding. And from these the initial themes were derived. You need to remember that these themes and encoding arise from one interview; more may emerge from other interviews.	
Themes	Charting
Mentor's role	
Reflective practice	
Clinical education	
Professionalism	
Assessment	
The next stage is **charting**, where the indexed/encoded items are related to the themes.	
Themes	Charting
Mentor's role	3.Gradual 4.Experience 5.Unofficial 8.Gradual process 10.Teaching 90.Providing opportunities 92.Encouraging them 99.Encourage 100.Provide opportunities 115.Constant use of theory 146.Future practitioners 148.Volunteers 149.Quality of mentor
Reflective practice	35.Informal discussion 38.Question practice 44.Not formally recorded 48.Active reflection 81.Thinking 83.Reflection 86.Time barriers 104.Learner responsibility 107.Difficult to manage
Clinical education	13.Co-workers 14.Each other 16.Observation 17.Being there 19.Different/Individualised 21.Pace of learning 28.Support individuals 33.Hectic 51.Practical 52.Challenges 54.Individual needs of patients 77.Confidence 117.Not explicit

Fig. 16.1 (**b**) Thematic framework and indexing/encoding

Professionalism	60.Role model 60.Good examples 64.Exemplify 66.Learning to work 67.Behaviours 70.Autonomy 109.Appropriate responses 112.Reduce prejudice/Right and wrong 135.Enhance learning experience 138.Best practice 138.Responsibility 142.Not teaching 145.Responsibility
Assessment	73.Capability 74.Competence 119.Active tool for learning 120.Intimidating 125.Reflect on errors 129.Informally/Continuous 131.Written – formal 132.Insufficient feedback 133.Consideration

Fig. 16.1 (continued)

16.11.2 Indexing and Thematic Framework

Included here is a copy of the interview, which demonstrates the indexing process, and from this we developed the themes and codes. We started by reading through the interview and thinking about the questions. We then highlighted the issues and concepts (Fig. 16.1a). We then wrote a word or phrase against each issue and concept. Look at all the words and phrases, including those highlighted in Fig. 16.1b (these will be referred to later in the analysis). How did these match to your indexing? The next stage is when the original questions are reviewed, and the indexed/encoded entries are related to the research questions.

16.11.3 Mapping and Interpretation

It is difficult with just the one interview in this example to undertake a full analysis and interpret the findings fully. Nevertheless, we can see that some links and perceptions have emerged.

With regard to the question: 'What do radiographers perceive as challenges to students learning in the practice setting?' The mentors indicated that assessment can be intimidating, but they also recognised that they may have not provided sufficient feedback. They recognised it is an active tool for learning. Observation was seen as an important aspect of clinical education. You need to look at all of the indexing/encoding and charting for all of your interviews and map responses to the relevant questions. You also need to review the emerging themes. You need to be aware that it is not always black and white; there may be overlap of categories. The numbering helps you go back to the original location of the data to seek clarification if required.

16.11.4 Key Points

- Interviewing is probably the most common form of qualitative data gathering. Remember that an interview requires good preparation and considerable skill in conducting.
- The qualitative approach looks at the 'how' or 'why' rather than the 'how many' and has an emphasis on understanding.
- There are advantages and disadvantages to choosing interviews and/or focus groups as data collection tools.
- Interviews are used to explore deeper meanings of, for example, experiences, ideas, or attitudes and are usually conducted on a one-to-one basis.
- Focus groups usually involve the participation of several persons in a discussion or exploration of a specific set of issues.
- Many methods can be combined to provide a rich picture of a particular phenomenon.

16.12 Questionnaires

A questionnaire can be a useful tool to extract data from a wider population. The field of questionnaire design is vast. The aim here is to give some basic guidance on the development of a questionnaire that can be used in a qualitative research design. This approach seeks to explore experiences and feelings as discussed earlier. The design of a questionnaire thus needs to be such as to capture an individual's perceptions. It is important to say that the structure of a questionnaire differs depending on whether a study is a qualitative or quantitative design. Under certain circumstances both methods can be used to analyse the same questionnaire. Going back to the example of the impact of a touch-screen health information system, questions could be asked on frequency of use, reasons for use, other information sources consulted, etc. Here one is seeking quantitative data: how many people use the system, how often, and for what purpose, as categorised in the survey. Their responses need only be in the form of a tick box. However, if one is interested in individual perceptions, ideas, and attitudes, the questions are framed differently, and the response options might include allowing free text.

Pay attention to questionnaire structure and style. Remember that respondents are subjects and not objects of research. Indeed, particularly in the case of research with vulnerable groups, the term 'participant' has come to be used to help foster inclusiveness [8]. Issues such as the informed consent of participants and their right to withdraw have to be considered.

There is much debate about whether open or closed-ended questions should be used for data collection. A problem with closed-ended questions is that they have the potential to create false opinions: only limited options are open to a respondent. They are often perceived as easier to answer. An open-ended question gives respondents the opportunity to express their feelings and perceptions in their own words. But as a general principle the larger the sample size the more structured, closed, and numerical the questionnaire is likely to be. Questionnaires, particularly those that

are self-administered, are used for convenience and speed rather than in-depth analysis of individual responses. 'Depth' is achieved by sometimes very sophisticated statistical analysis beyond the scope of this chapter.

The method can be useful when exploring phenomena in a population that is distributed across a wide geographical area. This means that large population sizes are available and therefore expand the potential sample size to be included in a study.

An example of an open question is presented below.

- How do you feel about the care you received during your examination/treatment on your recent visit to the imaging/radiotherapy department?

Expressed in a closed-ended format it could be the following.

- How do you rate the care you received prior to your examination/treatment on your recent visit to the imaging/radiotherapy department? Please circle one number.
  Excellent = 1
  Very good = 2
  Good = 3
  Fair = 4
  Poor = 5

16.12.1 Semi-Structured and Open-Ended Questionnaires

As we are looking primarily at language data the focus of this section is on the use of semi-structured and open-ended questionnaires. Open questions do generate more detailed responses from participants; as a consequence a great deal more effort is required to encode their responses. These questions take longer to complete (so more effort on the part of the respondents too) and this should be taken into consideration when compiling a questionnaire. On the other hand, it could be considered an economical approach in terms of the time spent distributing and collecting the data, and has the potential to capture honest personal comments from respondents.

The wording and design of the questions do need careful consideration; this aspect cannot be rushed and is time-consuming. Otherwise you could potentially write unsuitable questions. Examples are presented below.

- Double-barrelled questions: these are questions that essentially ask two questions at once, therefore a respondent could answer either part of the question. An example might be: 'Do you think the new appointment system is easy to use and what effect on waiting times do you think this will have?' It would be better to ask these questions separately as the answers given may elicit a positive answer to the first, but a negative answer to the second.

- Complex questions: such as 'Would you prefer to undertake a short programme of study, e.g. 3 or 4 sessions, which does not carry any award, and is delivered on a Wednesday evening each week, or a longer award bearing programme that is designed to be undertaken during the day rather than the evening?'
- Irritating questions or instructions: such as 'Have you ever attended a personal tutorial during your undergraduate programme', or 'Have you attended any continued professional development activities during your career?'
- Ambiguous questions: this is where the words could be interpreted in different ways, such as: 'Do you regularly undertake self-managed study while you are on placement?' What do we mean by 'regularly'? Is this once a day, once a week, once during your placement?
- Biased and leading questions are those that are worded in such a way as to suggest to a respondent that there is only one acceptable answer. An example might be: 'Do you prefer to plan your studies well in advance or to leave it until the week prior to submission?' If this was asked by your academic tutor, I think we can guess which response you might make!

An open-ended questionnaire can be administered face-to-face or self-administered. Clearly if undertaken face-to-face, this can be time-consuming and may limit potential participants. However, it does give a researcher an opportunity to encourage participants to expand on their responses. See Sect. 16.11.1 as an example is provided involving exploration of imaging practitioners' perceptions of their role as mentors. Initially interviews were carried out to get a feel of how imaging practitioners perceived themselves in that role. From this data an attitudinal questionnaire was developed. Open-ended questions regarding aspects of the mentor role that participants found most challenging were asked as well. Questionnaires were used so that perceptions of a larger number of imaging practitioners could be captured.

Radiotherapy practitioners have a role as health educators in preparing their patients for treatment. A questionnaire could be used to explore the perceptions of practitioners and patients of that role using open-ended questions.

16.12.2 Reliability and Validity

When using or creating a questionnaire, reliability and validity of the tool should be considered.

– Reliability of a tool is the ability to reproduce the results.
– Validity is more concerned with the accuracy of the test used and asks whether it measures what it is intended to measure.

For example, a questionnaire may be considered to be more reliable than interviews, as a respondent is anonymous and therefore may give more honest answers

to the questions. When considering the validity and reliability of questionnaires in particular you must think about the sample size; a small sample may skew the results or be unrepresentative.

16.12.3 Key Points

- Questionnaires can be a useful way to collect large amounts of data from a large number of participants.
- Questionnaires are versatile and can be used in both qualitative and quantitative research designs.
- Questionnaire structure can incorporate both open- and closed-ended questions or a combination of both.
- When using a questionnaire to collect data, pay attention to the criteria that may affect the validity and reliability of a study.
- Questionnaires are the most popular data collection tool used for cross-sectional surveys, for student research projects.

16.13 Observation

Observation is often a preferred method or key component of case studies or action research. It is often used to supplement other methods, such as interviews, and also when the phenomenon we wish to study is not well known to us. Here, observation is inductive. We begin with specific or general observations to detect patterns and regularities. We then formulate some tentative hypotheses, and finally end up developing some general conclusions or theories.

You may be surprised to know that the apparently simple act of looking at something can be done in a great many ways. The two extreme of observational methods are: (1) a form of participant observation where a researcher becomes one of the people he or she is observing, and (2) to being as invisible as possible as a 'non-participant' observer. The former research is favoured by sociologists and anthropologists. A 'classic' example is that of observers of a religious sect whose leader had prophesied that the world would end that year. Researchers investigating the cult pretended to be followers for a considerable time. Unlike many such 'immersion' studies, however, they deliberately interfered with what they were observing by suggesting to cult members that their leader might be a fraud. Observational research in which observers try to immerse themselves in a culture or environment is known as 'ethnographic'. It is not however the only type of research for which observation may be used, as some of the examples below suggest.

By contrast, the extreme non-participant observation is where an observer is as unobtrusive as possible. A seminal example is that of a sociological study of infant schools, in which the researcher ignored children who came to him with their work, etc. to such an extent that eventually they learned to ignore him as much as he was apparently ignoring them [9]. Apart from being physically present and as discreet as

possible, non-participant observation can be undertaken without the physical presence of a researcher (and here we generally part company with ethnography).

For example, researchers can use the following.

- Recorded behaviour, e.g., use CCTV to record footage of people's use of a touchscreen health information system as one of a number of data gathering techniques to look at system usage.
- Human trace behaviour, which involves examining the things people leave behind as they go about their daily routines or undertake the activities researchers are interested in. This is actually a very rare technique, the classic example of it being a study by Rathje [10], who examined household waste. However, be sure to consider all ethical implications of undertaking observational studies involving human participants.
- Computer log analysis, which entails observing computer usage through the transaction logs created by the system. As users in our health kiosk study were required to 'log-in' with these details, researchers were able to compare time 'online', pages accessed and navigational behaviour, and relate it to age and gender.

16.13.1 Advantages and Disadvantages of Using Observation

The advantages of using observational techniques include the following.

- Capturing non-verbal behaviour, which may tell more than may be elicited from interviews.
- Subjects do not have to do anything, except 'act naturally'.
- Researchers do not have to ask questions and no interpretation is required by a subject.
- Good for preparing the ground for other fieldwork, by familiarisation with a situation, process, or environment.
- Good for triangulation, i.e., for combining with other methods.
- Phenomena are studied as they occur and there is no need for the participants to rely on memory.
- A more intimate or informal relationship is permitted, where participant observation is involved.

The disadvantages of observation include the following.

- Observation could change what is being observed. This is as true of people as it is of subatomic particles and has been given a name: 'the Hawthorne effect'. This comes from a set of studies conducted in the 1920s and 1930s at the Western Electric Hawthorne Works in Chicago, by Harvard Business School Professor Elton Mayo, who was interested in productivity and work conditions at the works. Mayo found that the very act of being observed made the company

employees work harder [11]. The experiments led to the term 'Hawthorne effect' being applied to research generally where researcher observation leads to a positive change in the behaviour of those observed. Although the effects have been questioned, it is a concept worth bearing in mind.
- An observer may pose an imagined threat. This may be particularly true in a work context, where people may be worried about their performance.
- It is hard to track many activities at once.
- Interpretation of observations is difficult.
- Observation is time-consuming.
- A researcher can feel a bit awkward. The term 'wallflower' springs to mind. This may seem a trivial point, but the work will be your study thus you will want to do your best. You may feel that you just would not be able to do this by feeling self-conscious or awkward.

16.13.2 Collecting Data

Having weighed up the pros and cons and decided to go ahead with some form of observational process, you need to decide how to collect and record data. There are several techniques, which are often relevant to both participant and non-participant observation. The main ones, in addition to those not involving you as a direct observer, are the following.

- Protocol analysis or 'think aloud' method. Here participants describe what they are doing as you observe. This is common in looking at the usability or accessibility of information technology systems.
- Take field notes; decide first on what you want to record. You can probably make notes that equate to approximately 500 words per hour. You may wish to record specific events, or to see what is going on generally. These can be described as:
  - event sampling: an example might be to record communication activities between hospital staff,
  - time sampling: this is where you record what is happening at given times during an observational session. This was common in education at one time, where interactions in terms of frequency of 'teacher talk', 'pupil talk', 'silence', and 'confusion' among others were analysed.

16.13.3 Recording Tips

Much observational recording is undertaken, by necessity, at the time of an observation. This is particularly true where specific instances, events, or times are being recorded. Even where you just want to get an overall picture of an environment, without recording chronologically, you may like to follow these recommendations.

- Record during or straight after the event otherwise you will forget the details.
- Do not start a new observation session until you have recorded the last one.

- Recording can take as long as actual observation if you are meticulous.
- Include everything important, or which might be important.

16.13.4 Analysing the Data

As with interviewing, people get really worried about analysing observational data. 'What does it all mean?' one might ask. As done with interviews (see Sect. 16.8) consider the aims and objectives of your study, and how observation fits in with those. We present some tips.

- Look for similarities and differences and try to explain them. You may need to consult interview data if you have any, in this, or simply ask the people observed to explain certain actions.
- Watch for re-occurrences. Why do the same things keep happening?
- Formulate 'rules' based on repeated occurrences, but look for exceptions.
- Try to explain the exceptions you find. Do they mean your rule is wrong? If so, change it; or is the exceptional case really different in some way? If so, in what way, and how does it inform your study?
- To systematise observations, it may be helpful to devise categories. If you did not do this for the actual observation, doing it afterwards is still valuable, in a way even more, as you will make the categories fit what you have observed not the other way around. With luck patterns will emerge.
- If the observation is a preliminary 'staking out' of the field? Consider what questions it raises that could be incorporated in any interview or survey you may be considering.
- There is the potential to observe practitioners or students undertaking a specific radiographic examination or intervention to explore the variation in approaches to techniques or tasks. It might be useful to observe how practitioners use different approaches to a task, which may well produce a similar outcome. An example could be a task analysis of how a lumbar spine examination is undertaken, handwashing techniques, moving and handling of radiographic or therapeutic equipment, observing the tools applied when evaluating images on DR or PACs system. Task analysis is a range of techniques used by operators to describe and evaluate interaction between humans and machines and is a way of investigating participants' behaviour in a specific context.
- Within a practice setting you could observe patients finding their way around the imaging and radiotherapy departments or their interaction with staff at the reception area. Again, remember the ethical implications of studies involving human participants.

16.13.5 Key Points

- Observational methods involve many approaches on a continuum from participant to non-participant, and the methods of data gathering, including 'event' and 'time' sampling.

- Observational research methods form a key component of case studies and action research.
- Observational research is often employed to supplement other research methods, e.g., interviews.
- Do not forget the 'Hawthorne effect' during observational studies.
- When analysing data, pay careful attention to the explanations drawn from the similarities and differences observed. In addition, participants can be asked to explain their actions.

16.14 Case Study

A case study is where a single instance is studied in-depth and can be considered as an approach rather than a method; several methods are employed in a case study. Examples may include where a patient, a group of specialist practitioners, a clique, a class, or an imaging department is studied as a unique case. A case study is used to explore and reflect what is happening in a unique situation and allow these specific situations to be explored in greater depth, which may not be captured using other data collection tools. One of the strengths of these studies is that they are drawn into the context of the case itself. Case studies aim to describe 'what it is like' to be in a particular situation and to give a rich description of the reality. Observation (see Sects. 16.13–16.13.5) is a frequently used tool in a case study. When will a case study be useful?

- When a randomised approach is not appropriate.
- When it is not possible to study a particular population as a group.
- When you need to evaluate intervention outcomes over a period of time.
- When pilot information is required.

Case studies are useful as a theory generating tools. When conducting a case study a number of factors need to be considered, such as negotiating access to people and how the data are to be collected. Data collection tools that could be employed in case studies include interviews using open or semi-structured questions; observations or narrative accounts. In addition, documents or diaries could be used to explore the uniqueness of a specific situation.

16.14.1 Key Points

- A case study can be considered an approach rather than a method.
- A case can be an individual or group of patients, practitioners, a class of students, or an imaging department.
- Case studies are used to explore a 'happening' in a unique situation.
- Observations are frequently used as tools to collect data in case studies.

16.15 Action Research

The purpose of action research is to improve understanding of practices, in a specific context, with a view to making changes for the better. It is a reflective activity. Action research is designed to bridge the gap between research and clinical practice. It should bring about improvement, change, and development to enable practitioners to have a better understanding of their practice.

There are three basic phases to action research.

1. Look: build up a picture of a situation and the context in which it occurs, thinking about what the practitioners, as well as the patients, are doing.
2. Think: this process requires you to interpret a situation and explain what is happening, reflecting on what the participants have been doing, and look for any deficiencies or issues.
3. Act: whereby actions for change are identified and put into practice.

Action research is about looking at a local issue or problem and exploring that and making changes as well as expanding knowledge. In addition, this research approach provides an opportunity for personal and professional development.

Methodologically it has a distinctive set of requirements. It should be the following.

- Collaborative: participants contribute to the overall project.
- Action oriented or participatory: intervention and change are a part of the process.
- Contextualised: it relates to a specific place, situation, or circumstance.
- Reflective: through a process of planning, action, evaluation and critique.

Action research relies on the following:

- communication of all group members,
- time to reflect on the process and outcomes of the project,
- verification of project: can it be replicated or reproduced?

It can be argued that the process of action research enhances a change process, a key agenda in the ever changing context of healthcare; particularly in the fields of medical imaging and radiotherapy. It is an approach that enables a researcher actively to participate in development in their specific area or field. It is suited to small-scale projects. Action research generates change through reflection, communication, cooperation and collaboration, and empowerment among participants.

An action research process consists of several stages, namely

- questioning existing practices and coming up with an idea,
- collaborative decision-making and planning,

- action, implementing the changes with ongoing evaluation and monitoring,
- critical reflection on the intervention and the process,
- re-evaluation of the original plan based on reflections, implementation of changes and continued monitoring,
- reflection on the knowledge generated and the reshaped practice.

The actual methodological approach employed depends upon the research question posed. As discussed triangulation (see Sect. 16.7) is often employed to increase validity and to identify convergence and obverse patterns. Whatever data collection process is employed the validity and reliability of the chosen method should be considered. There are some potential problems associated with action research. These are the level of skill of a research facilitator and the culture of an organisation in which the research being done.

A survey might elicit information from a patient group about waiting and changing facilities, thus the context in which this data are collected has to be taken into consideration. The data collected must be reflected on and discussed, a plan of action for any changes discussed and implemented. Follow-up focus groups, for example, could be employed to determine whether the interventions did affect change in clinical practice.

16.15.1 Key Points

- Action research is a reflective activity in which researchers aim to improve understanding of their practice within a given context.
- The three basic phases involved in action research are: look, think, and act.
- Action research should be collaborative, action orientated, contextualised, and reflective.
- Triangulation is often employed to increase validity and reliability of findings.

16.16 Content Analysis

Content analysis is used for studying the content of communication and documentary evidence. It is a careful, detailed, and systematic examination of large amounts of data [2]. Content analysis is used to determine the presence of certain words or concepts within literature. The literature used can be from a variety of resources, books, journals, research articles, professional journals, departmental or hospital protocols, newspapers, audio of video media material, etc. Any information, whether it is primary research or information, which is in the public domain, can be used.

Content analysis can be qualitative or quantitative and usually involves inductive reasoning. This methodology can uncover underlying meanings within a text and enables the content to be quantified by the use of a set reproducible method of data extraction. Researchers quantify and analyse the presence, meanings, and relationships of words and concepts within the literature. Categories and themes emerge

from the data. These identify the focus of the research and the extracted data are assigned to these themes and categories as analysis of the data occurs.

16.16.1 Content Analysis Is Generally Categorised into Two Types

Two types are discussed: conceptual and relational analysis, respectively.

Conceptual analysis (a.k.a. thematic analysis) involves quantifying the existence and frequency of words of phrases in the literature being studied. The focus is examining the occurrence of selected terms within a text or texts. The terms may be implicit or explicit. While explicit terms are easy to identify, coding for implicit terms and deciding their level of implication are much more complex, judgments are based on a more subjective system. The level of implication refers to how you have defined the words or phrases and must be kept constant throughout your analysis; using the example cited earlier about what is meant by 'regular', you need to define whether it means, for example, daily or weekly. To limit the subjectivity when coding such implicit terms use can be made of specialised software packages. These packages increase the reliability and validity of extracted data, adding rigor to a study and its findings. However, they are time-consuming to use and are not usually used by students because of this. Software packages are ATLAS and MAXQDA, amongst others.

- ATLAS is software for text analysis and model building. It handles graphical, audio, and video data files and text.
- MAXQDA works with a wide range of data types, including focus groups, surveys, webpages, Twitter, and other social media. You can transcribe, analyse, and code audio and video files in a so-called Multimedia Browser. It has a wide range of coding features, a facility to link within documents or to external sources. You can add variables to your data to keep track of demographic or other types of quantitative information.

This is not an exhaustive list. Other packages available include NVivo and Qualrus to analyse the data.

Relational analysis (a.k.a. semantic analysis) involves searching for meaningful relationships present in a given text or set of texts. Relational analysis explores the relationships between the concepts identified.

The difficulty with content analysis is not that of locating relevant information. It is analysing the often vast amounts of available data. This makes the process very time-consuming and labour-intensive. Content analysis cannot easily investigate implied meanings and is not a useful methodology for assessing subtle meanings within the literature.

Patient information leaflets could be selected from a number of Trusts or from just one. These are in the public domain, and you can find them online from the websites of a number of NHS Trusts [12]. Once you have collected your leaflets, you could randomly select a manageable number to review. These can then be

scrutinised for the type, quality, and level of language used within these texts. You could analyse the sentence length and count the number of syllables used; you could also ask readers to rate the readability of the information provided.

You could search historical archives of professional material looking at the development of advanced practitioners or the development of non-traditional radiographic skills such as counselling, for example. This could be a specified timeframe, say the previous 5 or 10 years. Then you could look for articles, references or editorial comments relating to counselling skills. You could also look for training and educational programmes. You could explore whether definitions have altered, and which areas of practice specific skills focused upon, such as ultrasound, radiotherapy practice, or mammography services.

16.16.2 Key Points

- Content analysis involves the study of the content of communication and/or documentary evidence.
- It can be both qualitative and quantitative and usually involves inductive reasoning.
- It can be categorised into conceptual analysis and relational analysis.
- Content analysis can be time and labour-intensive.

16.17 Critical Reviews

Critical or systematic reviewing is a research methodology that aims to review primary research evidence with rigor. High quality reviews identify all relevant studies in a particular area of practice, assess the studies, synthesise the findings in an unbiased way, and present the results in a balanced and professional manner.

Evidence-based healthcare relies on systematic reviews ensuring that healthcare practitioners have a clear understanding of available research and ensuring that their practice is based on the best available evidence. Healthcare professionals are turning towards current available research in order to aid in efficient clinical decision-making. Critical reviews contribute to evidence-based practice by using explicit methods to select, critically appraise, and summarise large quantities of information and literature thus aiding the decision-making process. See Chap. 12 for more information.

16.18 The Search Strategy

Carrying out a structured search is essential and of utmost importance when undertaking a critical review. It helps to fully understand the topic in a question. It enables awareness of existing research in the same area and ensures the intended project has not been undertaken before. Even if a study has been done before there is often a need to review the latest available information and studies to ensure all

evidence is current. This may necessitate undertaking a review that was done previously. This is of particular importance if the new research will add knowledge to the literature already available. The Cochrane handbook for reviewers [13] recommends that a variety of sources should be systematically searched to reduce the risk of bias and broaden the search base. A critical review should include all available evidence. See Chaps. 3 and 12.

The identification of the best evidence and selection of research literature requires the construction of an appropriate research question. A stepwise process named PICO (population/participants, intervention, comparison, and outcome) has been developed to achieve this. A research project (study) should have elements of a population, investigation, comparative investigation, and an outcome in its question. Formulating a question using these key components should assist in specifying the criteria used to select studies. The importance of developing a focused research question is crucial, ensuring it highlights the significance of the problem. Critical reviews in radiography often use the PICO system.

16.19 Exclusion Criteria

The method used for including relative literature is undertaken in three stages.

- Stage 1: is the inclusion of studies based on their title and abstract to decide whether they were relevant to the question posed.
- Stage 2: involves establishing if the studies met the inclusion and exclusion criteria.
- Stage 3: assesses the methodological quality of the study and extracts the data.

Ethical implications need to be considered and addressed when undertaking any systematic review or content analysis. Ethical release may be required from an ethics committee and higher education institution (HEI). An unbiased, objective approach should be followed, using published guidelines for critical reviews combined with good reflective judgement. This ensures the answers to a research question are based on available evidence rather than unsubstantiated claims that may potentially produce misleading results.

16.20 Writing Up Qualitative Research

The manner in which your project should be structured and presented is explored in Chap. 17. A number of considerations should be taken into account when writing up your study.

- Most importantly, you must refer to the original research question/aims and assess the extent to which your objectives have been reached.
- Try to be consistent with the data. In other words, do not try to make too much of one quote that seems to confirm what you already think at the expense of other data that do not.

- Treat all data 'fairly'. It is okay to say there was no apparent pattern in responses, and also to say that 'all the women in the sample thought X', as long as you do not try to imply that this suggests any generality.
- It is always good to say that 'further research is needed to establish whether this represents a general trend'.
- Try to include appropriate quotes in your write-up, and also some individualised accounts/stories, etc.
- Your conclusion should include an assessment as to whether your research questions have been answered, what unforeseen results arose and, if possible, some recommendations for further research or practical action.

Undertaking qualitative research can be both fascinating and satisfying, and that information unearthed is often completely unexpected, and can be of immense interest and importance. Good luck if you go down this route.

16.21 Key Points

- The aim of a critical review is to rigorously evaluate primary research evidence.
- It is important that healthcare practice is based on the best available research evidence.
- A robust search strategy is essential when conducting a critical review (see also Chap. 3).
- Critical reviews in radiography often use the PICO system as a useful strategy.

16.22 Conclusion

In this chapter a brief and practical definition of the term 'qualitative' was presented. Also covered was an extensive discussion on the many and varied methods by which qualitative data may be obtained, recorded, and analysed. A detailed investigation of a qualitative interview was presented. This is partly because of its prevalence in qualitative research, either as a 'stand-alone' method or as one of a suite of data gathering techniques used in action or case study research. In both your academic and professional career you will undertake both research and other interviews as part of your studies or your job.

The techniques, advantages, disadvantages of other qualitative methods were also outlined, together with the contexts within which each method would be most appropriate to adopt. In conclusion, it is necessary to consider exactly what your research is about and what it aims to explore before choosing to adopt a qualitative (or any) approach. Having done so, you must think very carefully about the actual qualitative data gathering method you decide to adopt. Hopefully, the chapter has provided a good guide as to the application and appropriateness of each, to equip you to make your decision.

References

1. Lund H, Brunnhuber K, Juhl C, et al. Towards evidence-based research. BMJ. 2016;355:i5440.
2. Jamshed S. Qualitative research method-interviewing and observation. J Basic Clin Pharmacol. 2014;5(4):87–8.
3. Besen-Cassino Y, Cassino D. Social research methods by example: applications in the modern world. London: Taylor & Francis; 2017.
4. Nicholas D, Huntington P, Williams P. An evaluation of the use of NHS touch-screen health kiosks: a national study. ASLIB Proc. 2002;54(6):372–84.
5. Mintzberg H. An emerging strategy of "direct" research. Adm Sci Q. 1979;24(4):582–9.
6. Richie J, Spencer L. Qualitative data analysis for applied policy research. In: Bryman A, Burgess RG, editors. Analyzing qualitative data. London: Routledge; 1994. p. 173–94.
7. Parkinson S, Eatough V, Holmes J, Stapley E, Midgley N. Framework analysis: a worked example of a study exploring young people's experiences of depression. Qual Res Psychol. 2015;13(2):1–21.
8. Nind M. What is inclusive research? London: Bloomsbury; 2014.
9. King R. All things bright and beautiful: a sociological study of infants' classrooms. Chichester: Wiley; 1978.
10. Rathje W. Rubbish!: the archaeology of garbage. New York: Harper Collins; 1992.
11. Mayo E. The social problems of an industrial civilization. London: Routledge & Kegan Paul; 1949.
12. Dawson D, Lighton B. A patient guide to the use of insulin for diabetes. Birmingham: Queen Elizabeth Hospital Birmingham; 2018. https://www.uhb.nhs.uk/Downloads/pdf/PiNewToInsulin.pdf. Accessed 04 Oct 2019.
13. Higgins JPT, Green S, editors. Cochrane handbook for systematic reviews of interventions, version 5.1 [updated March 2011]. http://handbook-5-1.cochrane.org/.

Further Reading

Smith J, Firth J. Qualitative data analysis: the framework approach. Nurse Res. 2011;18(2):52–62.

Part VI

Writing Up and Disseminating

Dissertation Structure and Presentation

17

Aarthi Ramlaul

17.1 Introduction

A conversational style is used in this chapter to 'speak' to students. However, later in the chapter the use of academic style writing is covered. The purpose of a dissertation is threefold: to complete a higher education degree, to evidence your intellectual ability, and to demonstrate that you can clearly and effectively communicate your findings. These three key purposes shape the structure and presentation of your research in a written format.

There is often some confusion over the terminology used to describe the process of undertaking a project and writing a dissertation. You can consider a project as a process that involves a planned activity designed to achieve a particular aim. A dissertation is a written document in which the background to, process of, and findings, conclusions and recommendations of a project are discussed. It is usually produced for an undergraduate degree or a postgraduate master's degree. A thesis usually refers to a dissertation produced at a doctoral level. In this chapter the term dissertation is used for simplicity and consistency.

Producing a dissertation is a major requirement of most higher education courses and is likely to be the largest single piece of work you will be asked to produce, especially at an undergraduate level. It involves undertaking some form of data collection where you usually have the opportunity to pursue in-depth a topic of your choice. Ultimately you will be expected to produce this extended documentation to demonstrate your ability to engage critically and analytically with appropriate literature, report on your own work, offer your own thoughts and interpretation on your findings, reflect on the research process and, in the case of doctoral studies, make an original contribution to knowledge. Essentially it is a document to demonstrate your

A. Ramlaul (✉)
Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: a.ramlaul@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_17

ability to present your research in a scholarly manner. As opposed to an essay, it requires more research in greater depth, more reading, more time, more independence, more planning and more writing. It is important, therefore, that you plan ahead, manage your time effectively and write up your project thoroughly. Key tips to ensure good planning and management include setting deadlines, starting to write early, writing regularly and in stages and allowing time for revision of draft work.

Whatever the nature of your project, a formally presented dissertation should be able to be read clearly for quick understanding. It should be logically organised and presented so that a clear story unfolds. It has to be accurately referenced to ensure that claims made are based on evidence.

Dissertations may vary in scope, but most are based on a systematic review of literature, clinical audit or primary research. A review of the literature involves carrying out a systematic critical appraisal and synthesis of the current state of knowledge relating to the topic under investigation. Since all published literature is in the public domain no ethical issues arise. A clinical audit involves collecting data on current practice and comparing this data with a locally, nationally or internationally agreed standard to find out what we are doing in an attempt to inform delivery of best care. You will normally need to obtain written department/hospital approval to collect such data (see Chap. 14). Primary research attempts to find out what we should be doing and to derive generalisable new knowledge. This involves collecting data, usually from hospital staff or patients, or using hospital equipment and resources.

As the Latin word *dissertare* means to debate, a dissertation requires you being able to examine and discuss a topic from different points of view and advance an original opinion. Thus, a dissertation shows that you are able to critically appraise relevant literature, apply principles of good project design and management and produce an academically rigorous document.

17.2 Structure

There is no best way to structure your dissertation: educational institutions and funding bodies produce their own specific requirements regarding the format and length of a work. However, all dissertations should contain similar elements. A suggested approach to structure and organise your work is presented below. This generic structure is used in most academic writing, where the process is quantitative and linear, with the review of literature preceding the collection of data. However, this may not be entirely suitable for qualitative projects; the literature and research may be more interconnected, or where you are looking at several themes that would be better dealt with in separate chapters. Whatever the nature of your project, structure is important and possible ways to achieve a sensible and consistent structure should be discussed with your research supervisor.

17.2.1 Title Page

This should be succinct, reflect the nature of your dissertation (hereafter the dissertation) and contain enough information to attract relevant readers. ‘Women’s attitudes to nuchal translucency screening’ is better than ‘Women’s attitudes to

undergoing the 11–14 week ultrasound scan to assess nuchal translucency where the risk of the foetus having Down syndrome is calculated': the first title contains enough keywords (women's attitudes and nuchal translucency screening) to encourage further reading. Make sure the title has no spelling mistakes or typographical errors: first impressions are very important. In addition, this page should include the degree for which the dissertation is presented, the name (or number) of the student, the awarding institution, date submitted and word count (excluding reference list and appendices). Figure 17.1 shows a typical sample title page.

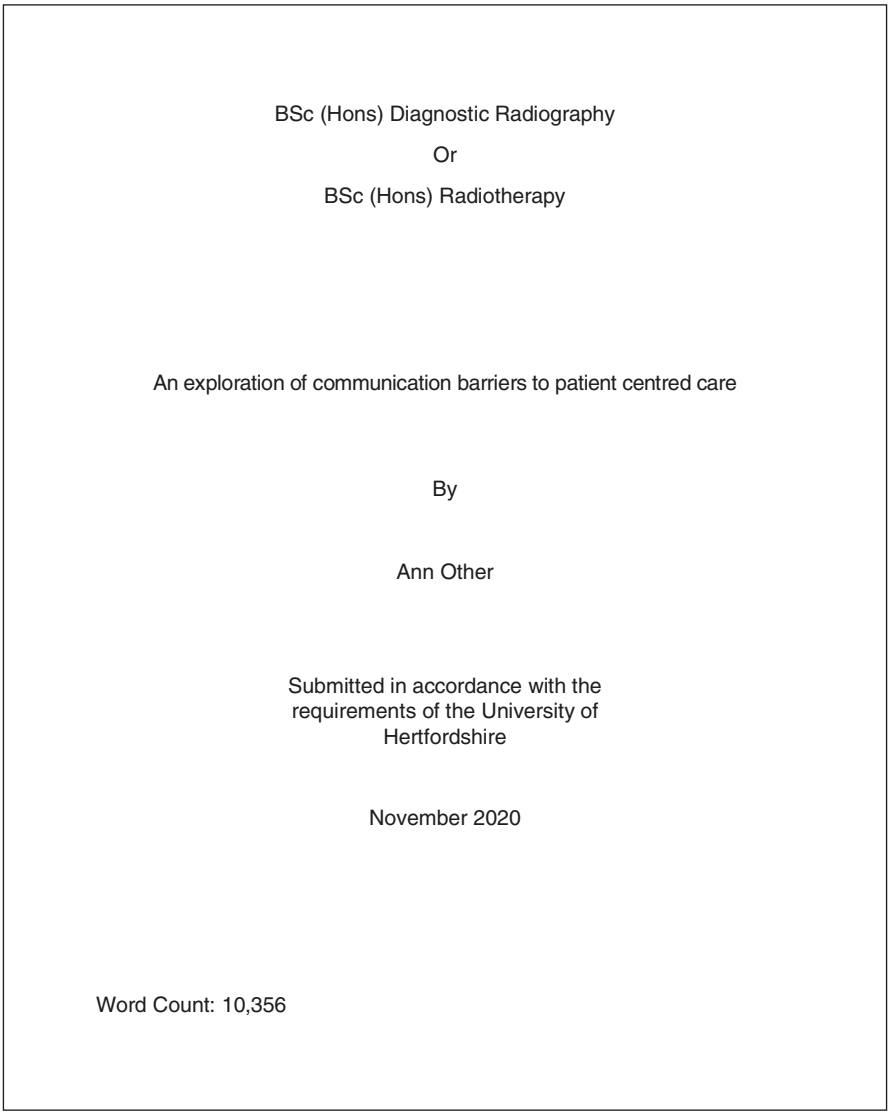


Fig. 17.1 Sample title page

17.2.2 Signed Statement

You must declare that the dissertation is your own work. This page must, therefore, contain a signed declaration such as: 'I hereby declare that the information contained in this dissertation is substantially my own work.'

17.2.3 Contents

This should be a list of the main chapters, subsections, figures/tables, reference list and appendices including page numbers. An example of contents is shown in Fig. 17.2. Do note that page numbers are usually indicated in Roman numerals before Chap. 1 and in Arabic numbers thereafter.

17.2.4 Abstract

The beginning of the dissertation should be especially clear and engaging for readers so that they are encouraged to explore further. The abstract should be a short summary telling readers everything they can expect to find in the dissertation, including a statement of what the dissertation is about, some background to explain the focus, scope and rationale for the project, the project question, how the data were collected, the key findings and clear conclusions/recommendations. All this information should be contained within about 300 words. Note that an abstract does not contain references.

17.2.5 Acknowledgements

Where appropriate, acknowledgement should be made for assistance given by individual supervisors and mentors or organisations that may have supported you, provided information or supplied equipment and materials. However, you may or may not wish to acknowledge persons; if you choose to, then remember to keep your messages of thanks professional. It is not appropriate to include very personal accounts of help or thank the family pet.

17.2.6 Glossary

This is a list of technical terms or abbreviations with definitions that can be included, if appropriate, in alphabetical order at the beginning of the dissertation. Examples are shown in Table 17.1.

Contents

Fig. 17.2 Sample contents page

Table 17.1 Example of a glossary

Technical terms	*Definition*
Bias	A subjective attitude or viewpoint which can cause distortion or deviation of research findings from its true value or meaning
External validity	This refers to the extent to which research findings can be generalised or applied to other population groups
Dosimetry	The measurement of exposure to radiation from a radiation-emitting source
Abbreviations	*Definition*
IR(ME)R	Ionising radiation (medical exposure) regulations
GCS	Glasgow coma scale
EBRT	External beam radiation therapy

Table 17.2 Example of a list of figures/tables

Figure/table	Page number
Figure 3.1 Search strategy for literature databases	29
Figure 3.2 Diagnostic referral pathway	32
Table 4.1 Results of dosimetry studies	42
Table 4.2 Results of cost-effectiveness studies	45

17.2.7 Contents Page with List of Figures and Tables

The contents page is generally produced at the end when the dissertation is written and contains all the chapters, subchapters, list of figures and tables. It is a useful guide to a reader. At a glance it clearly signposts the entire structure of the dissertation. It is a good idea to number the illustrations according to the chapter in which they occur. See Table 17.2 for an example of a list of figures and tables. Thus, in Table 17.2, Table 3.1 is in Chap. 3 (Methodology) and Table 4.1 is in Chap. 4 (Results).

17.2.8 Chapter 1

This is the scene-setting part of the dissertation and should explain background information. The nature of the project and its importance for clinical practice should be clear to readers. There should be a rationale for choosing the topic area. Definitions of any key concepts should be explained to help aid the understanding of, and context for, the study. You should assume that readers have some background knowledge of your topic, but you need to bring them up to date with developments in that area and explain the current situation. For example, for a dissertation entitled 'A comparative analysis of the effectiveness of magnetic resonance imaging (MRI) and computed tomography (CT) in the diagnosis of acute stroke' you could begin by outlining the aetiology, epidemiology and pathology of the condition, followed by a summary of treatment options, a review of the role of imaging in

diagnosis and the implications for practice. Also included in this chapter should be a statement of the project question/hypothesis and clearly defined aim(s) and objectives. See Chap. 2 for guidance on how to write these.

17.2.9 Chapter 2: Literature Review

This is a review of published literature and should give a reader a clear understanding of the topic under discussion, key issues that arise and the rationale for undertaking the project. You will be expected to demonstrate your ability to search for, and access, the breadth of literature informing the project and bring together past and present research and published work. The aim is to produce a distilled, critical appraisal of relevant literature, showing how your project links to previous research and existing knowledge, identifying any gaps in published work, and how you intend to take this forward.

The purpose here is to critically appraise the literature you have read, and not simply collect references, describe them and make a list of what you have read. Critical thinking involves higher order thinking. It requires you to assess the value of knowledge and information. This involves you looking at a piece of published work in an objective and structured way. It requires you to ask questions, such as those below.

- Is it an account of research or someone's views and opinions?
- How recent is the work and any other work discussed?
- What are the main points raised and are these supported in any way?
- Do you agree with the inferences and conclusions made?
- How does it fit with other work you have read in this area?
- What are its main strengths and weaknesses?

By considering these questions means you should be able to demonstrate that you are able to do the following.

- Make connections and see patterns in existing work.
- Organise your thoughts by being analytical and reflective.
- Explore issues and identify reasons for any conflicting information.
- Take a new perspective on an issue or challenge current perspectives.
- Consider all assumptions and alternate views.
- Justify your findings using evidence.
- Make a decision on whether practice should be changed.

For example, if you were undertaking a clinical audit to estimate how effective implementing the Ottawa knee rules in cases of acute trauma injuries is in reducing the number of radiographic examinations undertaken, then your literature review might include the issues identified in Table 17.3.

Table 17.3 Example of a structure of a literature review

Structure	Issues
Statement of question/ hypothesis	The use of the Ottawa knee rule (OKR) will decrease the number of patients referred, following acute knee injuries, for radiographic examination without decreasing diagnostic sensitivity
Background/ rationale	• Extent of the problem
	• Value of radiographic examination of the knee joint following acute injury
	• Radiation dose
	• Cost to health services
	• Importance of undertaking project
Critical appraisal	• Range and value of clinical decision rules
	• Derivation of OKR
	• Application of OKR
	• Comparison of effectiveness with other clinical decision rules
Summary	Your assessment of the effectiveness of OKR in reducing number of radiographic examinations based on the findings

Overall, the literature review is a piece of informative writing where you have the opportunity to 'tell a story' or 'paint a picture' about current knowledge related to your topic. How the literature has shaped or informed the project is, consequently, an essential component of the dissertation.

17.2.10 Chapter 3: Methods of Data Collection/Methodology

In this chapter you explain how the question/hypothesis you posed in Chap. 1 was investigated and why you employed the chosen methods and techniques. Your explanation should provide sufficient detail for a reader to understand what you did to collect and analyse the data. The design of the project is a vital part of the dissertation for both a reader trying to judge the validity of the findings and for anyone interested in replicating your study. It is suggested that this chapter could be structured to include to following.

- Restatement of the aim of the project.
- A discussion of the possible methodological approaches that you could have used to collect your data.
- An explanation and justification for the approach you did use.
- A description and validation of the ways in which the data collected were analysed.
- A discussion concerning any ethical issues and a statement that appropriate ethical/departmental approval was sought (written evidence that this has been obtained should be included in the appendices).

Details of any tools you used to collect data, for example, measuring instruments, questionnaires, interview schedules, etc., should be included in the appendices. This chapter is important as it allows a reader to be clear about how you met the objectives

identified in Chap. 1. For example, what steps were followed in collecting the data, why you chose this approach, how you selected your participants, the sampling technique used in recruiting them to your study and how you analysed your data. If using a qualitative approach, you must include information about the philosophy underpinning your chosen methodology. Due to the interpretive nature of qualitative research you must also demonstrate how you ensured validity and reliability of your findings. In qualitative research, these criteria are known as credibility and dependability and are important in order to demonstrate the trustworthiness of your research findings.

17.2.11 Chapter 4: Results/Findings

In this chapter you should provide an accurate and full summary of the data collected, possibly using illustrations (tables, figures, graphs and charts) to support the text. Quantitative data may include both descriptive statistics, for example, response rates, gender and age characteristics of the sample and inferential statistics (the results of any statistical tests). Generally, the text should be written after any illustrations so that it highlights significant aspects of the findings and not duplicate information. All illustrations should be numbered consecutively, titled and appropriate legends supplied (units, axes for graphs, etc.). Illustrations must be referred to within the text.

Qualitative findings are usually presented as key themes that emerge from the data collected, and may be supported by illustrations to demonstrate relationships between various components or sequences of events. Whatever the nature of the results, this chapter needs to be structured so that it allows a reader to clearly understand what was found in your study.

17.2.12 Chapter 5: Discussion

Here you should provide a broader and deeper interpretation of the findings and the possible implications they might have for practice. This is the section where you are required to critically evaluate and discuss your findings in light of published literature. Simply reporting findings is not enough; you must give some meaning to what has been found. However, while qualitative results are by nature interpretative, usually the results and discussion are combined in one chapter. Conversely, as quantitative results consist of some form of statistical analysis, the interpretation of these findings is best undertaken in a separate discussion chapter. Whatever the nature of the project, the aim is to draw together the findings of the study and discuss these in light of what the project set out to achieve and in the context of previous studies. This involves more thinking than any other part of the process. You should consider the following.

- Was the research question fully answered?
- Did you spell out acceptance or rejection of the hypothesis (if a hypothesis was stated)?

- Were there any weaknesses or limitations in the project such as design constraints, sampling limitations, data analysis problems and so on?
- Did you explain how problems encountered were overcome?
- Are there other approaches that might have been more appropriate with the benefit of hindsight?

Thus, this chapter should include an analytical interpretation of findings; a critical reflection of the process and to what extent your work has contributed to knowledge. For undergraduate and master's projects this will be in relation to the benefit to your and others' practice. For doctoral projects, there is an expectation of your study making an original contribution to knowledge and practice.

17.2.13 Chapter 6: Conclusions

This chapter should be written to the point. You must state clearly and succinctly what you found in your study. The conclusions should follow fairly obviously from the discussion and correspond to the original objectives. No new material should be introduced. However, only conclusions, based on the findings, should be made and you should avoid the temptation to add an opinion for which no evidence has been presented.

17.2.14 Chapter 7: Recommendations

This chapter may not always be appropriate or may be included in the conclusions section. Main recommendations include ways in which the findings of your project might improve practice or provide a stepping-stone for further research.

17.2.15 References

This is a list of sources (books, journals, websites, documents, etc.) that you referred to within the body of the text of the dissertation. The list should be compiled exactly in accordance with the referencing system requirements of the institution.

17.2.16 Bibliography

This is not an essential item but may be a list of published work that you consulted but not used in the text. Usually just the list of all references used is the preferred requirement.

17.2.17 Appendices

Appendices should be kept to a minimum. You should include, for example, a copy of a blank questionnaire, interview schedule or data collection sheet used and copies of letters granting ethical approval and permission for your project. Certainly, you should not include numerous copies of journal articles or downloaded internet pages. Do not include any confidential material within the appendices and any identifying details, such as names of persons or institutions (e.g., NHS Trusts) should be anonymised. It is a good practice to refer to all appendices within the body of the text.

17.3 Writing Style and Presentation

In addition to the general structure of a dissertation, the style of writing and presentation of material is important. Having spent so much time designing your project, reviewing literature and collecting data, it would seem illogical not to take care with the presentation of your dissertation. Therefore, your dissertation should be well organised and carefully presented. You should adhere to the usual conventions which, although they may seem rather pedantic, are useful because they provide a standard form of presentation.

It is a piece of work you should be proud of, and one which will be of use to future research. Remember, your dissertation is not an essay. It is a factual account of how and why a topic area was studied and what results were obtained. It is essential to write simply and clearly: the objective is to communicate your ideas and not to bury them in complex sentence construction or unnecessary jargon.

An excellent presentation makes the most of a good project. Although the dissertation is judged largely on its academic content, some marks are also given for good presentation. Each educational institution has its own presentation requirements; make sure you check details of specific requirements. In addition the following points need to be considered if you are going to make your dissertation attractive to look at, easy to read and logical to follow.

- It is imperative that from the outset you develop a consistent ‘house style’ where your chapter and section headings all have the same numbering system, font type and size. Normally three levels of headings/section headings are sufficient, as shown in Box 17.1. Make sure that sentences are well punctuated and not overlong. Paragraphs should be adequately developed. You should include linking words or phrases to guide a reader through the text.

Box 17.1 Example of ‘House Style’

Chapter 3 Methodology
- 3.1. Introduction
- 3.2. Research Design
- 3.2.1. Research Instruments
- 3.2.2. Ethical Considerations

- Issues of correct spelling, punctuation and grammar are crucial if your final dissertation is going to be taken seriously. You have invested a lot of time and effort in the project and it would be a pity if the impression or message resulting from your work was diminished or lost because of poor attention to these details. Therefore, use a spell/grammar checker and make sure the meaning is clear and the language is comprehensible. Paragraphs are best presented as fully justified (i.e., the left and right of each line of text line up). Make sure that contradictions, illogicalities and irrelevancies are avoided.
- Take time to ensure silly mistakes are not made: for example, paragraphs included twice, images indistinct, graphs poorly labelled and so on. Where viable, all illustrations should be created using appropriate software tools, as hand-drawn images detract from the overall appearance of the dissertation. As mentioned previously, refer to illustrations, for example, figures, tables, images or graphs within the body of the text.
- Avoid the use of the first-person singular such as 'I developed a questionnaire…', but rather write in the third person as 'a questionnaire was developed…'. You should also write in the past tense, as ultimately your dissertation will be read as something that has been completed and not as something that is yet to be done. Thus, for example, you should write that 'the data were analysed using inferential statistics' rather than 'the data will be analysed using inferential statistics'.
- It is usually a part of any university's equal opportunities policy to discourage the unnecessary use of gender related language. To overcome this problem it is easier to use the plural, for example, 'practitioners should always...', but if you need to use the singular, try words like a patient, a student, an individual, a practitioner and so on. Provide signposting by outlining the structure of the dissertation at the end of Chap. 1 and refer, in the text, to relevant points dealt with in other chapters/sections of the dissertation. It is a good idea to get into the habit of writing a brief introduction at the beginning of each chapter to give taste of what is to come, and a brief summary at the end of each chapter reflecting on key issues. In this way, you direct your readers to what you are about to tell them, then you provide the promised detailed information, and at the end you summarise what you have told them. This structure helps to keep chapters organised, logical and cohesive.
- All of this is most easily achieved by making sure you write drafts of chapters well in advance of the hand-in deadline, read them through yourself, get others to read through your text and make appropriate revisions after each reading. Rigorous proofreading of your dissertation allows it to be well crafted for its final version. Remember that the first draft is always for your eyes only. Do not submit your first draft to your supervisor for feedback guidance. Once you have completed a good quality first draft, read, re-read, edit and update your draft before sending the work for supervisor feedback.
- For students where English is not their first language, writing your dissertation can be a challenging task. Your educational institution should be able to provide you with English literacy skills support where you can discuss your academic writing needs with a skills support tutor. However, if this is not readily available

at your place of study, discuss your requirements for academic support with your research supervisor, so that additional help can be offered earlier on in the writing process.

- Present your aims and objectives clearly. Do not spend paragraphs telling your readers what you are about to do. Simply be direct and get to the point quickly. Avoid 'flowery' talk or overt descriptions. Your work should be written simply and must be clear. Avoid the use of long, convoluted sentences and jargon. Do not try to impress your readers by using difficult or fancy words if you do not fully understand their meaning. If used inappropriately, they can hinder the clarity of communication that you are trying to establish. The manner in which you write should tell your reader that you understand your study thoroughly and have taken the necessary measures to make it of the highest quality.
- Present your discussion points in a logical order, possibly using headings and subheadings. It may help to make a list of the points you are going to present. The quality of the writing is judged by the quality of evidence provided; therefore, do not include unsubstantiated opinions unless you are stating observations.
- The manner in which you present your information should be well organised. The main points of your work should be presented clearly. Pay attention to the overall structure of the dissertation and the chapters within it. Use paragraphs well to clearly present your ideas. Use different paragraphs for different ideas, but be careful to link them so that they provide a cohesive flow of information and draw a reader through your work. When you write, do not assume that a reader has any prior knowledge of the topic area. Consider the use of appendices to contain background information and then link this with your writing. It is a good practice to cover material from first principles, i.e., from basic to complex. Although this may seem elementary to mention, it is something that is often overlooked, especially by those who do not write very often.
- While it is important to get your point across you need to bear in mind the nature and context of your writing. In this case it is a research project so you need to write using a style to suit that format. In the same way, writing this chapter is suited to a more conversational style so it would be acceptable for me to write this as though I am speaking to you. Your writing style has to match the medium within which you are trying to express yourself.
- Writing in the first person as 'I' is not considered to be a good academic practice; therefore, writing in the third person is advised. However, if you are writing a reflective essay, for example, then writing as 'I' would be acceptable. Writing in the third person requires vigilance, as students tend to start off well and then by the sixth or seventh page of the dissertation, they have slipped into 'I' or 'my'. Pay attention to this as you do not want to lose valuable marks for avoidable mistakes like this one. It is a good practice to give your work to someone else to read. Choose this person carefully as they should act as a critical friend who is willing and objective enough to give you constructive criticism. It is not advisable to ask a classmate to read your work, however well you get along with each other.
- Presentation is not restricted to text. The manner in which you display your illustrations is also important. Try to choose the best fit for your data using the guid-

ance given in previous chapters; do not present the same information in more than one way. The manner in which you present this should be appropriate, clearly labelled with descriptive titles and understandable without tedious amounts of explanatory text. If information does not appear clear, then the chances are that it really did not make sense to you anyway. You may either rework it to make it clear or leave it out entirely, provided it is not the information critical to your study.

- Pay attention to the use and presentation of abbreviations. Only use them if you have written the words out in full first. It is a good practice to include a list of abbreviations as an appendix. This may be presented either alphabetically or in the order in which they appear in the text.
- Be careful of the information you present as appendices. Students often include many appendices that they do not refer to within the body of the text. If an appendix is not important enough to be mentioned in your written text, then leave it out. Items that you choose to present as appendices should be laid out in the order in which they would be referred to within the text and clearly marked. It is a good practice to attach a list of all the appendices that are sequentially numbered to your dissertation.

17.4 Pay Attention to Analysis

Students often carry out extensive literature searches and have seemingly exhaustive data extraction forms but one of the main areas in which students lose the most marks in their final write up is in analysis of literature, usually reflected in the literature review or discussion sections of their dissertations. These sections of your dissertation should be the areas that most clearly demonstrate your ability to evaluate and interpret the literature in relation to your original research topic and the findings of previous studies. Literature should be used in an integrative way to support an argument or justification. It should also discuss material that offers a conflicting view. Look for gaps in your arguments and evidence to determine whether you have sufficient information to support your arguments. In this way, you then will be attempting to provide a well-balanced view that can be built up into a discussion or an argument depending on the nature of the information. Remember to consider both theoretical (of concepts and theories) and empirical (of studies, e.g., randomised clinical trials) literature. The discussion chapter of your dissertation must link with the literature review chapter. Although you are encouraged to develop your thoughts and present original ideas, be careful of introducing new literature in this chapter it is a common pitfall.

While on the topic of the discussion chapter, a common pitfall occurs when students introduce new information towards the end of their dissertations. It is a good practice for ongoing literature searching and constantly checking for updates on information and news that will be ‘hot off the press’, however, it is important that this information is analysed in your literature review. You can then revisit this information in your discussion of findings later on.

It is a good practice to get into the habit of evaluating your work as you go along so that you become your severest critic. Think about how you would like to read a topic of that nature if written by someone else by putting yourself in a reader's position. It is harder to make yourself clear in written communication than it is during verbal communication.

During verbal communication, you are easily able to 'fill in the blanks' as your conversation proceeds; however, written communication can be tricky in ensuringthat every angle of the topic you are writing about has been adequately covered. Poor analysis of literature impacts on your review of literature and your discussion chapter(s) of your dissertation. It has the potential to rob you of a good mark.

One of the reasons students often lose marks for analysis of literature is poor strategy in literature searching. A robust literature search strategy helps ensure that you have adequately covered your topic area's research base to include a good range and depth of material in your analysis. There is no fixed number of literature sources, such as journal articles, recommended for writing a dissertation; however, it is expected that you would conduct exhaustive searches in your field of study. Guidance on literature searching is given in Chap. 3 and on literature evaluation in Chap. 4.

To analyse literature, try breaking down the information into component ideas. Then look for literature that supports these ideas or criticises them. Pull all pieces of information together and then write up the information you gathered in an all-encompassing manner. This is known as synthesis of information: building up of information using new ideas, concepts and theories, etc., from findings. Remember the quality of the online search engines used will determine the quality of information gathered. Google and Wikipedia are not considered to be credible sources of information for academic writing because what is published is unlikely to be peer reviewed.When evaluating and interpreting your results be careful of over-interpretation of the findings, especially in qualitative studies. It is not uncommon for researchers to analyse their way through an emergent theme from their findings and then talk themselves completely out of the same theme. If in doubt, or in need of clarification, it is a good research practice to go back to your participants and have them verify the transcripts and discuss your dilemma with your supervisor. It is easy to misinterpret or over-interpret transcripts. Being reflective and reflexive (see Chap. 5) throughout the process will help you acknowledge your own biases and opinions and will help in the accurate interpretation of research findings.

When you draw conclusions from your findings, remember to justify these from evidence. Recommendations made, for example, for further study, should be feasible and contribute to advancing practice.

17.5 Referencing

Referencing is an essential academic requirement. It demonstrates that extensive reading has taken place and that you properly acknowledged the work of others. Little or no referencing indicates insufficient reading to support the project or that you have copied ideas, data or facts from the source material. If you use text that is

unreferenced (generally work that is not your own is easy to spot), you may at worst be accused of plagiarism or at best poor academic practice. You should take particular care to make each reference full and accurate: include the editors and publishers of books and volume numbers of journals in which articles appear.

There is a multitude of referencing conventions used in academic writing: for example, Harvard, Vancouver, Ciba and American Psychological Association (APA) systems. Details of how each system is actually adopted for use with a variety of material including books, journals, chapters in an edited book, internet websites, electronic material, student dissertations, newspapers, etc., should be checked with your educational institution. Always use the referencing convention that is stipulated by the institution who you are submitting to.

Remember that presenting someone else's words, ideas or images as your own, without referencing them is plagiarism. If you included material, which is not your own without any adequate attempt to give appropriate credit, or you incorporated material as if it were your own when in fact it is wholly or substantially the work of another person, then this is essentially cheating. Therefore, pay attention to the academic integrity of your work by ensuring that all literature used is acknowledged through accurate citations and referenced appropriately.

Referencing accurately according to the required convention is an area in which students commonly lose valuable marks. Very often students leave the writing and editing of their referencing to the last minute, resulting in a rushed approach that is clearly visible during the marking process. Despite extensive and easy to follow guidance supplied by the educational institution students do not apply the referencing guidelines/technique to the required expectation. Losing marks in the referencing/citation marking criteria can sometimes make the difference in grade boundaries.

17.6 Length

Make sure that you know the lower and upper word limits acceptable for your dissertation. A written project word limit typically excludes title page, abstract, list of contents, list of accompanying material, acknowledgements, glossary, references, bibliography and appendices.

Usually a word limit is imposed (e.g., 10,000 words/þ/−10%), as it is believed that at this level the topic can be covered in sufficient depth and breadth in about 10,000 words providing it is well focused and structured. Thus, for example, if you write less than 9000 words, it is likely that you will not have dealt with the subject in sufficient depth or detail. Alternatively, if you find yourself writing significantly more than 11,000 words, it is possible that you have not sufficiently focused your ideas. While you may not be penalised for going under the stipulated word count (as it may be deemed that you have penalised yourself), a significantly exceeded word count (i.e., more than 10% in this case) may mean that penalties will be imposed. Diagrams, graphs, tables, etc., normally count as the word equivalence of the space they occupy.

17.7 Pagination, Margins and Spacing

Number your pages consecutively throughout the dissertation and locate them centrally at the bottom of the page, approximately 10 mm above the edge. Margins at the left-hand edge should not be less than 40 mm (to allow for binding) and other margins not less than 20 mm. Double or 1.5 line spacing should be used for all text, although indented quotations should have single line spacing.

17.8 Fonts

Arial and Times New Roman are the most common font styles: they are easy to read, thus avoid elaborate font styles. Font size should be about 11, although the size of chapter headings may be larger and bold, while subsections may be italicised or underlined. Since you will be building up an argument, ensure you break up the text so that key issues can be readily identified by a reader. The important point is that once you have chosen the font style you should apply it consistently throughout the dissertation.

17.9 Hard Copy and Online Submissions

For hard copy submissions, you should use A4 size paper of good quality and of sufficient opacity for normal reading. Type or print on one side only. The dissertation should be professionally bound using either loose-leaf spiral binding or hard covers. It is a good practice to include an acetate over the front cover to protect it from unwanted marks or damage.

Not all projects are submitted in hard copy format. Most educational institutions have moved towards online submissions only. If submitting online, take care with the format in which you save your work, for example, use portable document format or PDF, or if submitting a word file ensure that it is saved as a docx file. Depending on the intranet software that your educational institution uses, some formats may be distorted upon submission and you want to ensure that your dissertation is received by your supervisor/markers in the format you intended.

17.10 Conclusion

In this chapter guidance on how to structure and present your dissertation was provided. Some commonly encountered errors were discussed and suggestions on how to overcome those were presented. Also covered were good practice tips to follow. A well-structured dissertation should be clearly presented, logical to follow and coherently aligned drawing together the arguments made in relation to answering the research question.

Acknoweldgement Alan Castle is thanked for his work on the original chapter. Many illustrations have been retained within this new version.

Further Reading

Bell J, Waters S. Doing your research project. A guide for first-time researchers. 7th ed. London: Open University Press; 2018.

Writing for Publication and Presenting at Conferences

18

Julie Nightingale

18.1 Introduction

Radiography has an expanding evidence base which demands that its practitioners engage in continuing professional development in order to keep abreast with developments, thus ensuring high quality patient care. Radiographers need to be evidence-based practitioners who have a thorough awareness and understanding of recent research findings. They may do this by reading key professional journals, searching databases or setting up e-alerts in their specialist fields. In order to undertake critical appraisal of an article and identify its significance to one's own practice, a radiographer requires basic research skills that are provided in undergraduate and postgraduate programmes. However, while some radiographers work within a multi-disciplinary team and support the research of others, there are a range of opportunities available for a novice researcher to contribute to radiography-based research beyond a dissertation.

Conversion of academic work submitted for a university course may be the first opportunity for many novice authors to publish or present their work on a larger stage. Following graduation, engagement in local audits, service evaluations and research projects are an excellent way for radiographers to develop their research awareness. While these are often disseminated locally, many novice researchers lack confidence in making the next step on the publishing ladder, such as having an article published or presenting at a conference. Why is this? Perhaps authors do not think that they have anything interesting to say, or writing is not for someone like them. Perhaps they are confused by the options available or the process of submitting is a bit of a mystery. Or they have taken those first steps into publishing but have been disappointed or disillusioned with the feedback.

J. Nightingale (✉)
Department of Allied Health Professions, Sheffield Hallam University, Sheffield, UK
e-mail: J.Nightingale@shu.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_18

Effective research dissemination is just as important as undertaking the research itself. If nobody knows about your work then they cannot learn from it and potentially change their practice or enhance patient care. The different options available for novice researchers to disseminate their work via conference presentations and journal articles are explored in this chapter.

18.2 Peer Review

A peer review process is a well-established method of improving the quality of articles published within a journal, and of selecting credible and topical research for presentation at a conference. Applications to present your work at conferences (known as proffered abstract submissions) usually require you to state whether your preference is for a presentation or for a poster (visual hard copy display or an electronic poster). However, the number of proffered presentations is often limited, so occasionally you will be offered an opportunity for a poster submission as an alternative. Tips for preparing a poster are provided in another chapter.

When submitting articles to a journal, an editor briefly reviews your paper (known as a manuscript prior to publication) before sending it to at least two peer review experts in the subject or methodology. There are a number of different peer review systems. The one most commonly applied in the radiography field is that of double blind peer review. In this system the authors do not know the identity of the reviewers, and vice versa, promoting fairness and independence in decision-making. Very few manuscripts are accepted at first submission; the majority require either minor or major revision to bring them to the standard required of the selected journal. This often requires more work for the authors, but the published article will invariably be of better quality than the original submission. However, for some papers an editor makes a decision to reject a manuscript before sending it for peer review, or rejects it following peer review because of significant revisions required. Even with a rejection the feedback received should be seen as an opportunity to build upon for the next submission; gaining a few knock-backs are a reality of being a researcher, and resilience is a positive attribute. Some simple guidelines are offered to help authors to avoid the common errors and pitfalls in article preparation and abstract submissions. The chapter concludes with a reminder of an author's responsibilities in ensuring that a published article is promoted widely.

18.3 Ten Top Tips for Publishing Success

Whether submitting abstracts to conferences or manuscripts to journals, there are a number of potential pitfalls to be avoided. Figure 18.1 outlines ten tips for success that have to be addressed during the preparation and submission process. Each tip is explored below.

Tip	Tip for Success	Main Focus
1	Choose your platform carefully	• Journal and conference selection • Message, audience and reach • Journal access models • Journal metrics and altmetrics • Predatory journals and conferences
2	First Impressions	• Title and abstract • Selection of keywords
3	Stick to the rules	• Guide for Authors • Word counts and referencing • Scientific Article Format • Converting academic work into articles and presentations
4	Ethical Publishing	• Co-authorship • Permissions • Ethical approvals • Plagiarism • Duplicate publication
5	Proof Reading and Preparation	• Clear and succinct English • Spell check and grammar checks • Avoiding jargon • Presentation rehearsal
6	Rigorous referencing	• Referencing styles and accuracy • Referencing software • Currency • Adaptation to specific journal
7	Replicability	• Explicit ethical approval and consent • Clear and reproducible • Clear tables and figures • Analysed, not raw data • Quotations
8	Credible conclusion	• Alignment to abstract and results • Results placed in context • Recommendations
9	Don't take it personally!	• Dealing with rejection decisions • Responding to reviewer feedback • Responding to audience questions
10	Self-promotion	• Effective publication and impact • Journal responsibility • Author responsibility • Copyright and embargoes

Fig. 18.1 Ten top tips for publishing and presenting success

18.3.1 Top Tip 1: Choose Your Platform Carefully

The selection of the most appropriate platform for your work (the right journal or the right conference) is arguably the most important decision that you will make in disseminating your work. A high quality article or presentation will be of little value if it is never read or heard by people who can potentially use it to develop knowledge or effect change in the workplace.

Matching your work to the correct journal or conference (and hence the correct audience) is essential; make the wrong choices and your work is likely to be rejected. So how do we make the right choices?

The first step is to identify your message, audience and reach. Firstly, consider what type of message you need to convey (e.g., audit, service evaluation, reflection, original research, case report, discussion piece) as this will help to narrow down the available options. Then consider who will be your target audience. This may be radiographers, but are your findings applicable to all radiographers, or only to therapeutic or diagnostic radiographers? Is the message better suited to a particular subject or specialism, such as breast cancer, nuclear medicine or ultrasound? Many of these 'sub-specialisms' have separate journals and conferences. However, do not confine yourself to your own discipline; your work may be better suited to a wider multi-professional, education or policy platform outside of radiography. The geographical reach of your message often depends on the type of study undertaken; service evaluations and department audits are often of local or regional interest (your department or wider organisation), whereas some research studies and multicentre audits may have both national and global interest.

Well-established national and international conferences related to radiography and radiology are held annually, but each year they may request abstracts with different conference themes: it is important to be sure that your work aligns to the themes. Submission is usually online, and there are strict deadlines for abstracts to be received. These submission deadlines are often well in advance of the conference. Before submitting your work, ensure that if successful you will be able to access funding for the relevant conference registration, accommodation and travel, as well as time away from the workplace. Abstracts are withdrawn if at least one author does not register for the conference.

Finding an appropriate journal for your work takes a little time and preparation and involves both qualitative and quantitative judgements. While your work may be suited to submission to a professional magazine or practitioner newsletter, the following discussion relates to selection and submission to peer reviewed journals.

Start by drawing up a short-list of suitable peer reviewed journals in your field of interest (e.g. therapeutic radiography) and visit their websites. Read their 'aims and scope' carefully as this will provide a clear idea of the subjects and types of articles they receive, and of their target readership. Journal websites highlight which journal access model (see Fig. 18.2) they comply with; increasingly, research funded by government agencies or research councils is expected to be published 'open access' which incurs a publication fee, but is freely available to all readers.

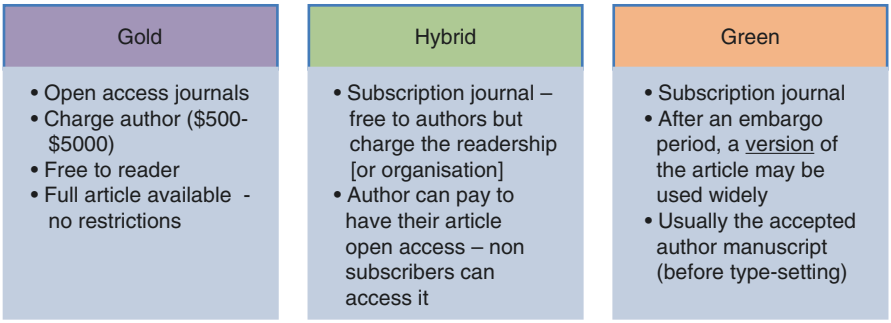


Fig. 18.2 Journal access models

Many subscription journals are free to authors, but some operate a hybrid system where an author can choose to pay a fee if they wish.

Review some recent issues to check the style of the articles, and most importantly ask experienced colleagues about their familiarity with the selected journal. This is important because there are increasing numbers of 'predatory' journals (and conferences) with a limited publishing and quality track record. These journals email novice authors regularly and persuasively with a promise of rapid review and publication. Lack of credible peer review and poor indexing means your work will be unlikely to be found by others; the publishing fee is often concealed at the time of submission.

Quantitative comparisons of journals can be made by comparing journal metrics. These are available on journal home pages and include information about journal turnaround times (e.g. speed of peer review), reach (readership and authorship) and citation metrics. Elsevier's CiteScore and Clarivate's Impact Factor are examples of citation metrics that provide an indication of how often articles in the journal are cited or referenced by other authors. Increasingly journals also provide information about altmetrics (alternative metrics) related to their articles, such as social media notifications or activity. A short-list of journals should be drawn up taking into account all of the above factors. However, remember that there is no 'best' journal in a particular field, only a 'best fit' for your research. Arguably the most important factor is whether the readership is the appropriate target audience for your work.

18.3.2 Top Tip 2: First Impressions

The most important elements of an article or conference presentation are the title, abstract and keywords. They are the first (and often only) part of your article which is read; they must therefore be designed to attract the interest of journal editors, conference reviewers and readers. Additionally, they are vital elements for choosing appropriate peer reviewers and ultimately for indexing an article or presentation; poor indexing may mean that readers struggle to find your article in literature searches.

A title should be succinct, interesting and ideally include reference to an article's subject, settings and methods. An abstract normally follows scientific sub-headings (see top tip 3) and usually has a 200–250 word restriction. It should not include any references. It should provide a brief justification for the study and some key findings. It should clearly align with the aims and scope of the selected journal or a conference theme. It is important to highlight anything that is new or unexpected in your findings, and/or the implications for practice, as this will peak the interest of conference reviewers and editors. Such an abstract is more likely to encourage a delegate to attend your conference presentation, or a reader to access the full text of your article. Four or five keywords should be selected to assist with indexing and peer review. They should ideally be different from those used in the title and should be selected to align with Medical Subject Headings (MeSH). MeSH is a controlled vocabulary of terms used for indexing and searching for content within the Medline® and PubMed® search databases.

18.3.3 Top Tip 3: Stick to the Rules

Some journals and conferences receive high numbers of submissions on semi-automated submission systems. It is essential to comply fully with submission deadlines and word count limits. Read the guide for authors thoroughly, ensure that all required elements are submitted (e.g. a manuscript without author names, and a separate 'title page' document with author details inserted).

Most common formatting mistakes include using an incorrect referencing style in the text and in the reference list or including more than a maximum stated number of tables, figures and images. These illustrations must be of publishable quality for either print publication or projection on a large screen in a Power Point presentation.

A journal's guide for authors outlines the required article or presentation format. While there are variations for some qualitative articles and review articles, most research is presented in a scientific article format (see Fig. 18.3).

One of the most challenging aspects of writing for publication and presenting at conferences is meeting the strict word counts. This is particularly difficult when a work to be published or presented originated with a different purpose in mind, for example, an undergraduate (~5000 words) or postgraduate dissertation (~15,000 words) or doctoral thesis (~50–100,000 words). It is difficult indeed to condense an academic work into an article, often between 2000 and 4000 words, or into a conference presentation of between 6 and 15 min.

One of the best approaches is to take a step back from your dissertation, and think about this 'elevator pitch' scenario. If someone 'important' at a conference asked you about your dissertation, but you only had 5 min because a lecture was about to start, what would you say?

Think of the highlights of your dissertation; you may decide that some aspects of what you did could be left out. The most important thing to keep in mind is what is your main message; remember that the message for your article or presentation may be different to the 'whole' dissertation.

Sub-heading	Comments
Title	Succinct, captures the essence of the article
Keywords	4-5 keywords, aligned to MeSH
Abstract	Within word count, presented in journal or conference format. No references
Introduction	Sets the scene and justifies the research. May include the literature review, or this could be another section. Finishes with the aims of the research and/or the research question or hypothesis.
Methods	Explicit and reproducible methods for how the research was conducted
Results	Key findings supported by analysis. For qualitative research, the results are sometimes combined with the discussion in a 'findings' section
Discussion	Discusses implications of the results in light of other literature Includes author-identified limitations of the study
Conclusions	May include recommendations and suggestions for further research
Acknowledgements	Not essential, but could include research funders, supervisors, gatekeepers or participants
References	All work cited within the manuscript. A bibliography of work read but not cited is not required.
Appendices	Supporting documents such as survey or interview questions.

Fig. 18.3 Scientific article sub-headings

Another strategy is to summarise your dissertation on a Power Point presentation (you may have already presented it at a conference). Try to summarise the information in bullet points on each slide. Use separate slides for the introduction, aims and objectives, methods, results, conclusion and recommendations. You must be really ruthless in deciding what goes in and what is left out. Once you have done this then you can expand each section to create your article (cut and paste some relevant sections from your dissertation).

18.3.4 Top Tip 4: Ethical Publishing

It is essential that all parties involved in the act of publishing agree upon standards of expected ethical behaviour: this includes authors, journal editors, peer reviewers, publishers and the society of society-owned or sponsored journals. Very few authors set out to compromise ethical standards, but there are a number of pitfalls that must be avoided when submitting work to journals or conferences.

The most crucial is that the work is an accurate account of the work performed and is an author's own research; any information that was sourced from other

authors should be properly referenced. This includes the reproduction of figures and tables from other work, where permissions may need to be sourced from the relevant publisher. Few authors knowingly cheat or copy, but poor academic practice in citing work is often highlighted by reviewers. Similar to universities, journals have sophisticated plagiarism software that flag areas of strong similarity between papers. However, a word of caution here about the notion of 'self-plagiarism'; this occurs when an author submits a similar paper to more than one journal (duplicate publication), or more commonly when several papers, all relating to the same study ('salami slicing'), are submitted. While there may be genuine reason for doing this, there inevitably is some duplication and potentially watering down of the results.

Co-authorship disputes in both article and conference abstract submissions are not uncommon; authors should agree inclusion and order at an early stage in a manuscript and abstract preparation. Authorship should be limited to those who made an important contribution to the conception, design, conduct or write-up of the research. For academic work being published or presented, this would normally include any academic supervisors who supported a student to develop the work. If there are multiple authors, the first and second author positions are usually assigned to those who contributed the most to a study; the final author position is sometimes reserved for a research lead who may have overseen the work. All co-authors should have seen and approved the final version of a conference abstract or manuscript and agreed to its submission for publication.

One of the checks that editors and reviewers make is what level of ethical approval was required for your study. Therefore be explicit about any approvals gained (list the institution and the ethics number). Pay particular attention to any approvals or explicit consent required when using patient images, for example, in a case study or poster presentation. Authors should disclose any financial (funding source) or other substantive conflict of interest that might be construed to influence the findings or interpretation of their results. Some conferences ask for presenters to include a 'disclosure' slide at the beginning of their presentation.

18.3.5 Top Tip 5: Proof Reading and Preparation

Editors and reviewers are busy people. Conference reviewers screen high numbers of abstracts in a short decision window. Too many errors in a manuscript or conference abstract (e.g. English, grammar, spelling, tenses) lead to it being rejected. The use of computer spelling and grammar check software is very helpful, but beware of medical terminology that is often not picked up by the software. Also check whether the journal or conference requires English or American spelling and be consistent in its use. If English is not an author's first language, then request a native speaker to proofread the work prior to submission.

One of the best checks for grammar is to read the manuscript out aloud. If you are gasping for breath then a comma or full stop is required. Reading out aloud is essential when preparing a presentation. You could make use of 'rehearse timings' software in Power Point. You are likely to be over the time limit initially; further editing

of slides and content helps. Inexperienced presenters are advised to use cue cards to write down what they will be saying for each slide. As you gain experience you can then use the information on each slide as your prompt, which is a more natural approach. However, be careful not to simply read verbatim from the slides; add additional detail or explanation to the visual displays. If you are presenting at an international conference be aware of potential language barriers in the audience; simple words are always better than complicated terminology and jargon.

18.3.6 Top Tip 6: Rigorous Referencing

Referencing the work of others is an essential component in journal publishing; it is also expected (to a lesser degree) within oral presentations.

The selected journal's guide for authors covers the referencing convention used; an article submitted with an incorrect referencing style will be returned to an author. Attention to detail is paramount. Other authors' work must be cited correctly both in the text and in the reference list. Reviewers should not have to proofread your work. Many authors use automated referencing software, but occasionally adding a reference can create problems so do double check your final manuscript.

Reviewers will be familiar with the literature in the field, and especially in their journal. Prior to submission undertake a keyword search of the journal and make sure you have included any key references from within the journal. Reviewers are likely to comment where they think referencing is insufficient (unsupported statements) or the use of old references without an explanation of their use.

Referencing within a presentation is always tricky. It is best to avoid using a numerical system as the citation details only appear on the last page and you may not have time to display it. Ideally display the author and date on the relevant slide, and possibly the full citation details in the footer of the slide. If you are referring to a document or well-known article, you could display a screen-shot of the document instead of a traditional reference. Remember that a presentation is best served by visual aids rather than lots of text.

18.3.7 Top Tip 7: Replicability

In research we often talk about validity and reliability. In other words was a study accurate and truthful in its reporting, and can it be replicated by other authors? A study can only be reproducible if sufficient information is given regarding its methodology (research approach) and methods (how it was done and analysed). Selected research tools (e.g. a questionnaire) need to be referenced, or if new tools are developed they should be included in appendices. All selected methods have advantages and limitations, so it is important to explain the issues around validity, reliability, and, if required, ethical approval.

The results section should show analysed results (descriptive or inferential statistics and charts, qualitative themes) rather than raw data. Ensure that you use the

minimum number of unambiguous, easily interpreted tables, figures and images (referenced in text), or quotations to illustrate themes emerging from qualitative work.

In the event of any queries, you should keep your raw data (interview transcripts, audit data) for a minimum period of time after publication. Check the prescribed storage requirements of your institution.

18.3.8 Top Tip 8: Credible Conclusion

The conclusion is arguably the most important part of your work. A busy researcher often reads the title and abstract, then proceeds to the conclusions, before deciding whether to read on. The audience most likely focuses on the conclusion slide of a presentation, as this is often the last slide which is displayed while questions are invited.

However, be careful what you claim. Conclusions should be what you have found, not what you would like to have found. They must be based on your results placed in context of other literature. Any recommendations should be noted in this section.

Often, following revision of a manuscript, either the abstract or the conclusion section may have been amended. Prior to resubmission, check that they both are aligned and say the same message.

18.3.9 Top Tip 9: Do not Take It Personally

Responding to questions and feedback is an inevitable part of publishing and presenting. Presentations are often followed by an opportunity for questions from the audience, and while this can be a little daunting, from my experience people do not set out to pose really difficult questions. Sometimes, however, particularly in international conferences, you may struggle to understand a question or a person's accent. Do not hesitate to ask them to re-phrase the question, or ask the chairperson for assistance as they are there to support you. While there may only be time for one or two questions, do stay in the lecture venue for a short time after your presentation, as some audience members prefer to approach you individually with their questions and comments, and indeed congratulations.

No article is expected to be perfect. The feedback of an editor or reviewer highlights any serious gaps prior to publication. Most of these issues can then be resolved or acknowledged as limitations of the work. When an author submits a manuscript, an editor decides whether to send it for peer review. Taking into account a peer reviewer's comments, an editor may then either reject the manuscript or ask the author to revise the work. If you receive a reject do not despair; use the comments to improve your work and then consider other publishing options. Rejections are a reality of research, but the feedback received usually results in a much stronger article. Following a revise decision, an author will have to re-work the manuscript and respond to reviewers' comments on a 'line by line' basis. You are not obliged to

implement everything that they suggest (and indeed some reviewers may disagree on what is required). You are, however, expected to justify your responses if you do not follow the approach suggested.

18.3.10 Top Tip 10: Self-Promotion

Once a conference abstract or article has been accepted, how do you spread the word about it? The specific journal or conference has a responsibility to publicise the journal issue or conference programme, but it is an author's responsibility to publicise their own paper or presentation and ensure that their message is heard widely.

Increasingly professional social media (particularly Twitter) are being used by authors to reach a wider audience beyond the confines of a conference or journal issue. Publishing and presenting your work is not the end point. It is the beginning of promoting research impact via effective publishing.

Once an article is accepted, a publisher usually holds the copyright. However, authors can share their respective article PDFs internally in their organisation and use their article in their work (e.g. to support a lecture). They can paste a link to the published article from their own website and can publish the final accepted manuscript (before any copy-editing and typesetting is performed). However, most publishers issue an electronic article link to authors for free distribution to their colleagues (for a limited time period).

Copyright and publishing embargoes mean that unless an open access fee has been paid, an author is not able to post the full copy-edited PDF version on social media or on personal/institutional websites (e.g. university repository or ResearchGate).

Once you have had an article published, you could set up a citation alert (easily done on most platforms) for your article. This automatically sends you an email if your article is cited by other authors, so that you can track the longer-term impact. For older articles that have attracted a number of citations, authors can review the citation data via several freely available citation analysis tools.

18.4 Conclusion

Publishing and presenting your work to a wider audience takes time and preparation. It is incredibly rewarding and is an expectation of radiographers in leadership roles such as advanced and consultant practitioners and research radiographers. If you are considering publishing or presenting some of your academic work, do not hesitate to seek support from academic supervisors or more experienced authors and presenters as they will be happy to guide you. Similarly look out for opportunities to engage in audits, service evaluation and research projects in your organisation. Make connections with more experienced practitioners and learn dissemination hints and tips from them. Similarly learn from others beyond your organisation.

When attending a conference as a delegate, or reading the latest issue of a relevant journal, consider not only what may be 'good' research, but also what contributes to effective dissemination. Taking learning points from a well-written article, and a carefully crafted presentation, helps you to develop your own dissemination style for the future.

Further Readings

Committee on Publication Ethics (COPE). https://publicationethics.org/guidance/Guidelines [Guidelines on ethical publishing for authors and publishers].

Elsevier. Publishing Campus / Researcher Academy. https://researcheracademy.elsevier.com/ [free e-learning modules for researchers and authors on publishing].

National Library of Medicine. Medical Subject Headings (MeSH) tutorial. https://www.nlm.nih.gov/bsd/disted/meshtutorial/introduction/ [tutorial explaining the use of MeSH].

Society and College of Radiographers. Research Strategy 2016-2021. Date published: 30 September, 2015. https://www.sor.org/learning/document-library [most recent UK professional body research strategy].

Applying for Research Funding

19

Karen Knapp and Fiona Mellor

19.1 Introduction

Research is an essential element of all professions and it is no different in radiography. Research facilitates the building of a profession's evidence base. As a young profession, with rapidly moving technological advances, this is especially important in radiography [1]. Improving the evidence base must always hold patients at the centre; the ultimate goal is to improve their care and outcomes. High impact studies in the field of radiography serve to raise the profile of the profession as well as improve patient diagnostics, therapeutics and experiences through a robust evidence-based practice [2].

Without exploring new ideas and development of novel treatments, diagnostic and therapeutic techniques and applications, we would not have seen the advances in practice today. In an era when artificial intelligence (AI) is providing exciting opportunities for enhancements in patient outcomes through improved diagnostics and therapeutics [3, 4], it is more important than ever for radiographers to be actively involved in developing the evidence base for their profession. For advanced and consultant practitioners, research is one of the four pillars of practice, demonstrating how research is integral to high-level practice [5, 6]. In the following sections of this chapter, the various tasks and activities that should take place to get a good proposal together are laid out.

K. Knapp (✉)
University of Exeter, Exeter, UK
e-mail: k.m.knapp@exeter.ac.uk

F. Mellor
AECC University College, Dorset, UK
e-mail: fmellor@aecc.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_19

19.2 From Conception to Grant Proposal

The cornerstone of research is having a novel idea or investigating an established area from a different perspective. Writing a grant proposal is an important part of research. Obtaining funding enables research to happen. Research funding is competitive and consists of grants, which are awarded in response to investigator-initiated projects, or contracts under which a research topic is proposed by a funding agency [7]. It may be necessary to submit a proposal to more than one agency in order to obtain sufficient funding to operate and support the required research infrastructure. Industrial partnerships are likely to play an increased role in funding or part-funding research in the future. There is a lot of competition for research grants; to avoid rejection of an application, one cannot afford to make mistakes in a grant application. Many worthwhile projects will, however, be rejected because only 10–20% of all grant proposals are generally accepted, though some research councils do report slightly higher success rates [8].

These guidelines apply for quantitative, qualitative and mixed methods proposals. Applicants should choose topics that are of interest to them in order to maintain momentum and some of the best research relates to clinical research questions which frequently arise from practice. Radiographers should be undertaking research directly aligned to their areas of practice or speciality since this is likely to lead to the most beneficial research questions. It is also important to write a grant proposal that captures the uniqueness of your organisation and shows the correct fit between your organisation and the funding body [9]. Sufficient time should be allowed to complete and application to minimise errors and optimise the quality of a grant proposal. Typically, it takes up to 12 months from the generation of an initial idea to the actual submission of a grant proposal.

It is essential to document sufficient information to convince funders that the proposal is worth funding [7]. Bear in mind that grant writing should not be a lonely pursuit and all members of the research team should be involved; there are many sources of help, for example, books, videos, colleagues, consultants and the World Wide Web [10, 11]. However, for first grant proposals, it is always advisable to have a more experienced researcher as part of a research team to act as a mentor [12]. The checklist in Table 19.1 can be used to assist with writing a funding application. It usually takes 6 months to be informed of the outcome of a submitted application. Although it can absorb a lot of an applicant's time, overall the development of a grant proposal should be enjoyable. It should also be seen as an opportunity for researchers to crystallise an idea and to critically appraise their research plans. This is an essential exercise, since it is likely to enhance the quality of a study. The scope of the proposed work can be evaluated and altered, and aspects including methodology and analysis can be thought through critically. At this stage potential follow-up studies should become apparent, which should effectively create continuity.

A comprehensive literature review should be done to test a hypothesis in terms of current available data and knowledge in the area of interest. An additional simple yet effective method is to share your idea with colleagues and encourage constructive feedback. Typical questions to establish the potential impact of your research: What am I trying to test/explain? What are the possible causes? What causes will I explore/ what are the possible mechanisms? and 'so what'? [11].

Table 19.1 Checklist for grant applications

Section of application	Item to check
Eligibility	Check if you fit the requirements for eligibility Ensure the proposed study fits in with the funding body's priorities—especially when applying for a themed call for proposals Explain why the planned work is novel and necessary
Hypothesis and objective	Check if the research has not already been done before; the need for the research must be justified
	Clearly define the hypothesis
	Place the proposed study in the context of the current knowledge and evidence-base on this topic
Public and patient involvement (PPI)	Patients, carers or the public should be consulted and involved in developing your research question and methods. Patient and public involvement in the research monitoring processes and dissemination plans are also recommended
Methodology	Explanation of the procedures, outcome measures and testing reliability and validity
Data analysis	Include an appropriate data analysis plan
Timeline	Ensure you have developed an appropriate timeline using a Gantt chart or other method
Finances	Justify the amount you are requesting Carefully calculate the total amount requested; double check all aspects of this, from number of hours for wages to prices for equipment Include the finance department from your institution. Consider the costs of the dissemination and publication strategy, and any PPI
Communication	The research team will need regular meetings for project monitoring and a dissemination plan is required
Finishing touches	Review and check for grammar and spelling errors Adhere to the guidelines: do not exceed the maximum number of words allowed Ask someone to proofread your application, particularly if some sections have to be intelligible for lay people. The public and patient group can often assist with this

This approach should be applied throughout an application process. At the literature review stage it should become clear whether an idea is simply a matter of building on current knowledge or whether it goes against the grain of what others think. If there is a systematic review, which has already provided the answer to the research area of interest, then moving onto recommendations from the review can be a good way to extend the work. The international prospective register of systematic reviews (PROSPERO) is a useful place to check whether a systematic review is currently underway on the proposed research area [13]. It is also worth consulting the Cochrane Library for completed systematic reviews and trials [14]. It is fair to say that a research grant of an evolutionary nature stands a better chance than a plan involving a revolutionary hypothesis. However, a research proposal that promises very little in terms of added value will probably fail to impress. Advice on developing a research question and literature searching is given in Chaps. 2 and 3.

Looking at current 'hot topics' in the field of radiography or researching themed calls for proposals will give an indication of what experts in the field think the priorities for research are. The Society and College of Radiographers (SCoR) in the UK recently published a list of research priority areas, [15] which are updated periodically. It is therefore recommended that the list should be checked on the SCoR website because if a proposed project is tailored towards addressing these areas, it will further enhance the chances of success. Moreover, it also means a grant proposal will better fit the eligibility criteria.

19.3 Added Value of Preliminary Work and Pilot Data

The majority of, if not all, organisations, which fund research, apply peer review to select the best grant applications. It is easier to convince reviewers of the merits of a proposed project if there is already some promising data accompanying an application. Obtaining preliminary or pilot data serves two purposes. Firstly, it would show that the hypothesis to be tested may be correct, or that the aims set can be met. Secondly, producing pilot data highlights to reviewers the active ability of the applicant to undertake the research proposed and indicates the capacity to handle larger, full-scale projects. An important task is to explain the implications of the preliminary data in terms of the aims of the proposed full project, and to cite past work by peers in the field. One important thing to remember is that peer reviewers may have produced the data that your proposal is based on, so accurate citation and representation of the data are essential.

Demonstrating the viability of a research project through pilot data or through statistics to establish the ability to recruit participants or the acceptability of a protocol to the clinical environment and participants is essential. Sometimes it is impossible to gain these data without funding for this stage of the research. Funding does exist specifically for small-scale pilot and feasibility studies, often called 'pump-priming' grants (for instance the Radiological Research Trust http://radiologicalresearchtrust.org/) or the College of Radiographers Industry Partnership Scheme (CORIPS) https://www.sor.org/about-us/awards/corips-research-grants. Many funders accept that feasibility and pilot studies are an essential part of working up larger trials; without investment in this crucial first part of research, a larger trial would be too high risk to fund [16]. The complex interventions framework can provide a useful aid when considering where the proposed research sits in relation to the development and implementation of research. While the complex interventions framework was developed for mental health research it is also useful for wider healthcare research [17].

19.4 Membership of the Research Team

To further increase the chance of being successful, it is important to have correct collaborators on board. This certainly applies to less experienced researchers who wish to apply for funding. For professionals who want to become involved in

research, or who want to start writing grants, it is best to join forces with researchers who have a track record in their chosen field. Having a senior co-author on an application increases the faith reviewers have in a grant proposal. Undoubtedly, a senior researcher should be able to give invaluable advice on how to write and develop a proposal. The temptation is for junior researchers to try and do too much within the proposed time. A seasoned researcher is able to evaluate whether the planned work will fit in an allocated period, thereby avoiding a regular criticism by reviewers of a project being too ambitious [18]. Another sensible option is to link up with an established team of radiography researchers in a hospital, university or institute that provides good infrastructure and support.

A research team that only includes radiographers may not be the most efficient group of members. Increasingly there is a move to interdisciplinary research to address the potential of those from different backgrounds working together to enable creative solutions to the complexity of research in healthcare [19]. The merging of expertise, from different disciplines working together on a radiography problem, can create an environment that can provide impactful outcomes using techniques that may not have previously been utilised within the radiography field [20, 21]. Writing a radiography grant proposal may be aided by consultation with a radiologist/oncologist and/or physicist, depending on the precise nature of the bid. Co-investigators who complement your own background and training should be chosen [7]. If you require significant input from another person, it should be considered whether they should be included as a collaborator on the bid. A research team also benefits by the addition of the necessary methodologists to ensure all aspects of a proposed study are rigorously designed and completed. Methodologists such as statisticians, qualitative researchers, operational researchers and health economists can all enhance the research team if you do not have the expertise in these areas yourself. Radiography research is best undertaken in teams. Building the right team is essential.

19.5 Which Grant and Where to Find It?

Once an idea has been created, the aims and outcome measures have been formalised, and a team of people has been assembled, the next stage is to identify a source of funding. The types of grants, which are available, and the different funding bodies, are discussed below.

19.5.1 Grant Types

Grants come in different shapes and sizes, just like projects. There are different types of grants: research grants (money for one specific project), programme grants (large collaborative efforts that encompass a number of projects), studentships (to fund a research-based MSc or PhD degree) and fellowships (to fund career development). Pump-priming grants may also be available as previously mentioned. Studentships and fellowships are awards made to a specific person.

Naturally, variations on each type of grant exist. For example, certain grants promote collaboration between industry and the public sector, and certain fellowships are intended specifically for certain professions.

The type of grant for which one should apply depends entirely on the project and the applicant's circumstances such as employment status, qualifications and research track record. Many funders have guidelines on what can be claimed for in terms of salary and study costs. Increasingly funders are looking for industrial partners to be included on grant applications, so it is important to consider how to make these links.

For those wishing to establish themselves in research, career development grants are a good option. These grants are allocated to applicants who can demonstrate that they have the potential to become successful independent researchers; a track record is not an essential prerequisite. Apart from the need for a sound project proposal, other requirements have to be met to satisfy the reviewers. For the grant proposal, it must be evident that a candidate has a strong desire and commitment to work in research long term. The organisation that a candidate works for also has to be committed to support and develop this person. It is therefore essential that an infrastructure exists to provide that support, both through the presence of a mentor and adequate research facilities [22]. Even if a candidate shows promise and a research plan is of a high quality, an employer has to match this level of potential with sound back-up support.

19.5.2 Small Grants

Small grants are designed to support early stage projects with the main purpose of collecting preliminary data to underpin a large project. A small pilot grant is mainly designed to collect the feasibility or pilot data and should not be used for larger studies as this will leave many costs not covered. Pilot or feasibility studies may provide proof of concept or a small study to collect sufficient data to underpin a power calculation for a larger study [7, 23]. The level of innovation and the originality tend to be high in these grants as there would not be data published to underpin the larger study already. It is therefore important to highlight this in the application. These grants should be highly focused, and have a limited scope. They are usually time limited to a short period of time, which must be reflected in the work proposed [7]. Many NHS Trusts have small grant funding available, as do many universities. You should therefore explore opportunities within your organisation as well as external funders.

19.5.3 Funding Bodies

There are several bodies that currently fund research for radiography and radiology. Although other countries also have charities and governmental research councils, for conciseness this chapter concentrates on the UK and USA. In the UK, the SCoR offers research awards for smaller projects and doctoral fellowships through the College of Radiographers Industry Partnership Scheme (CORIPS) (see www.sor.org). This may be a good initial funding body for those researchers just starting out,

since the funding is available specifically for professionals in the fields of radiography and radiotherapy. Larger grants are available from the National Institute for Health Research (NIHR), and from the more generic research councils, such as the Medical Research Council (MRC), Engineering and Physical Sciences Research Council (EPSRC) and the Wellcome Trust.

Depending on the topic of a project, specific organisations may be approached. For example, if a project looks into aspects of mammography then a charity such as Cancer Research UK may be an appropriate funding body. Additionally if a project is a pilot or feasibility study then a smaller charity such as the Radiological Research Trust, which funds pilot projects using new radiographic technology, may be approached. The Centre of Sport, Exercise and Osteoarthritis research, which funds researchers across all disciplines (http://www.sportsarthritisresearchuk.org/fundingopportunities/pump-priming-grants.aspx) may be appropriate. Other organisations that award pump prime grants include the Royal College of Radiologists (https://www.rcr.ac.uk/clinical-radiology/awards-and-prizes/pump-priming-grants), but the principal investigator must be a radiologist, thus the research team would need to identify the suitable P.I. based on their profession. It is worth noting that pump prime grants rarely cover the full costs of research, such as an investigator's salary, but they can be a useful step up to a full proposal, indicating that the research has already been peer reviewed and thought of as valuable. Individual universities may also have internal pump prime grants so it is always worth consulting with academic colleagues. In the USA, the Radiological Society of North America (www.rsna.org) provides support for professionals in radiography and radiotherapy. The main financial backer of research in the USA is the National Institute of Health (NIH); they have a plethora of different grants available.

For all funding bodies, the types of grants and themed calls for proposals are subject to change so check the websites for grant news on a regular basis. The SCoR also holds a list of potential funders on its web pages in the UK (see https://www.sor.org/career-progression/researchers/finding-funding).

Once a decision has been made to apply for funding from a certain funding body it is useful to get in touch with the selected organisation. People are employed by funding bodies to provide guidance and information to applicants. Always make sure to check your own eligibility or the appropriateness of your project with the programme or organisation's administrator. There is nothing worse than going through the whole application process and finding out that your application cannot be considered. Different funding bodies use different software and formats for applications: an application therefore can often not simply be cut and pasted to fit another call for proposals.

19.6 Writing a Good Grant Application

A good grant application uses most of the elements discussed earlier in this book. It requires an understanding of what research is (see Chap. 1), how to formulate a research question (see Chap. 2), extensive literature evaluation (see Chap. 4),

selecting the appropriate outcome measure for the results (see Chap. 11), and an appreciation of the ethical implications of the proposed research (see Chap. 6). In addition, a research project should apply the most fitting methodology available. Data should be recorded and analysed by the selected outcome measure and by applying an appropriate statistical test if it is quantitative data. Qualitative data on the other hand are usually analysed via themed analysis or quasi-statistics. These areas are covered in Chaps. 15 and 16.

Writing a grant application is time-consuming. Below are the main tasks and activities involved summarised from start to submission.

- Conceptualise the project and review the literature.
- Identify potential collaborators or mentors.
- Approach patients or the public about their involvement.
- Start work on any pilot or feasibility studies (ensuring compliance with governance and ethics).
- Further define the research question and methodology/analysis tools.
- Decide on the type of grant and funding body. Contact the programme officer and review application forms and instructions.
- Outline and draft the proposal.
- Consult the finance department, and Research and Development (R&D) office of your institution, if you have one.
- Consult experts in statistics or other disciplines.
- Develop and finalise the budget.
- Determine the dissemination strategy, identify suitable journals and conferences.
- Review and draft the second version of your proposal.
- Critical appraisal by research team members, collaborators and lay persons.
- Final revisions and submission of grant application.

A full research proposal is typically several pages long. Strategically restating the key questions in the project enables reviewers to maintain focus on the proposed research. For a reviewer, who reviews numerous proposals in the space of days and who is not necessarily a specialist in your field, it is helpful to be gently reminded by an applicant how the proposal fulfils the key points in the proposed research. Likewise, each section of an application may benefit from a summary containing the key points. When constructing a grant application candidates have to sell themselves to the reviewers and funding body. For applicants there is no harm in highlighting the fact that they have worked on a similar project in the past or attended a specialist course relevant to the proposed research project. Reviewers most probably would not know anything about an applicant's background. They therefore need to be informed about any skills that make an applicant qualified to deliver on the promised work, such as specialist training. An applicant's research track record will also be taken into account, especially when applying for larger

grants, so it is important to record previous grants and their publications on one's curriculum vitae. Some funders do ask for previous publications relevant to the proposal. it is important to include those most relevant to the research in question. Also is can be useful to provide evidence that a research team has previously worked together.

Although it sounds obvious, the key is to adhere to the guidelines for submission. If an abstract is limited to 200 words do not exceed this word limit as your application may be rejected. Stick to the instructions in a grant application, from budget to bibliography. It is not compulsory to write to the nearest maximum word count; a grant application needs to be concise, clear and complete [24]. Any grant proposal must communicate the study clearly. It must be easy to read, concise and well-structured. Above all, it must follow the required guidelines from the relevant funding body [7].

More and more funding bodies require input from patients or lay people in the design and dissemination of a research project. User involvement during the development of a research bid can have a positive effect on the design of a study [25]. Involving PPI in the development of participant information sheets and questionnaires can highlight the use of too much jargon, acronyms and/or abbreviations. Similarly, discussions with patients may shed light on important ethical issues, such as how and when to approach patients for participation in a study. It can also help identify outcome measures that are important to the participants (patient reported outcome measures PROMs), which can increase the usefulness and impact of a research project.

Patient and public involvement (PPI) must be costed in including travel expenses and payments for time. There are recommendations for the funding of PPI in the UK and these can be found on the INVOLVE website (https://www.involve.org.uk/). Undertaking sufficient PPI prior to a funding application should always be considered. The Research and Development Service (RDS) at many NHS Trusts in the UK can assist with this. Public and patient involvement does not require ethical approval since they are working with the researchers to influence and help develop the research. However, it is important to ensure that PPI does not cross the line into research; the input of the PPI members cannot be analysed as research data [25].

It is imperative to decide the appropriate sample size for a study. For analysis of generated data, it is also important to choose the correct type of analysis to fit in with the methodology (quantitative, qualitative or mixed methods) [26]. It is currently obligatory to obtain authorisation from a statistician or methodologist when applying for ethical approval, such as the importance of sound appraisal of a project. An in-depth explanation of the application of analysis in radiography studies is beyond the scope of this chapter. It is covered in Chap. 15. If a proposed project observes a reduction in the length or number of treatments, or any other change that impacts on the costs and potentially impacts on quality of life, there may be also a requirement for the input of a health economist.

19.7 Financial Considerations

Undertaking independent research with funding from a charity or research council means getting to grips with calculating the costs of a project. Depending on the size and length of a project this part of a grant proposal can vary from being manageable to requiring a specialist. When writing a first grant application it is vital to ask someone with previous experience to look at the projected finances. There are different costs incurred when carrying out a project and these costs need to be categorised. Directly incurred costs are the salaries for the people stated on the grant application, i.e. those who will be working specifically on the project and therefore will have to be paid directly from the grant. Consumables purchased for the project also fall under this category. Directly allocated costs are those inherent to conducting research in a department or an institute. These include charges for radiographic staff and overheads for research infrastructure such as the use of imaging or therapeutic equipment time. Finally, there are indirect costs: costs for general back-up support staff such as library, human resources and finance staff.

There are certain things to bear in mind when preparing a budget for a grant proposal. Wages are calculated with inclusion of the on-costs an employer incurs: pension and national insurance costs should be added to the gross wages. If a project runs for more than one year then salary increases, based on inflation or annual increments, need to be included. For equipment and consumable costs it is important to find out whether these figures should include value added tax (VAT). Ask for quotes from different companies when an expensive item is listed to get an idea of the costs involved. It has to be noted that some grants do not allow asking for capital expenditure like equipment or machines. Finally, other costs associated with a research project must be considered. There are travel and registration costs associated with presenting research outcomes at a conference, and open access costs for publishing a manuscript in a peer-reviewed journal. As with equipment, it is best to identify beforehand what conference the proposed project will most likely be presented at; in this way, costs can be calculated more precisely. Sharing the outcomes of a project at conferences and in articles helps to gain a reputation based on the work carried out. Chapter 18 covers the processes involved in optimising dissemination of data.

19.8 Intellectual Property

Increasingly radiographers are undertaking research that yields intellectual property (IP); this may be due to involvement in potential product design or know-how [27, 28]. Some funding bodies require a section on IP to be completed. It is important that radiographers understand the know-how that they bring to projects, particularly if this involves innovations. It is important to have support from IP experts in developing these sections for funding applications and in drafting collaboration agreements for IP ownership prior to commencing a study. Many universities have IP experts to assist staff with this. The NHS research and development departments should also be able to assist. The IP agreements should also contain the potential

split of any future profits to be shared between the institutions involved from the product or tool. It is important to discuss IP prior to submitting a funding application; for novel innovations there are strict guidelines on what can be shared prior to applying for a patent [29].

19.9 Project Management

Developing a timeline and project management plan is also necessary for some research applications and good practice to ensure the research is kept on track. Careful consideration is needed for the time taken to complete the necessary approvals, such as ethics and R&D approvals; under-estimating this can eat into the time for the completion of a study. Breaking down the project into months, with targets for recruitment or milestones to be completed within each month, is very useful. This helps to plan a project to ensure it will be achievable within the timeframe and funding envelope applied for. It is important to consider milestones for your project; include these in project monitoring meetings to ensure you are keeping on track. Some funders also want stop/go decision points to be built into your grant and project monitoring. These are particularly important in efficacy studies; a study may be stopped early due to safety concerns or due to benefits demonstrated in interim analysis [30].

A Gantt chart can provide a useful tool for project planning. A Gantt chart for a small pump-priming study is shown in Fig. 19.1. It is important to consider all

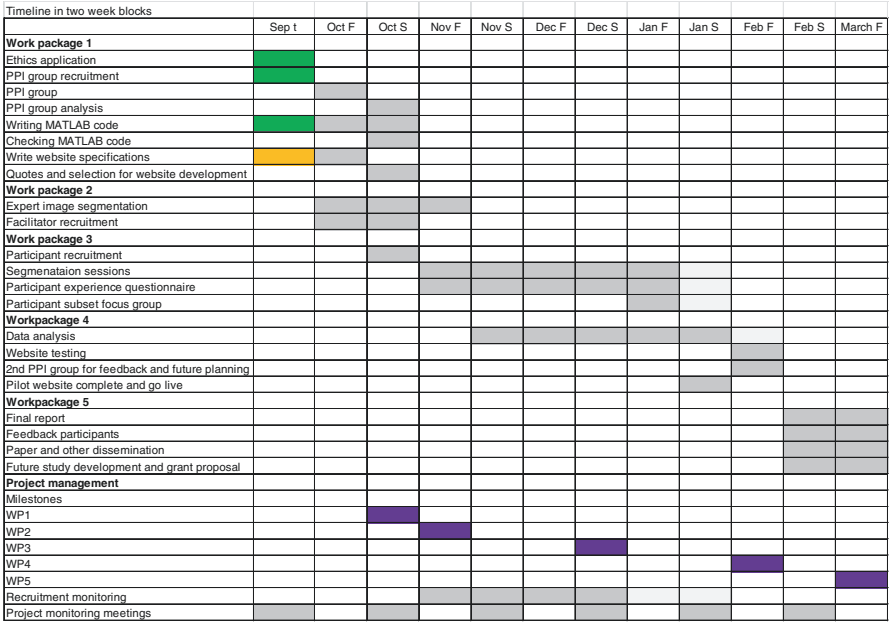


Fig. 19.1 An example of a Gantt chart for a 6-month pump-priming type study taken from the CITSEG study (https://www.exeter.ac.uk/tree/ourresearch/seedcornfunding/autumn2016projects/)

aspects of the study in your Gantt chart and to break the timeline into appropriate blocks. For a short project, the blocks are likely to be smaller than for a longer project. Including project management in the Gantt chart is a useful addition. This demonstrates to a funder that you have considered this as an integral to the study.

19.10 Common Shortcomings of Grant Applications

Grant applications may be rejected for various reasons. An application may be turned down because the candidate does not have a strong enough track record, or the university, hospital or organisation, where the applicant works, is not renowned for hosting research. If this is the case, it will be fairly tricky to turn this around in the space of a few months and resubmit an application.

Content-related errors are fairly common and these could lead to a reviewer deciding to turn down a request for funding. In the past, funding bodies have been approached about why some grant applications were funded and others not [18, 31]. A common omission in small grants for pump-priming type studies is the omission of how a study will lead to further work. Other common errors include an unclear study design, inadequate statistical power, issues with originality, hypothesis problems or a lack of aim and how the proposed study will create impact.

If your application is rejected, it is important to address any issues and not to try your luck with another funding body. Be willing to share your grant rejection with others. Even professors have grants rejected; this is a normal part of research. The peer review process can be fickle; sometimes minor tweaks can mean the difference between funding and rejection [32]. This is especially so when dealing with niche subjects including radiography and radiology where there is a significant chance that the same reviewer may be approached. If your perseverance and efforts pay off, this should certainly be celebrated. Obtaining a research grant is a prestigious feat and together with a positive outcome of the actual project work should lead to more successful grant applications and worthwhile collaborations.

19.11 Chapter Points

- Obtaining funds for research through a grant application is an essential activity to support research which will extend the evidence base of radiography practice.
- A successful funding application often takes up to a year to fully develop an idea and project plan.
- Apart from getting the science right, writing a grant proposal is a meticulous process.
- Patient and public involvement and engagement are an essential part of developing a funding application.
- Ensure the research team applying for the grant has the expertise to deliver on the project. Include appropriate methodologists such as statisticians, health economists or qualitative researchers where appropriate.

- Ensure that the grant is feasible within the time period and that there is sufficient staff time and resources costed within it to deliver the project on time.
- Ensure you check all the items detailed in Table 19.1.
- The current success rate of grant applications is 10–20%. It is far more likely that a grant application is rejected than accepted. If this happens, it is important to learn from the feedback.
- Most of the reasons for rejection can be addressed and worked on. If there is criticism regarding the statistical aspects of a study, this can be reviewed. If reviewers believe that an applicant does not have a strong enough track record, more senior peers can be contacted to propose collaborating and mentoring the applicant.
- Do not give up after the first rejection of a funding application. Address the feedback and try again.

19.12 Conclusion

Applying for research funding is an essential part of a research process since the majority of studies require financial support. Sufficient time needs to be put aside to work on a funding application. The right team of collaborators is essential. There is fierce competition for research funding, thus the majority of applications for grants are rejected. It is important to address the feedback from reviewers and try again, either with the same funding body if this is allowed or elsewhere. Once funding is successfully achieved, it is important to deliver the project within the specified timeframe and financial envelope, ensuring that reports are submitted and the project outcomes disseminated.

References

1. Snaith BA. An evaluation of author productivity in international radiography journals 2004-2011. J Med Radiat Sci. 2013;60(3):93–9.
2. Probst H, Gallagher HL, Harris R. Research and the radiography profession: a strategy for research 2010–2015. Radiography. 2011;17(4):268–9.
3. Al Arif SMR, Gundry M, Knapp K, Slabaugh G, Global localization and orientation of the cervical spine in X-ray images. International workshop on computational methods and clinical applications for spine imaging; 2016: Springer, Berlin.
4. Gundry M, Knapp K, Meertens R, Meakin JR. Computer-aided detection in musculoskeletal projection radiography: a systematic review. Radiography. 2018;24(2):165–74.
5. Milner RC, Snaith B. Are reporting radiographers fulfilling the role of advanced practitioner? Radiography. 2017;23(1):48–54.
6. Snaith B, Milner RC, Harris MA. Beyond image interpretation: Capturing the impact of radiographer advanced practice through activity diaries. Radiography. 2016;22(4):e233–e8.
7. Gholipour A, Lee EY, Warfield SK. The anatomy and art of writing a successful grant application: a practical step-by-step approach. Pediatr Radiol. 2014;44(12):1512–7.
8. Council MR. Funded research success rates: MRXC. 2019. Available from: https://mrc.ukri.org/research/funded-research/success-rates/.
9. Sayer B. Writing organization and funder profiles for a grant proposal. Nurse Author Ed. 1999;9(2):7–9.

10. Reif-Lehrer L. Applying for grant funds: there's help around the corner. Trends Cell Biol. 2000;10(11):500–4.
11. Gambling T, Brown P, Hogg P. This is not the end, nor is it the beginning-but it is the end of the beginning-getting to grips with the research process. Radiography. 2003;9(2):161–7.
12. Steiner JF. Promoting mentorship in translational research: should we hope for Athena or train mentor? Acad Med. 2014;89(5):702.
13. PROSPERO. International prospective register of systematic reviews. National Institute for Health Research, 2019. Available from: https://www.crd.york.ac.uk/prospero/.
14. Cochrane. Cochrane Library. 2019. Available from: https://www.cochranelibrary.com/.
15. Radiographers SaCo. The College of Radiographers Research Priorities for the Radiographic Profession: A Delphi Consensus Study. Available from: https://www.sor.org/system/files/article/201702/research_priorities_170117.pdf.
16. Lancaster GA. Pilot and feasibility studies come of age!: BioMed Central. Pilot Feasibility Stud. 2015;1:1.
17. Craig P, Dieppe P, Macintyre S, Michie S, Nazareth I, Petticrew M. Developing and evaluating complex interventions: the new Medical Research Council guidance. BMJ. 2008;a1655:337.
18. Inouye SK, Fiellin DA. An evidence-based guide to writing grant proposals for clinical research. Ann Intern Med. 2005;142(4):274.
19. Young HM, Siegel EO, McCormick WC, Fulmer T, Harootyan LK, Dorr DA. Interdisciplinary collaboration in geriatrics: advancing health for older adults. Nurs Outlook. 2011;59(4):243–50.
20. Rachuba S, Knapp K, Ashton L, Pitt M. Streamlining pathways for minor injuries in emergency departments through radiographer-led discharge. Oper Res Health Care. 2018;19:44–56.
21. Knapp KM, Welsman JR, Hopkins SJ, Fogelman I, Blake GM. Obesity increases precision errors in dual-energy x-ray absorptiometry measurements. J Clin Densitom. 2012;15(3):315–9.
22. Gill TM, McDermott MM, Ibrahim SA, Petersen LA, Doebbeling BN. Getting funded. J Gen Intern Med. 2004;19(5):472–8.
23. Arthurs OJ. Think it through first: questions to consider in writing a successful grant application. Pediatr Radiol. 2014;44(12):1507–11.
24. Bourne PE, Korngreen A. Ten simple rules for reviewers. Public Libr Sci. 2006;2(9):e110.
25. Liabo K, Boddy K, Burchmore H, Cockcroft E, Britten N. Clarifying the roles of patients in research. BMJ. 2018;k1463:361.
26. Ettarh R. Common descriptive and analytical statistics in investigative studies. Radiography. 2004;10(4):299–302.
27. Arora A. Licensing tacit knowledge: intellectual property rights and the market for know-how. Econ Innov New Technol. 1995;4(1):41–60.
28. Burk DL. Intellectual property in the context of e-science. J Comput-Mediat Commun. 2007;12(2):600–17.
29. Sterzi V. Patent quality and ownership: an analysis of UK faculty patenting. Res Policy. 2013;42(2):564–76.
30. Goodman SN. Systematic reviews are not biased by results from trials stopped early for benefit. J Clin Epidemiol. 2008;61(1):95.
31. Agarwal R, Chertow GM, Mehta RL. Strategies for successful patient oriented research: why did I (not) get funded? Clin J Am Soc Nephrol. 2006;1(2):340–3.
32. Graves N, Barnett AG, Clarke P. Funding grant proposals for scientific research: retrospective analysis of scores by members of grant review panel. BMJ. 2011;d4797:343.

Research Writing: Tips and Common Errors

20

Leonie Munro and Aarthi Ramlaul

20.1 Introduction

The most important part of research is to select a topic that you are interested in because it should hold your interest from start to finish [1]. There are several issues that need to be borne in mind when one starts to work on any form of research writing; whether it is a proposal, reporting on what was done, writing a paper for submission to a journal, or an abstract of a paper for a congress. Each of these needs careful planning so that unambiguous sentences address the aim, objectives, and relevance of a research study. Tips on how to avoid common errors in research writing are presented in this chapter.

20.2 Reference Method

It is important that you know which reference method to use. If you write a paper to submit to a journal, make sure that you use the correct reference method. Carefully read the instructions to authors including the reference method that must be used: Vancouver, Chicago, Harvard, or American Psychological Association (APA). Take time to check that each cited author or publication is listed in your references.

In this chapter the Vancouver reference style (author-number system) is used for in-text citation of actual publications because it is used by the publisher: Springer International Publishing AG. The author-number system is used in most medical

L. Munro (✉)
Formerly School of Radiography, King Edward VIII Hospital, Durban, South Africa

A. Ramlaul
Diagnostic Radiography and Imaging, School of Health and Social Work, University of Hertfordshire, Hatfield, Hertfordshire, UK
e-mail: a.ramlaul@herts.ac.uk

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5_20

journals/publications; in-text citations are numbered. On the other hand, most universities require students to use APA or the Harvard method when writing up research findings. In view of this the Harvard method is used for fictitious examples of in-text citations in this chapter. The references only list the names of actual cited authors or publications.

20.3 Plagiarism

Plagiarism means to copy the ideas, words, or writing of someone else without citation [2]. It has legal and ethical consequences, thus many universities, peer reviewers, and journal editors use plagiarism detection software. You must cite all your sources, including your own work, to avoid the risk of plagiarism. Each university and journal has its own referencing method, which you must use to accurately present your references.

It is a good practice to paraphrase. In other words use your own words when you discuss published research. Be careful of using verbatim text as if it is your own words. If you use verbatim text you must cite the source of the text; if you use two or more consecutive words you must place the verbatim words in double inverted commas and cite your source. The findings were that "the majority of healthcare professionals in Newcastle" do not know how to keep accurate patient records (Smith 2016, p. 25). Note that this is an example of how to apply the referencing technique when using a verbatim quote.

Many countries have strict copyright legislation for use of information including information on the internet. If you use information on the web, it too must be acknowledged. For example, if you copy a photograph of a MRI scanner from the web, you may be breaching copyright if you do not acknowledge your source.

20.4 Be Consistent

Make sure you use the same spelling for words. Do not use a mix of American and English spelling. Also when you report on your study the same wording for the aims, objectives, research questions, hypothesis, and significance of the study must be used.

Discuss literature chronologically. For example, discuss earlier studies to lead up to current ones. This will allow readers to follow the thread of your argument. Your discussion must be logical. When you summarize refer to key points in the cited literature in terms of the aim and objectives of your study. Keep to the point under discussion. Refrain from using words that may show subjectivity. For example, I was thrilled to discover that the participants' comments supported my point of view in the study. This should be reworded. The participants' comments were in keeping with the aim of the study.

20.5 Punctuation

English is a dynamic language. Over the centuries punctuation rules have changed. In many instances there is an ongoing debate in the correct use of commas [3]. Some universities provide punctuation guidelines. There are many good examples on the internet of the correct use of punctuation in academic writing [4–7]. It is beyond the scope of this chapter to discuss all types of punctuation. Note that an exclamation mark (exclamation point) is not recommended in academic writing [8]. The use of an apostrophe is discussed below as it often is not used correctly in research writing.

20.5.1 Use of Apostrophe

A common error in writing is the incorrect use of an apostrophe [3]. An apostrophe has two main functions: to denote possession (e.g., patients' records); to indicate a contraction when an apostrophe indicates missing letters (e.g., let's, can't). The use of an apostrophe before the 's' is to show possession of a single noun (e.g., a patient's records). When a word is possessive and plural, then the apostrophe is placed after the 's' (e.g., two weeks' records were accessed). An apostrophe is placed after the 's' of a noun that ends with an 's' to denote possession (e.g., Jones' study). Do not use an apostrophe to show a plural. The use of an apostrophe in the 1980's is incorrect. It would be correct in the following sentence. What was the 1980's dress code for radiographers? Consider this sentence. We are interested in researching radiographers' dress code in the 1980s. Can you see the different meaning in this example?

20.5.2 Capitalization

Overuse of capitals (upper case) is very common in writing [9]. A capital letter is used at the beginning of a sentence (e.g., The study was conducted in 2018). Names of institutions have capital letters (e.g., The study was conducted in 2018 at the Smith Hospital). A capital letter is used for names of people (e.g., Jones and Brown conducted a study in 2018 at the Smith Hospital).

It is not acceptable to use upper case for the first letter of a word for emphasis in a sentence. Jones and Brown (2018) reported that Radiographers participated in the study. This is incorrect and should be: Jones and Brown (2018) reported that radiographers participated in the study. Do not use capitals for emphasis. In the following sentence the use of capitals for emphasis is a definite no-no in academic writing. A researcher must obtain INFORMED CONSENT from participants. Informed consent in the sentence should be in lower case. In addition, a capital is not used after a colon (:) unless the word is the name of a person, place, or the beginning of a direct quote.

Capitals are used for names of vendors (e.g., Siemens), for acronyms (e.g., HIV), and abbreviations (e.g., CT, MRI, PET). It is not correct to use a capital letter for the word cancer unless at the beginning of a sentence; this also applies to radiologist, student, team, hospital, and so on. The use of upper case for the first letter of a word in sentences is a common error in academic writing. However, the first letter of a scientific name of a bacterium is in capitals (e.g., *Escherichia coli*) and is in italics.

20.6 Italics

Some writers use italics for emphasis in a sentence. This can be confusing if use is also made of italics for verbatim comments. If emphasis is essential, then add a clause to underscore what is being emphasized. For example, the researcher will investigate patients' perceptions of *quality* service delivery with an emphasis on quality. Use italics for verbatim comments of participants in a study. In addition, the title of a book, journal paper, and newspapers/magazines in a sentence is in italics. *The Sunday Times* covered poor service delivery at the Smith Hospital in 2018; Jones underscores research ethics in his book *How to conduct research*. The scientific names of bacteria are in italics: tuberculosis is caused by *Mycobacterium tuberculosis* as previously mentioned.

20.7 Anthropomorphism

Anthropomorphism is when we ascribe human activities and behavior to non-human objects; research study, data, and findings, for example [10–12]. It is also called personification. You need to be clear about who did the action so that what you state makes sense. Pause and question whether the noun you have used can perform human actions.

For example, does the following sentence make sense? This study investigated how many patients reported that they received poor service delivery. Is it possible for a study to perform a human action? This should be reworded. In this study the number of patients who reported poor service delivery was investigated. It is however acceptable to state, for example, that hospital management reported that the number of patient complaints had increased. Hospital management comprises groups of people; they can report on complaints. The same applies to organizations as they too comprise groups of people. It is also acceptable to use anthropomorphism when referring to the legends of tables and figures in your text. For example, Figure 2 presents the total number of patients who complained about poor service delivery. This chapter does not include figures; Figure 2 was used for discussion of verb tenses. To avoid anthropomorphism be clear who did a human action, or will do a human action.

20.8 Use of Verbs

Verb tenses tell us when something happened in time [13–15]. We can convey this information to indicate whether it still can happen in the future; whether it is currently happening (present time), or whether it has already happened (in the past). We make use of future, present, and past tense verbs. For example, "consent will be obtained from the participants" indicates that this will be done in the future. "Consent is being obtained from the participants" indicates that this is currently happening, thus it is in the present. "Consent was obtained from the participants" indicates it has already occurred: a past tense verb is used.

It is a standard practice in academic writing to use future tense verbs in research proposals. They provide information of what will be done by you. When writing up your research you must use past tense verbs to inform readers about what you did. However, if you include recommendations, then use verbs that indicate what should be done in the future. Be careful in your choice of verbs though. If you state that role-players will use your findings, then in this context "will" means they have to do this as a future action. Bear in mind that they are not obliged to use your findings, thus you should indicate they have a choice. For example, role-players could use the findings. Similarly, if you make recommendations for future research based on the findings in your study, then you should state: future research should (could) be done to address identified problems in this study.

In addition, verb tenses are important when citing literature. When there is one author we need to use the correct verb. For example, Smith (2013) reported that most patients did sign informed consent. According to Jones and Bennett (2019) very few patients in their study had signed informed consent forms.

When reviewing literature make sure that you use the correct verb tenses. Some examples are as follows. According to Smith (2000) the most common examination in imaging is chest radiography. This implies that the statement is generally accepted as being true. According to Smith (2000) the most common examination in imaging has been chest radiography. This implies that this is generally accepted as being true and that chest radiography is still the most common examination. According to Smith (2000) the most common examination in imaging was chest radiography. This implies this was true in the past.

20.8.1 Reporting Verbs in a Thesis/Dissertation

- Abstract and introduction

  It is a standard practice to use the past tense in an abstract as it covers what was done by a researcher. The introduction (usually chapter one) tells readers about the study and its relevance. The present tense is used as you need to explain to the readers why your study is important. When you discuss the study in terms of literature you are providing information which you believe is true. If you cite authors,

then what they said, for example, is in the past, thus past tense verbs are used. For example, Smith and Jones (2015) found in their study that most patients were unhappy with service delivery, and, Evans (2016) reported that healthcare professionals had been trained in the management of patient complaints.

If you intend conducting a study on patients' perceptions of the attitude of radiographers while imaging them, then you need to cite relevant studies. For example, most orthopedic inpatients have at least two imaging examinations during their period of hospitalization (Smith and Howell 2016). According to Jones (2017) the majority of inpatients are asked to only comment on service delivery by nursing personnel. Evans (2017) conducted a study on hemiplegic patients' perceptions of physiotherapists. There seems to be a gap in the literature because no studies have been done on patients' perceptions of the attitude of radiographers.

- Literature review chapter

  This chapter should cover literature chronologically in terms of aim, research questions/hypothesis, and objectives of the study. For example, Smith (1999) found in his study done in Bristol that most inpatients are not asked their opinion about the attitude of healthcare personnel. James and Jones (2004) did a similar study in Edinburgh and they concur with Smith. It would be incorrect to state that Smith (1999) concurs with James and Jones (2004). The latter is a common error in theses/dissertations. It would not be possible for an author in 1999 to concur with findings in a 2004 study.
- Methods

  In this chapter you are reporting on what you did, thus the past tense is used. Also ensure that the discussion is logical so that anyone, who wishes to do a similar study, will be able to use each step of the method you used. Keep to simple language and explain clearly what was done. Passive voice is usually used, but it would not be incorrect to use active voice. For example, twenty inpatients were invited to participate in the study. They were purposively selected. This was a quantitative study; the research tool comprised 15 closed-ended questions.

  Refer to tables and figures in the present tense if used to explain your method or what you did (see the discussion on anthropomorphism below). For example, Table 1 illustrates the main questions. Figure 1 shows a pie-chart of the responses of the participants.
- Discussion

  In this section you explain the significance of your results; the present tense is used. For example, the participants' responses show that most had negative perceptions in terms of the attitude of radiographers. The present tense is used to explain or unpack the results. Use the past tense when you summarize the findings and results. For example, the majority of participants had negative perceptions of radiographers.

- Conclusion

  This is usually a combination of past and future tenses. For example, although the participants had negative perceptions of radiographers, the sample was small and limited to one imaging department. Further studies, with a bigger sample, are needed to determine whether the findings are applicable in all imaging departments in the country.

20.9 Incorrect Words

The incorrect use of words is a common mistake in both formal and informal writing [16–18]. Table 20.1 provides some common mistakes of words used in academic writing. The list is based on my observations of the incorrect use of words in academic writing. The definitions of the words in Table 20.1 in this chapter are not exhaustive; many of the words have several denotative meanings [19].

Table 20.1 Examples of incorrect words in academic writing

Words	Examples	Examples	Comment
A/an versus the	A is an indefinite article. It is used before a consonant • A researcher may undertake a qualitative, quantitative, or mixed method study • A patient may need additional tests An is an indefinite article and is used before a word that begins with a vowel • An abdominal CT study was performed in this study	The is a definite article and refers to a specific thing or person. In academic writing it refers to a specific person or object • The researcher conducted a study on patient care • The patient in this case report required additional tests	Use the (definite article) when referring to your study including what you did • The questionnaire was handed to the participants by the researcher Use an indefinite article to clearly spell out the person is not you or not a specific person • A researcher could use a semi-structured questionnaire to interview a cancer patient
Affect versus effect	Affect as a verb means to influence or act upon something • Poor service delivery will affect optimal healthcare	Effect as a noun means a consequence of something • The effect of poor infection control is well documented in the literature	

(continued)

Table 20.1 (continued)

Words	Examples	Examples	Comment
Assure versus ensure	Assure as a verb means to confirm, make sure, or guarantee something • The researcher assured the respondents that their names would not be shared with others	Ensure as a verb means to make safe or safeguard • The researcher ensured that the collected data were password protected	
Attain versus obtain	Attain means to achieve a task • The aim of the study was attained	Obtain means to get or acquire something • A signed consent form was obtained from the participants	
Allude versus elude	Allude means to refer indirectly to something • The researchers alluded to possible causes of poor follow-up visits	Elude means escape, avoid, or baffle • The cause of the artifacts on the images continues to elude the researcher	
Being versus been	Being as a noun means something that exists • He is a human being It also forms the passive voice of all transitive verbs • The research is in the process of being conducted	Been is a verb; it is the past participle of be • The researcher has been conducting a study at the hospital	
Complement versus compliment	Complement means a person or object that completes something. Also means a total number of persons in a team, etc. • The staff complement in the nuclear medicine department is twenty	Compliment means to express admiration • The professor complimented the students on their academic results	
Complementary versus complimentary	Complementary is an adjective and means that things are different from each other but make a good combination • CT and MRI have complementary roles in brain examinations • The study included the role of CT and MRI as complementary modalities in staging of colon cancer	Complimentary means flattering. It also means given for free as a courtesy • The outpatients who were rebooked were given three complimentary meals	

Table 20.1 (continued)

Words	Examples	Examples	Comment
Ethics versus ethical	Ethics is a noun. It is the moral principles that govern behavior. It is a code of conduct • The ethics of patient care in radiography include do no harm • An ethics committee of a university assessed the research proposal to ensure that the principles of research ethics would be adhered to in the study∗	Ethical is an adjective and refers to morally acceptable behavior • The researcher conducted an ethical study by adhering to patient confidentially • The ethics committee found that the researcher's study was unethical as patients were coerced to participate∗	∗It is incorrect to state that a study was approved by the ethical committee of the university. The binary opposite of ethical is unethical. Thus the use of ethical committee implies the university may have an unethical committee
Its versus it's	Its is a possessive adjective and noun and never has an apostrophe • Research is complex, hence researchers have used different models to explain its process	It's is a contraction, hence the use of an apostrophe to show a missing word. For example, it's is a contraction of it is • It's cold in the ward	
Overtime versus over time	Overtime is a noun. It refers to work done outside of regular times • The researcher investigated the overtime pay rate of healthcare professionals at the Smith Hospital	Over time means gradually in a period • The patient was assured her mobility would improve over time	
Prevalence versus incidence	Prevalence is a noun. It means widespread or pervasiveness • A quantitative study was conducted in 2017 to determine the prevalence of workplace violence in all imaging departments in Australia • The prevalence of breast cancer worldwide is well documented in the literature	Incidence is a noun. It means the extent or frequency of occurrence of something • A quantitative study was conducted to determine the incidence of workplace violence in an imaging department in London during January and February 2017	

(continued)

Table 20.1 (continued)

Words	Examples	Examples	Comment
Principal versus principle	Principal is a noun. It means main or chief • The principal granted the researcher permission to interview the teachers	Principle is a noun. It is a standard of conduct, rule, or code • The study adhered to the four principles of medical ethics: respect for autonomy; non-maleficence; beneficence; and justice	
To versus too	To can be used in a sentence as a preposition before a noun/pronoun It is part of an infinitive verb: to run, to conduct, to go, for example∗∗ • The researcher handed a questionnaire to him • The objective of the study was to determine the incidence of reported downtime of the PET scanner from June to December 2017	Too is an adverb. It means also or excessively • A researcher should pilot a questionnaire to check that it will not take too long to answer • The researcher invited more teachers to participate in the study because several were too busy and could not commit to spending time answering the online survey	∗∗Guard against splitting infinitives The specific objectives of this study were to: (1) Identify factors that cause poor service delivery at the Smith Hospital (2) Determine appropriate strategies that could be used to reduce poor service This is an example of splitting infinitives. It is preferable to use to at the beginning of each objective. To identify… There are other uses of to in English. For example, to show equality: convert inches to centimeters Tip to check the use of to in a sentence. If to can be replaced with, for example, also, as well, or very, then use too and not to
Were versus where	Were is a verb • The participants were randomly selected • The findings of this study could be tested if future research were to be conducted	Where is a adverb and means, for example, what place • Smith Hospital is where the study was conducted	We're is a contraction of we are

Table 20.1 (continued)

Words	Examples	Examples	Comment
Which versus that	Which is a relative pronoun and is normally preceded by a comma • The hall, which has a high ceiling, was selected for the study site	That is a demonstrative pronoun • The hall that has a high ceiling was selected for the study site	It is incorrect to use which and that interchangeably. If the clause's meaning will not be affected, then use which as a non-restrictive element [17]. Thus in the sentence "the hall, which has a high ceiling, was selected for the study" the meaning will not change if the clause "which has a high ceiling" were to be removed. A non-restrictive element is an aside [18]. Whereas in the sentence "the hall that has a high ceiling was selected for the study site" use is made of that: a restrictive element. It limits the meaning to a specific hall being selected

20.10 Circumlocution, Tautology, and Clichés

Effective academic writing is precise; therefore use words sparingly. In this point in time is an example of circumlocution: several words are used to discuss something instead of being precise. Use now instead of in this point in time. Tautology is the repeated use of words or clauses in a sentence that mean the same thing. I personally interviewed the participants is an example of tautology. If personally is removed in the sentence the meaning does not change. Another example, "the participants' comments on service delivery problems were an added bonus." "Added" is redundant because bonus is something additional. Guard against using clichés in your writing. These are overused phrases. They lack original critical thinking, which is the foundation of academic writing. The researcher used Smith's (2018) questionnaire in order to not reinvent the wheel. The interviews were put on the back burner while waiting for permission to conduct the study. The pilot study led to going back to the drawing board. These sentences all include clichés: reinvent the wheel, put on the back burner, and back to the drawing board.

20.11 Proof Reading Is Essential

Two common errors are the incorrect use of another, and concur. Can you identify the errors in the sentences below?

- In their study among 30 radiographers in Germany, Tomas et al. (2013) found that only 30% could demonstrate proper handwashing methods. This figure increased to 53% at the first follow-up and was 47% at the second follow-up (Smith 2010). Another study of 100 radiographers found that only 15% confirmed they did use soap for handwashing (Evans 2018).

The cited fictitious studies are not listed chronologically. The study by Smith (2010) preceded the 2013 study by Tomas et al. Another study of 100 radiographers means that Evans had conducted a previous study of 100 radiographers. The use of another is often incorrectly used to mean a study by someone else.

Can you identify the error in the sentence below?

- Only two (20%) hospitals, one in an urban area and another in a rural region have operating theaters.

This needs to be revised. Only two (20%) hospitals have operating theaters: one is in an urban area and one is in a rural region.

Can you identify the errors in these sentences in this fictitious 2018 thesis?

- The findings of the current study concur with those of a study done among Australian healthcare professionals that revealed that nearly all respondents had never undergone infection control training (Smith 2003). The findings from this study also concur with Higgins and Murray (2010) who stated that most healthcare professionals do not undergo infection control training. Another study conducted by Khan (2013) also concurs with the findings of this study; most healthcare professionals had not received training in infection control.

How can a current study concur with the findings of a study done in 2003? This would only be correct if the objective was to use the 2003 results as a baseline to compare with the results of the current study. This also applies to the study by Higgins and Murray (2010). The use of another study means that Khan had done a similar study before the one in 2013.

Careful proofreading should be done for all drafts of your document to ensure that the meaning of the text makes sense.

20.12 Answer the Research Question

Not answering your research question, or leaving it up to a reader to decide whether you have achieved this, is another common pitfall. This could happen by losing focus during the course of your study and not maintaining clear communication with your

supervisor. The conclusion chapter of your dissertation (as mentioned in Chap. 17) is where you should pull together your findings and summarize them in light of the aim and objectives of your study. The conclusions drawn from this summary should enable you to adequately answer your research question. It is possible that you may not have been able to answer all of your questions by the end of your study. Be overt about this, as research is seldom perfect. You may need to make recommendations for further work (see Chap. 18). Recommendations for further work can then be picked up as a master's level study. It may even lead to a doctoral thesis.

20.13 Time Management

Many of you may ask the following question. What about time management? Universities today have diverse student populations and many students have other responsibilities: part-time employment, and families to care, for example. Your study therefore has to be carefully organized and well managed.

Time management is a valuable skill and an attribute to develop. Once you have it well established, it becomes a transferable skill that can be applied in all situations. Time is something we often take for granted. It is said to be human nature to leave things to the last minute. However, while you may be able to write up a 1000 word essay in a few days, you will have difficulty applying this strategy to a 10,000 word research project. The importance of a carefully planned study, which has been divided into small manageable chunks, mapped against a definitive time scale, cannot be overemphasized. Even a small-scale study involving just a few participants needs careful time management.

As you will have read in previous chapters, undertaking a piece of research requires a systematic approach. The merits of a completed study depend on your commitment to the steps within that process. A good practice measure is to draw up a study plan or keep a research diary that has dates indicating practical time scales according to how you envisage moving your study along. This plan then becomes a working document. It can act as a monitor to show which tasks you have completed. There are time management templates on the internet if you need some guidance on creating your own plan. Ticking off completed tasks has an added therapeutic benefit of a 'feel good' factor that says 'Well done!' at each stage throughout the process. It is important to reward yourself for these little accomplishments along the way; in so doing this helps to keep you self-motivated and focused on your next target.

Keeping a well-designed study plan also helps to identify additional time constraints: for example, deadline to apply for ethics approval from a relevant committee if your study requires this. Ethical approval is required for all projects that involve the use of human participants (see Chap. 6). These include academic or clinical staff, students, patients, the general public, and written or visual records that potentially allow identification of individuals. Remember that ethics, in a nutshell, is a critical reflection on morality. It is considered poor research practice to recruit participants in close contact with a researcher, such as good friends or family members. It is, however, considered acceptable practice as part of a course, in the interests of the

experience of the research process. Writing up an ethics application form, and gathering supporting documentation, can take several weeks, depending on the nature of your study: allow yourself ample time to get this done properly if you are going down this route. Pay attention to the dates on which committees meet; target a meeting early in the year to allow enough time for data collection. You (the researcher) must consider the ethical implications for the research participants (see Chap. 6 for more information on ethical considerations). Diarize hand-in deadline dates of your work. Make sure you visibly display these dates so that at a glance you will be able to note deadlines. Set aside ample time for putting your dissertation together, including organizing appendices, printing, and binding of the final document.

20.14 Research Supervision

All undergraduate and postgraduate student projects are supervised by staff who are academic tutors, or clinical tutors, or both, depending on the research topics. The role of a research supervisor is to provide guidance and support throughout a research process. Your research supervisor is the first point of contact in all research project-related matters and is there to answer any query you may have in that regard. Once you have been allocated your supervisor, do not hesitate to contact him or her and arrange your first meeting. This first meeting is the most important of all of those to come. It allows both you and your supervisor to mutually set the ground rules for communication and the manner in which supervision will proceed. More importantly it gets the ball rolling. Your supervisor would normally expect you to have a proposal of your intended study at hand for discussion at this meeting. It is a good practice to keep a log of meetings and a recording of the proceedings, again with target dates. These can then be used to inform and update your own personal action plan, as recommended above. Also to inform your supervisor of your progress.

All too often students meet with their respective supervisors at the beginning of the year, get advice, and then do not get in touch until the very last minute when the hand-in date looms. Do not let this happen to you. Be sure to schedule regular meetings with your supervisor for a steady drip-feed of guidance throughout your study. This will ensure that any problems or errors are picked up and dealt with efficiently and effectively and do not hinder the development of the final project. When you arrange to meet with your supervisor, ensure that it is at a mutually suitable time. It is your responsibility to initiate and maintain contact with your supervisor. Supervisors act as mentors and motivators; they will not undertake the work for you. You have to be committed to follow their guidance and work to set targets within agreed timeframes. It is a good practice to submit some work prior to a meeting. You could do this either electronically or drop off a hardcopy. This will give your supervisor some time to read through the text and note points for discussion. Your dissertation is unlike any other work you have undertaken for your course; it provides a unique opportunity to actually work with your supervisor.

Make sure that you also prepare for each arranged meeting with your supervisor. Draw up a list of questions, and have writing material to write down notes. Do not rely on your memory. You may prefer to record the proceedings of the meeting

using a voice recorder. Ensure that this is agreed with your supervisor beforehand. Supervision guidance should be reflected on and used in a positive and constructive manner to improve your work. Using feedback positively should also 'snowball' into other aspects of your work so that you do not make the same mistakes over again. In this way you will begin to 'work smart' rather than work hard.

A research process has a certain degree of flexibility within it and the various stages are merely offered as guidelines to follow. These have been produced by experts in the field and are tried and tested recipes that have yielded successful outcomes. Find your own measure among them, but be careful of deviating too much from the straight and narrow. It is a good practice to read the guidelines after every little bit of literature searching or writing that you have completed. This helps you to keep focused on the aims and objectives for your research project.

Remember to use the guidance you receive from your supervisor, no matter how intelligent you deem yourself to be. Even experienced research students need assistance. If stumbling blocks appear in your research process, do not be disheartened. Modify your work and carry on. Very little research is conducted exactly as it was envisaged.

20.15 Key Points

- Verb tenses are important. A research proposal is used to inform readers what will be done by a researcher. A research report includes what was done, hence past tense verbs are used.
- Discuss literature in a logical manner: early studies to current studies. Use the aim and objectives of your study to link to your discussion of literature. If an objective includes comparing the results of your study with those of a previous study, then you can use concur or disagree in your discussion, for example. If not then consider using the following: in keeping with, similar to, or in accord with. The use of synonyms adds to the richness your text.
- Read your drafts carefully to check spelling and grammar.
- Refrain from excessive use of adjectives to reduce the risk of being subjective.
- Be consistent in your use of words, and citing of sources.
- Keep a log of meetings with your supervisor.
- Plan for each meeting with your supervisor and act on feedback.
- Time management forms the backbone of your study as you need to remain on track and meet deadlines

20.16 Conclusion

In this chapter tips on how to avoid common errors in research writing are presented. Effective academic writing is precise [20, 21]. Good grammar and correct spelling underpin a well written proposal, report on a study, or journal paper. Research requires careful planning and rigorous proof reading to ensure the text is error-free. Strict management of each step in research is important to ensure that deadlines are met. Did you notice the incorrect use of both American and English spelling for some words?

References

1. Ramlaul A. Good practice tips and pitfalls to avoid when writing up. In: Ramlaul A, editor. Medical imaging and radiotherapy research: skills and strategies. London: Churchill Livingstone; 2010. p. 265.
2. Calvano B. Plagiarism in higher education. 2011 [cited 2019 May 10]. http://www.examiner.com/adult-education-in-pittsburgh/plagiarism-higher-education.
3. Truss L. Eats, shoots and leaves. London: Profile Books; 2005.
4. Punctuation guidelines [cited 2019 May 10]. http://en.fel.zcu.cz/AE%20III%20Guidelines%20for%20Academic%20Writing/Punctuation/Punctuation%20guidelines.pdf.
5. Punctuation. A brief overview [cited 2019 May 8]. https://www.up.ac.za/media/shared/Legacy/169/hwc.zp12816./Punctuations/punctuation.zp12841.pdf.
6. Basic punctuation rules [cited 2019 May 8]. https://www.apu.edu/live_data/files/288/basic_punctuation_rules.pdf.
7. University of Kent. Grammar, spelling and punctuation [cited 2019 May 10]. https://www.kent.ac.uk/learning/resources/.../grammarspellingandpunctuation.pdf.
8. Woods G. English grammar for dummies. 2nd ed. Indianapolis: Wiley; 2010.
9. Vinz S. English mistakes commonly made in a dissertation. 2015 [cited 2019 May 10]. https://www.scribbr.com/academic-writing/english-mistakes-commonly-made-in-a-dissertation/.
10. Barrass R. Scientists must write. A guide to better writing for scientists, engineers and students. London: E & FN Spon; 2000. p. 31.
11. Barrass R. Students must write. A guide to better writing in coursework and examinations. 3rd ed. London: Routledge; 2005.
12. Anthropomorphism in academic writing - Enago Academy [cited 2019 May 10]. https://www.enago.com/academy/anthropomorphism-in-academic-writing.
13. Tenses in writing [cited 2019 May 10]. https://depts.washington.edu/engl/askbetty/tenses.php.
14. University of Nebraska–Lincoln. Writing about your research: verb tense [cited 2019 May 10]. https://www.unl.edu/gradstudies/connections/writing-about-your-research-verb-tense.
15. University of Melbourne. Using tenses in scientific writing [cited 2019 May 10]. https://services.unimelb.edu.au/__data/assets/pdf_file/0009/471294/Using_tenses_in_scientific_writing_Update_051112.pdf.
16. Conrad J. 35 Mistakes e-book [cited 2019 May 10]. https://www.writingexplained.org/35-Mistakes-to-Avoid-Wrting.pdf.
17. Klems BA. Which vs. that [cited 2019 May 10]. https://www.writersdigest.com/online-editor/which-vs-that.
18. Restrictive and non-restrictive elements [cited 2019 May 10]. https://grammar.yourdictionary.com/grammar-rules-and-tips/when-to-use-which-or-that.htm.
19. Butterfield J, editor. Fowler's modern English usage. New York: Oxford University Press; 2015.
20. Murray R, Moore S. The handbook of academic writing: a fresh approach. Berkshire: Open University Press; 2006.
21. Perneger TV, Hudelson PM. Writing a research article: advice to beginners. Int J Qual Health Care. 2004;16(3):191–2.

Glossary

Accuracy How close a measurement represents the true value of something. It is mainly affected by 'systematic errors' which are referred to as bias.

Aim of study Purpose or goal of research plan.

Analysis The process by which sense and meaning are derived from the information gathered. It involves breaking down information to extract a deeper meaning from the key elements that make up the information, for example, analysing literature for a review or analysing findings from research results.

Anonymity Where the identities of participants are replaced by codes which the researchers use during the study. In this way, subjects cannot be identified by the reader or sometimes by the investigators themselves.

Attitude Refers to beliefs, views or feelings about something specific, such as interprofessional learning. There is an emotional attachment which may be positive or negative.

Bayes theorem Also known as the 'inverse probability principle', which relates the probability of the occurrence of an event to the occurrence or non-occurrence of an associated event; or a theorem that updates the probability of an occurrence following new research findings or evidence.

Bias A subjective attitude or viewpoint which can cause distortion or deviation of research findings from its true value or meaning.

Blind study When researchers are unaware which participant is allocated to or belongs to which group, as during a RCT (randomized controlled trial).

Clinically significant Refers to a result that is large enough to be of practical importance to healthcare providers and patients, for example, a new form of treatment or imaging technique. This has a direct effect on clinical practice and is not the same as being statistically significant.

Coercion Urging participants to participate in a study, by manipulation or intimidation, or by pressurizing or tricking them. Coercion forms an important ethical consideration during data collection.

Collusion Unauthorized collaboration; most commonly occurs during group work.

Confidentiality An important code of conduct whereby participants' identity and personal information is not made public knowledge.

A. Ramlaul (ed.), *Medical Imaging and Radiotherapy Research: Skills and Strategies*, https://doi.org/10.1007/978-3-030-37944-5

Confounding variables A variable other than the independent variable that may have certain effects on your participants' behaviour. These need to be taken into account when drawing conclusions from your research findings.

Contrast To contrast is to draw comparisons or point to differences in data.

Control group The group that has not been exposed to the intervention being tested.

Critical analysis The development of argument when reviewing literature. It is an evaluation of the positives and negatives or the agreements and disagreements with the author's point of view and has been grounded in theory or justified using additional supporting evidence from literature.

Delphi technique This is a forecasting tool that does not require face to face participation. It is a useful way to get public opinion on issues or ideas for problem solving.

Dose fractionation Refers to exposing tissues to smaller (fractionated) doses of radiation resulting in less tissue death, rather than one large exposure, over a period of time.

Dosimetry The measurement of exposure to radiation from a radiation-emitting source.

Effectiveness Refers to how well an intervention works in practice, i.e. does it do what it says it is going to do?

Efficacy Refers to the measurement of how well an intervention works, i.e. does it produce the intended result?

Empirical Refers to research that has been based on evidence from observations or experiments.

Epidemiology The study of the health of whole communities and populations, not just of particular individuals.

Epistemology The theory of knowledge and how one comes to know about a specific topic.

Ethics A set of moral principles which researchers have to abide by in order to protect their participants and themselves.

Evaluate A thought-provoking process of asking and answering critical questions, collecting appropriate information, and then analysing and interpreting the information for a specific use and purpose.

Evidence Findings to support claims from research that would inform practice.

Exclusion criteria This refers to a specific set of conditions that participants may be subjected to in order to determine whether they are eligible to participate in a research study, for example a particular study might exclude persons under the age of 18 years.

Experiment A set of observations undertaken to solve a research question or problem in order to prove or disprove a hypothesis.

Explore To investigate or to examine.

External validity Refers to the extent to which research findings can be generalized or applied to other population groups.

False negative A false finding that a person does not have a particular condition, when in fact they do.

False positive A false finding that a person has a particular condition, when in fact they do not. See Specificity.

Feasibility Refers to the possibility, capability and practicability of, for example, a plan working.

Generalizability See External validity.

Gold standard A set of guidelines for a specific aspect of practice, such as a diagnostic test or protocol, that is regarded as definitive.

Hypothesis A statement which is tested through research. This can be an experimental hypothesis (positive) whereby the researcher states, for example, that the public have good knowledge about the hazards of radiation; or it could be a null hypothesis (negative) which states, for example, that the public do not have good knowledge about the hazards of radiation.

Inclusion criteria This refers to a specific set of conditions that participants must have in order to take part in a research study, for example females aged 18–30 years with a history of migraine headaches.

Inductive reasoning Refers to making generalizations based on observations to support the conclusion of a study.

Informed consent An agreement by participants to take part in the study after receiving a full explanation of the study procedure and any risks involved, etc. Consent may be given in an oral, written or implied format. The manner in which consent is gained from research participants is an important ethical consideration.

Internal validity Refers to the manner in which the study was con ducted free of bias, i.e. the extent to which the study actually did what it intended to do and the truthfulness of the results.

Interpretivism A research methodology that is based on humanistic qualitative methods.

Intervention Any measure put into place with the primary purpose of improving health or altering the course of disease, for example, testing a new drug or therapy.

Justify To support with evidence any assumptions, statements or assertions made.

Methodology The theoretical analysis of the research methods employed within a study.

Negligence The behaviour or conduct of a reasonable person who fails to protect another person from foreseeable harm or risks, for example, those associated with certain interventions.

Objective To provide a balanced view taking into account the pros and cons of the issues involved.

Paradigm A model, pattern or a set of ideas that influences how one looks at something.

Phenomenology A research approach that focuses on the lived experiences of particular groups of people.

Placebo An inactive drug or intervention that is administered to 'blind' participants, for example in a randomized controlled trial where the effect of the actual drug or intervention is compared to the placebo. Participants sometimes feel a

therapeutic, beneficial effect of receiving the 'treatment like' placebo and this effect is known as the placebo effect.

Plagiarism Passing off other authors' work as your own and not acknowledging the work of others.

Population All people who share certain specified characteristics, for example males aged 25–35 years with a history of lower back pain. A sample of participants is usually then selected from this population.

Positivism A research methodology that is based on scientific quantitative methods.

Precision The degree of similarity among a study's results, e.g. how close the measurements are to each other. Precision is mainly affected by 'random errors'.

Probability The likelihood of an event occurring. Probability is measured numerically usually from 0 (never occurs) to 1 (always occurs).

Qualitative Refers to methods whereby data collected is based on participants' experiences, attitudes and ideas; from these themes are drawn in order to extract deeper meaning and understanding of the specific issues.

Quality Refers to the strength or appreciation of the study in relation to any methodological flaws and/or presence of bias.

Quantitative Refers to methods whereby data collected is based on numbers; uses statistical analysis to draw meanings from which generalizations can be made.

Randomization Refers to the process of randomly allocating participants to research groups, for example, to the control group or experimental group of a randomized controlled trial.

Randomized controlled trial (RCT) An experimental design which tests the effectiveness and efficacy of an intervention, for example, therapy, health service or health technology. A sample of participants is randomly allocated into two or more groups so that the possible effects of the intervention can be compared with non-intervention under controlled conditions.

Reflection A process whereby concrete experiences are reviewed in depth and analysed together with relevant theories to provide a plan of action to be implemented in practice.

Reflexivity Refers to an acknowledgement or awareness of the involvement or influence that the researcher has in his or her own study, for example how the researcher's own values and experiences have influenced the study, and also how undertaking the study has impacted on the researcher.

Reliability The extent to which the study is able to yield the same result if conducted under the same conditions or using the same measurements over again. The more consistent the result is, the greater the reliability of the result.

Reproducibility The ability of the study to be done again (reproduced) in the same way, elsewhere.

Sampling Method of drawing participants for your study from an identified population.

Sensitivity This is the measure of 'positives' identified, i.e. a test's ability to correctly identify persons with disease.

Specificity This is the measure of 'negatives' identified, i.e. a test's ability to correctly identify persons without disease.

Statistically significant Refers to the probability that a result is not due to chance alone. The level of significance determines the degree of certainty or confidence with which we can rule out chance.

Subjectivity Refers to the researcher's personal, introspective view.

Synthesis 'Building' of information using the key elements that have been 'broken down' (analysed) together with new found information (usually from literature searches). Where analysis is the breaking down of information, synthesis is the building up of information. Synthesis usually follows on after information has been analysed.

Transactional analysis Analysis of the emotion observed during communication in interpersonal relationships, where there are conflicting egos, for example, in the doctor–patient relationship.

Triangulation A technique that uses two or more data sources to verify findings of a study.

True negative A correct claim that a person does not have a particular condition.

True positive A correct claim that a person does have a particular condition. See Sensitivity.

Validity Refers to the extent to which the results or measurements are accurate, reliable and free from bias. See also Reliability and Bias.

Variables This refers to anything that can affect or change the results of a study.