Medical Imaging in Clinical Trials

Colin G. Miller
Joel Krasnow
Lawrence H. Schwartz
Editors

Springer

Medical Imaging in Clinical Trials

Colin G. Miller • Joel Krasnow
Lawrence H. Schwartz
Editors

Medical Imaging in Clinical Trials

Springer

Editors
Colin G. Miller, BSc, PhD, FICR, CSci
Medical Affairs
BioClinica Inc.
Newtown, PA
USA

Lawrence H. Schwartz, MD
Department of Radiology
Columbia University Medical Center
New York, NY
USA

Joel Krasnow, MD, MBA
Department of Strategy
and Regulatory Affairs
TherapeuticsMD
Boca Raton, FL
USA

ISBN 978-1-84882-709-7 ISBN 978-1-84882-710-3 (eBook)
DOI 10.1007/978-1-84882-710-3
Springer London Heidelberg New York Dordrecht

Library of Congress Control Number: 2014931376

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

This book is dedicated to my family, who has stood beside me as I have moved us between continents and who has allowed me to pursue my career and ambitions beyond expectations: To my beautiful wife of 25 amazing years – Angie; To my sons, Joe and Elliot, who have kept me young at heart and helped me "update" many life skills; To Katie, my late "adopted daughter," who has brought me fresh perspectives; To Biggles, my faithful golden retriever of 13 years who walked by my side and sat for many hours as I worked on this book, and to Toby my new puppy who wants me to play rather than write.

There are many mentors and team members who helped develop the concepts and theories over the years: Without their friendship and intellectual discourse, this book would not have had significant content. A deep and heartfelt "Thank You" to each one.

Colin G. Miller

There are many people without whom this book would not have been possible. First, I would like to thank the researchers and educators who have mentored me over the years. I would also like to thank those I collaborated with on this book as the final product could have never been reached without our combined efforts. Foremost, I would like to thank my wife, Lisa, and my children, Michael and Emily, for their love and support that made this book possible.

Joel Krasnow

To Hilda, Martin, Amanda, Rachel, and Alyssa – in order of entrée into my life! You have inspired and taught me, and I hope you will continue to do so. There are big "writing shoes" to fill here; Dad, I hope to make you as proud of me as I am of you.

Lawrence H. Schwartz

Foreword

Quantitative and functional medical imaging play an increasingly important role in the development of new therapeutics. There is rising demand for quantitative imaging-based biomarkers because they noninvasively detect disease and predict the likelihood of response to therapy and subsequent patient outcomes. Imaging findings, alone and in combination with other markers, are used to make decisions about trial eligibility, to assess response to therapy as either efficacy or safety endpoints, to monitor patients for potential relapse during follow-up periods, or to provide important mechanistic insights. And, as imaging technologies become more widely validated as biomarkers of disease, imaging will play an even larger role in clinical trials. Yet, despite this increased focus on imaging in clinical trials, few clinical trialists possess the detailed knowledge required to optimize imaging contributions. This book provides an essential resource to address that problem.

In the areas of oncology, cardiovascular disease, brain disorders, musculoskeletal disorders (especially arthritis and osteoporosis), and infectious diseases, as well as an array of metabolic, gastroenterological, and inflammatory disorders, imaging plays a vital role in clinical trials to determine the effectiveness of new therapies. The market for imaging analysis in clinical trials in 2009 was approximately $550 M in total annual revenue. Conservative estimates for future annual growth are 5–10 %, although some analysts project more rapid growth. Thus, there is a critical need for imaging expertise to ensure that such an investment returns information that is reliable and meaningful.

Because the use of medical images in clinical trials has accelerated rapidly, government (e.g., NIH) and commercial sponsors are requiring more complex and comprehensive imaging services, which require changes in study design and data interpretation, as well as an expanded knowledge of competing modalities and technologies. Sponsors and investigators are increasingly reliant upon recognized experts to implement complex imaging in clinical trials. Such expertise is essential in study design, for example, in defining inclusion/exclusion criteria and imaging endpoints. These requirements create challenges for researchers who are unfamiliar with complex imaging. Similarly, regulatory agencies such as the US FDA are increasing their expectations and requirements for rigor in the imaging components

of clinical trials. Knowledge of these regulatory guidelines is another necessity for imaging members of clinical trial teams. These issues are extensively addressed in Part I of this book.

Clinical trials of cancer therapies are the largest single area for imaging in drug development and will likely continue to gain share. However, the use of imaging in other therapeutic areas also continues to increase because of the same attributes that are advantageous for imaging's use in cancer. For example, because of the large cost of phase III trials, it is increasingly important to measure tumor response at a relatively early time point so that trials could be terminated or modified if the investigational therapy is not working as expected. Additional considerations that influence the desire to monitor and measure tumor response effectively relate to potential toxicity and cost issues. It is desirable to terminate trials of ineffective, toxic, or expensive therapies as early as possible. Similar considerations apply to other therapeutic areas, especially to brain and heart disorders where biopsy is even less feasible than it is for cancer. Part II of this book provides individual chapters on several of the disease-specific issues that must be considered in therapeutic clinical trials.

There is widespread agreement that extracting objective, quantitative results from imaging studies will reduce the variability inherent in subjective, qualitative interpretations and thereby improve the quality of imaging-related endpoints in clinical trials. Such imaging in clinical trials today necessarily draws on a variety of expertise including, but not limited to, clinical medicine, informatics, computer science, statistics, biology, chemistry, physics, and engineering. Multidisciplinary collaborations are essential. This textbook provides an indispensable resource of fundamental principles and information for the success of these multidisciplinary clinical trial teams.

Daniel C. Sullivan, MD
Professor of Radiology
Director, Clinical Research
Medical Director of Radiology Site-Based Research
Codirector, Radiation Oncology and Imaging Program,
Duke Cancer Institute
Duke Radiology
Duke University School of Medicine
Durham, NC, USA

Preface

There are a number of books written about clinical trials [1–5]. They take a biostatistical approach providing generic information about how imaging can be used as clinical endpoints or biomarkers. This book is written specifically to address the questions around the application of medical imaging in the complex and highly regulated environment of clinical trials. It has also been timed to coincide with a new set of guidelines issued by the US Food and Drug Administration (FDA) entitled "Guidance on Standards for Clinical Trial Imaging Endpoints" which we have included verbatim, with permission, as Appendix 1.

Medical imaging has made dramatic advances in the last 30 or 40 years with the advent of higher computing power and new technologies. The magnitude of data that clinical trialists and radiologists have to manage has grown exponentially as have the skills required to accurately evaluate and interpret these images.

The development of new therapeutics as well as devices within the framework of the FDA and other international regulatory authorities has become more challenging. Yet, there is a drive to get new medications to suffering patients to relieve disease and prolong life. Clinical trial methodology has, by the very nature of the statistical evaluation required, been a very quantitative science with the so-called hard endpoints (e.g., death, myocardial infarct, fracture). Radiology has historically been an interpretative discipline with images being read qualitatively. This has led to the challenge we face today of bridging the "divide" between a quantitative and qualitative or descriptive science.

The quantitative application of medical imaging in clinical trials has really only been in existence for about 20–25 years. Dual-energy X-ray Absorptiometry (DXA) was probably the first modality where this process was fully described and thus could be utilized by pharma in selected clinical trials [6, 7]. As the need for quantification has evolved, the development of the semiquantitative or pseudo-quantification endpoints has grown, especially in the therapeutic area of oncology. Table 1 shows a complete listing of many of these criteria with references, which will surely change over time.

The goal of this book is to present key concepts of medical imaging in clinical trials by assembling the thoughts, concepts, and understanding of key thought

Table 1 Listing of semiquantitative or pseudo-quantification endpoints/criteria of medical imaging in clinical trials

Response criterion	Pathology	Year	Link to paper
Choi criteria	GIST	2008	http://www.ncbi.nlm.nih.gov/pubmed/18434631 Choi H. Response evaluation of gastrointestinal stromal tumors. Oncologist. 2008;13(Suppl 2):4–7. doi:10.1634/theoncologist.13-S2-4.
Choi criteria	GIST	2007	http://www.ncbi.nlm.nih.gov/pubmed/17470865 Choi H, Charnsangavej C, Faria SC, Macapinlac HA, Burgess MA, Patel SR, Chen LL, Podoloff DA, Benjamin RS. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumor treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. J Clin Oncol. 2007;25(13):1753–9.
RECIST	Malignant mesothelioma	2004	http://www.ncbi.nlm.nih.gov/pubmed/14760119 Byrne MJ, Nowak AK. Modified RECIST criteria for assessment of response in malignant pleural mesothelioma. Ann Oncol. 2004;15(2):257–60.
RECIST	Pediatrics	2003	http://www.ncbi.nlm.nih.gov/pubmed/12857700 McHugh K, Kao S. Response evaluation criteria in solid tumours (RECIST): problems and need for modifications in paediatric oncology? Br J Radiol. 2003;76(907):433–6.
Macdonald and Rano criteria	Supratentorial malignant glioma	1990	http://www.ncbi.nlm.nih.gov/pubmed/2358840 Macdonald DR, Cascino TL, Schold Jr SC, Cairncross JG. Response criteria for phase II studies of supratentorial malignant glioma. J Clin Oncol. 1990;8(7):1277–80.
Macdonald and Rano criteria	High-grade gliomas	2010	http://jco.ascopubs.org/content/28/11/1963 Wen PY, Macdonald DR, Reardon DA, Cloughesy TF, Sorensen AG, Galanis E, DeGroot J, Wick W, Gilbert MR, Lassman AB, Tsien C, Mikkelsen T, Wong ET, Chamberlain MC, Stupp R, Lamborn KR, Vogelbaum MA, van den Bent MJ, Chang SM. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. J Clin Oncol. 2010;28(11):1963–72.

CRPC, PCWG2	Castration-resistant prostate cancer	2011	http://jco.ascopubs.org/content/29/27/3695 Scher HI, Morris MJ, Basch E, Heller G. End points and outcomes in castration-resistant prostate cancer: from clinical trials to clinical practice. J Clin Oncol. 2011;29(27): 3695–704.
Immune-related response criteria Wolchok criteria	Melanoma	2009	http://clincancerres.aacrjournals.org/content/15/23/7412.full.pdf Wolchok JD, Hoos A, O'Day S, et al. Guidelines for the evaluation of immune therapy activity in solid tumors: immune-related response criteria. Clin Cancer Res. 2009;15:7412–20.
Image-guided tumor ablation	Solid tumors	2009	http://www.ncbi.nlm.nih.gov/pubmed/19560026 Goldberg SN, Grassi CJ, Cardella JF, Charboneau JW, Dodd 3rd GD, Dupuy DE, Gervais DA, Gillams AR, Kane RA, Lee Jr FT, Livraghi T, McGahan J, Phillips DA, Rhim H, Silverman SG, Solbiati L, Vogl TJ, Wood BJ, Vedantham S, Sacks D; Society of Interventional Radiology Technology Assessment Committee and the International Working Group on Image-guided Tumor Ablation. Image-guided tumor ablation: standardization of terminology and reporting criteria. J Vasc Interv Radiol. 2009;20(7 Suppl):S377–90.
RECIST 1.1	Malignant lymph nodes	2009	http://www.ncbi.nlm.nih.gov/pubmed/19091550 Schwartz LH, Bogaerts J, Ford R, Shankar L, Therasse P, Gwyther S, Eisenhauer EA. Evaluation of lymph nodes with RECIST 1.1. Eur J Cancer. 2009;45(2):261–7.
RECIST 1.1	Solid tumors	2009	http://www.ncbi.nlm.nih.gov/pubmed/19097774 Eisenhauer EA, Therasse P, Bogaerts J, Schwartz LH, Sargent D, Ford R, Dancey J, Arbuck S, Gwyther S, Mooney M, Rubinstein L, Shankar L, Dodd L, Kaplan R, Lacombe D, Verweij J. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228–47.
Cheson	Malignant lymphoma	2007	http://jco.ascopubs.org/content/25/5/579.full.pdf Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, Coiffier B, Fisher RI, Hagenbeek A, Zucca E, Rosen ST, Stroobants S, Lister TA, Hoppe R T, Dreyling M, Tobinai K, Vose JM, Connors JM, Federico M, Diehl V. Revised response criteria for malignant lymphoma. J Clin Oncol.2007;25:579–86.

(continued)

Table 1 (continued)

Response criterion	Pathology	Year	Link to paper
Hallek criteria	Chronic lymphocytic leukemia	2008	http://bloodjournal.hematologylibrary.org/content/111/12/5446.full.pdf Hallek M, Cheson BD, Catovsky D, Caligaris-Cappio F, Dighiero G, Döhner H, Hillmen P, Keating MJ, Montserrat E, Rai KR, Kipps TJ. Guidelines for the diagnosis and treatment of chronic lymphocytic leukemia: a report from the International Workshop on Chronic Group 1996 guidelines lymphocytic leukemia updating the National Cancer Institute Working Group 1996 guidelines. Blood. 2008;111(12):5446–56.
Multiple myeloma	Myeloma	2006	http://www.nature.com/leu/journal/v20/n9/pdf/2404284a.pdf Durie BGM, Harousseau J-L, Miguel JS, Blade'J, Barlogie B, Anderson K, Gertz M, Dimopoulos M, Westin J, Sonneveld P, Ludwig H, Gahrton G, Beksac M, Crowley J, Belch A, Boccadaro M, Turesson I, Joshua D, Vesole D, Kyle R, Alexanian R, Tricot G, Attal M, Merlini G, Powles R, Richardson P, Shimizu K, Tosi P, Morgan G, Rajkumar SV, on behalf of the International Myeloma Working Group. International uniform response criteria for multiple myeloma. Leukemia. 2006;20:1467–73.
PERCIST	Solid tumors	2009	http://jnm.snmjournals.org/content/50/Suppl_1/122S.full.pdf Wahl RL, Jacene H, Kasamon Y, Lodge MA. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumor. J Nucl Med. 2009;50:122S–50S.
Sarcoma	Sarcoma	2008	http://theoncologist.alphamedpress.org/content/13/suppl_2/32.full.pdf Schuetze SM, Baker LH, Benjamin RS, Canetta R. Selection of response criteria for clinical trials of sarcoma treatment. Oncologist. 2008;13(Suppl 2):32–40.
Metastatic urogenital cancer	Solid tumors	2010	http://content.karger.com/ProdukteDB/produkte.asp?Aktion=ShowPDF&ArtikelNr=000318985&Ausgabe=254445&ProduktNr=224282&filename=000318985.pdf Heidenreich A, Albers P, Classen J, Graefen M, Gschwend J, Kotzerke J, Krege S, Lehmann J, Rohde D, Schmidberger H, Uder M, Zeeb H. Imaging studies in metastatic urogenital cancer patients undergoing systemic therapy: recommendations of a multidisciplinary consensus meeting of the Association of Urological Oncology of the German Cancer Society. Urol Int. 2010;85:1–10.

Hepatocellular carcinoma	Solid tumors	2010	http://www.ncbi.nlm.nih.gov/pubmed/20175033 Lencioni R, Llovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010;30(1):52–60.
Hepatocellular carcinoma	Solid tumors	2010	http://annonc.oxfordjournals.org/content/21/suppl_5/v59.full.pdf Jelic S, Sotiropoulos GC, On behalf of the ESMO Guidelines Working Group. Hepatocellular carcinoma: ESMO clinical practice guidelines for diagnosis, treatment and follow-up. Annals of Oncol. 2010;21 (Suppl 5): v59–64.
Cheson	Malignant lymphoma	2007	http://jco.ascopubs.org/content/25/5/579.full.pdf Cheson BD, Pfistner B, Juweid ME, Gascoyne RD, Specht L, Horning SJ, Coiffier B, Fisher RI, Hagenbeek A, Zucca E, Rosen ST, Stroobants S, Lister TA, Hoppe RT, Dreyling M, Tobinai K, Vose JM, Connors JM, Federico M, Diehl V. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25:579–86.
Cheson	Lymphoma	2007	http://bloodjournal.hematologylibrary.org/content/110/10/3507.full.pdf Seam P, Juweid ME, Cheson BD. The role of FDG-PET scans in patients with lymphoma. Blood. 2007;110:3507–16.
PERCIST	Lymphoma	2007	http://jco.ascopubs.org/content/25/5/571.full.pdf Juweid ME, Stroobants S, Hoekstra OS, Mottaghy FM, Dietlein M, Guermazi A, Wiseman GA, Kostakoglu L, Scheidhauer K, Buck A, Naumann R, Spaepen K, Hicks RJ, Weber WA, Reske SN, Schwaiger M, Schwartz LH, Zijlstra JM, Siegel BA, Cheson BD. Use of positron emission tomography for response assessment of lymphoma: consensus of the Imaging Subcommittee of International Harmonization Project in Lymphoma. J Clin Oncol. 2007;25:571–8.

leaders in this discipline. While the key concepts in this text will not change, we recognize that many of the details will. Therefore, we designed the book to be read as needed and not necessarily from beginning to end. This book is broken into two main parts. Part I includes chapters on the design and concept of blinded reads as well as the details of how to write an imaging charter. Each chapter can be read in isolation; on the other hand, for example, Chap. 1, a basic chapter on medical imaging, may be skipped by the experienced radiologists. Part II includes chapters on each of the main therapeutic areas where imaging is employed in clinical trials. This portion of the book has been developed to provide greater detail of the biologic and clinical specifics in each therapeutic area. Part III leads us to the future of imaging in clinical trials, with a pharmaceutical industry perspective regarding imaging techniques. Finally, we end with three appendices to bring some of the key information together in one location. These are Appendix 1, the FDA Guidance for Industry on Standards for Clinical Trial Imaging Endpoints; Appendix 2, a glossary taken from www.ClinicalTrials.gov and a Lexicon developed specifically for Medical Imaging in Clinical Trials in conjunction with the FDA, DIA, and PhRMA; and Appendix 3, Information from the Quantitative Imaging Biomarker Association (QIBA) web site, a group which is looking at the evaluation of new quantitative biomarkers initially for clinical trials but also for clinical use.

This book has been written to be useful to the imager as well as the clinical trialist without any imaging experience. The editors hope that this book will be a useful contribution to the field of medical imaging in clinical trials and consolidate many different concepts into one location.

Newtown, PA, USA — Colin G. Miller, BSc, PhD, FICR, CSci
Boca Raton, FL, USA — Joel Krasnow, MD, MBA
New York, NY, USA — Lawrence H. Schwartz, MD

References

1. Portney LG, Watkins MP. Foundations of clinical research. 3rd ed. Upper Saddle River: Prentice Hall; 2008.
2. Friedman LM, Furberg CD, DeMets DL. Fundamentals of clinical trials. 3rd ed. New York: Springer; 1998.
3. Hulley SB, et al. Designing clinical research. 3rd ed. Philadelphia: Lippincott Williams & Wilkins; 2007.
4. Pocock SJ. Clinical trials: a practical approach. New York: Wiley; 1984.
5. Girling DJ, Parmar MK, Stenning SP, Stephens RJ, Steward LA. Clinical trials in cancer: principles and practice. Oxford: Oxford University Press; 2003.
6. Miller CG. Bone density measurements in clinical trials: the challenge of ensuring optimal data. Br J Clin Res. 1993;4:113–20.
7. Miller CG. Potential measurements in patients with osteoporosis. In: Nimmo WS, Tucker GT, editors. Clinical measurement in drug evaluation. London: Wiley; 1995.

Contents

Contributors

Harris A. Ahmad, MD Department of Medical Affairs, BioClinica, Inc., Newtown, PA, USA

Prabu Anandaraj, MBBS, MMST Advanced Clinical Applications, GE Healthcare, Chennai, India

Eric S. Forsthoffer Department of Business Development, BioClinica, Inc., Newtown, PA, USA

Frederik L. Giesel, MD, MBA Department of Nuclear Medicine, University Hospital Heidelberg, Heidelberg, Germany

George P. Heyrich, MD, FACC, FSCAI Mercer Bucks Cardiology, Nowtown, PA, USA

Joel Krasnow, MD, MBA Department of Strategy and Regulatory Affairs, TherapeuticsMD, Boca Raton, FL, USA

Amit Mehndiratta, MBBS, DPhil (Oxon) Department of Engineering Science, Institute of Biomedical Engineering, Old Road Campus Research Building, University of Oxford, Hedington, Oxfordshire, UK

Colin G. Miller, BSc, PhD, FICR, CSci Department of Medical Affairs, BioClinica Inc., Newtown, PA, USA

Sumit N. Niogi, MD, PhD Department of Radiology, Weill Cornell Medical College, New York, NY, USA

Derek Pearson, PhD, BSc Medical Physics and Clinical Engineering, Nottingham University Hospitals NHS Trust, Nottingham, UK

Andrea M. Perrone, MD, BS Finance Department of Imaging, Merck, West Point, PA, USA

Lawrence H. Schwartz, MD Department of Radiology, Columbia University Medical Center, New York, NY, USA

Robert C. Sergott, MD Department of Neuro-Ophthalmology, Wills Eye Hospital, Philadelphia, PA, USA

John A. Shepherd, BS, PhD Department of Radiology and Biomedical Imaging, University of California at San Francisco, San Francisco, CA, USA

Apostolos J. Tsiouris, MD Department of Radiology, Neuroradiology Section, New York-Presbyterian Hospital, New York, NY, USA

Cornelis van Kuijk, MD, PhD Department of Radiology, VU University Medical Center, Amsterdam, The Netherlands

Jennifer A. Villetard, BA Department of Imaging, Merck, West Point, PA, USA

Hendrik von Tengg-Kobligk, Priv-Doz, Dr med Institute for Diagnostic, Interventional, and Pediatric Radiology (DIPR), Inselspital—University Hospital Bern, Bern, Switzerland

Department of Diagnostic and Interventional Radiology (DIR), University Hospitals Heidelberg, Heidelberg, Germany

Department of Radiology, German Cancer Research Center (DKFZ), Heidelberg, Germany

Hui Jing Yu, PhD Department of Medical Affairs, BioClinica, Inc., Newtown, PA, USA

Christian M. Zechmann, MD Department of Diagnostic Radiology and Nuclear Medicine, Rinecker Proton Therapy Center, Munich, Germany

Binsheng Zhao, DSc Department of Radiology, Columbia University Medical Center, New York, NY, USA

Part I
Overview and Trial Management

Chapter 1
Medical Imaging Modalities

Harris A. Ahmad, Hui Jing Yu, and Colin G. Miller

Abstract Medical imaging is now utilized extensively in clinical trials for eligibility, efficacy, and safety evaluations. The uses of imaging span from a qualitative assessment of disease findings to quantitative assessments, each resting on diagnosis of the condition or change in the severity of the condition. This introductory chapter is designed for the novice with a limited or no background in radiological techniques and aims to briefly review the different imaging techniques, technology, terminology, and optimal imaging uses.

Keywords Radiology • Planar imaging • Tomographic imaging • Nuclear medicine • Ultrasound techniques

Introduction

Medical imaging is now utilized extensively in clinical trials for eligibility, efficacy, and safety evaluations. The uses of imaging span from a qualitative assessment of disease findings to quantitative assessments, each resting on diagnosis of the condition or change in the severity of the condition. Several imaging modalities have emerged as the mainstay techniques for evaluating such evaluations in clinical trials across several therapeutic areas. The later chapters in this book will go into these therapeutic specific details.

This introductory chapter is designed for the novice with a limited background in medical imaging and aims to briefly review the techniques, technology, and

H.A. Ahmad, MD (✉) • H.J. Yu, PhD • C.G. Miller, BSc, PhD, FICR, CSci
Department of Medical Affairs, BioClinica, Inc.,
826 Newtown-Yardley Road, Newtown, PA 18940, USA
e-mail: harris.ahmad@bioclinica.com; huijing.yu@bioclinica.com;
colin.miller@bioclinica.com

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_1,

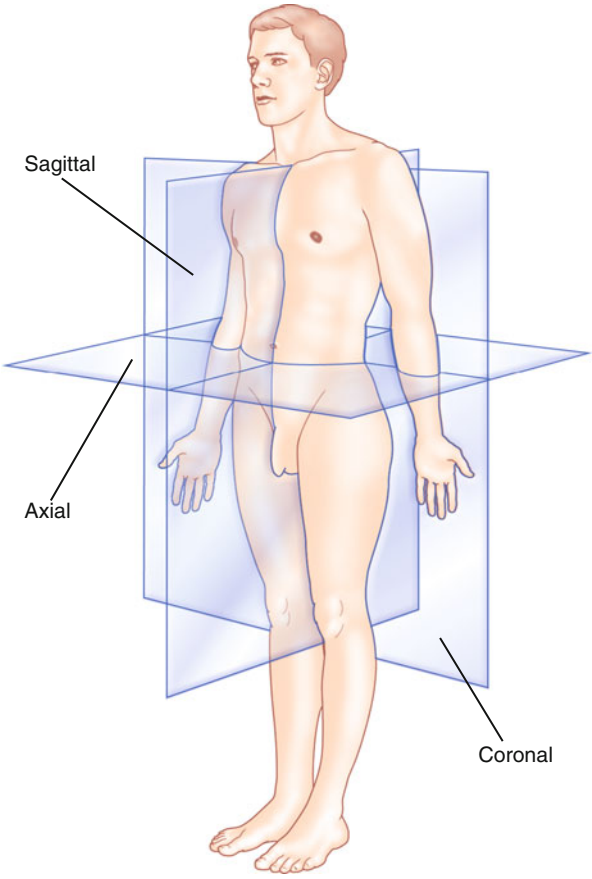


Fig. 1.1 Three orthogonal directions of the medical imaging of the human body

terminology. It is not designed as an in-depth evaluation of any specific technique nor is it designed to provide the reader anything more than the basic set of pros and cons of each technique and its general applicability.

Image Orientation

Before discussing the different imaging modalities, it is vital to understand the different orientations of which there are mainly three: axial, coronal, and sagittal. These are demonstrated in Figs. 1.1 and 1.2.

In medical imaging, the axial plane refers to the X–Z plane which divides the human body into superior and inferior positions, i.e., the head from the feet. In other words, each image in axial orientation is similar to a horizontal slice (Fig. 1.3).

The coronal is the X–Y plane which remains perpendicular to the ground and divides the human body into dorsal and ventral regions or front and back slices. This

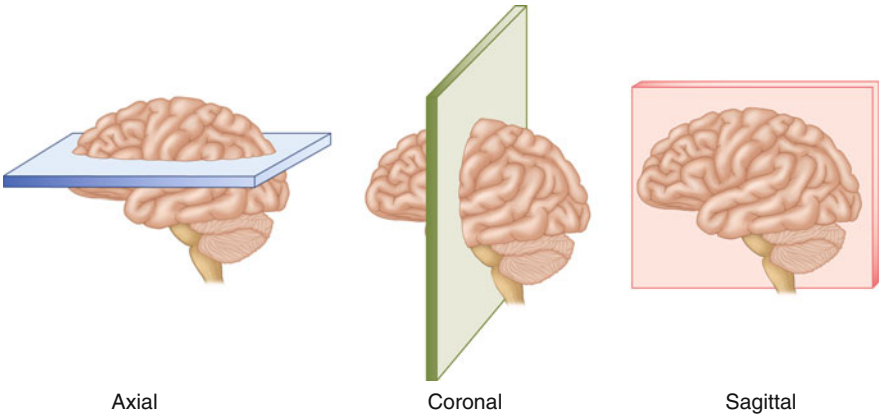


Fig. 1.2 The three orientations for imaging (Modified with permission from http://users.fmrib.ox.ac.uk/~stuart/thesis/chapter_3/section3_2.html)

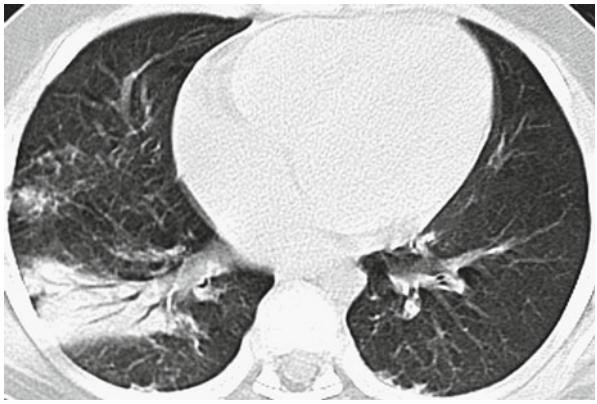


Fig. 1.3 Computed tomography (CT) of the chest in axial view (Used with kind permission of Springer Science+Business Media from Levine et al. [17])

can also be termed the anterior-posterior or posterior-anterior view in modalities such as X-ray. The more colloquial term is a frontal view.

The sagittal plane, or lateral view, is the Y–Z plane and can be commonly referred to as the side view. It is also perpendicular to the ground and distinguishes the left and right side of the body. The midsagittal plane passes right through the center of the body to create equal halves with this side view. In radiographs, the sagittal view could be termed the lateral view because it is the side angle view of the patient's anatomy. Figure 1.4a–d demonstrates a lateral view of a chest radiograph.

Finally, there is the oblique plane where the beam or radiation passes diagonally through the body and divides it into two diagonal halves or in other words images at a slight angle to that of the traditional view. For example, an oblique coronal would be a front view sliced at a slight angle.

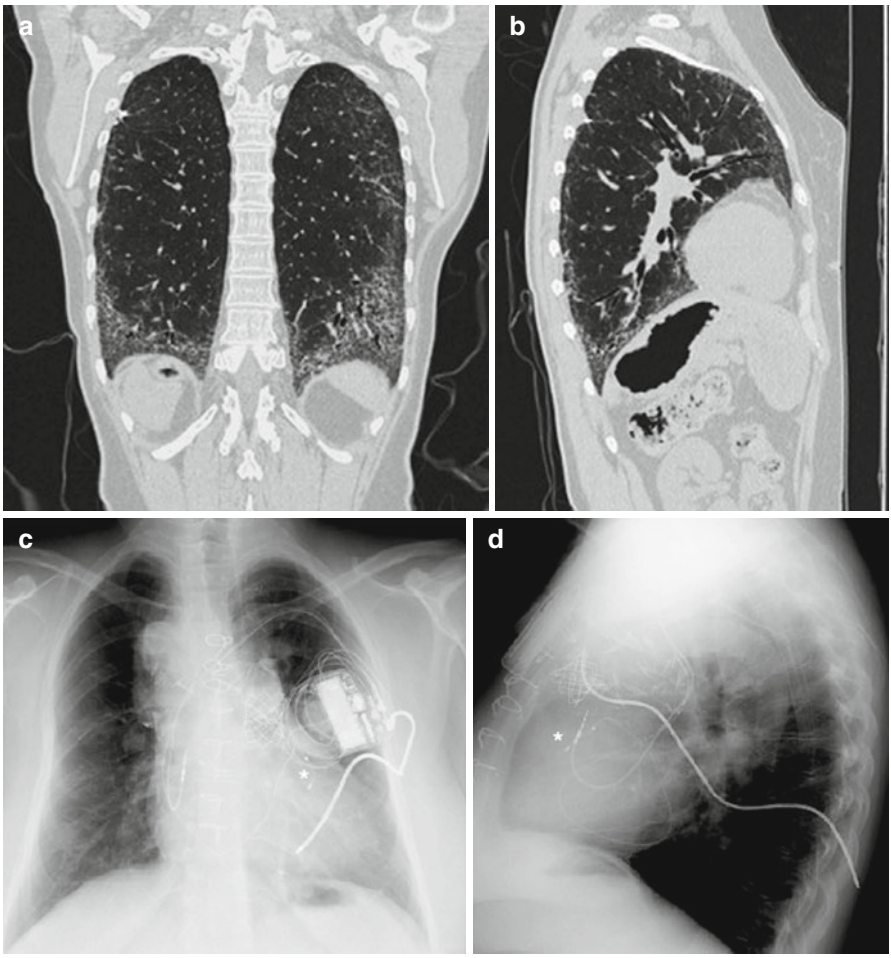


Fig. 1.4 Chest CT in coronal (**a**) and sagittal (**b**) view. Chest radiograph in a posterior-anterior (**c**) and lateral view with heart indicated by * (**d**) (c, d: Used with permission of Springer Science + Business Media from Gupta et al. [18])

Planar Imaging: X-Ray Techniques

Radiography/X-Ray

The earliest form of medical imaging was the radiograph or X-ray. This was originally discovered in 1895 by Wilhelm Conrad Roentgen and rapidly became the mainstay of imaging assessment of clinical diseases where applicable for almost a century [1, 2]. Even with all the new complex imaging techniques available, radiography is still an invaluable tool, particularly for the imaging of the skeleton. Further,

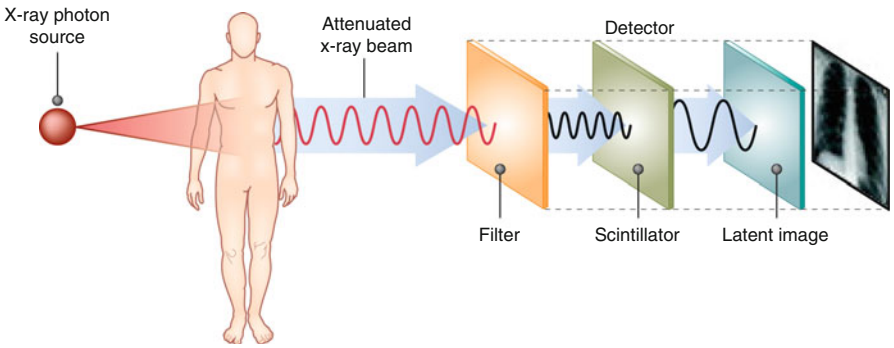


Fig. 1.5 X-ray from source to image (Modified with kind permission of Springer Science + Business Media from Aberle et al. [19])

it continues to be heavily relied upon by the FDA for ongoing and future trial endpoints as a consistent comparison to historic data, e.g., rheumatoid arthritis (see Chap. 11).

In radiography, the production of an image starts with a high-voltage electric current which creates a stream of electrons which are fired at a metal plate. The resulting interaction is the creation of X-rays which are collimated into a beam. This source produces X-rays which are directed towards the desired object to be imaged such as the patient. Three results of this X-ray beam are possible and as a consequence produce an image. The X-ray could pass through the patient, be absorbed by the patient, and/or be scattered or in other words the beam is attenuated. In the original and basic form, the X-rays are detected on a sheet of film in an X-ray cassette. The film is developed and the resulting image is a negative image of the attenuation [2]. Nowadays, most radiology departments use a digital system using a detector and hence digital X-ray or DXR, as shown in Fig. 1.5.

The X-ray beam is attenuated more of the material through which it is passing. Hence bones, predominantly consisting of calcium, attenuate the beam to a much higher degree than soft tissue [2]. Any X-rays that are attenuated do not obviously expose the film and therefore appears as white or radiopaque. The density of the tissues among the patient can vary and therefore be the determining factor in how much of the X-ray beam is attenuated [1]. Figure 1.6 shows how the density of these tissues and their respective atomic weights can result in either a radiopaque or radiolucent appearance on X-ray. This difference creates the image as those tissues with a high density such as enamel of teeth or bone result in a radiopaque image, while those with a very low density such as air result in a "black" area or radiolucent area of the film. Air is the least dense patient area followed in ascending order by fat, water, bone, and metal [3].

Despite its limitation as a 2D image with only a spectrum of black to white, X-ray remains one of the most useful imaging techniques in clinical practice with the major advantages, disadvantages, and applications listed in Table 1.1. The low

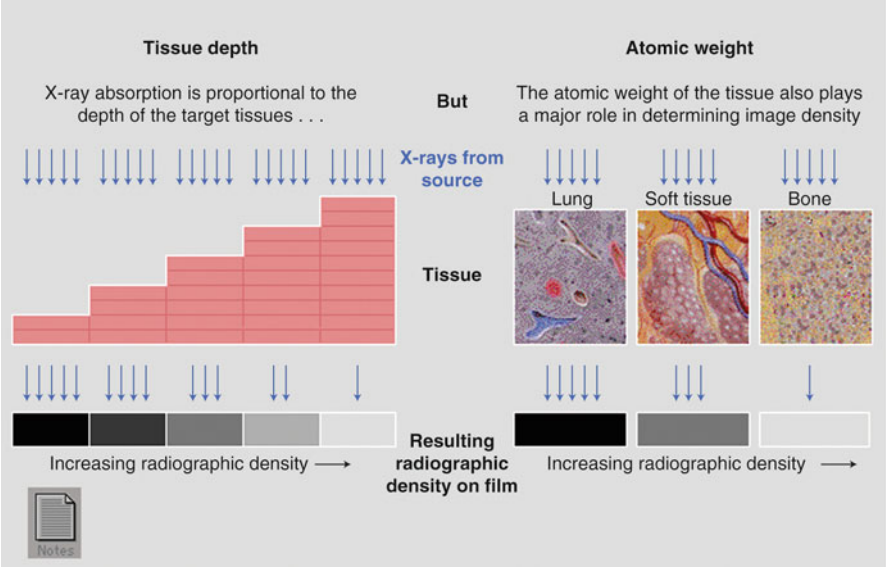


Fig. 1.6 Relationship of radiographic density as a gray scale verses atomic weight (Modified with permission of Patrick Lynch, Yale University, from http://www.yale.edu/imaging/techniques/radiographic_density/index.html)

Table 1.1 Radiography: applications, advantages, and disadvantages

Applications	Fractures, bone diseases, pneumonia, pulmonary edema, intestinal obstructions, renal or gallbladder stones
Advantages	Low cost, widely available, portable, bedside
Disadvantages	Radiation, limited color spectrum, 2D information

cost of equipment and acquisition is very attractive in comparison to more involved methods such as magnetic resonance imaging (MRI), computed tomography (CT), and imaging using radioisotopes. Further, the radiation dose is a quite a small fraction as compared to CT. Another advantage of course is the mobility of the X-ray acquisition at bedside of the patient, in the emergency room, or in a small outpatient practice. In clinical trials, this cost-effective, widely available, and well-practiced technique among radiology technologists contributes to its continued use as an efficacy endpoint in therapeutic areas such as rheumatoid arthritis, osteoporosis, and osteoarthritis. The chest X-ray continues to be of particular use for diagnosis and management of pneumonia, pulmonary edema and detection of calcified masses, while the abdominal X-ray can help detect and manage intestinal obstructions and associated pathology such as gallstones or renal stones. However, in clinical trials for oncology in which a volumetric or cross-sectional diameter assessment of lesions is paramount to determining response to therapy, the modalities such as CT and MRI clearly outperform planar radiography.

Dual Energy X-Ray Absorptiometry

Dual energy X-ray absorptiometry (DXA or DEXA) was first described by Cameron and Sorenson in 1963 [4]. In this first publication they not only described the concept of single photon absorptiometry but also developed the basic underlying equations that are the core of DXA measurements. The basic operational concepts are that an X-ray beam of two discrete energies (or two X-ray beams) is passed over the body or region of interest and the attenuation of the X-ray beam(s) calculated, since the number of X-ray photons being emitted is a known quantity. The next underlying assumption is that the body consists of three compartments: fat tissue, lean tissue, and bone. In the area of soft tissue, there are two components – fat and lean – and so with two compartments of known attenuation coefficients at the two discrete energies, simultaneous equations can be built. With two unknowns (the amount of fat and lean) the equations can be solved, and the quantities of both tissues derived. This provides the information for body composition. A second major assumption is then made that the soft tissue composition juxtaposition to the bone remains consistent where over the bone and called the r or k value depending on manufacturer. This constant is then used to define a second set of simultaneous equations for soft tissue and bone mineral content (BMC). The quantity of bone can then be derived. The key measurement that is required is the bone mineral density (BMD) and the underlying equation is:

$$BMD = BMC / Area$$

The area of bone can be identified by an attenuation threshold methodology and hence the BMD of bone calculated. As can be appreciated by this definition, DXA is a 2D measurement technique and creates a so-called areal density of the bone and body composition. With all the inherent assumptions and calculations, DXA has been shown to be remarkably precise and accurate. Precision for spine and total body BMD and body composition measurements in healthy individuals is around 1 %. The precision measurements around the proximal femur (the other key measurement site for BMD, besides the AP lumbar spine) are 2–3 %. Accuracy has some different issues, since there is debate as to how the accuracy of areal BMD should be defined. There is not the space here to go into this debate but enough to say that there is a calibration offset between the two manufacturers of between 10 % and 15 %, which means this has to be accounted for in clinical trials along with calibration shifts etc. This is well documented in other textbooks [4] and will not be discussed here.

However, DXA is well established as an imaging modality and as a surrogate for fracture, at least in prevention of osteoporosis with the measurement of BMD. It is also a good measure of fat and lean tissue and has been used in many clinical trials to demonstrate the change in body composition. It is therefore extensively used as a modality in trials evaluating therapies in osteoporosis, obesity, diabetes, and sarcopenia. The body composition assessments using DXA are detailed further in Chap. 12. The BMD assessment is covered in more depth in Chap. 11. The two main

manufacturers of DXA equipment are GE Lunar (Madison, Wisconsin, USA) and Hologic Inc. (Bedford, Massachusetts, USA), and, unlike BMD values, they are more closely calibrated for body composition measurements.

Computed Tomography

In essence, X-ray and computed tomography (CT) are very similar to each other in the physics of the technique. X-ray beams are targeted at the patient and, depending on the physical properties of the patient's differing tissues, they are attenuated; however, unlike "plain film X-ray," CT is a tomographic technique [1]. An early CT scanner consisted of a single X-ray emitter and an in-line detector that could rotate around the object or patient that was placed within the tube that housed the emitter and detector. A single "slice" or image of the body was scanned and the body moved a centimeter or more through the tube and the scan was repeated, thereby building up a series of tomographic images of the object or subject [2].

As technology progressed more detectors were introduced into the system and then more X-ray emitters. As the complexity grew, the acquisition speed increased and the slice thickness decreased. The X-ray tube and the electronic detectors are now present in the gantry or the circular structure. As soon as this information is received by the detectors, they are passed on to the computer for the calculation of attenuation of X-rays as shown in Fig. 1.7a, b. This structure can be rotated in different angles to take images of various portions of the body from various angles thereby producing an image in multiple planes as shown in Fig. 1.1.

In the modern systems it is not unusual to have 64, 128, or even 256 detectors and emitters which allow for very rapid acquisition. Furthermore the system spirals around the patient without the need for discrete steps (hence spiral CT), since the reconstruction algorithms on the image processing side have become more complex and elegant [5].

The differences in the physical properties of the tissue again compromise the characteristic images but now in an axial dimension. With this technique, CT provides a cross-sectional view of the body and can produce views in the 3 dimensions as described previously: axial, coronal, and sagittal. The differences in the densities of tissues are displayed on CT as Hounsfield units (HU) with a range of approximately −1,000 to +1,000. Air has the lowest HU ranging from −1,000 to −200 with metal at +500 to +1,000. The lower the HU, the "blacker" the color is on the CT image. Therefore, from black to white the sequential order are air, fat, water, soft tissue, blood, bone, and metal (which are the same for plane film X-rays). This distinct difference on a black and white color spectrum on CT is very advantageous for distinguishing key anatomy and pathology.

A contrast agent can be given on CT to obtain further distinction of certain anatomical structures and pathology. A contrast agent is often an injected or ingested liquid that has a distinct density as compared to physiologic tissues [3]. This allows

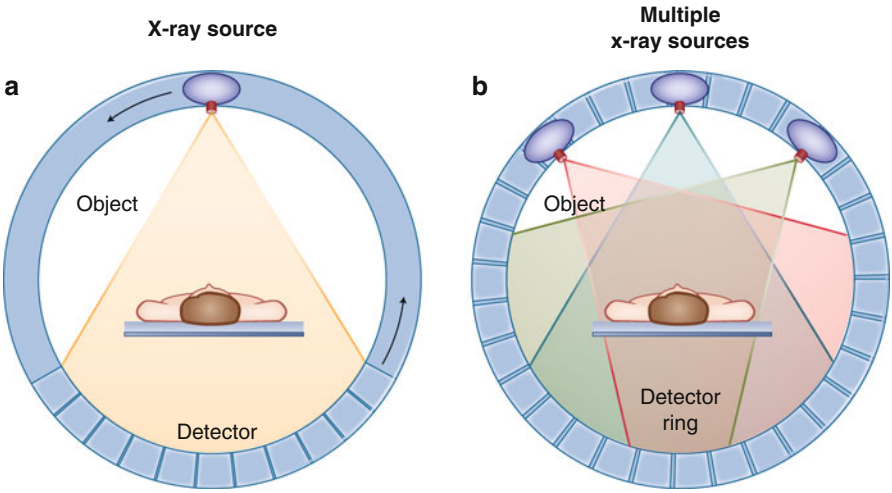


Fig. 1.7 (**a**, **b**) Multiple X-ray sources (**b**) arranged in a configuration to produce a CT scan (Modified with permission from Zhang et al. [20])

Table 1.2 Computed tomography: applications, advantages, and disadvantages

Applications	Lesion assessment, trauma evaluation, evaluation of nearly all organ systems (gastrointestinal, neurologic, bone, vascular etc.)
Advantages	Cross-sectional view, tissue contrast, rapid acquisition
Disadvantages	Radiation, contrast allergy, cost

for differentiation or "highlighting" of internal organs and structures for evaluation. For example, the function of an injected vascular contrast material is to raise the density of vascular structures and organs and delineate any pathology such as a mass in the bowel wall or aneurysm of the vascular wall. Bowel anatomy and associated pathology can also be distinguished through oral ingestion of the material before the scan. Proper timing and dosage is key to an accurate scan [5].

Numerous applications and advantages of CT as listed in Table 1.2 have made the modality one of the most clinically robust imaging techniques. A cross-sectional view as described previously with the delineation of different tissues with and/or without contrast has proven to be major advantages at all stages of clinical care. Examples include assessment of lesion size in oncology studies, cardiac disease detection and management, gastrointestinal disease diagnosis and management, and other numerous applications such as traumatic injury evaluation.

However, the disadvantage of radiation dosage and possible carcinogenic effects of the dosage have resulted in some concerns of overuse of the imaging modality. Further, contrast medium risk particularly in those patients with renal failure or allergic responses to the agent is also of concern. CT carries with it some of the less attractive features for the patient such as being in a closed machine and the adverse reactions to contrast administration such as nausea, vomiting, pain at the injection

site, as well as further compromise of renal function. These allergies and renal contraindications can be life threatening, and therefore, assessment of each patient's clinical status through proper history and lab work is often required delaying an otherwise urgent scan. Lastly, there is a high cost of acquisition and maintenance of a CT scanner in comparison to radiography.

Magnetic Resonance Imaging

Magnetic resonance imaging (MRI) is used to image nuclei of atoms (e.g., ^{1}H, ^{13}C, ^{14}N, ^{23}Na, ^{31}P) inside the body based on the principles of nuclear magnetic resonance (NMR). The NMR phenomenon was first reported in 1946, and the use of NMR was then established as a technique for in vivo imaging in the early 1970s, known as MRI today. Since then, several Nobel Prizes have been rewarded to the field of NMR, demonstrating the importance of such technology.

The majority of clinical MRI focuses on imaging hydrogen nuclei (^{1}H) which are abundant in the human body and have a relatively large magnetic moment. In the absence of an external magnetic field, the hydrogen nuclei in the body are randomly oriented, and the net macroscopic magnetic moment is zero. In the presence of an external magnetic field (i.e., a patient placed in a MR scanner, Fig. 1.8), water becomes polarized such that hydrogen nuclei are oriented in the direction of the applied magnetic field.

To obtain a MR signal, a radio frequency or RF pulse is applied. Protons absorb energy from RF excitation that brings them out of equilibrium. When the RF pulse is turned off, the system of protons relaxes back to its equilibrium while dissipating the absorbed energy to their surroundings (Fig. 1.9a, b). The spins return to their equilibrium usually by two spin relaxation mechanisms known as T1 or longitudinal relaxation and T2 or transverse relaxation (Fig. 1.10a, b). T1 relaxation is caused by the protons giving up their energy to the surrounding environment. The T1 relaxation time describes the time constant for restoring the net magnetization to 63 % of its original strength in the direction parallel to the applied field (i.e., longitudinal magnetization). T2 relaxation is caused by protons exchanging energy with their neighbors, resulting in the loss of magnetization perpendicular to the external field (i.e., transverse magnetization). The T2 relaxation time represents the time it takes for the transverse magnetization to decay to 63 % of its original strength. Since the physical properties of the tissue affect the T1 and T2 relaxation times, tissue contrast can be generated [6, 7].

Tissues differ in relaxation constants and thus measuring the MR signal during the relaxation period provides image contrast which translates into grayscale visualization, unlike CT, grayscale intensity reflects tissue density. Table 1.3 shows the list of water relaxation time (in ms) at 1.5 T [8].

By changing the imaging parameters, the images can be "weighted" to reflect one type of relaxation more than another. Within the MRI pulse sequence, the echo

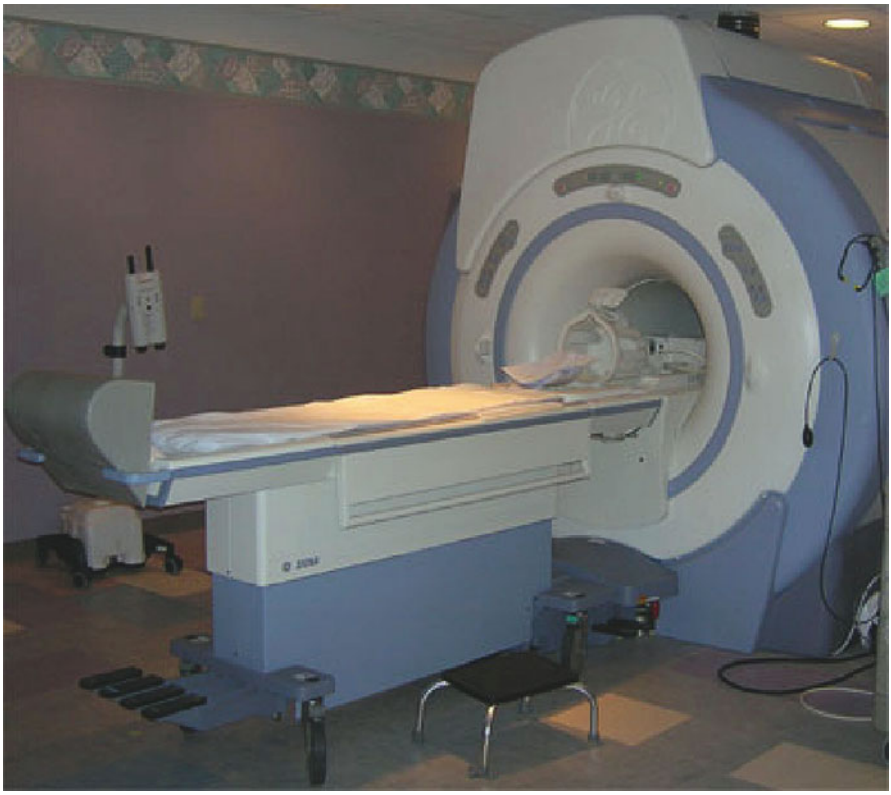


Fig. 1.8 Picture of a MRI scanner (Used with kind permission of Springer Science + Business Media from Semrud-Clikeman et al. [21])

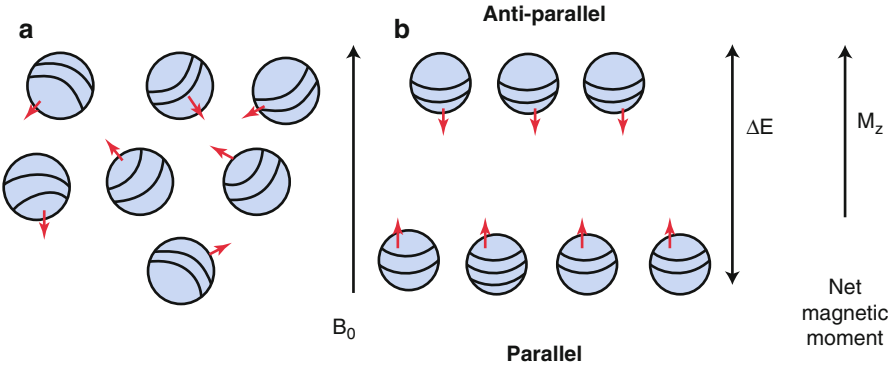


Fig. 1.9 (**a**, **b**) Alignment of protons with the magnetic field (Used with kind permission of Springer Science + Business Media from Saha [22])

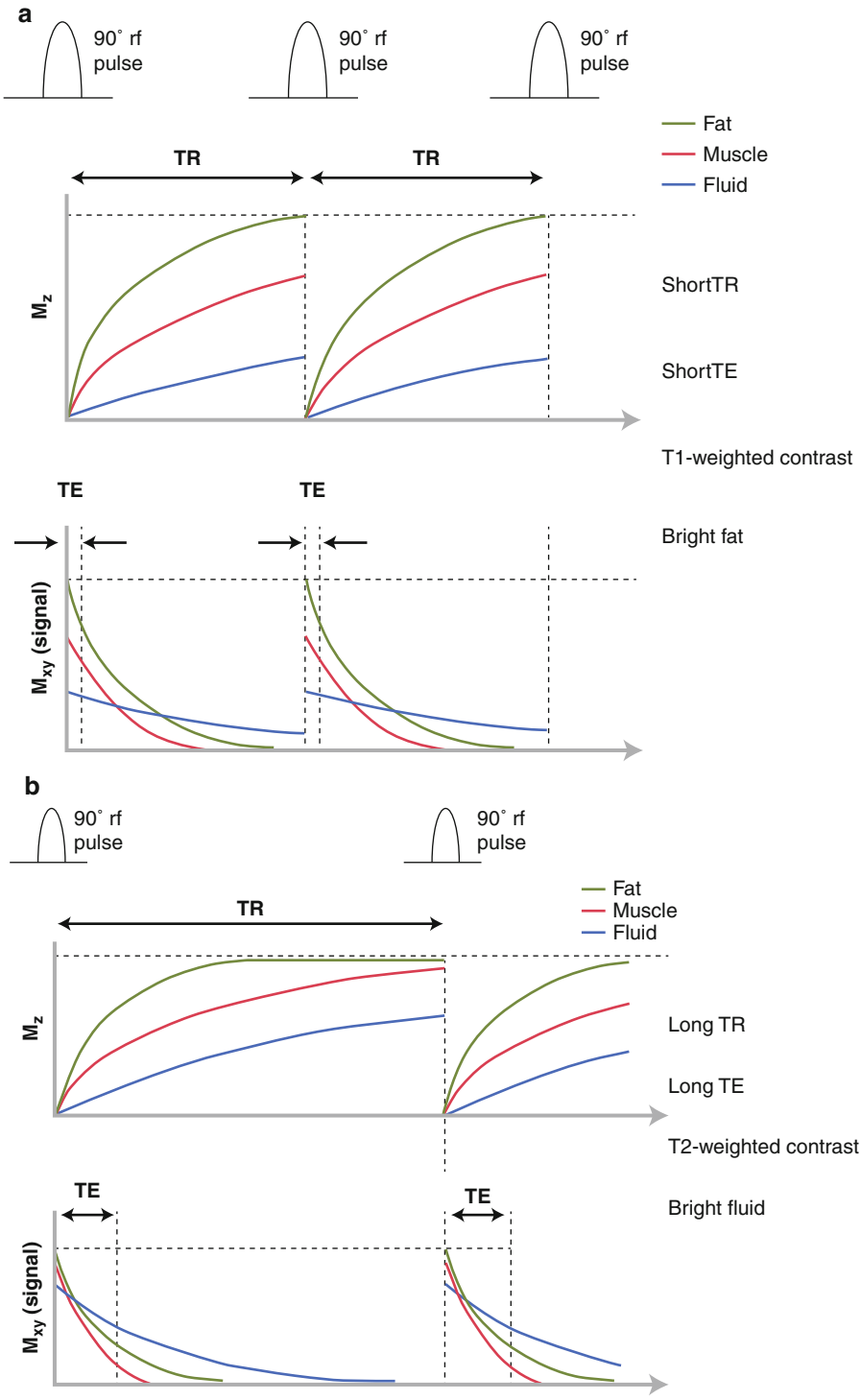


Fig. 1.10 T1 (**a**) and T2 (**b**) relaxation properties (Modified with kind permission from Ridgway [23])

Table 1.3 List of water relaxation time (in ms) at 1.5 T

Tissue type	Approximate T1 value	Approximate T2 value
Adipose tissues	240–250	60–80
Whole blood (deoxygenated)	1,350	50
Whole blood (oxygenated)	1,350	200
Cerebrospinal fluid (similar to pure water)	4,200–4,500	2,100–2,300
Gray matter of cerebrum	920	100
White matter of cerebrum	780	90
Liver	490	40
Kidneys	650	60–75
Muscles	860–900	50

Based on data from Wood et al. [8]

time (TE, the period between the start of the RF pulse and the peak of echo signal) and the repetition time (TR, the period between two RF pulses) are used to determine how the resulting image is weighted. To sum up, typically, a short TE and long TR sequence is called proton density-weighted; a short TE and short TR sequence is called T_1-weighted; and long TE and long TR sequence is called T_2-weighted. In general, fluids have long T1s and T2s and thus appear dark on a T1-weighted image and bright on a T2-weighted image, whereas fat have short T1s and T2s and thus appear bright on a T1-weighted image and dark on a T2-weighted image [6, 7]. Figure 1.11a, b shows different T1 and T2 relaxation curves and corresponding examples of T1-weighted and T2-weighted images.

Both T1-weighted and T2-weighted images are acquired for most clinical examinations. An important addition to help further delineate tissue abnormality was the introduction of MR contrast agents. The most commonly used is gadolinium, a paramagnetic contrast agent that influences the local magnetic field to markedly shorten the T1 of neighboring water protons, thus locally increasing the signal on a T1-weighted image. This is especially useful for brain MRI because these large size contrast agent complexes cannot pass through the cell layers that comprise the blood-brain barrier, unless the barrier is compromised, thus providing high sensitivity for tumor and lesion detections [6, 7].

Over the past two decades, several MR imaging techniques have been developed and are utilized for monitoring and assessment in clinical trials [9]. These include diffusion weighted imaging (DWI), magnetic resonance spectroscopy (MRS), and functional MRI.

Diffusion Weighted Imaging

Diffusion MR imaging is based on phenomenon of Brownian movement. In water, individual molecules are in constant random motion in all directions, resulting in isotropic diffusion. When diffusion is restricted by structural components such as cell membranes, it becomes anisotropic. Diffusion MRI allows quantitative measurement of this molecular motion of water. In basic diffusion weighted imaging (DWI), the amount of water diffusion in tissue is usually quantified by the apparent diffusion

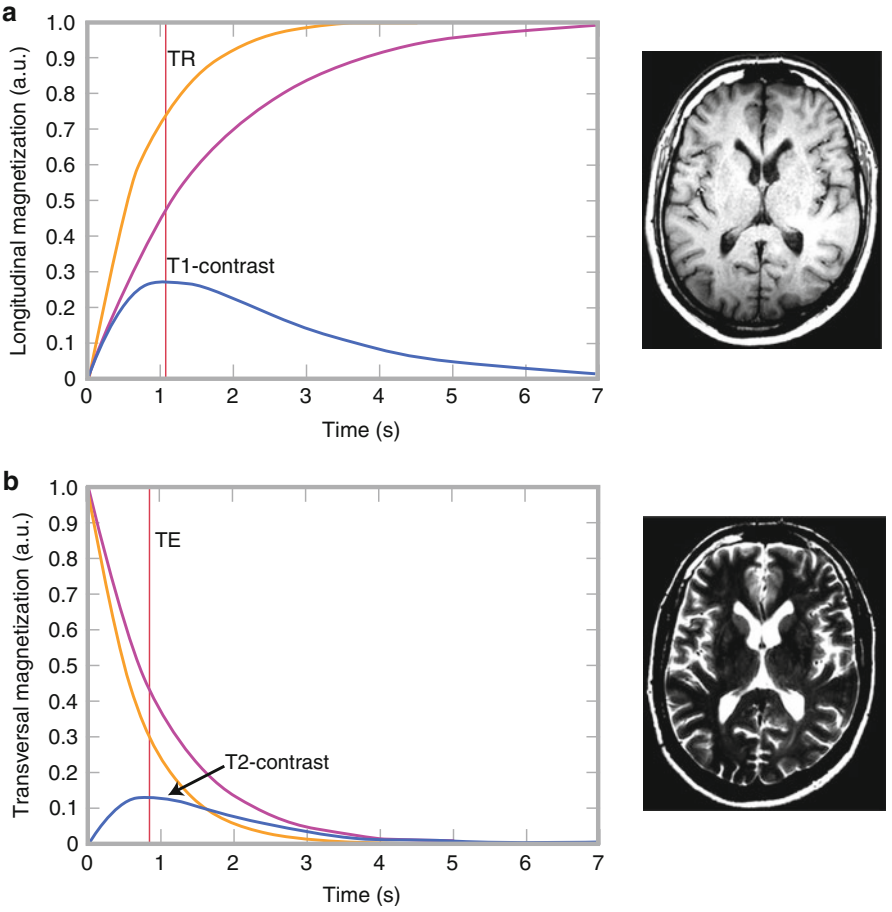


Fig. 1.11 T1-weighted (**a**) and T2-weighted (**b**) contrast in brain (Modified with kind permission of Springer Science+Business Media from Scheef and Frank [24])

coefficient or ADC. Restricted diffusion, such as ischemia, results in a decreased ADC and brighter signal on DWI. DWI has been used in a wide range of therapeutic areas and has been incorporated as endpoints in clinical trials.

Magnetic Resonance Spectroscopy

Magnetic resonance spectroscopy (MRS) is a method to obtain metabolic information from tissues in vivo. The metabolic information is based on differences in resonance frequency of nuclei depending on their chemical environment, a phenomenon known as chemical shift. MRS signals can be localized to one or several tissue volume of interest, using anatomical images as reference. The acquired signal in the

time domain can be converted to the frequency domain to obtain a MRS spectrum. Chemical shift determines the resonance frequency position of each peak on the spectrum, which is expressed as the shift in frequency in parts per million (ppm) relative to a standard. The resulting spectrum can be quantified as either the absolute value or normalized signal intensity.

MRS is commonly used for neurological and liver studies. For example, a decrease signal at 2.02 ppm corresponds to a decrease of NAA, a neuronal marker, indicating neuronal dysfunction or loss. Although the utility of MRS in diagnosis and evaluation of treatment response has been widely documented, there is a need for standardization both in the acquisition and the analysis in order to extend its benefit to multicenter clinical trials.

Functional MRI

Functional MRI (fMRI) uses conventional MRI equipment to measure the degree of brain activity in relation to the level of oxygen consumed in the blood of each brain region. The technology rests on the scientific rational that an area of the brain is more active if there is more oxygen consumption. The differing concentrations of deoxygenated and oxygenated hemoglobin in the blood in these areas can be detected by the magnetic field in the MRI scanner. With sophisticated software and post processing of the raw imaging data, a color spectrum gradient representing the more active versus less active brain areas can be produced. As one can understand, the applications of the technique are far ranging as subjects in the MRI scanner can be given a variety of stimuli or be requested to speak or produce a movement to determine which area of the brain is more active in the response or action. The resulting research in this field has amounted to an evolution in the understanding of the brain's function in intricate cognitive processes such as deception [10].

MRI Summary

One of the biggest advantages of MRI is that MR systems do not use ionizing radiation as opposed to other popular imaging modalities such as X-ray and CT. In addition, MRI has the ability to acquire and produce images in any plane. Moreover, information from a MR image is dependent upon the intrinsic properties of tissue, thus by varying different imaging parameters, one can produce desired images conveniently and effectively.

Although MRI is useful for noninvasive examinations and has become popular worldwide, it is not without disadvantages. The strong magnetic field is dangerous in the presence of ferromagnetic materials, such as some pacemaker implants. Other safety precautions include that the MRI scanner is very noisy during the process and the RF-induced thermal effects. Lastly, subject is required to hold still during scans,

and total scans time can be lengthy, and therefore not optimal for patient comfort, especially for those who are claustrophobic.

Nuclear Medicine Imaging

Positron Emission Tomography

Positron emission tomography (PET) is a nuclear imaging modality for producing images based on the spatial distribution of biochemical tracers within biological systems. It provides the ability to noninvasively and quantitatively determine metabolic activity by evaluating the location and uptake of a radioactive isotope and thereby indirectly behaving as an internal radiation source. The radiotracer consists of a radioactive isotope that has been chemically attached to, or incorporated into, some pharmacological relevant molecule. The most commonly used radiotracer in current clinical practice is fluorine-18 fluorodeoxyglucose (^{18}F-FDG). The ^{18}F-FDG is a glucose analog that accumulates in regions having high metabolic activity such as the brain, liver, and malignant tumors. Therefore, the accumulation (or uptake) of ^{18}F-FDG is directly related to the tissue's metabolic state, and an abnormal increase in uptake would indicate the presence of malignant tumor cells. Another increasing use of PET is amyloid imaging in neurology. Florbetapir ^{18}F is the first FDA approved PET amyloid tracer for the evaluation of Alzheimer's disease [11].

To obtain a PET image a radiotracer is injected into or inhaled by a subject prior to laying down in the scanner. The radiotracer has to be carefully evaluated and developed to have a "reasonable" half-life during which a sufficient number of positrons will be emitted as the product decays. Each radioactive atom in a PET radiotracer will spontaneously emit positrons over time. A positron emitted from a decaying nucleus travels a short distance and before colliding with a nearby electron. This results in an annihilation process that produces two gamma photons, each with energy of 511 keV, traveling at 180° apart. The pair of photons, after traveling in opposite directions along a straight line through the human body, can be recorded by radiation detectors [12].

The detectors are usually arranged in a circular fashion (i.e., a ring scanner), which fully encircle the patient lying in the axial field of view of the scanner. If two detectors receive photon signals within a very narrow time interval (i.e., within several nanoseconds), it is regarded as a coincidence event from the annihilation process. For each coincident pair, a line connecting the two detectors is then recorded as a line-of-response (LOR). The number of coincident events along the LOR between two detectors corresponds to the amount of radioactivity along the path. As the scanning process continues, many LORs are collected [12]. After collecting enough data, which usually means that millions of coincidence events are detected, the 2D or 3D image of radioactivity distribution within the body can be reconstructed (Fig. 1.12a, b).

Since PET provides functional images that have the ability to reveal changes in biological processes within the human body which usually precede any anatomical

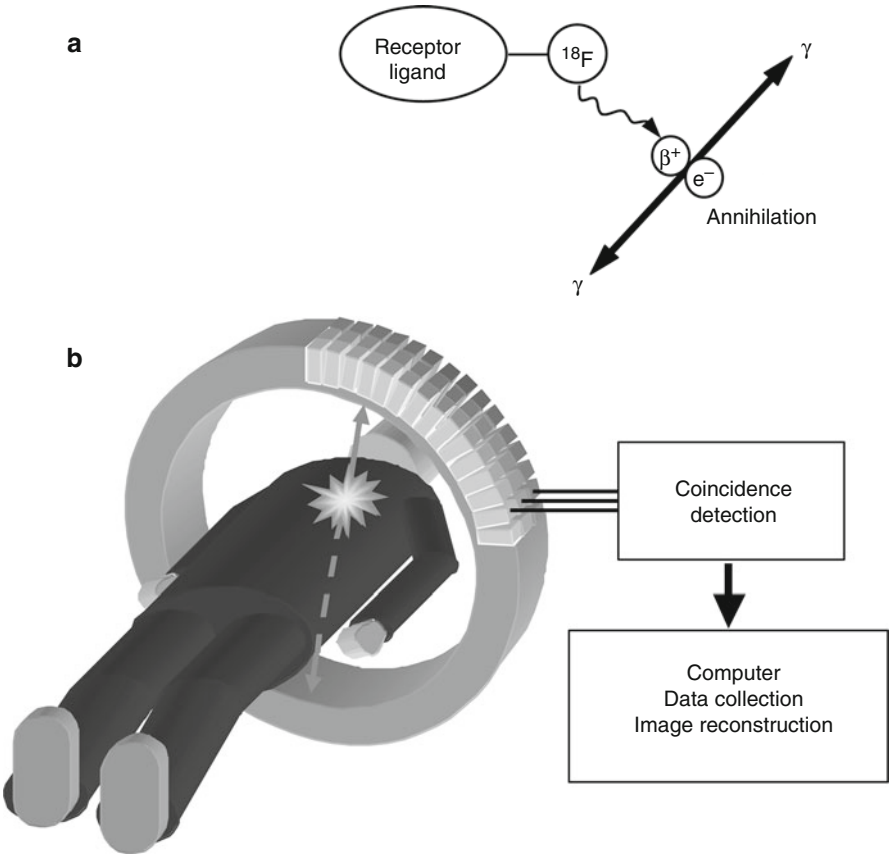


Fig. 1.12 (**a**, **b**) Illustration of a PET detector (Used with kind permission of Springer Science + Business Media from Johnström et al. [25])

evidence of abnormality, PET has a unique advantage over other anatomical or structural imaging modalities such as CT or MRI. PET imaging has been applied to fields including oncology (e.g., capturing increased metabolism in cancer and tumor), neurology (e.g., labeling amyloid plaques in Alzheimer's disease), and cardiology (e.g., detecting myocardial ischemia in coronary heart disease).

However, PET has limited spatial resolution as compared to CT and MRI. Factors that affect PET image spatial resolution and quality include detector size, random events, photon scatter, and attenuation. To address these issues of lacking accurate anatomical information, multimodality hybrid imaging systems such as PET/CT scanners have been developed. A PET/CT scanner (Fig. 1.13) is a PET scanner integrated with a multi-slice CT scanner. The patient receives both PET and CT scans within the same session without any transportation between scanners. This minimizes patient motion between scanners and facilitates better co-registration between the acquired PET and CT images. The CT image can also be used to correct for artifacts in PET data. The integration of PET and CT into one scanner provides complementary functional and

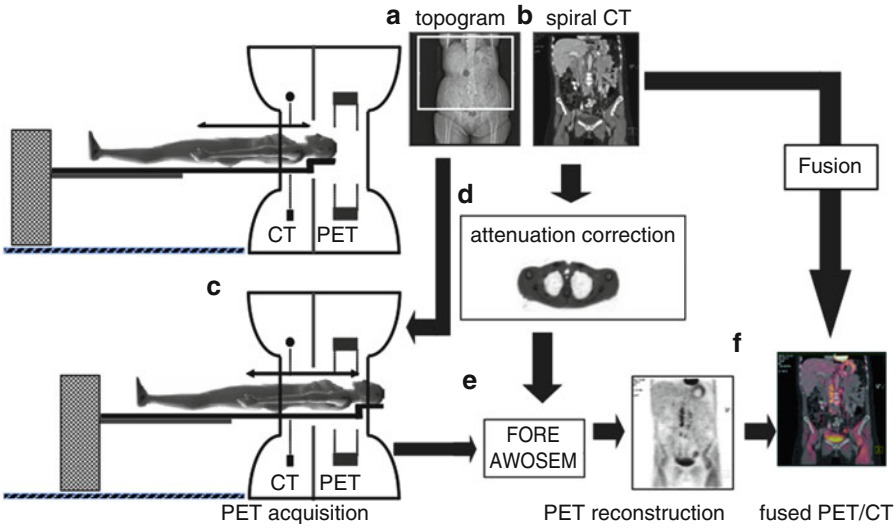


Fig. 1.13 A typical imaging protocol for a combined PET-CT study includes: (**a**) a topograph or scout image is produced for positioning, (**b**) a full volume CT image is produced, (**c**) a PET image is produced by use of a PET scanner covering the same volume as the CT image, (**d**) the CT is then corrected for attenuation. (**e**) a reconstruction algorithm (such as Fourier rebinning via attenuation weighted ordered-subsets expectation-maximization FORE+AWOSEM) is applied producing a reconstructed PET image, (**f**) which is then fused with CT to result in the final product of a PET/CT. (Used with kind permission of Springer Science + Business Media from Townsend [26])

anatomical information and has motivated the development of other multimodality imaging techniques such as PET/MRI. Figure 1.14 illustrates a typical PET/CT scan.

It should be kept in mind that there are a number of steps that require attention for the PET imaging in practice, including choice of isotope, isotope transportation, patient preparation, dose to be administered, and many aspects of image acquisition, reconstruction, and correction.

Single-Photon Emission Computed Tomography

Single-photon emission computed tomography (SPECT) is a nuclear imaging technique similar to PET in its use of radioactive tracer and detection of gamma rays. In a SPECT scan, a radionuclide is injected to the patient's bloodstream and absorbed by tissues. SPECT isotopes emit one gamma ray, while PET isotopes emit two. As the gamma camera rotates around the patient, it detects the emitted photon (gamma rays) and translates them into cross-sectional image. A 3D dataset can be then reconstructed from multiple 2D projections, similar to other tomographic techniques.

The radioactive tracers typically used in SPECT include iodine-123, technetium-99 m, xenon-133, and thallium-201. The most commonly used is technetium-99 m because it is considered a pure source of gamma ray with relatively short half-life

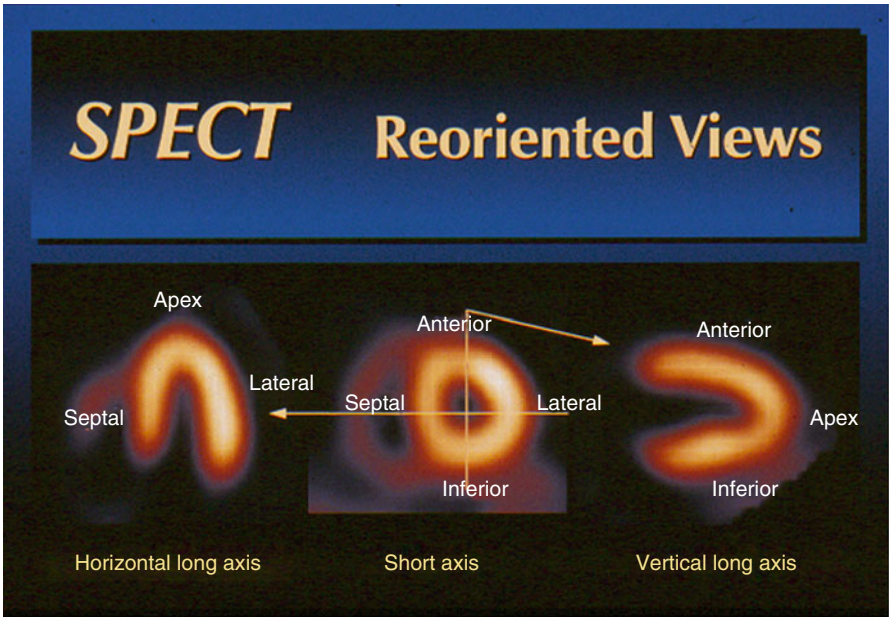


Fig. 1.14 SPECT: Myocardial perfusion scans (Used with permission from Braunwald [27])

(6.6 h) and can be chemically bound to a variety of compounds. Since the tracer travels in the bloodstream, it shows areas of blood flow. Concentrations of the tracer are highest in regions with the most blood flow, or highest level of metabolism. Normal tissues have high metabolic activity and higher blood flow, thus they appear brighter on a SPECT image [12].

A SPECT scan is less expensive than a PET or CT scan. SPECT have been utilized to detect diseases with changes in blood flow to the brain, bone, or heart, such as cardiovascular disease, brain injury, and epilepsy. Figure 1.14 shows reoriented views of myocardial perfusion images.

Ultrasound Techniques

Grayscale Ultrasound

Grayscale ultrasound, as the name suggests, is nothing more than a sound wave that is propagated through an object to a detector. For medical ultrasound, the normal frequency range is 2.5–12 MHz depending on the depth wishing to be penetrated and the material through which it is passing. In its basic format a transducer probe emits sound waves produced by the piezoelectric effect after a shot of electrical energy is generated. The sound waves are transmitted and reflected back towards the same probe which acts as an emitter and receiver. The sound waves are converted back to

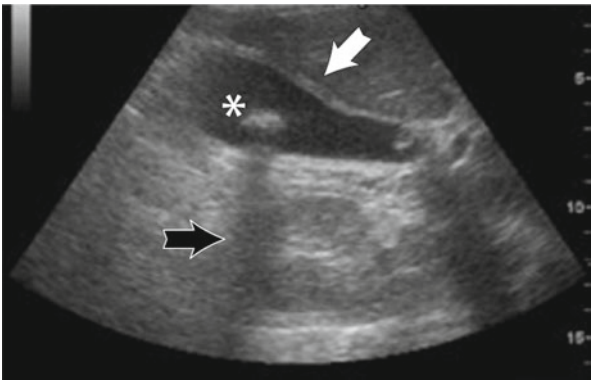


Fig. 1.15 Gallbladder with gallstones. The solid *white arrow* indicates the gall bladder wall, the * indicates a gallstone, and the *black arrow* indicates a shadow as a a result of the echogenic area produced by the gallstone. (Used with kind permission of Springer Science + Business Media from Duncan and Riall [28])

Table 1.4 Applications, advantages, and disadvantages of grayscale and Doppler ultrasound

Applications	Cholecystitis, appendicitis, pancreatitis, ectopic pregnancy, pelvic masses, aortic aneurysms, deep vein thrombosis
Advantages	Avoids use of ionizing radiation and therefore no biological damage, portable, mobile use; can be performed at bedside, less expensive than CT or MRI, can depict bodily organs in motion (i.e., heart), flexible imaging orientation view (transaxial, oblique, sagittal), no contraindications to use, real-time images
Disadvantages	Operator dependent, poor resolution, interpretation requires specialist, requires use of unpleasant gel on patient body

electrical signals which can then be processed to produce an ultrasound image based on physical density of the tissue and the time for the reflected signal to return to the detector [2]. The modern transducers are now an array of transducers lined up on the probe and can be set up to fire and receive in a pattern, and the resulting image analysis provides the operator with significantly more clarity than the early "A," "B," and "C" mode scanning [1].

Grayscale ultrasound waves, like the sound waves, undergo the same basic physics of deflection, reflection, and attenuation (or decrease in signal). Changes in density between two different materials provide a boundary from which sound is reflected (e.g., sound in air bounces off a wall and hence an echo). Likewise, echogenicity, or brightness on ultrasound, varies according to the medium to which the sound waves reflect upon [4]. Soft tissue will produce medium echogenicity, while fat is more echogenic than soft tissue. However, fluid will appear much less echogenic and actually appears dark or anechoic; it transmits ultrasound very well. Bone is not easily penetrated by ultrasound waves, and therefore only a thin rim outline of the bone can often be depicted. When ultrasound waves strike calcification or metals, the sound waves are blocked from going any further. As result, a highly echogenic area is seen as a shadow seen beyond the area (just like a shadow is cast in bright sunlight). This is commonly seen in the diagnosis of gallstones (Fig. 1.15).

Air, as contained in the bowels or lungs, is a second obstacle to ultrasound imaging as it does not reflect the sound waves back to the probe, and therefore an anechoic

image is depicted. Common bodily obstacles to ultrasound imaging include bowel gas and lung tissue.

Doppler Ultrasound

Doppler ultrasound uses the same physics principle to determine the presence, direction, and velocity of a moving fluid. If a sound wave is moving, the sound wave gets "stretched" when it is moving away from a receiver and compressed when it is moving towards the receiver. This is the so-called Doppler effect. We can observe the Doppler effect when an emergency vehicle passes us by on the road with its siren going. The tone of the siren changes from when it is coming towards you compared to when it is retreating. In medical imaging, this is most commonly applied to blood flow through vessels or organs such as the heart. Color can be applied to the direction of flow with red denoting flow towards the point of the probe and blue denoting flow away from the point of the probe. A common application to this technology can be seen in carotid ultrasounds which measure the disturbance of normal flow by stenosis [1].

Echocardiography

Echocardiography is an evaluation of the cardiac structures and motion by insertion of a probe through the esophagus (transesophageal) via the chest wall surface (transthoracic). Given the close proximity of the heart to the esophagus or chest wall, an ultrasound beam or variation thereof can produce a still image or video that displays not only cardiac wall, valves, and vasculature but also the motion of these structures as the organ works in unison to pump blood. Therefore, the cardiac cycle can be examined for pathology such as valvular regurgitation of blood, through Doppler ultrasound, or cardiac-associated pathology such as the accumulation of fluid around the heart. Advancements in the field have led to the use of 3D echocardiography which utilizes a multiple set of transducers and advanced processing system to yield a 3D view of the heart in multiple planes [1].

Bone Ultrasonometry

Bone ultrasound was a technique developed in the 1980s to assess bone fragility or osteoporosis [13–15]. Unlike the other ultrasound techniques previously described, this was a transmission methodology, at a relatively low frequency, nominally 1 MHz or lower. At this frequency it was found that sufficient signal was produced to measure and the attenuation of the signal correlated to the bone quality and quantity. Most systems assess the calcaneum (heel bone) and measure the broadband ultrasound attenuation (BUA) from 0.2 to 0.6 MHz or the speed of sound (SOS) through the bone. While there have been a number of variations, the mainstay has been

devices centered on the calcaneum using the techniques described. However, due to the precision being poorer than that of DXA, it has not been an acceptable technique for use in clinical trials nor has the FDA embraced its use.

Radiation Dosages Among Imaging Modalities

As mentioned previously, ionizing radiation, or radiation, is a disadvantage among the modalities of CT, X-ray, DXA SPECT, and PET/CT, while MRI and ultrasound carry the advantage of no radiation due to their technical nature. Medical professionals are trained to consider the clinical care benefits of an examination involving radiation in comparison to the risks. In the setting of a clinical trial, patients must be carefully informed of these risks versus benefits through the process of informed consent. Radiation dose and the relative risks are considered in more detail in Chap. 3.

Conclusion

Medical imaging has widely developed to be a particularly effective and necessary technique for clinical care among medical professionals. Its vast advantages throughout the multiple modalities have spurred its use as a biomarker in clinical trials for numerous endpoints for efficacy and safety and as an eligibility requirement. Regulatory authorities have recognized this benefit and acceptable imaging endpoints. This is highlighted by the recent draft guidelines that have been released on the role of medical imaging and the recommendations that govern its use in clinical trials [16]. This guidance will be discussed in more detail in Chap. 4. Medical imaging is continuing to play an increasing role in the development of new therapeutic agents. The so-called molecular image of PET, MRS, and other novel techniques is being used to assist in the early stage development for "go/no-go" decisions. In Phase III and IV there are certain techniques which are now a "mainstay" and continue to be refined and rest primarily on the modalities described in this chapter given their historical value.

While this chapter provides the clinical trial professional with a very high-level summary of the techniques available and the basic terminology, further texts with more advanced and detailed content are available for the imaging specialist or those involved more intricately in imaging clinical trial design.

References

1. Brant WE, Helms CA. Fundamentals of diagnostic radiology. 3rd ed. Philadelphia: Lippincott, Williams & Wilkins; 2007. p. xix. 1559 p.
2. Novelline RA, Squire LF. Squire's fundamentals of radiology. 5th ed. Cambridge: Harvard University Press; 1997. p. xiv. 621 p.

3. Mettler FA. Essentials of radiology. 2nd ed. Philadelphia: Elsevier/Saunders; 2005. p. xi. 416 p.
4. Pearson D, Miller CG. Clinical trials in osteoporosis. 2nd ed. Nottingham/Newton: Springer; 2007. p. xiii. 292 p.
5. Erkonen WE, Smith WL. Radiology 101 : the basics and fundamentals of imaging. 3rd ed. Philadelphia: Wolters Kluwer/Lippincott Williams & Wilkins; 2010. p. vi. 376 p.
6. Haacke EM. Magnetic resonance imaging: physical principles and sequence design. New York: Wiley; 1999. p. xxvii. 914 p.
7. McRobbie DW. MRI from picture to proton. 2nd ed. Cambridge/New York: Cambridge University Press; 2007. p. xii. 394 p.
8. Wood ML, et al. Physical MR desktop data. J Magn Reson Imaging. 1993;3(Suppl):19–24.
9. Yu HJ. Imaging markers of cognitive deficits in RRMS using a multimodal MR approach. In: Biomedical engineering. Stonybrook: Stony Brook University; 2011.
10. Mohamed FB, et al. Brain mapping of deception and truth telling about an ecologically valid situation: functional MR imaging and polygraph investigation–initial experience. Radiology. 2006;238(2):679–88.
11. FDA. FDA approves imaging drug Amyvid April 10, 2012; Available from: http://www.fda.gov/NewsEvents/Newsroom/PressAnnouncements/ucm299678.htm.
12. Bushberg JT. The essential physics of medical imaging. 3rd ed. Philadelphia: Wolters Kluwer Health/Lippincott Williams & Wilkins; 2012. p. xii. 1030 p.
13. Langton CM, Palmer SB, Porter RW. The measurement of broadband ultrasonic attenuation in cancellous bone. Eng Med. 1984;13(2):89–91.
14. Porter RW, et al. Prediction of hip fracture in elderly women: a prospective study. BMJ. 1990;301(6753):638–41.
15. Hans D, et al. Ultrasonographic heel measurements to predict hip fracture in elderly women: the EPIDOS prospective study. Lancet. 1996;348(9026):511–4.
16. FDA. Guidance for industry standards for clinical trial imaging endpoints. 2011. Available from: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm268555.pdf.
17. Levine D, Akikusa J, Manson D, Silverman E, Schneider R. Chest CT findings in pediatric Wegener's granulomatosis. Pediatr Radiol. 2006;37(1):57–62.
18. Gupta N, Moore JP, Shannon K. A novel approach to eliminate intraventricular lead placement in patients with congenital heart disease. J Interv Card Electrophysiol. 2012;35(1):115–8.
19. Aberle D, El-saden S, Abbona P, Gomez A, Motamedi K, Ragavendra N, Bassett L, Seeger L, Brown M, Brown K, Bui AAT, Kangarloo H. A primer on imaging anatomy and physiology. In: Bui AAT, Taira RK, editors. Medical imaging informatics. New York: Springer Science + Business Media; 2010.
20. Zhang J, Yang G, Lee Y, Chang S, Lu J, Zhu O. Multiplexing radiography for ultra-fast computed tomography. Lay-language papers. AAPM meeting. Minneapolis; 23 July 2007.
21. Semrud-Clikeman M, Fine JG, Bledsoe J. Neuroimaging in women. In: Fletcher-Jansen E, editor. The neuropsychology of women. New York: Springer Science + Business Media; 2009.
22. Saha GB. Physics and radiobiology of nuclear medicine. New York: Springer Science + Business Media; 2013.
23. Ridgway JP. Cardiovascular magnetic resonance physics for clinicians: part I. J Cardiovasc Magn Reson. 2010;12(71):12–71. doi:10.1186/1532-429X-12-71#_blank.
24. Scheef L, Frank T. Theoretical background of MR imaging. In: Boecker H, Hillman CH, Scheef L, Strüder HK, editors. Functional neuroimaging in exercise and sport sciences. New York: Springer Science + Business Media; 2012.
25. Johnström P, Fryer TD, Bird JL, Richards HK, Davenport AP. Dynamic in vivo imaging of receptors in small animals using positron emission tomography. In: Davenport AP, editor. Receptor binding techniques. New York: Humana Press; 2012.
26. Townsend DW. Design and operation of combined PET-CT scanners. In: Shreve P, Townsend W, editors. Clinical PET-CT in radiology. New York: Springer Science + Business Media; 2011.

27. Braunwald E (Series Editor), Beller GA (Volume Editor). Atlas of heart diseases, vol. 5: Chronic ischemic heart disease. Philadelphia: Current Medicine; 1995.
28. Duncan CB, Riall TS. Evidence-based current surgical practice: calculous gallbladder disease. J Gastrointest Surg. 2012;16(11):2011–25.

Chapter 2
The Metrics and New Imaging Marker Qualification in Medical Imaging Modalities

Colin G. Miller

Abstract There are four key uses for medical imaging: screening, diagnosis or prognosis, monitoring the natural history of the disease, and monitoring therapeutic intervention. There are eight key metrics when evaluating a biomarker or imaging technique which has to be put into context with the key use of the system. This chapter will describe these key metrics in this context with particular emphasis of the use in clinical trials.

Keywords Sensitivity • Specificity • Precision • Accuracy • Reliability • Safety • Cost-effective • Acceptability

Introduction

Recent developments in technology have changed medical imaging from a basic X-ray system evaluating three dimensional anatomy on a two dimensional X-ray film to multidimensional assessments (4D) with computer-aided diagnosis (CAD). This is challenging radiologists to evaluate data in unique ways that had not been envisioned even a decade or two ago. However, the same basic fundamentals or metrics of medical research have to be considered and evaluated when assessing which mode of medical imaging will be used in a clinical trial. These principles apply whether the imaging is a 2D X-ray film or a four dimensional reconstruction on the latest picture archive and communications system (PACS) or CAD system.

C.G. Miller, BSc, PhD, FICR, CSci
Medical Affairs, BioClinica, Inc.,
826 Newtown-Yardley Road, Newtown, PA 18940, USA
e-mail: colin.miller@bioclinica.com

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_2,

The aim of this chapter is to explore how to fully evaluate an imaging modality before it is introduced into a clinical trial and to understand the primary metrics that are driving the application for the end point selected.

The first fundamental principle is to understand that all imaging can be used in one of four primary ways:

- Screening
- Diagnosis and or prognosis, (or disease predictor or assessment of disease severity from which one can determine a prognosis)
- Monitoring the natural history of the disease
- Monitoring therapeutic intervention

The understanding of the concept of the four uses of imaging is paramount in clinical trials. Most healthcare practitioners are primarily utilizing imaging for disease screening or for diagnostic purposes in the clinical setting. The clinical trial research setting requires that medical images may be used for diagnosis if there is a radiological eligibility criteria, but the primary objective in the clinical trial is in assessing the therapeutic intervention. Furthermore, in placebo-controlled trials, imaging will also be used for monitoring the natural history of the disease for the patients in the placebo group. This distinction in use of radiology and the evaluation of the images is a fundamental difference that is often poorly understood and leads to confusion, i.e., the appreciation that there is a difference between clinical practice and clinical trial practice.

There is, furthermore, a key semantic difference in the word "screening" that requires clarification at this stage: Screening in clinical practice is applied to procedures such as mammography where patients are being evaluated for the presence of atypical images characteristic of an underlying malignancy. In clinical trials screening refers to the procedures that are conducted to determine a subject's eligibility for participation in a clinical trial. This may include radiological evaluation to verify a particular diagnosis and to determine disease severity.

In routine clinical practice medical imaging is used predominantly in a qualitative manner. Does the patient have the disease or anatomical variant that correlates with the symptoms? Is the patient's condition improving? Do the images correlate with the clinical symptoms being described? Are the lesions getting bigger or smaller? Is the fracture healing? etc. Evaluations are performed to treat the individual patient. Clinical practice focuses on a series of encounters with single patients.

While many different types of clinical trials are performed to enhance medical knowledge, the conduct of clinical trials to support a new drug application (NDA) or Biologics License Application (BLA) requires adherence to specific standards.

It is these types of clinical trials that will be discussed. One requirement for these clinical trials is that the results have to be quantitative, which produces further requirements on the methodologies employed. Medical imaging can then be split into two groups of techniques:

Type 1: Techniques which are quantitative at the point of collection, e.g., DXA, PET, and QCT
Type 2: Techniques which are qualitative and from which quantitative measurements are obtained, e.g., X-ray, CT, ultrasound, and MRI

Type 1 measurements are obtained where the instruments or imaging equipment are providing a direct quantitative output. This therefore requires some form of instrument calibration, and the use of phantoms for monitoring this aspect of the output to ensure any changes seen in the patient measurements are real and not an artifact caused by instrument calibration changes.

Type 2 measurements are often linear measurements of some part of the anatomy, such as area, length, and volume. The challenge as technology has developed is that these basic measurements may be more complex, and post processing software algorithms may not all produce the same absolute results. Furthermore each manufacturer has developed their own internal processes that provide an already highly processed image by the imaging equipment.

Therefore one of the challenges in clinical trials is how to ensure that the measurement, whether it comes directly from quantifiable data (Type 1 discussed previously) or from the imaging in a derived parameter, is accurate. This is one of the concepts that will be considered in more detail later in this chapter.

By whatever means the measurements are obtained; eight basic criteria have to be weighed and balanced before any assessment from medical images are evaluated in a clinical trial (or detailed in the protocol). The instruments have to be:

1. Discriminative between health and disease or have acceptable sensitivity and specificity
2. Acceptably precise and accurate
3. Reliable
4. Relevant
5. Acceptable to regulatory agencies
6. Of acceptable cost to the trialist
7. Acceptable to the subject
8. Safe for the subject and operator

For many imaging modalities that are used in trials, these parameters are not routinely evaluated. In some cases the measurements have become so

commonplace, and the aforementioned parameters are known and have been acceptable to the regulatory agencies. An example would be plain film radiographs of the hands or feet for the evaluation of rheumatoid arthritis.

However, a closer review of these criteria may ensure that trials are conducted more efficiently, even with known instruments. A case in point arose a few years ago with the random zero sphygmomanometer used to measure blood pressure that caused error in several clinical trials [1]. In this example the Hawksley sphygmomanometer was provided to each site to standardize the enrollment and follow-up of patients in a clinical trial in hypertension; however, it systematically underestimated blood pressure. Only after the study was completed was it apparent that the sites were not trained in the use of an initial calibration of the sphygmomanometer, and therefore, the patients being recruited into the study were not all consistent, and the eligibility criteria were not met in many cases. The data from the clinical trials were therefore suspect. Care on behalf of the trialists and the contracting out to specialists in the field may have prevented the problem. However, the precedent has been set for proper quality control and quality assurance for instruments in clinical trials, which in some instances is now the expected approach by the regulatory agencies [2, 3]. The same questions have arisen around intima-media thickness (IMT) measurements after the results of the Vytorin study came in negative [4].

Discussion of these eight criteria will allow the trialist to understand a little more regarding the factors affecting clinical measurements. Adherence to these principles should also reduce the potential error in instruments in clinical trials.

Discrimination or Sensitivity and Specificity

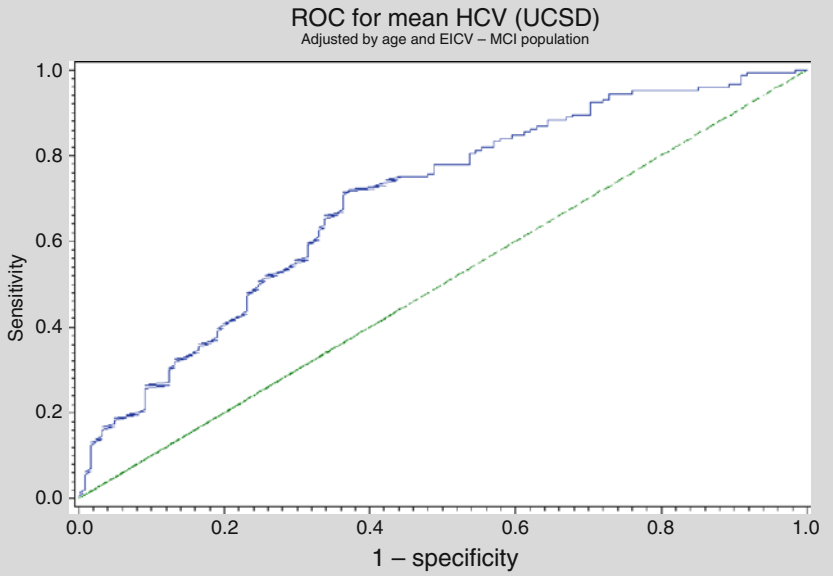
This is the ability of a technique, be it imaging or a biochemical marker, to distinguish between disease and non disease, or between stable disease and progression of disease. The assessment of this is usually evaluated in terms of sensitivity (true-positive) and specificity (true-negative) and can be shown graphically in a ROC analysis or Receiver Operator Characteristics curve.

ROC Analysis

Receiver Operator Characteristics or ROC analysis was originally developed by the British in World War II as signal detection theory to evaluate radar. For the application of radar, it was needed to determine the difference between enemy aircraft and other signals. Essentially with any technique designed to distinguish between disease and health or "normality," there is a need to evaluate how good the technique in differentiating between the two states, in its simplest form. If we consider a binary process, disease or no disease, then a truth table can be built.

Predicted outcome		
Actual outcome	*Positive*	*Negative*
Positive	True-positive	False-negative
Negative	False-negative	True-negative

With this concept, we can then calculate the sensitivity and specificity of the technique. The sensitivity is the true-positive rate, and the specificity is the true-negative rate. A ROC curve is a presentation of sensitivity v 1-specificity or false-positive rate. A perfect test will have a 100 % true-positive and true-negative value. A test which has a poor discrimination will have 50 % false results and 50 % true results. This is seen as the line at 45° on the ROC curve. The better a technology or measurement, the closer the curve to the top left corner of the graph. One way two techniques can be compared is to calculate the area under the curve (AUC).



A free web-based application for evaluating ROC curves can be found at: http://www.rad.jhmi.edu/jeng/javarad/roc/JROCFITi.html

However, before we discuss this further, a second differentiation has to be considered: whether the assessment of disease is based on truly quantitative assessments (like a linear measurement, such as joint space narrowing in osteoarthritis or brain volume in Alzheimer's disease) or a radiological interpretation of an image into a semiquantitative scoring system, like the Genant score for vertebral deformity [5] or Sharp score for rheumatoid arthritis [6].

In general (and this is a generalization, since there are examples that do not follow this distinction), the greater the numerical spread, the "better" the instrument at differentiating between treatments whether it be between active comparators or between active and placebo, particularly for measurement techniques. The definition of normal is different for each therapeutic area. For example, when considering bone densitometry as assessed by dual energy X-ray absorptiometry (DXA), bone ultrasonometry, or quantitative computed tomography (QCT), "normality" is defined by a population of young healthy individuals aged between 20 and 40 without any history of bone disease or medication usage likely to affect bone. The "normal" population should also be drawn from a geographically diverse population to avoid local regional differences. Owing to phenotypic variation, a separate population needs to be assessed for each of the major ethnic groups (e.g., Caucasian, African, and Asian) which has to be further divided by gender. The accepted definition of osteoporosis as defined by the World Health Organization (WHO) is an individual having bone mass less than 2.5 standard deviations (−2.5 SD) below the young normal mean [5]. The definition is also specific to the anatomical area being imaged with different values meeting the criteria for osteoporosis at the hip and at vertebrae. This has come about due to the inverse relationship between DXA measurement and the age-related increase in fracture risk. Once past the menopause there is a decrease in bone mineral density caused by an uncoupling of bone remodeling due to the decrease in circulating estrogens in women. There is a more gradual but age-related loss of bone in men due to the decrease of circulating sex steroids.

By the very definition of the disease being based on a normal population curve, as individuals stay alive they develop age-related bone loss, and a larger percentage of the population will become osteoporotic. However, it is critically important to appreciate that normality is gender, race, and anatomy specific, which creates further challenges when combining populations in clinical trials. However, the concept still remains – the greater the difference between normal and abnormal, the better the overall ability to diagnose with clarity but more importantly in clinical trials, to see a change caused by therapeutic intervention. This is further related to the precision of the measurement which will be discussed further in the next section.

With the so-called semiquantitative measurements, a scaling and graduation scoring system has to be used which is generally a numerical system with 0 being normal or healthy. The common ones are often used in musculoskeletal disease, as already mentioned. The challenge is to have a scoring system which has sufficient granularity to distinguish change by therapeutic intervention or to demonstrate worsening disease without being so complex as to make the scoring system unusable. A great example is the Sharp scoring system, modified by van der Heijde for

the evaluation of joint space narrowing and erosions in rheumatoid arthritis [7]. On initial view, it seems very complex, with a total score of 428 (280 for maximum erosion score and 168 for maximum joint space narrowing score), with a score of 0 being a subject with no observable disease radiographically on hands and feet radiographs. However, the score is broken down into a 0–5 (in most instances) for each of 26 joints per hand and feet being evaluated for erosions and again for joint space narrowing, with a slightly different evaluation. On a joint perspective, this makes a relatively simple and reliable scoring technique. Furthermore while there are reader interpretive differences on a per visit basis, these are decreased by evaluating the "change score" between time points. An alternative method that has been used in a number of trials is the Genant score, which has many similarities to the novice evaluator [8]. However the scoring system is reduced to a total of 312 (176 and 136 for maximum erosions and joint space narrowing, respectively). On initial review this would suggest a decrease in sensitivity to detect change, but a potential increase in the reproducibility of the readers to score. There is a subtle twist – the score can be in 0.5 increments! Debate regarding the merits of each scoring system has existed and will continue to do so, and it takes the trialist carefully evaluation as to which is the most pertinent to their trial for other reasons. After all, these scoring systems are not used for routine clinical practice. The final say probably now goes to the regulators; the European Medicines Agency (EMA) has come out with a new guidance for rheumatoid arthritis and stated that the van der Heijde modified Sharp scoring system should be used unless there is a good rational to use another methodology [9].

One further concept has to be considered when evaluating end point in clinical trials: the smallest detectable difference (SDD) or smallest detectable change (SDC). "The SDD expresses the smallest difference between two *independently* obtained measures that can be interpreted as "real" – that is, a difference greater than the measurement error" [10]. The SDC is a concept that has grown out of the field of rheumatoid arthritis, where there is a dichotomous group of subjects that can be observed in a clinical trial. The majority of subjects show little or no change in the 2 years of observation in their radiological scores, where there is a small subgroup which undergoes a large change in joint damage. Furthermore, since the reads are conducted with a number of time points presented all at one time, the images are not truly independent in the statistical sense. The read design and methodology will be covered more fully in Chap. 5. However, the key difference is that SDD is based on the concept that a patient's disease is progressing and SDC is based on the concept of how the much less disease is progressing. Either assessment is the determination of the degree of progression above the measurement error that can be statistically determined. In rheumatoid arthritis trials, Bruynesteyn and colleagues conclude that SDD over estimates the number of patients required for the study [10]. SDC has now been used for in other rheumatoid diseases such as ankylosing spondylitis [11]. However, when reviewing the literature using Google Scholar, SDD is the preferred methodological approach in all other therapeutic areas when it comes to radiological interpretations.

Precision and Accuracy

Precision is the term use to describe the reproducibility of the measurement. This is usually assessed by performing multiple repeated measurements using the same measuring instrument on the same patient, for example, measuring a patient's height or weight on an office scale. It is reported as the percentage coefficient of variation. The lower the %CV, the better the precision and the easier it is to detect small changes in the measurement. To perform true estimates of precision requires the repeat imaging of approximately 20+ subjects or patients. Each image must then be assessed by multiple readers. Due to the ethics of radiation dose for many imaging modalities except perhaps ultrasound and MRI, true precision measurements cannot be acquired. However, repeated measurements of the same images can be performed in order to acquire a reasonable estimate of precision.

Precision is not to be confused with accuracy, which is how close the measurement is to the actual quantity being measured. An example of the difference between precision and accuracy with target shooting is shown in Fig. 2.1. For clinical trials where the measurement is the primary inclusion or exclusion parameter, then the baseline measurement calls for high accuracy. At enrollment, a comparison is made of the individual to a normal reference population to assess the degree of disease. Precision then becomes more important for all future measurements to ensure they compare to baseline. Even if the assessment at baseline was inaccurate, but the

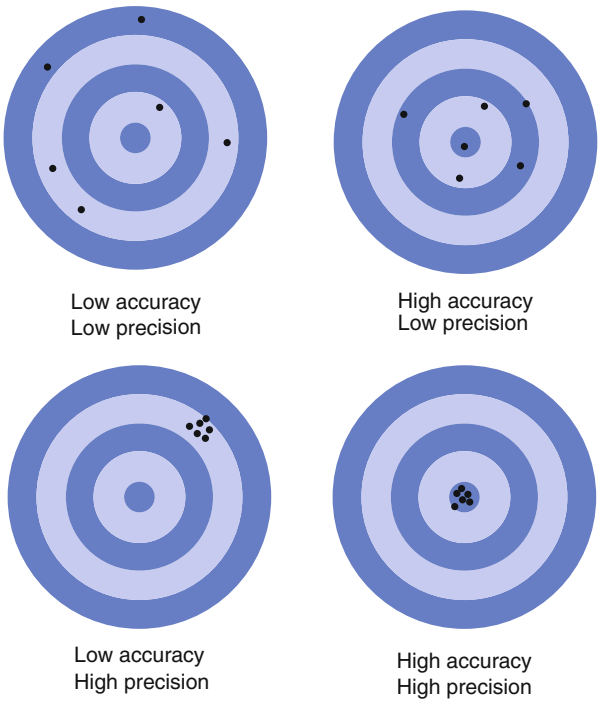


Fig. 2.1 A representation of precision and accuracy using targets

patient is included in the study, all future assessments need to be acquired as closely to the baseline measurement – precision is at the cost of accuracy.

Accuracy is very difficult to assess in many imaging situations. For example, with plain film radiographs there are beam divergence and magnification issues which means there has to be a reference or caliper in the field of view which is of known size. So taking the example of vertebral deformity, a washer or ball bearing should be placed on the skin at the level of a vertebral spinous process (T12 is often used, since it is common to both the lumbar and thoracic lateral spine films). The washer will allow an accurate measure of T12, but at the further proximal or distal column, measurements will be made where there is beam divergence which will create a more significant measurement error, beyond the 4 mm threshold which is usually used to define fracture. This issue is further compounded with patients with significant scoliosis which places the vertebrae higher up the spine in or out of the vertical plane from the caliper and changes the orientation from a true lateral image to a more oblique image for several vertebrae. Accuracy is truly difficult.

MRI is another example where accuracy is difficult to determine field distortion is just one of the issues that occurs. For measurements requiring high precision and accuracy, such as brain or hippocampal volume in Alzheimer's disease, distortion of the field may create artifactual changes, but which measurement is the most accurate will be unknown until death and biopsy can be used to correlate to the measurements. However, in this Alzheimer's disease example phantoms should be employed for the trial to monitor and evaluate these changes.

If there are no eligibility criterion requiring imaging, assessment then precision or reproducibility is the overriding parameter to consider. There is an inherent assumption that having selected the use of a particular imaging technique that the manufacturer has ensured that it measures precisely and accurately: this supposition may not be correct. With the sphygmomanometer example earlier, the problem would not have arisen if a calibration check had been performed prior to the start of the study.

Having separately discussed the concepts of precision and in the previous section discrimination, the two parameters cannot be evaluated in isolation. The poorer the discrimination, the higher the precision that is required to distinguish between cohorts or populations. Taking it to the extreme, an instrument with a precision of say 10 % would be of little value if the difference between normal and disease state was only 10 % or even 15 %. With that said, one of the characteristics of clinical trials is the use of groups of subjects to provide a good signal to noise ratio to detect a therapeutic signal amongst the biological noise. Precision has to be factored into the power calculations, which will determine the study size. If the precision, however, is very much less than standard deviation of population mean change, then precision is not the predominating factor. An appreciation for the difference between monitoring a subject's change to therapy or disease and the group effect of the treatment or disease, where in effect, there is signal averaging to elucidate the change in each treatment group. For example, in the field of osteoporosis, using DXA and the assessment of bone mineral density (BMD) the standard deviation of BMD in a group of subjects will be around 0.1 $g \cdot cm^{-2}$ with a mean of 1.0 $g \cdot cm^{-2}$, i.e., about 10 %. The long-term precision of most DXA equipment at the lumbar spine is between 1 and 3 % depending on the study population. While 10 % may not seem significant, this can

introduce a large cost burden, so this has to be factored into the study design [12]. Also this example was one where the precision is very much less than the standard deviation of the patient population. If precision is a larger percentage of the SD, e.g., with carotid intima-media thickness (IMT) measurements, then precision has the overriding influence and will significantly drive up patient numbers [13].

A parameter combining dynamic range and precision was developed in the field of bone ultrasonometry, called the standardized coefficient of variation [%SCV] [14]. This is the %CV multiplied by the dynamic range divided by the mean of the measurement. The smaller the %SCV, the better the measurement. Since the measure of precision is affected by the scale being used, it is impossible to compare instruments or imaging techniques using different scales. The %SCV allows for this comparison. Other statistical methodologies have also been proposed [15, 16] to overcome these issues.

Reliability

Reliability has two components: It refers to consistency in the properties of the imaging system (hardware, software, and radiologist or reader) that will provide a reproducible outcome with repeated use and the reliability of the surrogate or imaging biomarker on the end point in the clinical trial. These two aspects will be considered separately.

An imaging technique that cannot be reliability acquired is useless (a liability) in a clinical trial setting. Reliability is the assessment of the change in calibration, however caused. An example of calibration shifts is seen in Fig. 2.2. In this example it is of a DXA instrument calibration, but could be for any other system being assessed over time, e.g., reader calibration and change in MRI inhomogeneities. In this graph each point is a measurement of the same calibration phantom on the DXA scanner. Around September there is a downward shift in the mean calibration of the instrument by an average of 3 %. The calibration remains constant until January where there is an upward shift of 3 % followed by a gradual downward drift in calibration for the remainder of the graph. A subject who is scanned every 6 months starting in May will show a loss of bone mineral density (BMD) of 3 % at the November visit, followed by a nearly increase in BMD of 3 % the following May. After that, the drift is more difficult to characterize, and the effect on the subject BMD is uncertain.

Reliability as an evaluation of measurement consistency is therefore the assessment of minimal equipment failure and/or minimal change in calibration, the latter being particularly pertinent in Type 1 imaging equipment. The modern equipment that is used in any major radiology center is certainly going to be reliable since hospitals and clinics cannot afford to purchase equipment that is going to need constant service.

However, some academic groups may request the “latest and greatest” equipment which may not be the optimum methodology for clinical trials (cutting-edge technology for which more maintenance and calibration can be expected), e.g., the use of 7 T magnets which are not standard practice at the time of writing, and need significantly

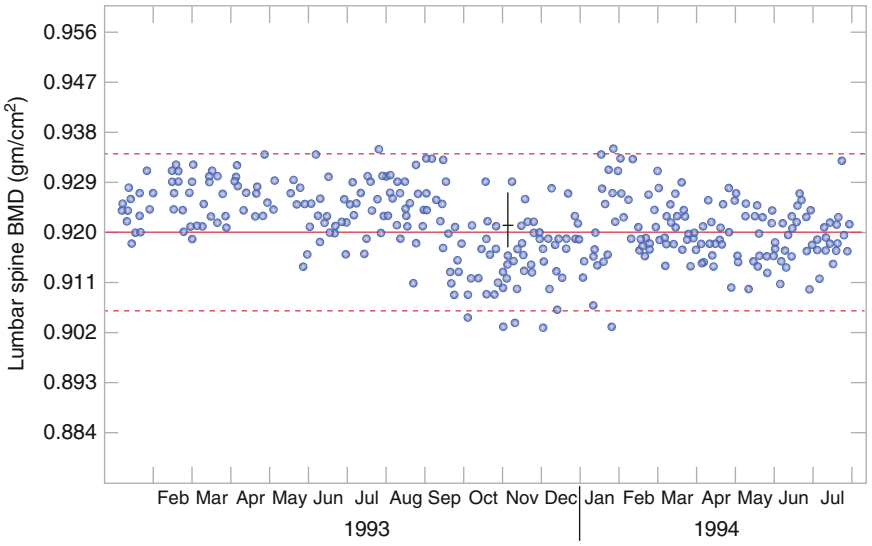


Fig. 2.2 Reliability of surrogate end points in assessing clinical outcomes (Courtesy of Bioclinica)

more maintenance than the standard 1.5 T or even 3 T magnets. All imaging technology will require service and eventual replacement, but the concept here is that the reliability is such that one has the confidence in the equipment that it will be working >95 % of the time and patient visits will not be lost due to service issues.

For quantitative imaging techniques, consistent calibration of quantitative imaging modalities (Type 1 imaging equipment) is a necessity to ensure the information coming from the images. This is conducted using calibration phantoms, generally measured on a daily basis. Merely measuring the phantoms is insufficient, and evaluation of the calibration output has to be performed. In a clinical trial the imaging core lab or central imaging vendor should be providing this as standard practice.

Reliability of the surrogate end point has been elegantly described by Fleming and colleagues [17] where the need for the imaging surrogate is to be on the causal pathway of the disease pathway and also the pathway in which any intervention will show change. If the surrogate (be it imaging or any other surrogate biomarker for that matter) lies on another pathway, then the inferences drawn will be incorrect. This is shown diagrammatically in Fig. 2.3a–e. This aspect of reliability is a critical determinant of the use of any imaging technique as an end point in clinical trials.

Relevant

Is the measurement going to provide useful clinical information? Does the instrument being used have the capabilities of detecting clinically relevant change over the specified time point in the study protocol? In situations where the study design

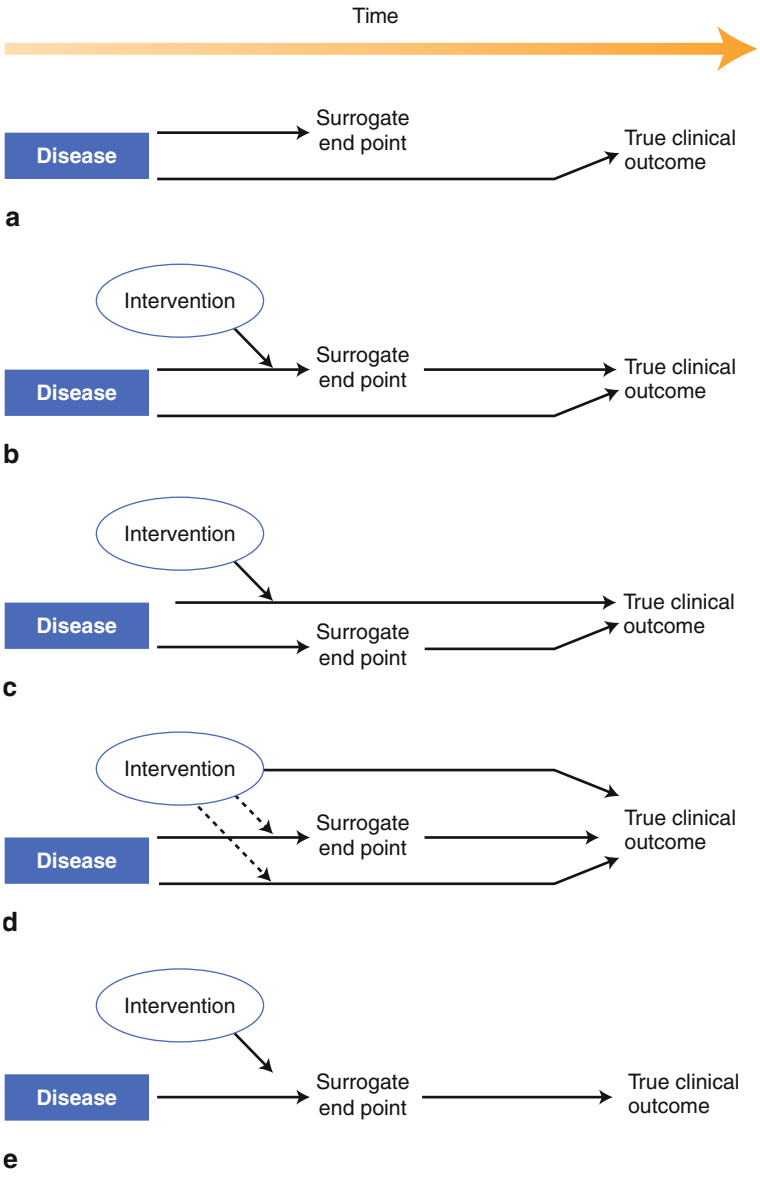


Fig. 2.3 Reason for failure of surrogate end points. (**a**) The surrogate is not in the causal pathway of the disease process. (**b**) Of several causal pathways of disease, the intervention affects only the pathway mediated through the surrogate. (**c**) The surrogate is not in the pathway of the intervention's effect or is insensitive to its effect. (**d**) The intervention has mechanisms of action independent of the disease process. *Dotted lines* = mechanisms of action that might exist. (**e**) The setting that provides the greatest potential for the surrogate end point to be valid (Used with permission from Fleming and DeMets [17])

is driven by non-imaging end points, modifications to the initial draft protocol can result in situations where the resulting imaging data will not be clinically relevant. For example, a draft protocol for a phase II trial in rheumatoid arthritis incorporates X-rays to assess the joints over 2 years with images being acquired every 6 months. Subsequently the trial is later shortened to 6 months since this is sufficient time to assess clinical signs and symptoms. Since very few subjects will demonstrate any change within 6 months, the study will be grossly underpowered for this end point and will not provide relevant information. Subjecting patients to additional radiation when there is virtually no constructive information to be gained is ethically inappropriate. Had the imaging modality been switched to MRI, then relevant data would have been generated over the 6-month study period.

Another situation where irrelevant data can be generated is when imaging occurs at anatomical locations that are not pertinent to the scientific question under investigation, or the assessment is statistically underpowered for the number of subjects. This has obvious ethical implications.

Accepted by Regulatory Agencies

If the study is for registration purposes and the data being collected are essential, then confirmation of the acceptability of the measurement is appropriate prior to study start, not at filing. Many measurements, however, can produce data that are useful supporting documentation, so that their use is appropriate in a well-designed clinical program. Most study sponsors are seeking to meet the rigorous requirements of the US Food and Drug Administration (FDA) and/or the EMA and will use methodologies that are licensed for diagnosis and monitoring by the relevant agencies. Again using examples from the field of osteoporosis, DXA and QCT are fully accepted methodologies for most agencies, although QCT is not the methodology of choice because of higher radiation dose to the patient. Historically the FDA has not accepted evidence based on QCT alone although it may be acceptable in a Phase II trial and an indication to move to the next stage of development, providing additional DXA data are available in the Phase III program.

The other aspect is that there may be imaging techniques that are fully approved for use in the clinical setting for certain indications, but the FDA will not accept the data for drug or biologic approval. Two such examples are (1) bone ultrasonometry for the identification of subjects at risk of developing osteoporosis and (2) MRI for the assessment of rheumatoid arthritis. For bone ultrasonometry, the precision is poorer than for DXA, but also it is not clear exactly what parameters of bone are being assessed by the technique. For the MRI example the FDA has stated that the clinical correlation is still not clear, and longer-term studies are required to show that MRI findings correlate closely with future disease and symptoms. However, many companies will use MRI in early Phase II rheumatoid arthritis studies to make

go/no go decisions on the new therapy or to elucidate the anticipated mode of action on both the soft and hard tissue.

Acceptable Cost

Acceptable cost is difficult to define since it will depend on the drug, its stage of development, and the proposed market in which it is to be used. An additional factor in Phase III studies is the likely cost of reimbursement when the drug is on the market. Historically, clinic/hospitals used clinical trial sponsors to obtain funding to purchase new equipment, sometimes directly for the particular trial. This is now fairly rare. One exception to this is ultrasound equipment or optical coherence equipment (OCT) in ophthalmology, both of which are still very manufacturer specific with a lot of variability between equipment. Therefore, if one wants to obtain quantitative output, a study sponsor may need to provide the equipment to the clinical sites. These are one of the few areas in 2012 where the author was involved in equipment procurement either for joint evaluation with ultrasound or ophthalmological evaluation in neurodegenerative diseases. With this exception, generally there is no reason for the sponsoring company to purchase such imaging systems.

The cost of imaging does have to be carefully considered within a clinical trial budget. If it is an early phase I or II study with a novel imaging end point, then it is highly likely only a few key academic centers will have the ability to obtain the end imaging required and so careful selection investigators can be chosen with access to such technology. This usually comes at a significant cost. An example was dynamic contrast-enhanced MRI or DCE-MRI. Until relatively recently this was a technique that only a few academic centers would consider providing. With the advancement of this technique and more sites being trained to conduct this technique, mainly by core labs offering to provide this service, DCE-MRI is becoming a little more "main stream."

The subject of cost of the assessment takes on another facet when considering investigator payments. Historically the radiology team was not involved in the trial and often did not know about the trial until the first patient turned up for the visit. It also meant that the radiology department did not receive reimbursement for the imaging or obtained funding through insurance, if it was a routine assessment. While this might have appeared to be a very cost-effective solution on the part of the sponsor, the short-term gain usually ended up with bigger long-term losses as subjects' images were not acquired correctly, and many time points were lost. To have good imaging, the radiology department needs to be involved in the study, at least to be informed and know that payments are associated with obtaining the images. If unusual or more images are required than would be used for routine clinical assessment, then the head of radiology needs to be aware of the study and assign one or two technologists to be the primary source of contact. One wants to minimize the number of technologists at each site to one or two to acquire all the scans on the patients to ensure standardization of images.

Acceptable to the Subject

There is only so much inconvenience and measurement that a subject will tolerate. This will vary between subjects considerably, but the way they are treated at the investigator site will also have a significant influence on the acceptability of the procedure. During a Phase I or Phase II trial where a battery of tests are being performed, the investigational team are usually highly involved with the trial and spend a great deal of time with each subject. In these situations, subjects are more likely to tolerate discomfort, particularly when they believe they are being altruistic for mankind. However, this is not the case in the vast majority of Phase III or IV studies, or in the routine clinical setting where the new molecular entity will be the anticipated treatment of the future. Therefore, it is essential that the measurements are acceptable to the subject who is normally required to undergo repeat evaluations at each visit. Poor tolerability to the measurement will lead to increased subject dropout and leave the results of the trial questionable. Good investigational staff at the site can make or break a trial in terms of acceptability for the subject. The point in the previous section about having a technologist or two dedicated to the study will aid this point considerably. Subject dropout in a study is an expensive proposition; therefore the cost of helping to motivate the sites imaging technologists, particularly in an image intensive study or where the technique is challenging, is probably a very good return on overall investment.

This scenario can very easily be demonstrated in rheumatoid arthritis studies. The standard requirements are radiographs of the hands and feet, every 6 months. This involves careful positioning of the limbs each time and will require 4 images per time point. If MRI scans are being conducted using a standard 1.5 or 3.0 T MRI scanner, then the subject will have to lie in a scanner for 30 min. Not only that, but they will have to lie in the scanner with their hands above their heads in the "superman" position and remain very still for this time. Until you have been in an MRI scanner, it is impossible to describe the noise and level of claustrophobia that a subject has to sustain. Plus the disease is painful, and the requirements for the scan are such that the subjects' hands have to be held in an uncomfortable positioning device. It is only a few studies that require subjects to have more than two or three scans, due to the unpleasant nature of the image acquisition. A good sympathetic technologist working with the subjects will be key to ensuring good subject compliance and good scans.

Safe for the Subject and Operator

With anything we perform in life, there are increased risks for injury and harm. Having a measurement taken increases an individual's risk of harm and may expose the operator to increased risks. Having our height or weight measured is

one end of the risk spectrum. Towards the other end of the spectrum, we could include a complex CT image being performed where the subject receives a highly significant dose of radiation. Alternatively an invasive procedure like angiography has increased risks for the subject. Most operators are well trained and do not expose themselves to undue risk, but it should be considered. In the angiographic example, from the imaging perspective, operators and attending physicians have to wear lead aprons and do receive some additional radiation exposure compared to the normal background dose. The safety issue is one that should be part of the IRB/IEC deliberations before granting the conduct of the trial.

Using the example in the prior section, the MRI may be challenging from acceptability view point, but there is no radiation dose and so is very acceptable from an ethical view. Ultrasound is another technique with no ionizing radiation and can be used for multiple measurements. In body composition measurements, visceral fat assessments can be obtained from a CT scan, and special analysis or a general fat content can be obtained of the whole region using DXA. Each has a different radiation dose to the patient. Likewise with the assessment of sarcopenia, there are a number of assessments that are available, including PET, DXA, and MRI. Each measurement has a different risk, acceptance for the subject and information about the muscle that has to be carefully evaluated.

Subject safety is, of course, paramount. The biggest safety concern in medical imaging is the radiation dose to the patient. This is covered more fully in Chap. 3 and will not be discussed further here.

The ergonomics of the equipment also needs to be considered from the operator's viewpoint. Is it difficult to gain access to the patient for positioning purposes? Is there a C-arm that needs to be rotated by hand? For small technologists this can be a problem. Are the ergonomics of the workstation acceptable? In the last few years, equipment manufacturers have become very cognizant of ensuring the equipment is operator friendly, since the technologists have tended to gain an increased input into equipment purchases.

Development of New Biomarkers

The metrics described in the prior sections set the ground work for the development of new imaging biomarkers or end points in clinical trials. As technology continues to develop, new biomarkers or potential end points are created. The question then becomes, how do these new imaging end points become accepted as an end point? Essentially any new technology has to go through a proving system where these metrics are evaluated.

Many new imaging methodologies are an adaptation or change of something that already exists e.g., DCE-MRI is a prime example of this process. However, de novo technologies or the other regulatory agencies equivalents to go through FDA approval usually as a pre-marketing application or PMA process. This requires a

clinical trial to evaluate safety and efficacy or improved sensitivity and specificity compared to the standard diagnostic or prognostic procedure in place. This still does not mean that the technology is going to be acceptable or useful for clinical trials, an example being bone ultrasonometry which has already been discussed. With the de novo technology, essentially accuracy and discrimination have to be proven along with patient safety. The other metrics are really evaluated in the routine clinical situation.

For a new development from a standard technology, the pathway for acceptance is a little different. The two primary metrics that need to be evaluated are the discrimination (sensitivity and specificity) and precision. There are many examples of these developments, but one of the easiest ones to use as an example is the distal femur measurement in pediatric indications. The team at duPont hospital in Delaware, USA, had found that in children with cerebral palsy had bone loss, but it was difficult to measure their BMD with the standard techniques. This team developed the technique of assessing BMD of the distal femur using the standard DXA technology [18]. The next step was to create a normal database against which disease could be assessed [19]. The technique initially was only developed on one DXA manufacturer's software and was only used at a few highly trained centers. The author, in conjunction with a pharmaceutical company, identified the potential and medical need to develop the technology for use in clinical trials. The next steps were to enlist the support of the developers and train a site that used equipment of another manufacturer of densitometers. The final step was to develop a brief protocol to have ten patients scanned twice at two sites to obtain an assessment of reproducibility or precision. With this completed, it was then possible to develop the concept for a larger clinical trial. The technique is now being used in a number of other clinical trials. The final steps, which have not been conducted at the time of going to press, are either for a DXA manufacturer to further develop this as a formal acquisition and analysis technique with a respective 510 k approval or to have the process formally documented and submitted to the FDA as a possible new biomarker.

Identification of Systematic and Random Errors

Within the context of clinical trials, the errors of the measurements have to be identified and controlled. Identification of the largest source of errors moving down the error tree is often not really considered nor is standard practice for an imaging core lab. It is pointless controlling for a source of experimental error that creates less than 1 % error when there are larger sources of error that are uncontrolled. An example is the use of phantoms. Phantoms are needed for control and identification of error in Type I imaging systems (direct quantification). They may also be needed in Type II systems particularly where indirect quantification is being obtained, e.g., DCE-MRI cartilage or brain volume. (The use of phantoms is a topic outside the remit of this chapter. Phantoms are at best "patient mimics"

since all medical imaging equipment is developed for the assessment of patients, not phantoms. Therefore this has to be understood and each trial and phantom will have its own set of constraints.) However, in Type 2 situations where there is significant reader interpretation or there is a scoring system such as RECIST, then the use of phantoms is not warranted. The source of error being controlled or potentially observed by the phantom is small compared to the potential variation between radiologists.

Experimental error, both random and systematic, needs to be identified in all trials, and the imaging end points are no exception to this concept. The area of the largest error which affects either accuracy or precision, or both should be controlled for and designed into the imaging management of the study. If investigator sites are unable to provide good quality data, then extra training should be provided or the site removed from the study.

Conclusion

In conclusion, the use of any instrument in a clinical trial is a multifactorial evaluation process. This chapter, while theoretical, is the foundation of many of the concepts developed further in this book. Many aspects of the metrics have been taken and honed to make further improvements in clinical trial methodology. For example, the precision of a so-called routine measurement can be improved if rather than having multiple radiologists evaluate the images at their own site using an efficacy system which is outside current clinical practice, the images are handled centrally, the so-called blinded read. This will be discussed more fully in Chap. 5. The variability around the measurement is not only further reduced, but a certain amount of bias is also removed in this process.

The use of all instruments be they imaging or physiological assessments (e.g., lung function) in clinical trials should be carefully evaluated. For routine and well-established techniques, this will be no more than a momentary mental check. However, for the more complex equipment, particularly involving imaging, a full evaluation is a valuable investment of time compared to the ultimate cost of the trial. The use of a good quality assurance commercial contractor should be seriously entertained, and the use of academic centres to conduct this kind of work, particularly if they have no previous track record, should be carefully weighed and considered. In many respects this field is still in its latent infancy with some fascinating new technologies in development which have the potential to change the evaluation of disease and therefore the process of therapeutic evaluation. The pharmaceutical and biotech industry is in desperate need for a more rapid drug development process. One of the greatest potentials that is available to help this issue is medical imaging, particularly for the Phase II go/no go decisions. Molecular imaging is just starting to play a role in this aspect. More will be discussed in other chapters, but ultimately it comes down to an understanding of the metrics for each new technique and how this plays into the overall evaluation of the subject.

References

1. Conroy RM, et al. Measurement error in the Hawksley random zero sphygmomanometer: what damage has been done and what can we learn? BMJ. 1993;306(6888):1319–22.
2. Guidelines for preclinical and clinical evaluation of agents used in the prevention or treatment of postmenopausal osteoporosis. Available from: http://www.fda.gov/ohrms/dockets/dockets/98p0311/Tab0026.pdf.
3. Guidance for industry standards for clinical trial imaging endpoints. Available from: http://www.fda.gov/downloads/drugs/guidancecomplianceregulatoryinformation/guidances/ucm268555.pdf.
4. Kastelein JJ, et al. Simvastatin with or without ezetimibe in familial hypercholesterolemia. N Engl J Med. 2008;358(14):1431–43.
5. Genant HK, et al. Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res. 1993;8(9):1137–48.
6. Sharp JT, et al. How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis? Arthritis Rheum. 1985;28(12):1326–35.
7. van der Heijde DM, et al. Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. Lancet. 1989;1(8646):1036–8.
8. Genant HK. Methods of assessing radiographic change in rheumatoid arthritis. Am J Med. 1983;75(6A):35–47.
9. Guideline on clinical investigation of medicinal products other than NSAIDs for treatment of rheumatoid arthritis. Available from: http://www.ema.europa.eu/docs/en_GB/document_library/Scientific_guideline/2011/12/WC500119785.pdf.
10. Bruynesteyn K, et al. Deciding on progression of joint damage in paired films of individual patients: smallest detectable difference or change. Ann Rheum Dis. 2005;64(2):179–82.
11. Machado P, et al. Ankylosing spondylitis disease activity score (ASDAS): defining cut-off values for disease activity states and improvement scores. Ann Rheum Dis. 2011;70(1):47–53.
12. Pearson D. Standardization and pretrial quality control. In: Pearson D, Miller CG, editors. Clinical trials in osteoporosis. New York: Springer Science + Business Media; 2007. p. 292.
13. Stensland-Bugge E, Bonaa KH, Joakimsen O. Reproducibility of ultrasonographically determined intima-media thickness is dependent on arterial wall thickness. The Tromso Study. Stroke. 1997;28(10):1972–80.
14. Miller CG, et al. Ultrasonic velocity measurements through the calcaneus: which velocity should be measured? Osteoporos Int. 1993;3(1):31–5.
15. Moris M, et al. Quantitative ultrasound bone measurements: normal values and comparison with bone mineral density by dual X-ray absorptiometry. Calcif Tissue Int. 1995;57(1):6–10.
16. Langton CM. ZSD: a universal parameter for precision in the ultrasonic assessment of osteoporosis. Physiol Meas. 1997;18(1):67–72.
17. Fleming TR, DeMets DL. Surrogate end points in clinical trials: are we being misled? Ann Intern Med. 1996;125(7):605–13.
18. Henderson RC, et al. Predicting low bone density in children and young adults with quadriplegic cerebral palsy. Dev Med Child Neurol. 2004;46(6):416–9.
19. Henderson RC, et al. Bone density and metabolism in children and adolescents with moderate to severe cerebral palsy. Pediatrics. 2002;110(1):e5.

Chapter 3
Radiation Risks and Dosimetry Assessment

Derek Pearson

Abstract This chapter summarises the steps to take in order to ensure that the radiation risks associated with a clinical trial are dealt with at the trial planning stage and that the radiation risks are adequately explained as part of the informed consent process. It outlines the steps to take to assess the radiation dose, taking into account the potential variation between participating centres. It examines the risks associated with radiation exposure and the impact on trial participants. Communicating the radiation risk to patients is always difficult, and the chapter suggests some possible comparisons that may be useful in the patient information.

Keywords Radiation • Dose • Risk • Benefit • Patient information

Introduction

The principles of quantifying, justifying and communicating the radiation risks associated with the medical imaging used in clinical trial are those of undertaking good research enshrined within the Helsinki declaration and the Belmont report [1, 2]. These are:

1. Respect for persons. Individuals must be able to make informed choices about the radiation exposure in the trial as well as the other trial procedures. Researchers must provide clear information on the risks in language that the potential participants can understand that allows them to make an informed decision on taking part.

D. Pearson, PhD, BSc
Medical Physics and Clinical Engineering,
Nottingham University Hospitals NHS Trust,
Nottingham NG5 1PB, UK
e-mail: derek.pearson@no3.clara.net

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_3, © Springer-Verlag London 2014

2. Beneficence or “do no harm”. Researchers have an obligation not only to protect research participants from harm but to secure their well-being. They must ensure the design is such that the trial will provide an answer to the research hypothesis and balance the detriment of the risk against the benefits from the research, either in terms of direct benefit to the subject or, when the radiation risks are low, in terms of societal benefit.
3. Justice. The selection of trial subjects should be fair and be representative of the population who is likely to benefit from the research. The burden of research should not fall on any one patient group more than others. Investigators should ensure that subjects are not recruited to multiple trials at the same time.

These principles have been set in the context of international guidelines on radiation safety in research in ICRP 62 [3] that states:

- Researchers should consider whether the information required can be obtained using alternative methods that do not involve ionising radiation (e.g. ultrasound, MRI).
- Imaging techniques are optimised so that the radiation dose is as low as reasonably achievable (ALARA) consistent with obtaining images of satisfactory quality.
- Equipment and procedures associated with the research trial are subject to rigorous quality assurance.
- The number of subjects entered into the trial is the lowest consistent with obtaining an unequivocal answer to the research hypothesis.

The radiation risk assessment must demonstrate that the risks associated with the trial are outweighed by the benefits.

The primary radiation risk to consider is cancer induction, although skin effects can be observed in patients undergoing complex cardiological interventions under X-ray screening. In Case Study 1, the risk of cancer from the radiation exposure in the trial was 1 in 300. Whilst this additional risk is small in absolute terms relative to an individual’s lifetime risk of cancer of 1 in 3, the overall risk of cancer induction from such trials is often seen as worrying by subjects, and clinicians find it difficult to communicate radiation risk to subjects in a way that is meaningful. Equally, it is now being recognised that the population burden from tests involving radiation is becoming significant in terms of excess cancer death [4, 5]. For example, Berrington de Gonzalez and colleagues calculated that approximately 29,000 future cancers were caused by CT scanning of the US population [6]. In a separate study they showed that approximately 700 radiation-induced cancers occur annually in the UK [7].

Case Study 1
Mary diagnosed with non-Hodgkin's lymphoma. As part of the diagnostic workup, she had a CT scan to stage her disease. At this point she was approached by her physician and was then enrolled in a clinical trial of a new chemotherapy regime. The imaging associated with the trial protocol involved undertaking PET/CT at baseline, an interim PET/CT after two courses of chemotherapy and a final PET/CT after six courses of chemotherapy. At the end of the study, Mary's total radiation dose was 85 mSv of which 50 mSv (or 2 PET/CT scans) were additional to normal clinical practice as it was routine for subjects to have PET/CT at the end of chemotherapy. The risk of radiation-induced cancer from this radiation dose was about 1 in 300.

In this chapter, "research exposure" refers to any radiation exposure documented in a trial protocol. Just as all the non-radiation procedures associated with a trial need to be explained to the participants and the Independent Review Board (IRB) or Research Ethics Committee (REC), all the radiation procedures that are part of the trial, whether routine practice or not, need to be included in the dosimetry assessment, radiation risk assessment and participant information sheet. So, in Case Study 1, it might be normal practice to undertake a PET/CT scan after six courses of chemotherapy as this gives a good indication of prognosis [8, 9]. In some centres the staging CT scan will be combined with an initial PET/CT, but that is not true in all centres. All 3 PET/CT scans, however, are required by the research protocol so need to be considered as research exposures. The principal investigator (PI) will need to make a judgement as to what is normal practice and the variation that will exist in normal practice from centre to centre and account for that in the information submitted to the IRB/REC, but more of that later.

The purpose of this chapter is to provide clinicians with guidance on the level of radiation dose and risk encountered in clinical trials and discuss how best to communicate this risk to participants. An explanation of radiation units and risk estimates used in this chapter and the measurement and calculation of radiation dose are given in Appendix 3.1 at the end of this chapter.

How Do We Assess the Radiation Dose?

The best way to do this is to ask an expert! Involve a medical physicist and a radiologist who have expertise in the imaging modalities that you intend to use in the trial as soon as possible within the development of the trial protocol. Their role will be to:

- Make an estimate of the radiation doses involved in the trial
- Undertake a risk assessment of the radiation risks to the participants
- Understand the variation in radiation dose and normal practice between centres participating in a multicentre trial
- Provide a justification for the radiation exposure in the context of the scientific objectives of the trial, taking into account the risks and benefits to the participants

For common procedures the medical physics expert will rely on published data and knowledge of patient dose measurements from their own organisation. There are many sources of published data on radiation dose from X-ray and nuclear medicine investigations [10–15]. The radiation dose from a range of typical procedures in the context of other radiation risks is given in Tables 3.1 and 3.2 gives a range of procedures typical for the specialties covered in the chapters in this book.

Table 3.1 Radiation dose from a range of typical procedures in the context of other radiation risks

	Dose (mSv)		Cancer risk
Category I	0.006	GFR measurement (^{51}Cr-EDTA)	
	0.007	1 day of natural background radiation	
	0.01	Chest X-ray, DXA	
	0.025		1 in 1,000,000
	0.03	AP skull X-ray	
	0.04	Flight from London to New York (cosmic ray dose)	
	0.1		1 in 240,000
Category IIa	0.2	1 month of natural background radiation	
	0.4	AP thoracic spine X-ray	
	0.7	AP lumbar spine X-ray	
		Nuclear medicine renogram (MAG3)	
	1	Nuclear medicine lung scan	1 in 24,000
Category IIb	2	CT head	1 in 10,000
	2.7	1 year of UK natural background radiation	
	3	Nuclear medicine bone scan	
		Barium meal X-ray	
	7	Barium enema X-ray	
	10	CT abdomen	1 in 2,400
Category III	18	Nuclear medicine myocardial imaging (thallium)	
	20	Annual dose limit for radiation workers	
	30	CT follow-up for chemotherapy (multiple CT scans)	

Table 3.2 Typical examination doses for each specialty covered within this book

Specialty	Examination	Radiation dose (mSv)	BERT (approximate)	Cancer risk
Oncology	Head CT	2	8 months	1 in 12,000
	Nuclear medicine bone scan	3	1 year	1 in 8,000
	Abdominal CT	10	3 years	1 in 2,400
	F-18 PET/CT	25	8 years	1 in 1,000
Cardiology	Chest X-ray	0.01	1 day	1 in 2,400,000
	Myocardial perfusion scan (stress and rest)	12	4 years	1 in 2,000
	Coronary angiogram	5–16	2–5 years	1 in 5,000 to 1 in 1,500
	Angioplasty	10–60	3–20 years	1 in 2,400 to 1 in 400
Neurology	Skull AP	0.02	2 days	1 in 1,200,000
	Head CT	2	8 months	1 in 12,000
	Nuclear medicine brain SPECT	5	2 years	1 in 5,000
Musculoskeletal	DEXA (whole body)	<0.001	<9 h	<1 in 24,000,000
	Hand or foot	0.005	15 h	1 in 5,000,000
	Lumbar spine	0.7	3 months	1 in 34,000
Pulmonary disease	Chest X-ray	0.01	1 day	1 in 2,400,000
	Nuclear medicine lung scan (ventilation and perfusion)	1.2	5 months	1 in 20,000
	CT chest	8	3 years	1 in 3,000
	CTPA	12–32	4–11 years	1 in 2,000 to 1 in 750
Body composition	DEXA (whole body)	<0.001	<9 h	<1 in 24,000,000
Ophthalmology	Head CT	2	8 months	1 in 12,000
Nuclear medicine	Bone scan	3	1 year	1 in 8,000
	Myocardial perfusion scan (stress and rest)	12	4 years	1 in 2,000
	Brain SPECT	5	2 years	1 in 5,000
	Lung scan (ventilation and perfusion)	1.2	5 months	1 in 20,000
	Renography	0.7–2	3–8 months	1 in 34,000 to 1 in 12,000

The estimate of radiation dose included in the protocol and trial documentation must take into account the potential variation between centres. There are a number of sources of variation to consider:

- There are known to be significant differences in the radiation dose from the same investigations from different centres, e.g.:
  - Equipment variations
  - Variations in local imaging protocols (e.g. kV, mA settings, slice width, injected activity)

- Variations in normal practice, e.g. in the example earlier some centres will combine the staging CT with an initial PET/CT as normal practice as well as including a PET/CT after six courses of chemotherapy
- Subject size, recognising that larger patients may receive a greater radiation dose
- The population being irradiated (adult, paediatric, older people) in terms of their prognosis and longevity, e.g. children have their whole lives ahead of them so the risk of a radiation-induced cancer appearing is greater than in an adult by about a factor of 3

Where novel imaging techniques are used, the medical physics expert will need to use more fundamental methods for calculating the radiation dose (Appendix 3.1 at the end of this chapter). The radiation dose assessment should document the method used so that the calculation can be understood by an independent expert. With novel techniques it may be necessary to undertake a separate pilot study to assess the radiation dose and optimise the imaging technique.

The radiation dose assessment must document the total radiation dose associated with the trial protocol from research exposures and the additional radiation burden from those exposures over and above normal practice. There will not only be variation between centres on the radiation dose from the same procedures, but there will be variation in what is considered normal practice between centres. For example, for non-small cell lung cancer, it might be normal practice in one centre to have a staging CT scan and then proceed to PET/CT, whereas in another centre subjects may go direct to PET/CT to make full use of the diagnostic CT scan obtained as part of the PET/CT protocol and speed the subject through the pretreatment phase of their care pathway. When designing the trial, the lead radiologist will have to make a judgement as to the extent of the variation in normal practice between the participating centres to enable the medical physicist to properly convey the additional risk and its potential variation.

The radiation dose assessment should also quantify the risks and give a clear statement of the risk coefficients and assumptions used so that an independent expert can follow the calculation. Any adjustments made to the risks calculated because of the population being studied must be documented, e.g. for a paediatric population.

It is helpful if the medical physicist also expresses the radiation dose in terms of a diagnostic reference level (DRL) that can easily be measured in real life. The DRL is a level of radiation dose that would not normally be exceeded for a particular examination for the average patient [16, 17]. It might be expressed in terms of dose area product (DAP) for X-ray, screening time for fluoroscopy, CT Dose Index (CTDI) for CT or administered activity in nuclear medicine. This allows participating centres to audit the actual radiation dose given to trial subjects – an increasing requirement from regulatory authorities.

What Are the Risks?

The main effect from additional radiation is an increase in the lifetime risk of cancer occurring, potentially, many years after the exposure. This risk is given in the right-hand column of Table 3.1 calculated using the latest published risk estimates; see Appendix 3.1 at the end of this chapter. This may be compared with the lifetime cancer risk of 1 in 3. Table 3.1 also shows the radiation dose from typical X-ray and nuclear medicine tests and radiation doses from natural sources. For example, the annual background dose in the UK is about 2.7 mSv, coming from cosmic rays (14 %), food (12 %) and the ground (16 %) as well as from radon gas (58 %). Radon gas seeps out of the ground and accumulates in our houses. In some parts of the UK, there is much more radon gas coming from radioactive materials in the rock, and the annual natural background dose is up to three times higher than the average.

The risks from different examinations in Table 3.1 are split into categories suggested by international guidelines [3, 18]. Each category has a tenfold increase in risk as you move from Category I to Category III. The guidelines also describe the level of benefit required to justify the use of imaging in a research trial in each category. When the risk is trivial, for example, it is acceptable to use an X-ray in research just to find out something new. In Category III, where there is a moderate risk and repeated exposures may lead to an unacceptable risk, the use of an X-ray examination should lead to significant benefit to society before it can be justified. A CT scan, for example, just to show the size of a tumour as an outcome measure in the trial of a new chemotherapy drug may not be justified simply on the basis that it is a requirement of the trial sponsor. If it is an assessment of patient response that will change patient management, it probably is justified.

The risks are additive. So two CT scans as part of the study double the radiation dose and thus double the risks.

Other risks have to be considered. For example, a trial of a new X-ray-guided cardiology procedure may give a high dose, but this may be offset against the alternate risk of major heart surgery and a general anaesthetic.

The lead radiologist will need to consider a number of issues within the radiation risk assessment. These will include:

- The reason for the radiation exposure. Is it required for patient management or is it solely to provide data for the research trial with no direct benefit to the subject? Do they meet the objectives of the study and are they ethically acceptable?
- Do the radiation exposures in the protocol form part of normal clinical management, taking into account variation at different participating centres?
- Is the radiation dose from the investigations ALARA consistent with obtaining images of sufficient quality to meet the study objectives. Procedures need to be optimised so that image quality is balanced with radiation dose. Give too much radiation and there is little gain in image quality, and the subject receives unnecessary radiation dose. Give too little and the images are not of sufficient quality to be diagnostic – again unnecessary dose for the subject as it was not worth carrying out the test.

- The potential diagnostic benefits, including direct benefits to the participant and the benefits to society from the knowledge generated from the research trial. This is one argument for ensuring the quality of clinical trials and doing good research. In reality, for any particular research problem, there are a few key papers that document well-designed and well-carried-out trials that are the definitive studies. They are well documented and can be used in meta-analyses. On the other hand there are many that sink without trace in the ever growing morass of research publications, where the societal benefit is minimal.
- The detriment to participants that the exposure may cause. This is predominately in terms of cancer risk. In trials of interventional imaging techniques, however, where there is a significant risk of direct radiation effects such as erythema, the radiation dose to the most heavily irradiated tissue outside the target area should be considered. In these situations, effective dose is not the appropriate measure of radiation risk to use. Instead, the local absorbed radiation dose should be used and the risks to the heavily irradiated tissues used to calculate the radiation risks. The medical physics expert will be able to calculate these doses and the radiation risk, but should document the methods used so that an independent expert can understand and validate the calculations, e.g. skin dose in interventional cardiology, breast dose in CT pulmonary angiography.
- The availability of alternative techniques involving less, or no, exposure to ionising radiation. For example, many forms of chemotherapy have a risk of cardiac complications. Often, a pretreatment measurement of ejection fraction is made using a nuclear medicine investigation. The ejection fraction can be measured using ultrasound, which involves no exposure to ionising radiation so should be considered as an alternative.
- The study design. Is the number of subjects entered into the trial the lowest consistent with obtaining an unequivocal answer to the research hypothesis? Too low and that radiation is wasted as the research hypothesis is not answered, too high and unnecessary imaging procedures may have been carried out.
- The possibility that participants will be participating in other trials involving additional radiation. Local investigators need to ensure that participants are not over-researched, and it may be possible that the imaging required for one trial will be available from other trials the subject is involved in. Duplication of imaging should be avoided.
- The characteristics of the research population will include such factors as the age of the participants and their likely life expectancy. For paediatric subjects, for example, increase the risk given in Table 3.1 by a factor of 2 or 3 depending on age – the younger the subject, the greater the risk (i.e. at 1 mSv the risk is about 1 in 9,000 rather than 1 in 24,000). For subjects over the age of 50 reduce the risk by a factor of 5–10 depending on age – the older the subject, the less the risk (i.e. at 1 mSv the risk for elderly subjects is about 1 in 240,000 rather than 1 in 24,000). Don't forget that any radiation-based comparisons you make when explaining radiation risk will also be age adjusted. For example, in Table 3.1,

a bone scan gives a radiation dose equivalent to 1 year of natural background radiation with a radiation risk of about 1 in 8,000. In a 70-year-old person the risk from the bone scan may become 1 in 80,000 – but so does the risk from 1 year of natural background.

- The clinical prognosis of the study cohort. There is little point considering the long-term cancer risk of radiation exposure in a population who are terminally ill.
- Pregnancy and breast-feeding women. It is unlikely that pregnant women will be entered into a therapeutic research trial. As the baby in the womb is particularly sensitive to radiation, care must be taken during pregnancy. Subjects attending for X-ray and nuclear medicine examinations will be asked if they are pregnant or breast-feeding, and if they are, the radiologist will discuss the need to postpone the investigation.

What Do I Tell Participants?

Clear information is a priority if participants are going to make an informed choice about taking part in a trial. Many studies are rejected by IRB/RECs simply because of poor patient information. It is wise to spend time producing a good patient information sheet. Take advice from those who are involved in the production of patient leaflets and do not assume that something that is obvious to you will be obvious to the participant. Guidance is available from the Food and Drug Administration in the USA [19] and the NHS Health Research Authority in the UK [20]. These require that there is a description of any reasonably foreseeable risks or discomforts to the subject, including the risks associated with any investigations associated with the trial, including the radiation risk from any imaging investigations. Equally, there should be a description of any benefits to the subject or to others which may reasonably be expected from the research. Where there is no direct benefit to the participant be honest and tell them that, but also tell them why you are doing the investigation. Best practice is to inform the subject about the risks associated with all the tests involving radiation in the protocol as well as that part of the risk that is associated with the component of risk that is additional to normal practice. Where there are complex or novel imaging techniques being used as part of the trial, it might be helpful to have a separate information sheet that describes the test in detail and includes pictures of the equipment so that the participant can understand better what to expect when they come for their test.

Communicating radiation risk to participants is difficult because of their perception of risk is based on anecdotes, their own experience and their systems of trust and belief rather than the cold, numerical risks that we, as scientists, tend to understand and rely on. The language of communicating risk is also difficult. Case Study 2 demonstrates some of the problems associated with communicating radiation dose and risk to a patient.

Case Study 2
George was referred for a myocardial perfusion scan in nuclear medicine. The actual radiation dose the he received was about 6 mSv with an associated cancer risk of about 1 in 4,000. This is a small additional risk compared to the lifetime risk of cancer of 1 in 3. George used the Internet and found what appeared to be a reliable source of information that told them that the radiation dose was equivalent to 900 chest X-rays. This alarmed him and, had he known this level of risk, he would never have had the test done. When he contacted the nuclear medicine clinic, they were able to reassure him that the information on the website was wrong. The radiation dose was actually equivalent to about 300 chest X-rays – but this was still an alarming number to George. He complained that the patient information stated that the radiation dose was equivalent to many comparable X-ray techniques (which it is) but, in the light of the number of chest X-rays involved, he felt that been misinformed about the level of radiation dose he had received.

Experience tells us that subjects will perceive the risk as low risk if the information about the risk is communicated by someone they trust. Also they perceive they will get substantial benefit from the clinical trial and the associated imaging, they feel that they are entering the trial on a voluntary basis and do not feel pressured. The opposite is also true. A high risk is one where they do not trust what they are being told, they feel they are being pressured into taking part and they perceive that they will get little or no benefit from the trial.

When it comes to radiation risks, subjects tend not to feel in control and distrust the information they are receiving. Using the chest X-ray as a unit of radiation dose is clearly problematic unless the radiation dose is of the same order of magnitude. Using the risks subjects face in their daily lives from Table 3.3 is often not helpful because all comparators come with subjective bias, subjects feel in control and often underestimate or modify the risk in their own minds in line with their own experience of that risk. The risk from 1.5 mSv radiation dose is roughly the equivalent to the annual risk of dying in a car accident in the UK, but the subject knows they are not a young man in an old car with four friends, full of alcohol and driving too fast down a narrow lane!

Explaining the radiation risk in terms of cancer risk is also difficult because the word cancer will cause alarm bells to ring, however low the absolute risk. The patient in Case Study 1 had an increased cancer risk of 1 in 300, an increase of 1 % in the subject's natural risk of cancer of 1 in 3. Taken in the context of their condition, this is a small increase in risk, but trying to communicate to a vulnerable patient in terms of this risk is difficult.

At low radiation doses it is often helpful to use what is known as the BERT (Background Equivalent Radiation Time) [21] or the radiation risks from an air flight as a radiation risk comparator. For example, a skull X-ray is roughly equivalent to the radiation risk from a transatlantic flight, and an AP thoracic spine X-ray is equivalent to 2 months of background radiation (Table 3.2).

Table 3.3 Risks of daily living

Cause	Risk
Lifetime risk of contracting cancer	1 in 3
Lifetime risk of dying of cancer	1 in 4
Annual risk of death from smoking ten cigarettes a day	1 in 200
Annual risk of death from heart disease	1 in 300
Annual risk of death from cancer	1 in 400
Risk of death all causes aged 40	1 in 700
Annual risk of death from an accident in the home	1 in 15,000
Annual risk of death from a car accident	1 in 17,000
Annual risk of being murdered in the UK	1 in 100,000
Risk of maternal death from pregnancy in the UK	1 in 170,000

This may not work with higher-dose investigations. Case Study 1 gave a total protocol radiation dose of 85 mSv that is equivalent to almost 30 years of background radiation or over 2,000 transatlantic flights, which sounds rather too much! At higher doses, it might be better to use the annual radiation dose limits and explain that each PET/CT scan is roughly equivalent to the maximum radiation dose that a radiation worker can receive in a year.

The clinical prognosis of the subjects should also be taken into account when communicating radiation risk. For example, in trials where experimental treatments are being tested in subjects who may be terminally ill, the radiation risk becomes less of an issue. The participant information sheet should tell them the facts that they will be exposed to additional radiation risk but state sensitively that, given the circumstances of their illness, the additional radiation adds negligible risk.

Guidance on the wording of the information sheet is summarised against the radiation risk categories in Table 3.4. These can be compared risks of daily living (from the UK) that might be helpful in explaining radiation risk.

What Will the IRB/REC Want to Know?

As part of the IRB/REC submission, the PI will need to include reference to the use of radiation in the protocol. This should include:

- Information on all the radiation exposures associated with the protocol
- Those exposures that are, on average, additional to normal practice
- Some commentary on the expected variation in imaging protocol and radiation dose between participating centres
- The total radiation dose associated with the protocol

Table 3.4 Categories of radiation risk

Category	Risk	Benefit	Participant information sheet
I	Trivial	Research leads to an increase in knowledge	The risk from these X-rays is very small and is equivalent to less than 2 weeks natural background radiation
IIa	Very low	Research leads to increased knowledge leading to health benefit	The risk from these X-rays is very low and is equivalent to a few months natural background radiation. It is less than the annual risk of dying from an accident in the home
IIb	Low	Research aimed directly to improvements in the diagnosis, cure or prevention of disease	The risk from these X-rays is low and is equivalent to 1 year's natural background radiation. It is less than the annual risk of death from any cause when aged 40
III	Moderate – repeated exposure will lead to unacceptable risk	Research leads to direct patient benefit directly related to saving of life or the prevention or treatment of serious disease	The risk from these X-rays is equivalent to the annual dose limit for radiation workers and is similar to the annual risk of death from any cause when aged 40 If the subject is part of a trial of an experimental treatment in patients who are terminally ill, then state that the additional radiation risk is of no consequence given the circumstances of their illness

- The radiation dose that is additional to normal practice
- The radiation risk assessment with reference to the risk estimates used
- In the case of novel procedures, a calculation of the radiation dose with enough information so that it can be verified by an independent expert including any pharmacological model used in nuclear medicine dosimetry
- A clear and appropriate description of the radiation risk, including a risk comparator, in the participant information sheet

This will allow the IRB/REC to make an informed decision as to whether the radiation risk is justified and seek expert advice if necessary.

Training and Quality Assurance

The variation in radiation dose between participating centres can be reduced by the lead centre or CRO undertaking pre-trial quality assurance (QA) tests on key items of equipment at participating centres before they are allowed to join a trial. This

may be as simple as circulating a phantom or imaging test object to participating centres with a clear protocol for acquiring the images on the phantom to, at the other extreme, a visit from the lead centre to undertake a full evaluation of the equipment involved.

The use of standard imaging protocols can also be helpful in reducing variation between centres and assuring data quality. This may include equipment settings, software settings for reconstruction and analysis and the requirements for transmission of data to the lead researcher or CRO. It is useful in those circumstances for the PI or CRO to organise training for the imaging technicians from each participating centre to ensure a consistency of approach.

The variation in image interpretation is also important when seeking to obtain maximum benefit from the radiation exposure. Some trials reduce this by having all imaging reported by an independent CRO, others by training radiologists in the image interpretation and others by double reporting some investigations or auditing random samples of reports to ensure accuracy and consistency of interpretation.

Appendix 3.1: Units of Radiation Dose, Risk Estimates and Measurement of Radiation Dose

Radiation dose in this chapter has been expressed in terms of absorbed dose and effective dose. The absorbed dose is the amount of energy the radiation deposits in a unit mass of tissue and is measured in units of J/kg or Gray (Gy), named after Louis Gray (1905–1965), a British physicist who worked on the effects of radiation on biological systems. Absorbed dose is used to measure, for example, the entrance skin dose or the radiation dose to a particular organ in the body.

To compare the risks of a radiation exposure and particularly the risk of very different types of radiation exposure, e.g. an ankle X-ray compared to breast radiotherapy, effective dose is used. This quantity includes factors that take into account the biological effect of different types of radiation and the different sensitivity to radiation of different tissues within the body. For example, for the same energy delivered to a unit mass of tissue (i.e. the same absorbed dose), neutrons will have a far greater effect than X-rays by a factor of 5–20 depending on neutron energy. At the same time the gonads are more radiosensitive than the foot. So the effective dose depends on the tissue irradiated and the type of radiation involved. The unit of effective dose is the Sievert or Sv, named after Professor Rolf Sievert (1896–1966) who was a medical physicist who studied the biological effects of radiation. A single whole body dose of 5 Sv would kill 50 % of those exposed. In reality, the radiation dose from X-ray and nuclear medicine tests are measured in mSv and μSv, and the exposure is restricted to the area of interest.

The average natural background radiation in the UK is 2.7 mSv. The annual dose limit for a radiation worker in the UK is 20 mSv, but most radiation workers in the hospital setting receive less than 1 mSv per year.

The effective dose from a barium meal (3 mSv) is the same as that from a nuclear medicine bone scan. These are very different tests, different organs are irradiated, but the overall risk of the radiation exposure is the same. The organs at risk from a barium meal are the stomach and the liver, whilst from a bone scan they are the kidneys, bladder and red bone marrow, but the radiation risk from both tests is equivalent.

In the USA the unit of rem is often used instead of Sv. There is a factor of 100 between the two, with 1 Sv being equivalent to 100 rem. So 1 rem is equal to 10 mSv.

In order to calculate the risk, epidemiological studies of populations exposed to radiation such as atomic bomb survivors and patients who developed cancer after a course of radiotherapy have been studied [22, 23]. Such epidemiology is very difficult to carry out, but the risk factor for calculating cancer risk is currently accepted to be 4.2 % per Sv, i.e. for 100 people exposed to 1 Sv only 4 of them will develop cancer because of their radiation exposure, a risk of 1 in 25. Considering the annual dose limit for workers of 20 mSv, the risk of radiation-induced cancer associated with that exposure is about 1 in 1,200, a risk that is deemed acceptable by the

regulatory authorities for a worker who has chosen to work in the radiation industry. For members of the public, dose limits are lower because it is less acceptable to expose those who have no choice about their exposure to radiation. In the EU it is 1 mSv per annum, equivalent to a risk of 1 in 24,000 of cancer induction. Our example of a nuclear medicine bone scan or barium meal has a cancer induction risk of 1 in 8,000, the additional radiation risk over and above the dose limit being balanced by the benefit of a medical diagnosis.

Radiation dose is measured or calculated by a number of methods:

- Complex computer models can be used such as Monte Carlo models. These follow the path of many thousands of photons as they are distributed and interact within a mathematical phantom patient. Once the computer codes have been established, the exposure conditions can be easily changed to emulate a wide range of examinations thus avoiding the necessity of making many direct dose measurements.
- The entrance surface dose at the patient can be either measured directly or can be calculated from the measured radiation dose in air at that point.
- Measuring organ doses within an anatomical phantom and from that calculating the effective dose.
- Using dose area product. A dose area product meter is a large parallel plate ionisation chamber that is fitted to the X-ray tube, perpendicular to the beam axis. If set up at the tube window below the light beam diaphragm and collimators, the beam size will never exceed the area of the chamber. The response of the chamber to the charge collected during the examination is proportional both to the size of the irradiated area and the exposure. When the chamber is set up perpendicular to and centrally aligned with the beam axis, the response will be proportional to the product of the irradiated area and the exposure (dose area product or DAP). The DAP meter will accumulate charge during the whole examination and deliver a DAP at the end, whether due to screening (fluoroscopy) or single radiographs, in $Gy.cm^2$. The DAP can then be used to calculate organ doses and estimate the effective dose for each patient/examination.
- In CT the CT Dose Index (CTDI) and dose length product (DLP) are the commonest methods of reporting dose. CTDI represents the dose from the current slice plus the scattered radiation from adjacent slices. The most commonly reported variation of CTDI is the $CTDI_{vol}$ [24]. Firstly, the $CTDI_w$ is calculated, which is the weighted sum of two-thirds peripheral dose and one-third central dose measured over a 100-mm range in an acrylic phantom. The $CTDI_{vol}$ is $CTDI_w$ divided by the beam pitch factor [25], i.e. the table increment per 360° rotation divided by the slice width. The dose length product (DLP) is the $CTDI_{vol}$ multiplied by the slice thickness and the number of slices in centimetres and is in units of $mGy.cm^2$. DLP is independent of patient size, but can be used to calculate effective dose. $CTDI_{vol}$ and DLP are used to set DRLs that can be easily used within clinical trials to monitor the radiation dose from CT.

- CT dose can be calculated knowing the make and model of CT scanner and the scan parameters using dose calculators available on the Internet such as the ImPACT CT Patient Dosimetry Calculator available at www.impactscan.org.
- In nuclear medicine the radiation dose is usually estimated using the method developed by the Medical Internal Radiation Dose (MIRD) Committee of the Society of Nuclear Medicine [26]. For a range of radionuclides, MIRD have calculated the absorbed dose to various organs within the body using the Monte Carlo method and tabulated these, from which the effective dose can be calculated.

References

1. Protection of human subjects; Belmont Report: notice of report for public comment. Fed Regist 1979; 44:23191–7. Available at: http://www.hhs.gov/ohrp/humansubjects/guidance/belmont.htm. Accessed 8 Sept 2010.
2. WMA declaration of Helsinki – ethical principles for medical research involving human subjects. Seoul. 2008. Available at: http://www.wma.net/en/30publications/10policies/b3/index.html. Accessed 8 Sept 2010.
3. ICRP Publication 62. Radiological protection in biomedical research. Ann ICRP. 1993;22/3.
4. Johnson DA et al. CT and radiation-related cancer risk- time for a paradigm shift? Nat Rev Gastroenterol Hepatol. 2009;6:738–40.
5. Smith-Bindman R et al. Radiation dose associated with common computed tomography examinations and the associated lifetime attributable risk of cancer. Arch Intern Med. 2009; 169(22):2078–86.
6. Berrington de Gonzalez A et al. Projected cancer risks from computed tomographic scans performed in the United States in 2007. Arch Intern Med. 2009;169(22):2071–7.
7. Berrington de Gonzalez A, Darby S. Risk of cancer from diagnostic x-rays: estimates for the UK and 14 other countries. Lancet. 2004;363(9406):345–51.
8. Molnar Z et al. Prognostic value of FDG-PET in Hodgkin lymphoma for posttreatment evaluation. Long term follow-up results. Neoplasma. 2010;57(4):349–54.
9. Kobe C et al. Positron emission tomography has a high negative predictive value for progression or early relapse for patients with residual disease after first-line chemotherapy in advanced-stage Hodgkin lymphoma. Blood. 2008;112(10):3989–94.
10. Shrimpton PC, Hillier MC, Lewis MA, Dunn M. NRPB – W67 doses from computed tomography (CT) examinations in the UK – 2003 Review. NRPB 2005. Available at: http://www.hpa.org.uk/Publications/Radiation/NPRBArchive/NRPBWSeriesReports/2005nrpbw067/. Accessed 19 Nov 2013.
11. Hart D, Hillier MC, Wall BF. HPA-RPD-029 – doses to patients from radiographic and fluoroscopic x-ray imaging procedures in the UK – 2005 review. Health Protection Agency 2007. Available at: http://www.hpa.org.uk/Publications/Radiation/HPARPDSeriesReports/HpaRpd029/. Accessed 8 Sept 2010.
12. Smans K et al. Results of a European survey on patient doses in paediatric radiology. Radiat Prot Dosimetry. 2008;129:204–10.
13. CRCPD Publication E-09-5. Nationwide Evaluation of X-Ray Trends (NEXT) Summary of 2003 Fluoroscopy Survey September 2009. Available at: http://www.crcpd.org/Pubs/NEXT_docs/2003NEXT_FluoroscopySurvey.pdf. Accessed 8 Sept 2010.
14. CRCPD Publication E-07-2. Nationwide Evaluation of X-Ray Trends (NEXT) Tabulation and Graphical Summary of 2000 Survey of Computed Tomography. Available at: http://www.crcpd.org/Pubs/NEXT_docs/NEXT2000-CT.pdf. Accessed 8 Sept 2010.

15. Spelic DC. Nationwide surveys of chest, abdomen, lumbosacral spine radiography, and upper gastrointestinal fluoroscopy: a summary of findings. Health Phys. 2010;98(3):498–514.
16. ACR practice guideline for diagnostic reference levels in medical x-ray imaging. 2008. Available at: http://www.acr.org/Quality-Safety/Standards-Guidelines/Practice-Guidelines-by-Modality/Medical-Physics. Accessed 29 Apr 2013.
17. IPEM Report 88. Guidance on the establishment and use of diagnostic reference levels for medical x-ray examinations. IPEM 2004.
18. Radiation Protection 99. Guidance on medical exposures in medical and biomedical research. European Commission 1998. Available at: http://ec.europa.eu/energy/nuclear/radiation_protection/doc/publication/099_en.pdf. Accessed 8 Sept 2010.
19. FDA guide to informed consent. Available at: http://www.fda.gov/RegulatoryInformation/Guidances/ucm126431.htm. Accessed 29 Apr 2013.
20. NHS Health Research Authority. Information sheets and consent forms guidance for researchers and reviewers. Available at: http://www.hra.nhs.uk/resources/before-you-apply/consent-and-participation/consent-and-participant-information/. Accessed 19 Nov 2013.
21. Cameron JR. A simple way to explain radiation. Clin Radiol. 1991;44:140.
22. Wall BF et al. What are the risks from medical x-rays and other low dose radiation? Br J Radiol. 2006;79:285–94.
23. ICRP 103: The 2007 recommendations of the International Commission on Radiological Protection. Ann ICRP. 2007;37:1–332.
24. Medical electrical equipment – part 2–44: particular requirements for the safety of x-ray equipment for computed tomography. IEC Standard 60601-2-44 Ed 2 Amendment 1. Geneva: IEC; 2003.
25. Silverman PM, Kalender WA, Hazle JD. Common terminology for single and multislice helical CT. Am J Roentgenol. 2001;176:1135–6.
26. MIRD Publication – MIRD Pamphlet #11: S, Absorbed Dose per Unit Cumulated Activity for Selected Radionuclides and Organs. Available at: http://interactive.snm.org/index.cfm?PageID=2199. Accessed 8 Sept 2010.

Chapter 4
Imaging Review Charters and Operational Considerations

Andrea M. Perrone, Jennifer A. Villetard, and Colin G. Miller

Abstract The regulatory framework around radiological endpoints in clinical trials has changed dramatically in the last 10 years or so. At the current time there are FDA guidelines (albeit in draft at the time of writing and presented verbatim in Appendix 1 of this book), which detail the requirements and contents for the Imaging Review Charter (IRC). This chapter provides the historical perspective for these guidelines and details the contents that have to be discussed within an IRC. This is presented within the background of an oncology clinical trial, to provide a comprehensive understanding of the requirements and expectations from the FDA.

Keywords Imaging Review Charter • Oncology • Read design • Imaging core lab

Background

Historical Development and Use of Imaging Review Charters

An Imaging Charter is a regulatory document drafted to define and describe all aspects of the procedures that an imaging core lab follows when processing image data and conducting the independent read. While this originated as a regulatory requirement, it has evolved into a standard document that is developed for most clinical trials in which imaging is a primary or secondary endpoint in order to

A.M. Perrone, MD, BS Finance (✉) • J.A. Villetard, BA
Department of Imaging, Merck, WP42-301, 4, West Point,
PA 19486, USA
e-mail: andrea.perrone@merck.com; jennifer.villetard@bioclinica.com

C.G. Miller, BSc, PhD, FICR, CSci
Medical Affairs, BioClinica Inc., 826 Newtown-Yardley Road,
Newtown, PA 18940, USA
e-mail: colin.miller@bioclinica.com

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_4,

capture the depth of content that is not covered elsewhere. The personnel of the imaging core lab write the Imaging Charter with input and guidance from the sponsor and medical and scientific experts in the relevant field.

Pharmaceutical companies began writing Imaging Charters partly in response to Food and Drug Administration (FDA, or "the Agency") guidance [1, 2]. The purpose of adopting the use of Imaging Charters was to have sponsors define a priori the key components required to manage the imaging aspect of a clinical trial in an effort to minimize variability and ensure consistency across studies and therapeutic areas.

There are several key intended audiences for the Imaging Charter. These include the sponsor, the study team within the imaging core lab, and regulatory agencies. The sponsor uses the Imaging Charter to ensure that the plan for the imaging component of their trial accurately captures the study as defined in the protocol. Study teams within an imaging core lab use the Imaging Charter as the procedural guidelines for a particular trial. Submitting the Imaging Charter to regulatory agencies such as the FDA and the European Medicines Agency (EMA) for review and approval helps to minimize regulatory discrepancies related to imaging throughout the course of a clinical trial.

As the use of Imaging Charters became a more regular component of clinical trials, Imaging Charter development began to follow one of two strategies. The first approach is to submit the Imaging Charter to the Agency for review and approval prior to the start of the independent read. The second approach is to write an Imaging Charter but hold submission to the Agency until the completion of the study.

Imaging Charters that are submitted to the Agency for review and approval prior to the start of the independent read are typically submitted along with the study protocol and statistical analysis plan (SAP). Using this approach allows sponsors to solicit feedback on read design from the Agency prior to the start of the read, thereby increasing the likelihood that the Agency will not question the read design at the time of final data submission.

If study sponsors choose the second approach to Imaging Charter development, they create an Imaging Charter without the intent of sending it to the Agency for review until the submission of final study data. Such Imaging Charters usually serve as a communication tool between the study sponsor and the imaging core lab. The goal of this type of Imaging Charter is to ensure that both parties fully understand and acknowledge the study and its imaging requirements. This methodology also facilitates design discussions between key parties at both the sponsor and imaging core lab early in the project lifecycle, enabling more accurate assessment of the scope of work for the imaging core lab.

Current Use of Imaging Charters

The Imaging Charter has become an integral part of the services and deliverables that imaging core labs offer. What was once a document created for primarily large,

Phase 3 oncology and rheumatoid arthritis trials is now an essential part of most clinical trials that mandates an imaging component, including trials in the musculoskeletal arena, CNS trials, and others.

Much effort has been invested collaboratively by pharmaceutical companies, imaging core labs, professional imaging societies, and the FDA to standardize imaging for clinical trials. This includes initiatives to develop a consistent table of contents as a template to follow which ultimately improves both the efficiency of development and also of subsequent review by the FDA. This collaboration has resulted in the establishment of several consortiums, described in Appendix 4.1 at the end of this chapter.

As a result of this across-the-industry collaboration, these groups are well positioned to assist the Agency in developing new guidelines to enhance their shared mission of accelerating the delivery of new treatments for unmet medical needs.

For some disease indications, and specifically for trials using a diagnostic imaging agent, the use of a robust and comprehensive Imaging Charter with emphasis on reader training has actually increased in importance and relevance in the last several years. Failure to implement a robust Imaging Charter and to substantially document reader training can harm the outcome of a clinical trial, which is evidenced by the rejection of EPIX in their first submission for Vasovist in 2005 [1]. That rejection was based on a lack of documentation of reader training and handling of specific data issues.

State-of-the-Art Imaging Review Charters

Standard Content

The Imaging Charter illustrates the processes that the imaging core lab follows from the time it receives an image through to final read-data export to the sponsor. This includes

- Data collection.
- Processing.
- Blinding of subject and site identifiers.
- Archiving.
- Reasons for and handling data queries.
- Differences in handling film as opposed to digital data.
- Brief imaging acquisition overview. Specific acquisition details needed at the site level are covered in other documents.

In addition, the content of the Imaging Charter provides the following, as described in the next few sections:

- Reader selection criteria, enabling the imaging core lab to identify potential readers for the study sponsor

- Reader training specifics, encouraging sponsors to consider reader training and developing a plan for this training prior to the start of independent reads
- A plan for the "Mock" read, which provides the study sponsor a walk-through of the independent read application
- A basis for creating a plan for monitoring the read to ensure consistency of each reader from the beginning of the read to the end as well as consistency among readers
- Detailed adjudication criteria to settle discrepancies between readers

Of critical importance, however, the Imaging Charter sets forth the guidelines for two key topics, which require the most input from the sponsor, the imaging core lab, and medical/scientific experts: the response assessment/scoring (also known as response criteria) and the design of the independent read.

Defining the Response Criteria

Most of the indications addressed in clinical trials have industry standards regarding response criteria [3–10]. These standards are largely independent from clinical practice but must be followed closely by the central readers in order to identify a signal of response or lack thereof.

In addition, after a response criterion has been employed in clinical trials, it often becomes apparent that some parameters must be better defined or some scenarios arise that were not accounted for in the original response criteria. Therefore, it is typical for the industry to revise the response criteria and publish the updated version [3, 4].

The Imaging Charter must identify and detail the response criteria that a clinical trial will employ. It must also describe any modifications to the published criteria and must provide the rationale for such changes.

Designing the Read

In addition to clearly defining the response criteria as well as highlighting any modifications to the criteria, it is equally important to prospectively outline the read design.

The number of new therapies in clinical development is highest in oncology and cardiovascular areas, which is reflected in the overall volume of work in independent reads by core labs. Therefore, due to the variability in modalities for cardiovascular disease and more standardization in the oncology arena, this section will highlight the content for a standard read design in an oncology trial for solid tumors.

The same concepts are found in other therapeutic areas, but the nuances are different. Read designs for rheumatoid arthritis and osteoporosis are covered in greater detail in Chap. 5.

In contrast to the clinical setting where a radiologist provides a narrative of a patient's response to treatment (regardless of whether their interpretation is formed in part by input from an oncologist who is treating that patient), the read design for a radiologist to review and assess imaging data acquired as part of a clinical trial must be completely independent of any clinical data from the site. The read design must also clearly define the questions to be asked of the radiologist, the sequence of presenting images, and the ultimate set of response assessments to be included. There are a multitude of read designs, and they can vary widely depending on disease indication, whether the data will be used for an interim analysis, whether the data will be used to make internal go/no-go decisions, or ultimate submission to the Agency for approval. The most common read design focuses on efficacy when progression or response as captured objectively in an imaging read is a primary or secondary endpoint and usually includes the following sessions (or a similar structure of sessions):

- Baseline/screening lesion selection and measurement
- Sequential lesion selection and measurement
- Incremental radiological response assessment

In addition, some read designs include the following two sessions:

- Global: The global session is often the topic of critical discussions when the sponsor and the imaging core lab are designing the read.
- Clinical.

An additional consideration regarding a read design for an oncology trial is the inclusion of a clinical session. This depends upon several factors including the criteria being implemented, disease indication, and the existence of an accepted serum biomarkers or outcome measures that may be used to assess the status of disease. Oncology trials that utilize Cheson 1999 or 2007 criteria essentially include a clinical component as part of the evaluation [6, 7]. It is critical to include a clinical session for a melanoma trials implementing the Wolchok Criteria or RECIST 1.0/1.1 [3, 4] to enable the development of new cutaneous lesions to be considered in the overall response assessment. Finally, trials that require sites to obtain serum assays that are considered biomarkers of disease status such as prostate-specific antigen (PSA) for prostate cancer, carcinoembryonic antigen (CEA) for colon cancer, or cancer antigen 125 (CA-125) for ovarian cancer may include this data stream along with other outcome measures (e.g. ECOG status) to evaluate clinical deterioration.

Figure 4.1 illustrates the standard, five-session read for oncology trials. Note that it employs the 2 + 1 reader model with the addition of an oncology session and reader.

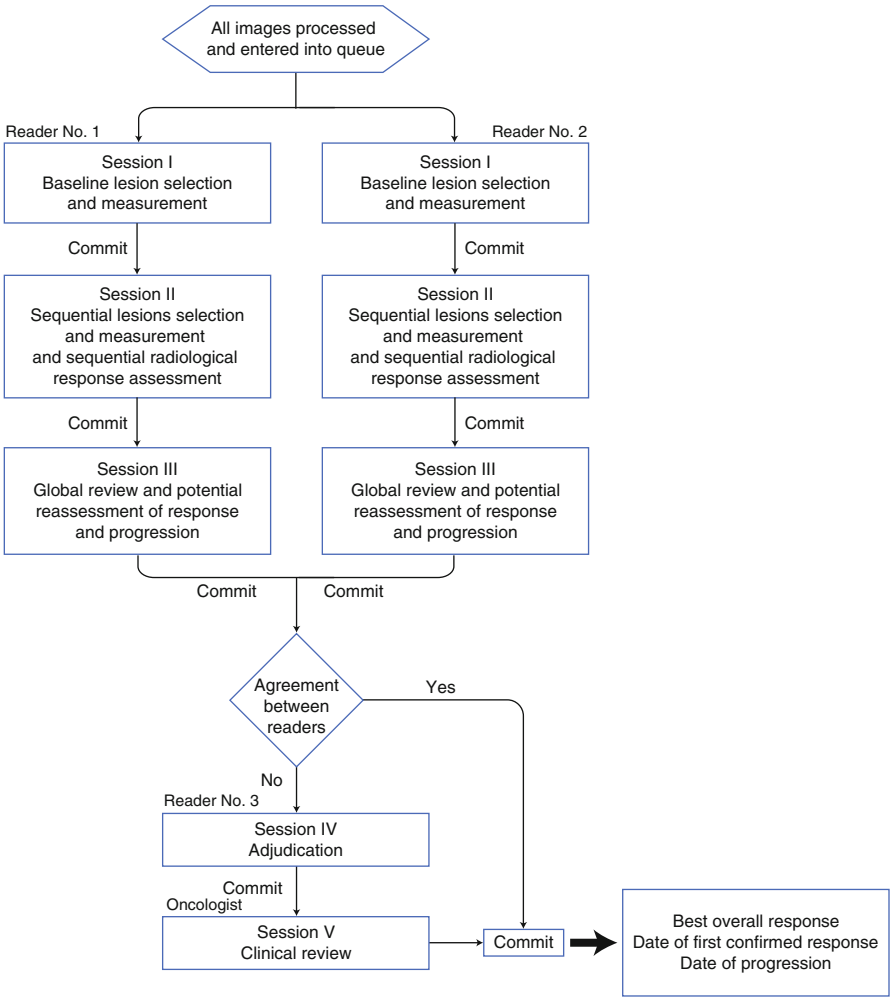


Fig. 4.1 Standard, five-session read design for oncology trials

Reader Allocation for Oncology Trials

The generally accepted industry standard reader allocation for a pivotal Phase 3 oncology trial in which progression or response as captured objectively in an imaging read is a primary or secondary endpoint is to employ two radiologists as primary readers who assess 100 % of the data. In the event that there is disagreement on the main points of assessment, a third reader reviews the primary readers' responses and selects the assessment with which he/she is in greatest agreement. This is referred to as the 2 + 1 reader allocation. Alternative read paradigms are presented in the read methodologies chapter (Chap. 5).

Purpose of the Global Session

An additional consideration regarding an oncology read design and reader model includes whether to include a session to enable the readers to review all of the images from baseline to the final time point at one time and to allow the reader to change previous assessments. This is known as a global session. Although the selection of target and nontarget lesions at baseline cannot be altered, this session allows the reader to modify a response assessment at a subsequent time point.

In the global session, the reader has the opportunity to review such critical progression- and response-related assessments, such as whether a lesion identified at a specific time point as a new lesion was in fact simply a benign lesion which may have "appeared" as new based on the phase of contrast (i.e., a hepatic hemangioma). Additionally, an evaluation of "unequivocal progression of nontarget disease" at a specific time point that ultimately is not consistent with target tumor burden and was likely a false-positive can also be noted in this session.

While the global session allows the reader to review and reevaluate key assessment criteria, identify changes, and provide the rationale for such modifications, both the original assessment and subsequent reassessment are captured in the read database and become permanent entries in the audit trail. However, if an original assessment is changed in the global session, only the final assessment is used to calculate overall progression or response.

The intent of implementing this session is to enable the readers to provide a more accurate assessment based on the overall burden of disease and the evolution of their radiographic findings over time. This affords the readers an opportunity to use their best clinical judgement during the independent read and more closely mirrors the clinical practice setting.

The Eligibility Read

Another type of read that pharmaceutical companies consider when planning a read design is the eligibility assessment. The eligibility read typically captures a subset of data and provides a high-level assessment of that data to determine whether potential subjects meet the inclusion criteria for a study. Sponsors of a clinical trial often employ an eligibility read if the personnel at the sites do not have as robust an understanding of the "measurability" of disease as the central readers or a sophisticated quantitative measurement has to be used, such as hippocampal volume in Alzheimer's disease. Sponsors also implement an eligibility read to mitigate the effect of PIs who may be overly eager to push patients with late-stage disease into a clinical trial.

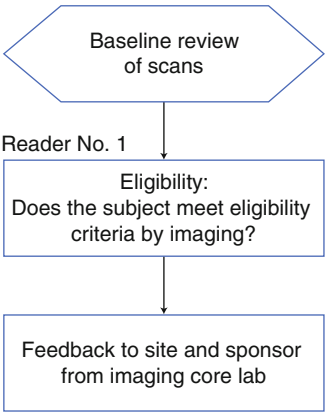


Fig. 4.2 Standard eligibility read design for oncology trials

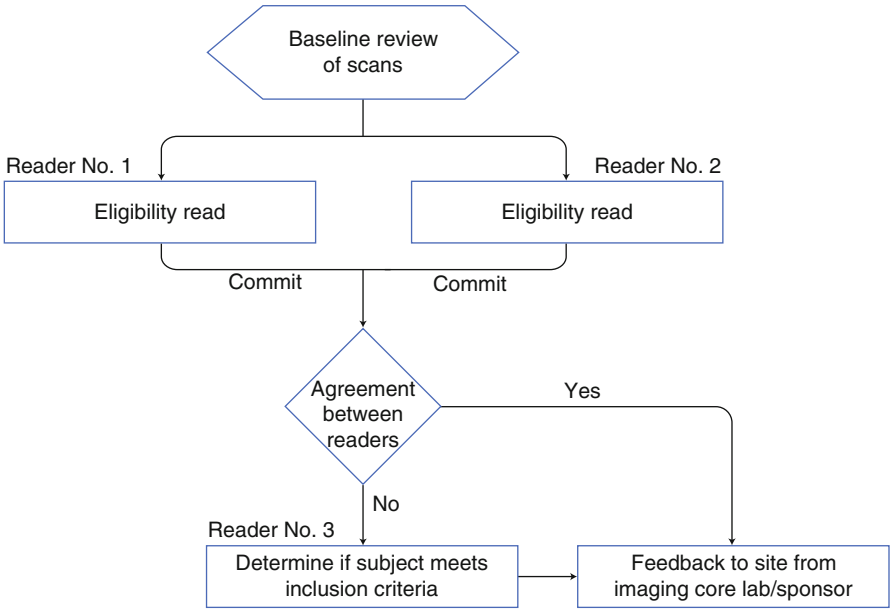


Fig. 4.3 Alternative eligibility read design for oncology trials

There are several different approaches to eligibility reads, but the two designs illustrated in Figs. 4.2 and 4.3 are the most prevalent. For both scenarios, it is important to establish in the Imaging Charter how the information is going to be conveyed to the sites and for the site PIs to understand that the assessment by the central readers takes precedence over assessments made at the site. In addition, the sponsor of a study must work closely with the imaging core lab to ensure that the data can be transferred quickly and that a timely assessment can be obtained and relayed back to the sites prior to patient randomization.

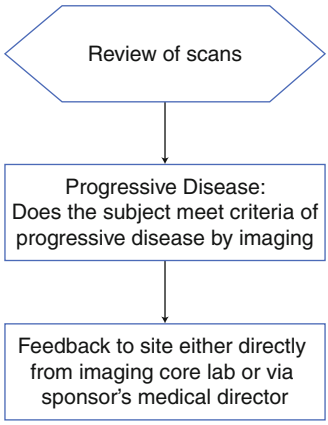


Fig. 4.4 Read design for confirmation of PD in oncology trials

Due to the rapid turnaround required for eligibility reads, usually a pool of readers has to be selected. The challenge is then to ensure all the readers evaluate images in the same manner or are "calibrated" to each other. A well-designed calibration program has to be developed, deployed, and continued throughout the course of the study.

Reading for Confirmation of PD in Oncology

In addition to these scenarios, sponsors sometimes request that the imaging core lab confirm progressive disease (PD) as part of the independent read (Fig. 4.4). This additional step is implemented when there is high risk of the sites declaring progressive disease prematurely, which often results in the patient being moved into another treatment arm of the clinical trial. When this happens, the patient's data set is "censored" by the sponsor and remains incomplete. When there is disagreement between the site read and the central read, the central read is considered to be truth. If the imaging core lab determines that PD has not been reached, that subject remains in the study and increases the likelihood that the data set will be complete.

Selecting and Screening Readers

Typically, medical and/or scientific staff at the imaging core lab with support from a clinical-operations type of department screens all potential readers to evaluate their past clinical trial experience. A reader's qualification depends upon their expertise and the type of study. The Imaging Charter for a particular trial will help the imaging core lab to understand what is expected of each reader. Reader qualification is based on formal training, performance on previous trials, experience with disease indication, image modality, and the reader's ability to meet timelines.

All readers should be board certified (or international equivalent) and, depending on the indication, may need to have additional qualifications. Imaging core labs will recommend that readers have a background check to ensure that the FDA has not debarred them from participating in clinical studies. Infrequently, sponsors may request the reader to complete a Financial Disclosure Form or a 1572.

A sound selection and screening model will identify readers who are considered experts in different therapeutic areas and/or modalities who are working in centers of excellence. These are often key opinion leaders who are eager to work in translational medicine, contributing to the clinical care of their patients as well as participating in clinical research.

Additionally, the medical and/or scientific staff at the imaging core lab should consider each project individually and work with a team to cross-reference a database of existing experts or to identify new readers, if needed.

After identifying potential readers, the imaging core lab contacts the reader and describes the nonconfidential details of the study as well as the projected time commitment. If the reader is able to commit to the study, their CV and reader rate is provided to the sponsor. If the study sponsor and reader are in alignment, the imaging core lab will execute a reader agreement and oversee the training, monitoring, and completion of the reader's participation in the trial. Study sponsors should approve all readers before they are assigned to a project.

Reader Training/Mock Read

Imaging core labs will provide training for the readers on a per-study basis. The selected readers often travel to the imaging core lab for a "Mock" read and are introduced at this time to the read system the imaging core lab has developed and to the read methodology. Training cases are evaluated during the session, both in a group setting if there is more than one reader, and individually by each reader, allowing the readers to come together at the end of the session to discuss methodology and rules for the reader. The imaging core lab provides documented read rules based on the read design outlined in the Imaging Charter, as well as the salient points discussed at the Mock read. The readers are then to follow the read rules to safeguard that a reader is consistent throughout the length of the study as well as among multiple readers.

Reader Monitoring

Imaging core labs often perform an evaluation of read results to ensure the adjudication rate (the percentage of cases between primary readers who disagree on an assessment and subsequently require the adjudicator to select one of the assessments) does not exceed the threshold that the study sponsor and the imaging core lab established. This assessment is also known as inter-reader variability analysis.

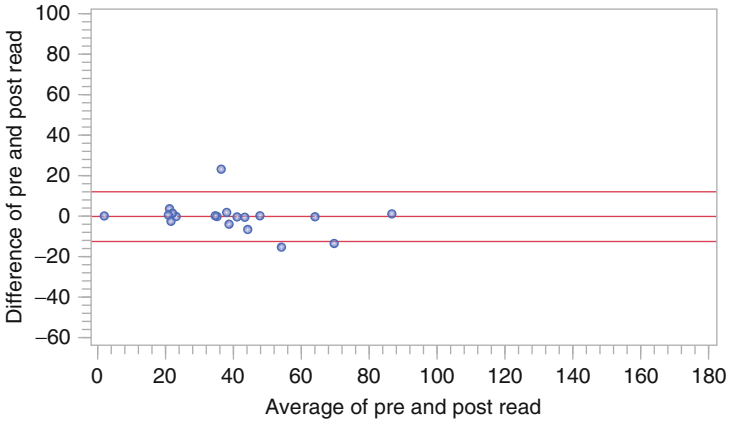


Fig. 4.5 Intra-read variability analysis Bland-Altman plot. Screening response diameter sum for Reader A. *Solid reference lines* include mean difference between pre- and post-read for one reader and mean difference ±1.96 SD

The same sort of analysis is often done to compare a reader's interpretation to their own earlier assessments of a selected set of cases. This intra-reader variability evaluation is performed to confirm that a given reader's rate of agreement with their earlier assessments demonstrates internal consistency in applying the response criteria and reader rules. This rate should similarly be monitored on an ongoing basis to ensure it does not exceed the threshold established by the study sponsor and the imaging core lab. The frequency of such evaluations should be determined at the inception of the read-monitoring plan.

Reader monitoring also safeguards that the readers are adhering to the reader rules and the criteria of the read. If the imaging core lab offers reader monitoring, the evaluation should be conducted on a regular basis and per the read design outlined in the Imaging Charter.

The methodology of reader monitoring usually includes an early assessment after a predetermined number of cases have been read by the primary readers and analyzed by the adjudicator, if necessary. The data are assessed based on the points of adjudication identified by the sponsor and the imaging core lab. Multiple statistical methods are applied as part of the reader-monitoring process. One example, the Bland-Altman plot, can be seen in Fig. 4.5.

If the read-monitoring assessment discovers that the concordance is not within the parameters that have been established, the data are further analyzed for trends or patterns to determine if the underlying cause is due to a lack of adherence to or a misinterpretation of the reader rules. If either of these causes is identified as the contributing factor for the lack of concordance, the imaging core lab provides a retraining session with the primary readers and adjudicator to review specific cases and then revises the read rules, if necessary, to document the points discussed in this reader retraining.

Adjudication in Oncology

Adjudication rates vary widely across trials, even of the same therapeutic area. The variance is based on the disease indication, the read endpoints being measured, number of readers, quality of data, the effect size of the therapeutic, and how many subjects lie on the boundary between discrete scores. For example, if the therapeutic has an effect size that is large, meaning response is dramatic, the adjudication rates will be minimized; the reverse is true if a relatively minor effect is expected as it is more challenging for readers to detect consistently the more subtle signals of response. The poorer the quality of the images, the tougher the interpretation of the images becomes, and this leads to potentially higher discordance between readers.

Adjudication trends are taken into account when developing an Imaging Charter for a specific trial, specifically when considering details for reader training and monitoring.

Use of the Imaging Charter in Adaptive Design Clinical Trials

In February 2010, the Center for Drug Evaluation and Research (CDER) and the Center for Biologics Evaluation and Research (CBER) released a draft version of a document titled *Guidance for Industry*: *Adaptive Design Clinical Trials for Drugs and Biologics* [2]. An adaptive design, according to the guidance, is one in which "adaptive features (i.e., changes in design or analyses guided by examination of the accumulated data at an interim point in the trial) that may make the studies more efficient (e.g., shorter duration, fewer patients), more likely to demonstrate an effect of the drug if one exists, or more informative (e.g., by providing broader dose–response information)" (2, pp. 5–6) can be included in the design of the clinical trial.

Pharmaceutical companies and the biotech industry can assess the benefits of improved accuracy, consistency, and efficiency in clinical trials in three main areas:

1. By decreasing the number of patients needed to reach statistical significance
2. By decreasing the length of the study
3. By potentially expanding the data gathered based on dose and response

Specific modifications that may occur throughout a trial based on an adaptive design and are defined a priori include:

- Study eligibility
- Randomization
- Treatment regimens
- Concomitant treatment
- Schedule of assessments
- Primary endpoint including outcome assessments
- Order of secondary endpoints and/or analytic methods to evaluate endpoints

A factor that is critical to the success of an adaptive design is the prospective nature of the modifications that are proposed prior to any unblinding of the data. In clinical trials that include an imaging component, the Imaging Charter is used to define this design and the impact of the independent read on making modifications to the trial.

Operational Considerations

In addition to clearly defining the imaging-related issues of a clinical trial and describing the read methodology and reader issues, the Imaging Charter also identifies other aspects of an independent review that must be defined to supplement the published criteria. This includes issues that arise due to data quality and missing data as well as providing instruction for how to apply assessment criteria when data are incomplete.

Similarly, most Imaging Charters specify how the imaging component of a read will comply with regulatory rules and laws, such as Health Insurance Portability and Accountability Act (HIPAA) requirements, the proper use of electronic signatures, and storage of data for potential audit reviews.

How to Handle Specific Data Issues

In virtually all studies, despite the best efforts of all parties, it is likely that some image data will be lost or improperly acquired, resulting in non-evaluable images. The Imaging Charter should define guidance on how to conduct the read with missing or suboptimal images. This includes guidance on how to handle missing or unreadable baseline image sets, missing or unreadable follow-up image sets, and incomplete time point image sets.

Handling Missing Image Sets

Readers should be made aware within the read software of any missing time points and prompt them to document that the specific missing time point is not readable and subsequently indicate that the response choice will be limited to assessment similar to “non-evaluable.” Depending on a sponsor’s planned analysis, subjects whose baseline imaging is missing but having follow-up time points can be either subjectively assessed (for progressive disease or non-evaluable) at follow-up time points or excluded from the independent read. Subjects missing all follow-up time points (“baseline only” subjects) may be included in the read with a limited assessment of baseline only or excluded from the read, depending on how the sponsor plans to use this data in their final analysis.

Handling Unreadable Image Sets

Image sets deemed to be of insufficient quality by the imaging core lab after attempts to acquire repeat imaging from the sites are still processed and loaded into the read system so that the independent central reader may make the final determination if the image set is readable. It is paramount to clearly define in the Imaging Charter the most likely reasons for an image set to be deemed unreadable, as this may vary based on disease indication and/or the modality acquired. Typical reasons for image sets to be deemed unreadable for an oncology trial with CT/MRI may include:

- Missing complete anatomical areas with documented or suspected disease
- Missing slices from examinations
- Missing images that contain a significant number of target and/or nontarget lesions
- Multiple images in an image window or one or more missing films (film-based images only)

Regulatory Compliance

As previously mentioned, the Imaging Charter should specify how the imaging component of a clinical trial will adhere to regulatory rules and laws. HIPAA compliance and 21 CFR Part 11 compliance are frequently the most critical regulatory issues the Imaging Charter addresses.

HIPAA Compliance for Digital Image Data

With the advent of the Picture Archiving and Communications System (PACS) and the global prevalence of digital data, the vast majority of image data for clinical trials is transmitted in direct digital format, not requiring any further digitization of the data. Following the receipt of digital image data on archiving media, an imaging core lab will Quality Control (QC) this data for protocol compliance and prepare the data for the independent read. During this process, the imaging core lab blinds all electronic header information (such as the subject identifier) within the digital data set to ensure HIPAA compliance. Similarly, if the core lab receives film data, they mask subject identifiers so that the actual pixels in the digital image are permanently replaced. In such a situation, the original film is left intact and is available as source data.

Supporting Internal Database Audits: Electronic Signatures and Records

Another regulatory concern the Imaging Charter usually addresses is how the read system will facilitate audits and comply with good-reading practices, as outlined in

21 CRF Part 11. The independent system should support the feature of an electronic signature. The read system typically employs electronic signatures by assigning a unique user ID and password to each user. This user ID and password authorize user access to the system. Whenever readers commit records to the database, they must enter their electronic signature. This ensures that their read responses cannot be modified by others.

To lock down the data associated with the images a reader assesses during a read, date and time stamps associated with the images and the reader's assessments are stored within the read system. These date and time stamps provide an audit trail to confirm the link between the displayed images and the evaluation for that particular electronic record.

Every electronic record should be captured in the database with a date and time stamp and should be duplicated in an audit trail. Any electronic record that is inserted, deleted, or clarified is thereby captured in the audit trail.

The Imaging Charter should also consider the following issues regarding electronic records in order for the independent read to be in compliance with 21 CFR Part 11:

- Security controls and system access
- Audit trails
- Protection of records for a given retention period
- Ability to supply copies of records during an inspection
- Developer and user qualifications
- System documentation control
- Controls over information transmitted over an open (external) network

To be compliant with 21 CFR Part 11 requirements, the imaging core lab and its read system should also support the following features of electronic signature:

- Unique identification of signers
- Appropriate display manifestation of signature on records
- Certification to the FDA of the imaging core lab's electronic signature equivalence
- Linking of signatures to records
- Periodic testing or rotation of signature components
- Monitoring of unauthorized use

Future of Charters

At the time of writing, the FDA has published a "Draft Guidance Document: Guidance for Industry Standards for Clinical Trial Imaging Endpoints" (Appendix 1 at the end of this book), which describes in more detail the requirements for Imaging Charters and the aspects of imaging that should be included in the document. This guidance was open for comments from August to October 2011. This guidance, if it remains unchanged, covers all therapeutic areas and describes the charter as having

the same significance as the clinical trial protocol. This has significant implications in clinical trial and program operations and development.

Perhaps the most significant impact to operations of a clinical trial with an imaging endpoint is the focus on Imaging Charter development at the time of clinical protocol development. Currently, the Imaging Charter is generally developed once an imaging core lab is contracted by the sponsor and clinical protocol development is in its final stages, if not complete. Moving the Imaging Charter development up in the process will require the imaging core lab services to begin much earlier in the process and key personnel from both the sponsor and the imaging core lab, including medical/scientific, operations, data management, and statistics to be involved in the read design from the very start.

In addition, considering the Imaging Charter to have the same significance as the clinical trial protocol necessitates a new level of oversight of the content of the Imaging Charter. This includes content describing procedures and operations at both the imaging core lab and the investigator site. While the imaging core lab has historically used the Imaging Charter as a guide for performing the imaging portion of a given trial, monitoring whether or not sites are adhering to the content of the Imaging Charter is an evolving challenge.

Conclusion

The clinical-study industry eagerly accepted the use of Imaging Charters even before the full benefits of the Imaging Charter were realized. The direct advantages of establishing a priori the key components of a clinical trial with an imaging component were quickly enhanced by the indirect value of encouraging study sponsors to address other critical imaging issues that affect their studies. A further indirect benefit is that of exposing regulatory agencies to potential modifications to read designs and especially scoring criteria that arise after they are tried and tested in practical application.

While Imaging Charters have changed considerably over the years, they still provide the basic content they offered when they were introduced. They also afford regulatory agencies and study sponsors greater efficiency in study planning and data review, thus leading to the collection of robust imaging data of high quality to support their submissions.

Appendix 4.1: Consortiums Established to Help Standardize Imaging for Clinical Trials

- Medical Imaging Stakeholders Call for Action: Standardization of Imaging Review Charters Task Force
  The Medical Imaging Stakeholders Call for Action conferences were established in 2007 and include participation of members of pharmaceutical companies, imaging core labs, the FDA, and various working groups. The Standardization of Imaging Review Charters Task Force was formed in 2007 and developed a draft IRC table of contents that was included in a Medical Imaging Standardization Technical Document provided to FDA in September 2007. The FDA reviewed this technical document prior to the October 2007 conference, and discussion at the conference resulted in the task force making minor modifications to the table of contents, as well as the establishment of a joint working group to continue development of a common IRC lexicon.
- Imaging in Clinical Trials/Uniform Protocols for Imaging in Clinical Trials (UPICT)
  Imaging in Clinical Trials/UPICT is a subgroup of the Clinical and Translational Science Awards Imaging Working Group (CTSA-IWG) and was originally formed in 2005 as a working group comprised of members of the Society of Nuclear Medicine (SNM) and allied organizations to create uniform imaging protocols for clinical trials. Group members are finding ways to reduce variance resulting from imaging used in clinical trials, as well as to improve the contribution imaging can make in a clinical trial. The group is working to develop, among other things, standard imaging protocol templates and to provide a web-based environment for those in the pharmaceutical industry, academia, and regulatory agencies working in the clinical trial space to better interact.
- Extended PhRMA (Pharmaceutical Research and Manufacturers of America) Imaging Group
  The Extended PhRMA Imaging Group is an open alliance of medical imaging professionals from numerous biopharmaceutical companies and imaging core labs devoted almost exclusively to biopharmaceutical product development. Its members work with stakeholders from imaging-device manufacturers, developers of image-analysis software, government-research institutes, regulatory agencies, academia, professional imaging societies, and other imaging-research alliances in an effort to standardize medical imaging in clinical trials.
- Quantitative Imaging Biomarkers Alliance (QIBA)
  QIBA is dedicated to improving the value of quantitative biomarkers in clinical trials. The group is comprised of researchers, healthcare and industry professionals, as well as imaging core labs. Focus is on best practices for use of quantitative imaging, across multiple imaging platforms.

References

1. Business Wire. EPIX Pharmaceuticals Receives Second Approvable Letter for Vasovist from FDA; FDA requires a re-read of images from previous studies and a new clinical study [document on the Internet]. FindArticles.com. 2005. Available from: http://findarticles.com/p/articles/mi_m0EIN/is_2005_Nov_23/ai_n15867545. Cited 15 Jun 2010.
2. US Department of Health and Human Services, Food and Drug Administration, Center for Drug Evaluation and Research (CDER), Center for Biologics Evaluation and Research (CBER). Guidance for industry: adaptive design clinical trials for drugs and biologics. Clin Med. 2010. http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM201790.pdf
3. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. European Organization for Research and Treatment of Cancer, National Cancer Institute of the United States, National Cancer Institute of Canada. J Natl Cancer Inst. 2000;92:205–16.
4. Eisenhauera EA, Therasseb P, Bogaertsc J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228–47.
5. Wahl RL, Jacene H, Kasamon Y, et al. From RECIST to PERCIST: evolving considerations for PET response criteria in solid tumors. J Nucl Med. 2009;50:122–50.
6. Cheson BD, Hroning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. J Clin Oncol. 1999;17:1244–53.
7. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25:579–86.
8. Macdonald DR, Cascino TL, Schold SC, et al. Response criteria for phase II studies of malignant glioma. J Clin Oncol. 1990;8:1277–80.
9. Bradnsma D, Stalpers L, Taal W, et al. Clinical features, mechanisms, and management of pseudoprogression in malignant gliomas. Lancet Oncol. 2008;9:453–61.
10. Wen PY, Macdonald DF, Readerdon DA, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. J Clin Oncol. 2010;28:1963–72.

Chapter 5
Defining the Radiological Blinded Read and Adjudication

Colin G. Miller

Abstract The term blinded read or independent read is used in literature to define a process by which medical images are read by an independent radiologist or team of radiologists and or physicians. There are three major kinds of blinded read: eligibility, safety, and efficacy, of which each of these is managed in a different manner. This chapter presents the various methodologies that are employed in blinded reads across phases of trials and between different therapeutic areas.

Keywords Blinded read • BICR • Blinding • Eligibility • Efficacy • Safety

Introduction

The evaluation of medical imaging in clinical trials has developed in the last few years to create a growing industry of imaging service providers. At the crux of the imaging core laboratory is the so-called independent read. Over time a number of different methodologies have been developed across therapeutic areas to remove bias and yet provide a meaningful radiological interpretation. The increasing size and complexity of imaging reads, such as the acceptance of adaptive clinical trial designs by the FDA, has led to the need for certain standards with methodologies that have been developed to present an approach to manage the read process more expeditiously while still maintaining the read integrity.

The term blinded read, central read, or blinded independent central review (BICR) within radiology is the practice of collecting all the images from a clinical trial acquired at the clinical sites and independently presenting the data to a

C.G. Miller, BSc, PhD, FICR, CSci
Medical Affairs, BioClinica Inc.,
826 Newtown-Yardley Road, Newtown,
PA 18940, USA
e-mail: colin.miller@bioclinica.com

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_5, © Springer-Verlag London 2014

separate group of radiologists or readers for evaluation. While this is generally understood, the nuances within this concept are vast and can change the powering of the study and outcome. These multiple read scenarios are usually required for the efficacy and safety reads, whereas eligibility reads are generally a "simple" single reader from a pool of readers. It should be noted here that the term reader is used throughout the chapter. In most instances, this term is synonymous with a radiologist, but in some indications this could be a rheumatologist, cardiologist, or orthopedic surgeon.

One aspect of BICR that is rarely discussed is the applicability to Good Clinical Practice (GCP) or Good Clinical Research Practice (GCRP). This is a set of standards and methodological implications that has been set down by the US Food and Drug Administration (FDA) and the Committee for Medicinal Products for Human Use (CHMP), as part of the European Medicines Agency (EMA) for clinical trials. While there is nothing specific for medical image management or reads in any of these documents, the overriding philosophy is that all data are reviewed by two pairs of eyes, e.g., a double data entry or source data verification. Therefore, the imaging corollary is that all images are reviewed by two pairs of eyes with agreement of the results. This becomes an important consideration as we review the read methodologies discussed next.

The BICR is in direct contrast to results provided by the local on-site radiologist(s). While there is usually only one principal investigator (PI) and study coordinator for every site, there may be many local radiologists. The radiologists will come from a pool in the radiology department, and it is unlikely there will be one person assigned to evaluate the images from the study. This methodology is accepted as best practices for point of care but is not adequate for clinical trials for several reasons:

- There is no audit trail for image evaluation particularly for measurements on images such as those for Response Evaluation Criteria in Solid Tumors (RECIST) criteria.
- The image display process will not be consistent within and between clinical trial sites.
- The available clinical information that is available and accessible to the local radiologist is variable.
- There is a bias caused by direct patient influence on the clinical staff.
- Many blinded reads use more than one reader which mimics the rest of the GCP process with two pairs of eyes reviewing the data. This is usually not available in a local situation.
- Clinical trials require the quantification by either direct measurement or the use of a scoring system. This is not a standard part of clinical methodology, where findings usually state that observations show increase or decrease in size or some other descriptive explanation.

There have been a number of papers that have drawn conclusions as to the appropriateness of the blinded reads and provided a meta-analysis of a number of reads without understanding the methodology involved [1–4]. This issue is further

complicated by the multitude of therapeutic areas being evaluated as well as the application of imaging modalities that continue to expand.

The BICR methodology needs to be defined a priori within a study and should be documented as within the methodology section of any paper describing the results of the read. However, it is now standard practice to include the completed read design in the Imaging Review Charter (IRC) as previously described in Chap. 4. The major differences can be described as Image Presentation and Radiologist or Reader Involvement.

The central review team will agree to a set of standards by which the images will be evaluated. This includes criteria for acceptable image acquisition, the process by which the images will be evaluated, and the generation of either a quantified output of the image interpretation, e.g., RECIST 1.1 criteria in oncology [5], or a semi-quantified or scoring system, such as the Sharp scoring system in rheumatoid arthritis evaluation of hand radiographs or modifications such as that by van der Heijde [6]. The methodology is detailed within the Imaging Review Charter and is monitored through quality control methods.

Within the clinical trial one aims to provide the best estimate of the disease activity (tumor burden, joint space narrowing, etc.) including more quantitative descriptors. This involves using current validated methodology in conjunction with modern equipment. It is paramount that the resultant measurements and imaging endpoint demonstrate good correlation with the clinical outcome of interest, such as disease progression, although it will never be 100 % because of the reasons previously described.

One additional constraint which needs to be considered is the regulatory environment. Health authorities have, in many therapeutic areas, used historical precedent in determining which radiological methods they will accept for registration of a pharmacological entity or medical device. By definition any newer method will differ in some aspects from the historical standards, and validation is required to compare the new modality (e.g., MRI in rheumatoid arthritis) with the historical standard (radiographs) and the clinical outcome. Efforts within the regulatory agencies such as the Critical Path Initiative within FDA are intended to address these issues. Sponsors are bound by the current regulatory rules for the selection of the imaging modality for their pivotal phase III registration studies, although they are free to add additional modalities from an exploratory perspective. This is discussed more fully in Chap. 2. Another key regulatory guidance is that an Imaging Review Charter detailing the imaging process is required prior to study initiation, as described in Chap. 4.

There are clearly significant differences between clinical care and the care provided within a clinical trial research protocol. One needs to anticipate that there will be differences in the read interpretation between local and central readers. It is anticipated that this may occur, though there is no definitive consensus on what level of disagreement is acceptable. Since patient care is driven by local decisions, this may result in patients being withdrawn from studies based on local reads. This needs to be considered in designing the read methodologies within clinical trials and in the statistical analysis plan.

Categories of Central Reads

There are three potential centralized read scenarios:

1. Eligibility reads: Reads to determine whether the study inclusion and exclusion criteria have been satisfied
2. Efficacy reads: Reads leading to the generation of efficacy endpoints
3. Safety reads: Reads to determine whether or not predefined deteriorations in disease activity have been met

Most of the work has been conducted with efficacy reads, but there is a growing number of eligibility and safety reads. The eligibility reads have historically been in the musculoskeletal arena where the primary disease might be incorrectly identified because of a subtle differential diagnosis or anatomical finding.

Eligibility Read

Sponsors often consider requesting that the central imaging core lab reviews the scans that are used to determine the subject's eligibility for inclusion/exclusion into the study. This is often employed if the investigator sites' understanding of "measurability" of disease is not as robust as the central core lab or in an effort to completely reduce the local bias or perception that site PIs may "round up" or "round down" on disease status to enable their patients to have access to the latest therapeutics.

It is important to establish how the eligibility information is going to be conveyed to the sites and for the site PIs to understand that the assessment by the core lab will take precedent over that determined by the site. In addition, much discussion is required with the core lab to ensure the data can be transferred quickly and a timely assessment can be obtained and relayed back to the sites, prior to patient randomization.

The obvious negative implications to an independent eligibility read is the delay in enrolling subjects since the evaluation has to be handled by a party that is remote and distinct from the PI and local study team. However, in many instances, this "gating" system pays off in the decreased cost of having incorrectly enrolled patients and also in the statistical cost of having incorrectly enrolled subjects. Eligibility reads, as stated, are most common in the musculoskeletal disease arena where there is significant subtle interpretation of the images. It is not unusual to have a 5:1 screen-fail-to-enrollment ratio, and the author has been involved in studies with the ratio as high as 10:1. This makes the eligibility read very cost-effective if one considers the financial impact of incorrectly enrolled subjects and/or a potentially failed study.

The ethics of conducting an eligibility read are sometimes questioned, since the initial thought is that the read is taking away some aspect of patient management and care from the PI. If the trial is correctly designed, there should be equipoise at the start of the study. With equipoise the PI should not be unduly influenced to enroll the patient compared to offering them current "state-of-the-art" therapies.

However, the PI is often under pressure from patients or their relatives requesting "the latest and greatest," so an independent third-party review provides an elegant manner for the PI to refuse a potential trial patient from entering the study, by referring to the independent read. The other pressure PIs are under is financial. A clinical trial, if well run, will be a good source of supplemental income to the PI or their institution. There is more often a competitive enrollment scenario to put pressure on the PI to recruit patients rapidly, thereby creating a further enrollment bias. This bias is not present with the independent eligibility reader, who will have no information as to the patients' details, location, or clinical background.

In conclusion, an eligibility read is merely ensuring the study criteria are met and not affecting the patient management.

Efficacy and Safety Reads

The following sections all describe the different read methodologies that have been employed predominantly in efficacy reads but also, to a lesser extent, safety evaluations. The read designs are descriptive, and the statistical concepts and evaluations will not be considered. However, in many situations it is standard practice to build in a 10 % over-read with approximately 10 % of the images put back into the read process, randomly, without the readers being aware of this action. This process allows for a thorough evaluation of inter- and intra-reader variation.

Consensus Read

This is the read performed by two or more readers (usually three) who review all study images together contemporaneously (Fig. 5.1). At one time, this could have only been achieved with all readers in one room using film. Given the technological developments, the readers can be at remote locations and reading connected via the phone or Web-X. Depending on the number of cases, this may require a number of sessions to accomplish the evaluation of all images.

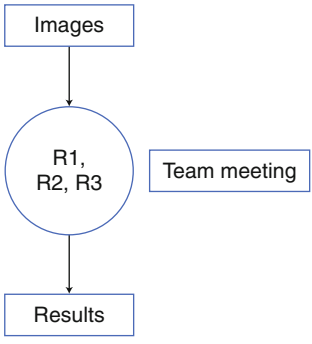


Fig. 5.1 Consensus read. *R1* = Reader 1, *R2* = Reader 2, *R3* = Reader 3

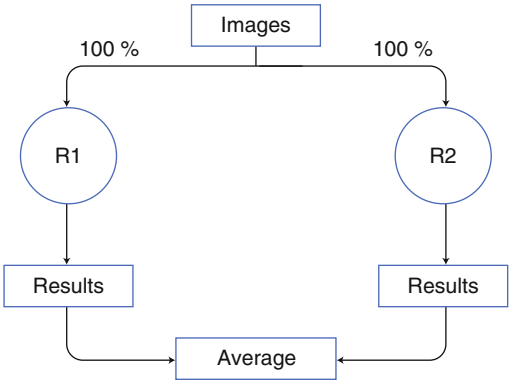


Fig. 5.2 Paired read, no adjudication. *R1* = Reader 1, *R2* = Reader 2

At face value this provides a straightforward methodology, with a single outcome of the read. However, there are a number of pitfalls, the primary one being that the so-called "alpha reader" who dominates the read decisions, and essentially this becomes a single reader's interpretation. This issue develops further the longer the read continues since reader fatigue increases with a greater incentive to complete the read and the less dominant readers coming in-line rapidly with the alpha reader interpretation. A second pragmatic pitfall is the challenge of getting the readers together to complete the reads, particularly for large numbers of patients.

Considering the pitfalls described previously as well as the onset of other read designs, consensus reads are employed less commonly.

Paired Read, No Adjudication

The next least complex methodology, which was used in early rheumatoid arthritis studies in particular [7], is to have two single independent readers read 100 % of all the images, without knowing the results of the other reader (Fig. 5.2). Where there is a discrepancy or discordance, the average of the two reads would be used as the outcome. This averaging smoothes out the variation and provides a single result, but does not address the issue of why there may have been a variation in the read. Therefore, there are two issues from a GCP viewpoint: the first includes that while two pairs of eyes have reviewed the images, where there is discordance, it is unknown whether this is a case of incorrect data entry, misinterpretation, or true disagreement. Secondly, averaging these results does not solve the first issue and from an audit trail perspective does not provide a complete data trail back to one image; there is no such thing as an average "read" by a radiologist.

Standard Paired Read, with Forced Adjudication

The two-reader paradigm with a third reader to adjudicate the discordances, is arguably the most common scenario (Fig. 5.3). When conducted in its simplest form, there are just two primary readers who both read 100 % of all the images and

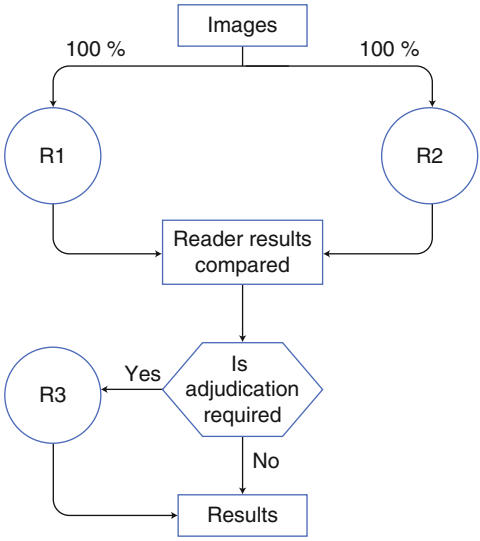


Fig. 5.3 Paired read with adjudication. Forced or open adjudication. *R1* = Reader 1, *R2* = Reader 2, *R3* = Reader 3 or "adjudication"

a third reader as an adjudicator. Within this paradigm there are two variations: the so-called forced adjudication and the open adjudication. In the forced adjudication scenario, the two agree with a set of results in its entirety, from one of the two primary readers. In this manner, all images are reviewed by two pairs of eyes with the same results, and it precludes the potential for a third set of discordant results from occurring.

There are some subtleties to the adjudication depending on therapeutic area; in oncology the adjudicating reader is normally asked to choose one reader over the other at the patient level. In osteoporosis, for example, the adjudicator should be able to choose at the vertebral level between readers.

Standard Paired Read, Open Adjudication

The alternative methodology to the previous paradigm is to allow the third reader not to truly adjudicate but to independently evaluate the images and provide a third set of results (Fig. 5.3). This now provides a further set of data from the interpretation without truly fulfilling the GCP requirements for two pairs of eyes to review the same data. Three people have reviewed the data, but with three different sets of results, there is no decision regarding which one should be considered "truth." While the third reader's results are often used, there is no logical reason why the first or second reader's interpretations should be used. The other variation is to average the results from the closest two readers. Oncology studies have generally favored the use of forced adjudication. Studies in rheumatoid arthritis have tended to use the open style of adjudication.

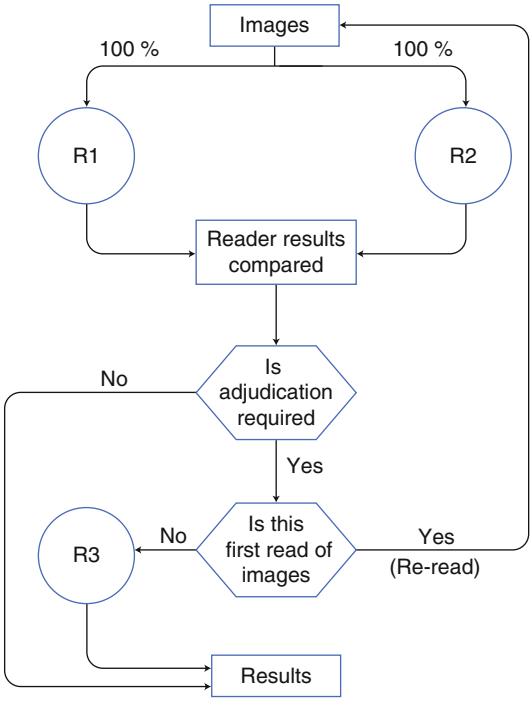


Fig. 5.4 Paired Read, reread then adjudication. *R1* = Reader 1, *R2* = Reader 2, *R3* = Reader 3 or "adjudication"

Paired Read, Reread Then Adjudication

In the field of rheumatoid arthritis, there is yet another variation on this theme that is being evaluated: the images that are identified as needing to be adjudicated are returned to the readers in a blinded manner, and they reread as if they were reading them for the first time. Only after this second double evaluation are they sent to an adjudicator if they still pass the adjudication threshold (Fig. 5.4).

Pseudo-Paired Read and Adjudication

Some organizations that provide a central review process have developed a methodology where there are several readers who are available for reading clinical trials. Each day they read the images from any of the trials that require reading. In this situation, the "pair of readers" can be any two radiologists who are available to read the data that day. In this situation there is no consistency to the pairing (Fig. 5.5). The adjudication that is used can be either of the two methods described but with any available radiologist, and again there is no consistency with the adjudicator. If a forced adjudication process is used, this methodology does provide a process that meets GCP requirements. However, the downside is that there is a higher degree of variability of the readers, more closely meeting the variability of the so-called local readers. One of the many reasons that central reads were developed was to reduce

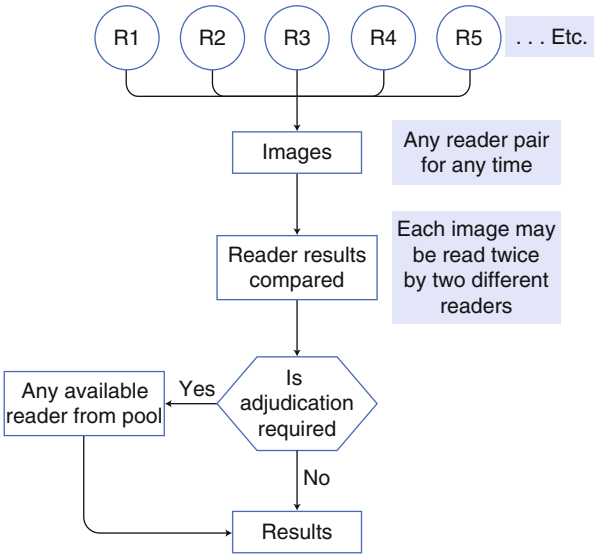


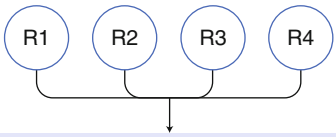
Fig. 5.5 Pseudo-paired read and adjudication. *R1* = Reader 1, *R2* = Reader 2, *Rn* = Reader n, etc

the variability of the reads, and this only increases the variability. A second factor that plays into this paradigm is that the readers all tend to be from the same institution. This provides more consistency but also bias, since they all tend to read the same way. Originally, the read designs were developed to have readers from different institutions, so that the inter-institution variability would be picked up, making the design more robust with the third forced adjudicator.

Multiple Paired Readers, Forced Adjudication

As pharmaceutical studies get larger, then it is impossible for two readers and one adjudicator to evaluate all the images in a timely manner. A number of different scenarios have been developed to accommodate this issue, from pairs of pairs to a more elaborate process recently described, with four readers. If we take a scenario of four readers being used for a study, there are six potential parings. If the images are sent in a randomized manner to two readers (pair), any adjudication that is required can be carried out by one of the other two radiologists not from that pair (Fig. 5.6). As a study progresses and a statistically significant number of images are read, the rate of concordance and the rate requiring adjudication between the readers can be evaluated. It is therefore possible to determine if one of the readers is evaluating the images differently by the higher adjudication rate observed with his/her pairings.

One can also assess the readers by determining the percentage of reads proceeding to adjudication for which the final determination is consistent with one of the original readers. If one assumes equal performance among all readers, the variation in adjudication will be 50 %, i.e., equal favoring to all readers in the



Approximate read volume	Six read pairs	Potential adjudicators
33 %	R1 + R2	R3 or R4
33 %	R1 + R3	R2 or R4
33 %	R1 + R4	R2 or R3
33 %	R2 + R3	R1 or R4
33 %	R2 + R4	R1 or R3
33 %	R3 + R4	R1 or R2

Fig. 5.6 Multiple paired readers, forced adjudication. This model is designed for large studies requiring extensive reading. Unlike the pseudo-paired approach, this is very tightly controlled so that there are limited pairings and adjudicated rates can be evaluated. *R1* = Reader 1, *R2* = Reader 2, *Rn* = Reader n, etc

pair. If this variation is significantly different from this value, it is suggestive of a reader that is calibrated differently to the other readers. If this process is performed correctly, high rates of adjudication or high rates of discordance with the final adjudicated result could be used to retrain, "recalibrate," or replace readers who are reading the criteria differently to the general consensus at predetermined thresholds.

To be more specific, if we take the example of readers A, B, C, and D, there are 6 pairings: AB, AC, AD, BC, BD, and CD. If we have say 540 sets of patient images that have been read, each pair will have handled 90 sets each. (This is utilizing the equal shared methodology). Let's assume that the adjudication rate is around the 10 % rate. We can then anticipate that each pair will have nine subjects' images that will be read by one of the other readers. If the generally accepted maximum of 20 % is used, 18 pts per pair could be anticipated to be adjudicated. However, if pairing AB, BC, and BD have a higher adjudication rate, then it might be anticipated that reader B is not reading to the same standards as the other readers. A review of the adjudication rate would confirm this: if more of the adjudication results went to readers A, C, and D in these pairings, then it would easily highlight that reader B requires retraining or recalibration of their read methodology.

A further nuance to this model would be the ability to expand the number of images read by readers who are demonstrating good performance during the trial and not requiring equal number of images to be read by all readers. One reader might therefore be allowed to read more subjects due to availability, but this balance would be maintained by the pairing of the results. The other advantage of this model over the now classical model is that it does not require a 10 % over-read, since the Inter-reader evaluation is built into the methodology. This therefore reduces the costs and read time while maintaining scientific credibility.

The challenge to this methodology is the need for an elegant read system and way of allocation of images and adjudication of images. A standard Picture Archival System (PACS) will not suffice, since the results need to be captured in a logical and electronic manner in real time. This requires a specialized read system, such as those now standardly used in clinical trials by the so-called imaging core labs (ICLs) [8].

One final design can be considered where, for a trial with a large number of sites, a group of readers based at a subgroup of trial centers become the central readers. This design requires some further operational considerations but has a significant impact on outcome as these radiologists would be well trained, thereby ensuring high-quality image data at their own site (peer pressure to not submit poor-quality data would be high). The operational considerations are that no reader could read images coming from their own site, due to bias. With a pool of readers, images could be read by the first available readers. With a larger number of readers, rather than have two plus one for adjudication, a larger sample of readers could be employed, e.g., 5 and take either an average or the blinded consensus. This has an interesting statistical sampling technique. As far as the author is aware, this methodology has not been employed in a large part because of the operational challenges. However, as technology continues to develop, this scenario is one that may be on the horizon.

Progressive Disease Read in Oncology

In addition to these scenarios, sponsors sometimes request that the imaging core lab functions to confirm Progressive Disease in on oncology studies. This additional step is implemented when there is high risk that the sites declare "Progressive Disease" prematurely, perhaps to cross the patient over to another treatment arm. If there is ultimately disagreement with the site read and the central read, the central read takes precedent. If not, the patient's cross over is ultimately a "censored" data point by the sponsor during statistical analysis, and the data set is incomplete. This read design paradigm is similar in many aspects to the eligibility read.

Image Presentation

Basic Blinding

A BICR assumes intrinsically that the first level of blinding is being maintained: the reader cannot see the subject's specific identifiers, the study arm of therapy administered, or the site where the subject has been enrolled. In contrast, it is common for local sites to strive for confidentiality, but subject identifiers are visibly present. Similarly, it is not uncommon for clinicians to come to the reading room at a site and request that a reader reviews the images while they share clinical information.

Blinding to Temporal Sequence

The next level of blinding is to remove all knowledge of temporal sequence from the reader. This can be achieved in a number of ways, but the most basic is that the reader reads each time point for each patient in isolation from all other time points. The next step in blinding is to provide all the images contemporaneously in a random timepoint manner. This methodology is used primarily in clinical trials of Rheumatoid arthritis and other inflammatory diseases.

The challenge with giving single time-point images to review is that by definition it does not provide the reader any information as to the preceding time points. Therefore, the alternative is to provide the reader with all the images from the patient in an organized chronological sequence. This is much more in line with a clinical read but does have some drawbacks, particularly in oncology where Progression-Free Survival (PFS) is the primary endpoint. If the reader can see all the images in one read session, there is the tendency to observe lesion nadir and identify the time to identified progression. All other images are therefore interpolated. A further subtlety is the fact that if a subject has not performed well on the trial, then there are fewer images to evaluate than a patient who has had a positive response to therapy.

Therefore, an alternative paradigm can be employed to have the baseline image always displayed and known. Each subsequent image can then additionally be displayed and scored without the future time points known. This removes the bias of interpolation and ensures the reader concentrates on just the latest image being shown and has no idea that there are more to follow. There is another version of this procedure, where the time points are displayed in single randomized manner but always compared to baseline. This has few advantages over the sequenced single time-point display but a number of disadvantages, since to ensure this kind of read is coherent, all the time points will have to be displayed in one go to ensure the reader can evaluate a true global response.

Incidental Findings

At the time of writing, the issue of what to do with incidental findings or other radiological observations has been documented without any unified clarifying methodology [9–17]. There are a number of subtle issues, including the ethical response, what the patient expects as part of the informed consent, the overreporting of observations, and what happens if the central reader misses an incidental finding or is not trained to identify the nonspecific items, e.g., EG a technologist or PhD scientist performing complex image segmentations, which do not require a radiological interpretation. It is not the remit here to go into all the finer details of this aspect of the concept but to provide the broad concept and considerations that need to be defined a priori to the read.

Most informed consents will and should explain to the subject that the images will be sent off site for a third-party evaluation. Ideally, the informed consent should have language that further states that any evaluation of these images or third-party finding will not be relayed back to the subject. However, this latter statement is rarely in the informed consent, and therefore, it is the subject's expectation that they will be notified of any third-party findings. It is part of the benefit they see of participating in a clinical trial.

With this expectation now the ethical requirements are that the PI, having the clinical management of the patient, should receive notification of any incidental findings. This may be some considerable time after the event was noticed if the efficacy readings are being conducted in bulk at a later stage. It is then up to the PI to evaluate this information in light of all the other clinical information and patient history at their disposal and inform the patient or filter it out. This is a very simplistic view of the process, as there are level findings that have more clinical significance or not.

Blinded readers are by definition removed from the clinical information of the subject and are reading to identify a primary endpoint, such as change in lesion burden, vertebral deformity, and gray matter atrophy. Unless specifically requested, the readers will not be looking for other disease and may in fact miss many other aspects of the read, since these are not clinical reads but central reads for clinical trials. Therefore, the central reader cannot be relied upon unless stated a priori that they are to look for other findings. However, if they have other observations, these should be reported.

It is hoped that, over time, the ethical challenges of reporting incidental findings can be reduced by having appropriate language in the informed consent. Until that time those involved in clinical trials have to consider how to handle incidental findings, be they from a medical image or blood test or genetic screen.

Pre-read Training and Reader Calibration

As the number of reader's increases, it is imperative to develop steps to improve the consistency of reporting. All readers must play an active role in reviewing and potentially modifying some aspects of the read criteria to ensure that they have a clear understanding of the trial objectives. Another aspect of the trial methodology that should be considered standard practice for all central imaging reviews is the so-called mock read or pre-read training or calibration. This is a round table meeting where all parties and especially the readers review the methodology to be employed in the study. This should also include a period of blinded review of non-study but similar patient sets of data. The results of the readers can then be compared and discussed, and agreements on how to interpret common findings are established before the study read is conducted. This step is especially important in clinical trials, where many of the methodologies being employed (RECIST criteria, Sharp scoring, etc.) are rarely or never used in the routine clinical setting. This mock read

then sets the standard and provides a "calibration" of the readers as to how the scoring or measurement system will be employed. With this methodology defined, intra- or intercalibration or reader drift can be evaluated either during the study if sufficiently elegant read schemes are employed or post-analysis conducted. One methodology to evaluate the calibration drift as previously discussed is by the use of repeat images being injected back into the reads. With large studies that are being read over a number of days or months, this is just a matter of randomly selecting images to back to each reader. However, the "random selection" of 10 % of images, depending upon the type of images selected, may not be representative of the overall images for the study. For most radiological studies there are a group of subjects that will be on the borderline of two categories of definitions of the disease. This phenomenon is most relevant if a semiquantitative scoring system is used which requires precise radiological interpretation, e.g., Sharp or Genant scoring system. It is at these boundary areas where there is the greatest variation in interpretation. Therefore, if subjects are selected for the over-read with results within the boundary area, there will be poorer inter- and intra-reader agreement than when the over-read is predominantly of subjects who are definitively within one of the categories.

Conclusion

Radiological reads in clinical trials have spawned a new industry in the last few years of imaging core labs that employ varying process in the acquisition, processing, and interpretation of images. For many of the processes employed, there has been minimal documentation in the scientific literature to support the methodologies. Of late, the FDA has taken this lack of documentation and worked with the industry to develop white papers [18] to create definition around the process and terminology. The lack of detail in the clinical trial literature is not surprising considering the relative infancy of this methodological approach to a subsection of clinical trials. The confusion in both the clinical trial and general radiological community is evident to the different read methodologies and requirements. Further development of these concepts needs to be reported, and the pharmaceutical and medical device industry should consider publishing the read methodology and design as part of the detailing of the clinical trials.

References

1. Amit O, Bushnell W, Dodd L, et al. Blinded independent central review of the progression-free survival endpoint. Oncologist. 2010;15:492–5.
2. Amit O, Mannino F, Stone AM, et al. Blinded independent central review of progression in cancer clinical trials: results from a meta-analysis. Eur J Cancer. 2011;47:1772–8.
3. Dodd LE, Korn EL, Freidlin B, et al. Blinded independent central review of progression-free survival in phase III clinical trials: important design element or unnecessary expense? J Clin Oncol. 2008;26:3791–6.

4. Meignan M, Itti E, Bardet S, et al. Development and application of a real-time on-line blinded independent central review of interim PET scans to determine treatment allocation in lymphoma trials. J Clin Oncol. 2009;27:2739–41.
5. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228–47.
6. van der Heijde D, Dankert T, Nieman F, et al. Reliability and sensitivity to change of a simplification of the Sharp/van der Heijde radiological assessment in rheumatoid arthritis. Rheumatology (Oxford). 1999;38:941–7.
7. Landewe R, Van Der Heijde D. Radiographic progression in rheumatoid arthritis. Clin Exp Rheumatol. 2005;23:S63–8.
8. Miller CG. Medical imaging core laboratories. Appl Clin Trials Acts. 2005. Available at: http://www.appliedclinicaltrialsonline.com/appliedclinicaltrials/article/articledetail.jsp?id=182766&pageid=2.
9. Berland LL, Silverman SG, Gore RM, et al. Managing incidental findings on abdominal CT: white paper of the ACR incidental findings committee. J Am Coll Radiol. 2010;7:754–73.
10. Berlin L. The incidentaloma: a medicolegal dilemma. Radiol Clin North Am. 2011; 49:245–55.
11. Booth TC, Waldman AD, Wardlaw JM, et al. Management of incidental findings during imaging research in "healthy" volunteers: current UK practice. Br J Radiol. 2012;85:11–21.
12. Lumbreras B, Donat L, Hernandez-Aguado I. Incidental findings in imaging diagnostic tests: a systematic review. Br J Radiol. 2010;83:276–89.
13. Morris Z, Whiteley WN, Longstreth WT, et al. Incidental findings on brain magnetic resonance imaging: systematic review and meta-analysis. BMJ. 2009;339:b3016.
14. Orme NM, Fletcher JG, Siddiki HA, et al. Incidental findings in imaging research: evaluating incidence, benefit, and burden. Arch Intern Med. 2010;170:1525–32.
15. Siddiki H, Fletcher JG, Mcfarland B, et al. Incidental findings in CT colonography: literature review and survey of current research practice. J Law Med Ethics. 2008;36:320–31.
16. Vernooij MW, Ikram MA, Tanghe HL, et al. Incidental findings on brain MRI in the general population. N Engl J Med. 2007;357:1821–8.
17. Wolf SM, Paradise J, Caga-Anan C. The law of incidental findings in human subjects research: establishing researchers' duties. J Law Med Ethics. 2008;36:361–83.
18. FDA. Guidance for industry standards for clinical trial imaging endpoints. 2011. Available at: http://www.fda.gov/downloads/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/UCM268555.pdf.

Chapter 6
Medical Imaging in Drug Development

Joel Krasnow

Abstract Imaging technology when appropriately employed can provide a competitive advantage in the development of pharmaceuticals. Key success factors include (1) a clear vision for the project that incorporates current and anticipated future treatment options for the primary disorder targeted by the proposed treatment, (2) understanding of standards of care across the world, (3) current and evolving imaging standards, and (4) regulatory authority precedent and emerging standards specific to the therapeutic indications being sought. Imaging biomarkers provide the ability to detect change in disease much earlier than standard clinical endpoints. They can also provide timely, functional information at the molecular, cellular, or tissue level regarding the impact of pharmacological intervention in a disease process. These properties can make imaging a valuable tool in preclinical as well as in clinical development.

Keywords Biomarker • Regulatory • Benefit to risk • Pharmaceutical market

Introduction

Over the past decade the pharmaceutical industry has been investing increasing funds into research and development, yet fewer new drugs or biologics have been approved by global health authorities [1, 2]. Scientific milestones during this time period include the sequencing of the human genome, advances in genomic technologies, and advances in medical imaging. The Critical Path Initiative (CPI) is FDA's national strategy for transforming the way FDA-regulated products such as

J. Krasnow, MD, MBA
Department of Strategy and Regulatory Affairs,
TherapeuticsMD,
Boca Raton, FL, USA
e-mail: jkras@optonline.net

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_6, © Springer-Verlag London 2014

Table 6.1 Phases of drug development and associated milestones

Development phase	Desired outcome
Basic research	Target identification
Discovery/candidate optimization/ prototype development	Molecule or prototype that has suitable properties to interact with the pharmacological target in vivo
Preclinical development	Demonstrate proof of principle in animal models
Early clinical development	Pharmacokinetics, pharmacodynamics, proof of concept, dose selection
Late clinical development	Demonstration that the product is efficacious and safe

human drugs, biological products, medical devices, and veterinary drugs – are developed, evaluated, and manufactured [3, 4].

One of its objectives is to improve the number of pharmaceutical and medical device products developed in order to improve the health of the nation. It is acknowledged that advances in imaging technology have not translated into improvement in clinical trial productivity. In an effort to close the gap between imaging potential and the use of imaging to improve the clinical development process, FDA in August 2011 released a draft guidance for industry titled: "Standards for Clinical Trial Imaging Endpoints." This chapter will review the development process for new pharmaceutical agents with a focus on the role of imaging. The same general principles apply to the development of medical devices.

New Product Development

Drug development involves a series of activities beginning in the research laboratory and culminating in the marketing of a new pharmaceutical agent (Table 6.1). This multidisciplinary process requires professionals with diverse skills to contribute to projects that span several years. Drug development begins with basic research into human physiology and pathophysiology. From this basic research, one or more hypotheses are formed which predict that increasing or decreasing a particular substance will have a beneficial effect on a targeted disease state. A strategy to modify the concentration or biological activity of this substance will be developed. One example is the inhibition of CNS neuronal uptake of serotonin, which is characteristic of the class of drugs known as serotonin selective reuptake inhibitors (SSRIs). Once a pharmacological target is selected, its activity may be modified with a monoclonal antibody, interference RNA, a recombinant protein, a small molecule, or other strategies. If a small molecule approach is selected, then potential molecular structures must be assessed for physical chemical properties, and known structure function correlations. The potential for both on target and off target safety effects must be assessed. In essence, the discovery group is responsible for discovering/developing a molecule that has the desired effect on the selected target, with minimal off target effects. Once a molecule is developed that meets these in vitro specifications, the manufacturing will be scaled up to enable preclinical or animal testing.

The objective in preclinical research is to demonstrate that the investigational product performs as desired. For a cholesterol ester transfer protein (CETP) inhibitor, this would mean that it increased HDL and lowered LDL. For a rheumatoid arthritis treatment, it may mean improvement of inflammatory biomarkers or improvement at the level of the joint. The criteria used to satisfy proof of principle and the animal model(s) selected can have a significant influence on subsequent steps. Whenever possible, it is recommended that the criteria used for proof of principle in animal models be similar to the criteria that will be used in phase II trials in humans. Once proof of principle has been demonstrated, distribution and metabolism of the pharmaceutical product is known, and appropriate toxicology experiments have been conducted, an investigational new drug application to test the drug in humans can be considered. The investigational new drug application is a major milestone in the drug development process which requires careful documentation of years of preclinical work. Guidance documents from health authorities such as FDA and EMA can be found on their websites and are helpful in the preparation of regulatory submissions.

The phase I or first in human studies are usually conducted in specialized facilities where the study subjects are closely monitored. Initially a single dose is administered to a single subject. Once a specified number of study subjects have completed a single dose and no clinically significant side effects are observed, the dose of the study drug can be increased. These studies are referred to as single ascending dose studies. The objective is to determine the dose range where efficacy is observed and side effects are minimal and within an acceptable rate. Next multiple dose studies are performed. In these studies, subjects receive multiple doses of the study medication. These are usually one daily dose for oral medications (dependent of the half-life) in order to better delineate the therapeutic dose range. For monoclonal antibodies, the rate of administration may be less frequent such as twice weekly, weekly, or monthly. Pharmacokinetic and pharmacodynamic assessments are performed as part of most phase I studies. The objective at the end of phase I is to have sufficient information regarding the dose range that will be required to demonstrate proof of concept in humans.

During phase II several dosing regimens will be assessed using a placebo-controlled or an active comparator experimental design. The objectives are to prove that the pharmaceutical product achieves the desired clinical effect (proof of concept in humans) and to determine the optimal dose or doses to be carried forward to larger phase III trials. In the design of the phase II trial, at least one dose should be higher than the anticipated optimal dose and at least one dose should be lower than the minimally effective dose such that the optimal dose or doses become apparent as a result of the study. In reality, this is seldom the case. Due to the limited number of subjects in these studies, surrogate endpoints are heavily relied upon to determine dose selection. Well-conducted imaging studies can add significant value during early clinical development [5]. This is because imaging studies can often accurately measure changes in pathophysiological processes, thereby providing valuable information for either efficacy or safety.

Phase III clinical trials require substantial strategic, technical, operational, and financial resources. The objective is to demonstrate that the new product in its studied route and frequency of administration provides a clinically meaningful benefit compared to the risks involved for the study population that has been investigated. The concept of benefit to risk ratio is paramount. Historically sponsors focused on demonstrating benefit while collecting adverse events in a routine fashion during phase III trials. Today, that strategy is unlikely to be successful in many therapeutic areas. Safety must be actively assessed by identifying potential safety risks and designing studies to evaluate the risk relative to placebo or active comparators. A recent example is the serotonin 2b antagonist lorcaserin for weight loss where echocardiography was performed to assess cardiac valvular function [6].

In retrospect, identification of a clinical target appears simple. We will use the example of hypercholesterolemia to demonstrate this concept. Basic research identified the key physiological steps in the pathway for cholesterol synthesis. This revealed several potential steps in which the synthesis of cholesterol could be inhibited to lower serum total and LDL cholesterol. Pharmaceutical developers targeted the HMG-CoA reductase enzyme and the products known as statins emerged.

A more recent target that was selected for the treatment of hypercholesterolemia is the cholesterol ester transfer protein. CETP transfers cholesterol from HDL cholesterol to very low-density or to low-density lipoproteins (VLDL or LDL). Inhibition of this process results in higher HDL levels and reduces LDL levels. Torcetrapib was the first molecule of CETP inhibitors that demonstrated a dose-dependent increase in HDL and a decrease in LDL with and without an added statin [7]. In the phase III trial, there was a 58 % increase in deaths among patients taking torcetrapib and atorvastatin versus those taking atorvastatin alone [8]. Some scientists believe that the increased mortality observed with torcetrapib was secondary to unintended increases in blood pressure [9]. These scientists and their organizations have continued to develop their CETP inhibitors by evaluating their prospective compounds for changes in blood pressure. Due to current limitations in the understanding of lipid physiology, there is uncertainty as to whether CETP will be a viable target for pharmaceutical intervention.

At the conclusion of a phase III program, there is a large amount of data that is available pertaining to the pharmaceutical product. Analysis of this data can be valuable in determining the potential for use of this drug in additional indications. Bevacizumab (Avastin®, Genentech, San Francisco, CA, USA) which is a monoclonal antibody to vascular endothelial growth factor will be used as an example of closely related additional indications. In February 2004 the FDA approved Avastin for use in combination with intravenous 5-FU-based chemotherapy as a treatment for first-line metastatic colorectal cancer. In June 2006, the FDA approved Avastin in combination with intravenous 5-FU-based chemotherapy for patients with metastatic colorectal cancer who have been previously treated for their cancer (or second-line metastatic colorectal cancer). Investigation of additional tumor types followed such that in October 2006, the FDA approved Avastin in combination with carboplatin and paclitaxel for the first-line treatment of patients with unresectable, locally advanced, recurrent, or metastatic non-squamous, non-small cell lung cancer. Subsequently approval for glioblastoma and metastatic renal cell carcinoma followed.

At times, the emerging data may indicate potential application for a very different patient population. Zoledronic acid will be discussed as a representative example of this situation.

Zoledronic acid is a bisphosphonate drug that works by inhibiting osteoclast-mediated bone resorption. It was first approved by the FDA in 2001 for the treatment of hypercalcemia of malignancy at a dose of 4 mg per infusion with retreatment permitted after 7 days. In 2002 zoledronic acid was approved for patients with multiple myeloma and patients with documented bone metastases from solid tumors at a dose of 4 mg per infusion every 3–4 weeks. During its development for oncology uses, it became apparent that zoledronic acid would also be useful in patients with metabolic bone disease. A development program for metabolic bone diseases was initiated. In 2007 it was approved first as a single 5 mg infusion for the treatment of Paget's disease of bone. Studies were performed to support additional indications within metabolic bone disease. It was approved as a 5 mg once-yearly intravenous treatment for osteoporosis in postmenopausal women. In 2008, zoledronic acid at a dose of 5 mg annually was approved for the prevention of fractures following a hip fracture and for the treatment of osteoporosis in men. In 2009, it was approved for the treatment and prevention of glucocorticoid-induced osteoporosis in patients expected to be on glucocorticoids for at least 12 months and for the prevention of osteoporosis as a single 5 mg dose that is effective for 2 years.

Marketing authorization is based on all the discovery, preclinical, and clinical studies performed to date in clinical trials through phase III. Phase IV studies are designed to provide additional data that are of value to patients and healthcare practitioners that were not collected during the phase III studies. These must be conducted within the current prescribing instructions. They may investigate specific populations, compare dosing regimens, monitor a safety parameter, or investigate a new efficacy endpoint. It is common for health authorities to make marketing authorization contingent upon the conduct of additional post-marketing studies to assess potential safety concerns. In conjunction with use of the pharmaceutical product outside of the clinical trial setting, it is also common for safety issues to arise. These safety issues will need to be evaluated using the clinical trial data as well as various epidemiological sources. Imaging studies within the phase IV environment are relatively common and add value by objectively measuring the impact of the pharmacological intervention on either efficacy or safety parameters.

Imaging as a Biomarker

Advances in imaging technology have enabled scientists to detect events at the cellular level. Hardware and software manufacturers have increased the resolution of their products such that detection sensitivity and resolution have improved markedly. It is clear that imaging technologies have revolutionized the practice of medicine over the past few decades. This has contributed to improved diagnostics as well as improved monitoring of response to therapy. The result is a quantifiable

improvement in the quality of care in most therapeutic areas resulting in improvements in both quality and duration of life. The percentage of the population living into their 70s, 80s, and beyond is among the most rapidly expanding segments of the population in many countries.

Many individuals who work in drug development view imaging endpoints as a biomarker analogous to C-reactive protein (CRP) for inflammation. A biomarker is a characteristic that is objectively measured and evaluated as an indicator of normal or pathological biological processes [10]. Biomarker programs within clinical development organizations typically assess serum markers, DNA, and RNA for *association with the disease process*. In an autoimmune inflammatory disease such as rheumatoid arthritis, CRP may be used to assess disease activity. Changes in CRP can be used to select doses of a therapeutic agent used to treat rheumatoid arthritis.

In contrast, imaging is the *measurement of an in vivo biological process*. For rheumatoid arthritis this would involve imaging of specific joints. While CRP is a systemic biomarker, imaging can provide additional information by demonstrating changes at specific anatomical locations (e.g., joints impacted by arthritis). Imaging is performed on the organ of scientific interest such as the heart, bone, or joint, while many other biomarkers are derived from the serum or circulating cells. During the development of a biomarker strategy, several biomarkers can be used together to derive a model predictive of a particular disease. Imaging may be considered as the net effect of both local and circulating factors on the disease process. As such, it is highly clinically relevant for many diseases.

A key property of imaging modalities is that they can provide critical data regarding disease progression. It is important to be able to detect progression of disease when patients are asymptomatic because often a disease process becomes less amenable to medical treatment as it becomes more advanced. From the perspective of drug development, demonstration of an improvement in patient signs and symptoms, while clinically relevant and necessary, is often not sufficient to gain regulatory approval. In diseases such as arthritis where pain is a key symptom, agents that reduce pain can be approved on this basis. From a payer perspective these agents will be considered alongside aspirin, acetaminophen, and NSAIDs. Pharmacological therapies which modify the disease process are much more highly valued. In order to gain approval for a product that modifies the disease, it is also necessary to show either an improvement or a reduction in the rate of disease progression which often requires the use of imaging.

The potential that imaging technologies have within clinical development has barely scratched the surface. While there are many reasons for this gap between potential and realized benefit, it is important to focus on the key elements required for success. The consensus opinion among imaging researchers and healthcare regulators is that the potential for the use of imaging in clinical trials for drug development hinges on appropriate use of standard principles. With appropriately rigorous trial conduct, the clinical trials will be more robust, require fewer subjects, and will be more likely to provide conclusive results that will allow for effective decision making. Simply incorporating imaging endpoints into a clinical trial does not guarantee that useful data will be generated. Rather it is the how the

imaging component is designed and implemented that is critical to the generation of high-quality data that will facilitate interpretation of the trial results. As will be discussed next, there is a trend to collect key data to determine whether the drug candidate will be likely to garner regulatory approval and, if so, whether it will be competitive in the marketplace earlier in the development process than may have occurred in the past.

Key Elements for Product Approval

Advancing a product from the laboratory into the clinic or from early to late development while necessary for product approval will turn out to represent a failed investment if the product does not get approved. Therefore, it is important for all individuals working in drug development to have a sound understanding of the elements that are required for marketing authorization of pharmaceutical products globally. Whether one is working to register a new chemical entity, a new biologic agent, or a new medical device, the same principles apply. Demonstration that a product is effective followed by a listing of the adverse events experienced during the clinical trials will rarely be sufficient for approval with a few exceptions. These exceptions include diseases with high unmet need where no treatments are available.

The key measure for health authorities in 2010 is the benefit to risk ratio (BRR) for the specific indication and specific population. For example, one may get approval for a product for hypertension, colon cancer, or rheumatoid arthritis. It may be approved for first-line therapy, or it may be approved for use after first- or second-line therapy. How the BRR of the product compares to other available alternatives will determine whether approval is for first-line, second-line, or salvage therapy. It is also a function of whether or not the sponsor has developed the product in a manner that highlights its benefits in specific populations, or relative to other products. These can be defined by means of demographic, disease, or genetic characteristics.

Approval truly boils down to whether the sponsor can genuinely demonstrate and communicate a sound understanding of how their product works. This requires building upon the foundation of the mechanism of action (MOA) to select a patient population that should benefit based on the MOA. The next step is to validate the working hypothesis by demonstrating efficacy in the target population. Finally, one must show a comparable or better overall BRR compared to other available treatments for the specific population to be used as first-line therapy.

These changes in regulatory attitude are part of the change in medicine towards individualization of therapy. While this is being led in oncology through the increased use of tumor antigens to determine prognosis and treatment, it is occurring in many other disease areas through the use of patient characteristics including laboratory and imaging variables [11]. Defining the patient population(s) that will benefit is a key part of any clinical development program.

How have these changes in health authority decision making impacted clinical development programs? These changes have resulted in the selection of molecules with very specific actions that maximize their effect on the desired pharmacological target, while minimizing unintended off target effects. Monoclonal antibodies have experienced a marked increase and represent an increasing percentage of new products reaching the market [12].

Earlier Decision Making

Unfortunately, pharmaceutical product development will have more products that do not make it to market relative to those that succeed. A product that does not make it to market generates no revenues. Within organizations there is often a desire by team members to continue projects despite extremely low probabilities of ever recovering the associated expenses. As the competitive landscape intensifies, successful organizations will be those who are able to generate scientifically robust and clinically relevant data early in the development process and those who make evidence-based decisions on therapeutic agents within the development portfolio [13].

During phase I studies individuals with the targeted disease are being studied earlier and more extensively than they were historically. This is because while healthy volunteers can be used to determine the basic pharmacokinetics, effects on vital signs, and routine laboratory parameters, they are not informative in providing data relative to the pharmacodynamics of the disease process. Phase I studies provide the opportunity to explore the effects of the drug on the pathophysiology of the disease. Proof of principle is often established in phase I. A good understanding of the impact of the drug on the various components of the disease guides the doses to be taken forward into phase II.

Phase II studies will prove that the scientific concepts leading to clinical improvement in the target disease have been met. They will hone in on the population(s) that will derive the greatest BRR, as well as the dosing regimen to be used in phase III. Increasingly, active comparators are being included in phase II studies so that sponsors can determine where (e.g., first versus second line) in clinical practice their therapy will be used.

Inevitably, at the end of phase II, there remain unanswered questions regarding the product and its potential to modify the targeted disease. Could these data have been collected earlier in the development process? In many cases, the answer is yes. Phase III studies are becoming increasingly larger and often represent major investments even for the large multinational companies. While it is relatively easy to make a good decision when the information quality is excellent, as the ambiguity of the data increases, the probability of making an investment error in the hundreds of millions of dollars for the initiation of a large phase III program increases. The take home message is to invest in understanding the MOA of the product and the impact of the drug on the pathophysiology of the targeted disease population by the end of phase II such that informed investment decisions can be made. Companies that do not do this are unlikely to remain viable in their current form.

Opportunities for Imaging

From a regulatory perspective, health authorities are looking for a logical sequence of events that give them a high probability that the product will perform as suggested by the sponsor in their marketing application. Health authorities do not want surprises. They are charged with protecting the public health which means that the public needs to be fully aware of any actual or potential safety issues. A thorough development plan will utilize early data such as the MOA and drug distribution to outline potential efficacy and safety effects. Potential safety issues if clinically significant will need to be assessed promptly. The use of imaging to evaluate potential safety issues is increasing. The observation of some fractures during a phase II study in a drug with a theoretical risk of impacting bone metabolism can be assessed by adding a bone mineral density sub-study into phase III. The risk of heart valve dysfunction can be addressed with echocardiography. The risk of direct CNS effects can be mitigated by demonstrating the absence of drug in the CNS using PET. Studies addressing specific safety concerns may be best performed during phase II since the presence or even the probability of certain safety issues can have a major impact on the approval of a product and may therefore influence the decision of whether to proceed into phase III.

In summary, imaging studies can be very useful in understanding how the pharmacological product modifies disease progression. This can lead to better decisions regarding the dosing regimen, the probability of clinically significant safety issues, and, ultimately, whether to progress further in developing the product. In many therapeutic areas imaging studies are required for initial approval or approval of specific indications.

Detection of Disease Progression

Many diseases remain asymptomatic until they are relatively advanced. Examples include atherosclerosis, osteoporosis, and certain malignancies. In many diseases including rheumatoid arthritis, patient symptoms both at the level of the joint and systemically may not correlate with disease progression obtained through imaging studies. Imaging may be considered more objective than documentation of clinical symptoms. This is not surprising since reporting of clinical symptomatology is dependent upon several factors which are difficult to precisely control in the context of a clinical trial.

One major challenge facing pharmaceutical companies is that in certain therapeutic areas, there are already high-quality medicines available to treat the disease. This may mean performing a head-to-head trial against existing options or comparing the existing therapy to a combination of the new and existing therapy. In both situations, the difference in therapeutic effect will be less than a comparison of the new agent to placebo. Since imaging is in essence “a sharper scalpel,” the use of this

instrument to demonstrate a relative improvement in either efficacy or safety can greatly modify the use of the product. An increase in progression free survival relative to existing standards of care in oncology is one example. For all late stage clinical development programs, imaging endpoints should be considered in evaluating the BRR relative to existing therapeutic options.

Do the Results of the Imaging Study Answer the Key Scientific Question?

Unfortunately the answer is often not to the degree that is necessary for health authorities. Why is this so? Suleiman and Gorovets in April 2010 presented their observation that the FDA desires scientifically robust evidence. However, many of the imaging trials lacked standardization, calibration, and reproducibility. They compared the standards for drug quality and purity required for chemistry manufacturing control and stated that similar rigor should be applied to imaging [14]. They note that several studies did not have sufficient power to detect a difference between treatment groups due to the large variability observed for the imaging parameter. This becomes particularly relevant when conducting non-inferiority studies for efficacy. It is also pertinent for safety studies because results which show no difference between a drug in development and placebo may not be considered sufficiently robust if there are no data on within subject variability of the imaging parameter to enable determination of study power. If a clinically meaningful difference is not detectable using the imaging technology due to inherent limitations in the technology employed in the study, or due to poor implementation of imaging standards, this will be considered an irrelevant study and will not diminish a safety risk.

Imaging studies are used to measure one or more variables associated with disease progression. This requires attention to detail for each of the steps involved in the process analogous to the manufacturing of marketed drugs. Lack of attention to critical details will result in a study that has similar value to a batch of a pharmaceutical product that does not meet desired product specifications.

Key Considerations for Successful Imaging Studies

Creating a Successful Imaging Team

The clinician(s) within the clinical development team are usually responsible for the task of clinical trial design with input from their statistical colleagues. The key component of a successful imaging trial is the recognition that specialized skills beyond those of the lead clinician and the core development team will be required. These imaging professionals, whether they are within the organization or external, should be involved early in the process of trial design. As a team leader, it is helpful to map out the key questions that will need to be addressed and to seek input from

individuals with the appropriate knowledge and experience to provide constructive input. Table 6.2 outlines some key questions that will need to be answered in order to develop the imaging component of a clinical trial and the skills that are required to answer these questions. For illustrative purposes a proposed study evaluating a novel agent for rheumatoid arthritis will be discussed. Management would like to know if this compound will modify disease at the level of the joints prior to investing in a phase III program which will cost in excess of 500 million US dollars.

Imaging Endpoints

In responding to management's request, there is a multitude of information that must be acquired and processed by the team in order to develop reasonable designs for their phase II study. This includes all of the parameters listed in Table 6.2. Generally, the first aspects of trial design that need to be agreed upon are the key endpoints and the associated imaging technology through which they will be measured.

In rheumatoid arthritis only a small percentage of patients will demonstrate progression of joint damage over a 12-month time period using radiographs. The size of a study using this imaging technology will require 300–500 patients per group [15]. MRI and ultrasound can detect changes in the joint that cannot be detected with standard radiographs. Therefore, changes can be detected at earlier time points and in a greater percentage of patients enabling a smaller sample size.

Another consideration is that as the imaging technology becomes more sensitive to detecting smaller changes, one must determine which specific imaging changes are temporary and reversible and which specific changes represent disease progression. This is part of the evolving advances leading to improvements in standards of care. In drug development, one must also have a firm understanding of what changes are predicted to be improved based on the mechanism of action of the molecule under development, over what time period, and in what patient population? The answers to these questions will be important in the design of the clinical trial program.

Invariably, there will be different answers to these questions based on which literature source is referenced and whose clinical opinion is sought. There will be variability in the imaging data reported depending on the patient population studied, the acquisition method, hardware, software, and reading methodologies. Professionals skilled from a clinical perspective in conducting research in this area can be helpful. Likewise, experienced musculoskeletal imaging professionals especially those who have conducted clinical trials with similar endpoints will provide significant value. Ultimately, one will be required to estimate the incidence of imaging changes, the rate of change over time, and the effect of the pharmacological intervention in the population under investigation in order to design and adequately power the studies.

Say that a preliminary decision is made to proceed with MRI as the imaging modality and the anatomical areas for evaluation include the hands, wrist, and feet

Table 6.2 Key elements in the design of imaging parameters within clinical trials

Key questions	Relevant parameters/examples	Skills required
Endpoint(s) to be measured and selection of imaging modality	Joint space narrowing and bony erosions by X-ray are the regulatory standard for approval in phase III. Synovitis, tenosynovitis, bone marrow edema and bursitis can be detected by MRI or ultrasound	Rate of change over time in the population of interest. Determination of the change that is most likely to be impacted by the treatment over the intended duration of the trial
Availability of validated metrics for the selected endpoint	Scoring systems validations conducted	Knowledge regarding test validation and regulatory standards in imaging
Imaging hardware	Acceptable hardware for imaging of key endpoints	Ability to detect changes in the selected parameters. Differences between available hardware and impact on imaging endpoints
Image acquisition	Protocol for image acquisition	Performance characteristics of imaging devices
	With or without contrast Image type and desired resolution	Experience with the pros and cons of different acquisition protocols
Precision (reproducibility)	Difference between two measurements from the same patient on the same day	Knowledge of the conduct, analysis, and interpretation of reproducibility studies
Accuracy	Comparison to gold standard (phantom)	Determination of whether phantoms are required for this study
Image analysis	Hardware and software	Knowledge of the clinical relevance of differences in hardware and software
Image interpretation	Read methodologies	See Chap. 5
Data management	Identification of key imaging metrics for both operational purposes and statistical analysis	Experience in the therapeutic area with the specific imaging technology and operational experience
True potential for detecting change	Following selection of the patient population, imaging modality, and image acquisition protocol an estimate of the change that will be demonstrated in the control group and in the treatment groups	Experienced individuals in translating potential for detecting change into an accurate estimate of expected change between groups in the clinical trial

joints. Are there scoring systems that are recognized for MRI in rheumatoid arthritis? Are these scoring systems validated and if so by what methods? Are they acceptable to health authorities? In addition to the professionals mentioned previously, individuals experienced in validation and with the evolving regulatory position on imaging endpoints will provide significant value to the team.

Imaging Hardware

Once the imaging modality has been identified, the question of which equipment to use for this clinical trial arises. Manufacturers of MRI scanners improve their products over time. Published literature may be based on single-center studies with scanners that are not currently used by many imaging facilities. Knowledge regarding scanners from different manufacturers and even models within the same manufacturer are relevant. Scanners in use at sites that will be considered for the study will need to be determined. Biomedical engineers can explain these differences. Discussion between the biomedical engineer and clinician will be helpful in making decisions regarding the tradeoff between hardware consistency and models available at potential sites. Manufacturers are usually very willing to have their engineers go through the specifications and performance characteristics of their products. Most will explain to you why their products are superior to those of their competitors. This process can be confusing as it can be difficult to determine how these differences in technical parameters will impact the images acquired for the study. It can also be quite challenging to determine the magnitude of the impact that these differences will have on the imaging endpoints proposed for the study. CROs with professionals experienced in imaging often have staff members who are familiar with manufacturer upgrades and understand the differences including the impact that these differences will have within your trial. Consulting them will save time and get you the information needed in a timely fashion.

Image Acquisition

For purposes of discussion, image acquisition includes all of the steps from when a patient enters the imaging suite until the images are digitally stored. Ideally image acquisition should be identical for all scans in the study. In reality, numerous factors that vary over time limit us to approximating this goal. A standard acquisition protocol must be developed. This includes all of the variables that may impact image metrics. These include patient positioning, slice thickness, image type (e.g., T1 or T2), and use of contrast. Having the same technician perform the scans is the most important variable. Incorporated into a particular technician's routine is not only positioning but also many other factors involved in their management of the patient through the process. It is good practice to speak with site technicians regarding whether the proposed acquisition protocol is easily understood and reasonable to conduct in their facility.

Precision

Reproducibility refers to the difference obtained between two scans, obtained with the same scanner, by the same technician on the same patient; see Chap. 2 for comparison of precision and accuracy. In some publications this is also referred to as the

precision of the measurement. Typically a scan is performed using the study protocol. The patient is instructed to get off the scanner table and a few minutes later the process of performing the second scan is initiated. The difference between the two scans represents the intra-subject variability. This should be performed on a group of patients with different degrees of disease to assess the intra-patient variability across the disease spectrum. These studies are usually performed at 1–3 sites. Good technicians will be able to point out sources of variability within the acquisition protocol. This will serve as the basis for site training that is required to qualify imaging personnel at the site and ongoing training and monitoring procedure to minimize variability between sites [16].

When conducting clinical trials, consistency is extremely important. Even the same scanner will generate different results over time. How does one detect and manage these changes? Also when we get a reading from a scanner, how close to the truth is it? Accuracy is the term used to describe how close a measurement comes to a "gold standard." Phantoms can be utilized to describe accuracy and to monitor consistency. If phantoms do not exist for the anatomical area under investigation, you may want to consider having one built. This will be a costly procedure, so it is best to discuss this with an experienced imaging professional. Typically phantoms are imaged at regular intervals during a clinical trial in order to detect variation in machine performance over time. Minor changes in machine performance can be managed by applying correction factors to the study images generated, but more significant changes may make some of the images unsuitable for reading. Images which do not meet the predefined study quality standards will require a repeat image for reading. If this does not occur within a specific time period as defined in the protocol, there will be no usable data for this patient. Since imaging endpoints are typically calculated as the change from baseline, the baseline and final images are the most critical. Statistical analysis will commonly be performed using a last observation carried forward methodology. Therefore scans which are missing will tend to reduce the change detected with the effect of reducing study power.

In addition to imaging hardware, the software provided by manufacturers is routinely updated. These software programs contain instructions for assessing the pixels acquired during the scan. These instructions result in a digital image or a numerical value.

When a digital image is generated, it must be quantified by readers trained according to study-defined prespecified criteria. Strategic thought is required to develop an appropriate read methodology. All data must be maintained with a full audit trail in compliance with ICH and CFR part 11.

Statistical Considerations

Several statistical inputs will be required in order to intelligently design the studies. The minimal detectable change refers to the minimum change that falls outside the measurement error for an instrument. These determinations are usually performed under idealized conditions with highly experienced imaging professionals. In

statistical terms for normally distributed data, this is defined as [$1.96 \times \sqrt{2} \times$ standard error of the mean]. The standard error of the mean is the standard deviation of the measurements divided by the square root of the number of measurements.

Another key variable is the minimal clinically important change which represents the smallest change that is clinically relevant. Ideally, one powers a study sufficiently to detect a change that is greater than the minimal clinically important change. While it seems obvious, this change must also be greater than the minimal detectable change. Studies have been performed where the variability of the imaging measurements were sufficiently large such that the minimal detectable difference was greater than what the study had been powered to test.

The design of imaging endpoints in clinical trials involves the estimation of the difference between treatment groups for the population under investigation. Additionally estimates of the variability in measurement must be performed. It is important to have productive discussions regarding tradeoffs between scientific precision and operational efficiency regarding patient recruitment. Investments in minimizing variability may be greater when imaging is the primary endpoint compared to when it is a secondary endpoint. Lastly, many of the team members involved in the study design should remain engaged in the project as the study is initiated. When a handoff of a protocol occurs to an operational team, it is easy for the operational team to focus on recruitment which can sometimes be at the expense of image quality and consistency. Therefore maintaining a degree of project history including the rationale for specific aspects of protocol design will increase the probability of a successful study.

Core Principles Pertinent to Imaging Studies

The practice of clinical medicine where patients are treated on an individual basis is very different from the design and conduct of clinical trials. When evaluating and treating a patient, the core information is their clinical signs, symptoms, comorbidities, lifestyle priorities, values, etc. One then uses your knowledge base as a healthcare provider, consisting of the literature and personal experiences to present treatment options to that particular patient. In contrast, clinical trials are performed with the objective of determining the impact of a specific intervention such as a new pharmaceutical agent on a specific treatment outcome. In order to achieve this objective, we standardize the patient population that can participate as well as the treatment regimen. While we allow some variability in the patient population, we are more stringent regarding the treatment protocol. We do this because we know that increasing the variability in the treatment regimen by permitting variations in the dosing regimen (e.g., varying the dose intervals or drug quantities, skipping doses) or variations in the assessments (e.g., morning versus evening, month 2 or 4 versus month 3 of the study) will make it more difficult to determine the effect of the treatment being investigated.

When we use imaging to measure a biological variable, we are often trying to detect relatively small but clinically significant changes. Therefore minimizing

variation in the conduct of the imaging assessments requires planning and attention to key details during trial execution. We have divided key considerations pertinent to the design and conduct of imaging endpoints in clinical trials into the following four categories: scientific, regulatory, financial, and operational. While these categories will contain items that overlap, they are broken out in this manner because they require a different focus and as such are often the responsibility of distinct team members within biopharmaceutical organizations. For early phase studies, we have integrated these four functional areas and subsequently describe them separately for clinical trials in phase II and beyond.

First in Man Studies: Role of Imaging

Despite extensive testing of new chemical entities in animal models, differences in bioavailability, pharmacokinetics, tissue distribution, and metabolism are significantly different in humans, resulting in modifications that can cause considerable delay or termination of a project. To address these issues, phase 0 also known as microdosing studies can be performed. Guidance for conduct of these studies can be found on the websites of the EMA and FDA. The dose administered must not have any pharmacological effect. It has been defined as the administration of 100 μg of candidate drug or 1/100th of the pharmacological dose determined from animal models and in vitro systems, whichever is lower. PET scanning is the most common imaging technique used to determine pharmacokinetics, pharmacodynamics, and tissue distribution in these studies [17].

PET scanning requires labeling of the compound with [^{11}C] which has a half-life of 20 min or [F18] which has a half-life of 110 min. Fluorine-18-labeled glucose (FDG) is widely used to measure glucose uptake in tissues. The use of [^{11}C] necessitates that the radiolabeling laboratory be within a few minutes of the imaging facility as the rapid decay will not usually permit accurate detection for determination of pharmacokinetics of the compound beyond 2 h (6 half lives) from the time of synthesis. These studies can determine whether a compound is getting to its intended location and also whether it is distributed to unintended areas. The use of FDG-PET or F-18-labeled investigational drugs with its longer half-life allows greater flexibility.

Phase 0 studies have been used for candidate selection [18, 19], for example, when there are 2–3 potential molecules that have the desired activity in animal models. Since humans and the animal models may differ significantly, administration of each of these molecules in a phase 0 study to 3–5 study subjects will provide data that can determine which of these molecules (if any) should be advanced further. Pharmacokinetic parameters and tissue distribution can aid in this important decision. These parameters as well as bioavailability, tissue distribution, and metabolism are estimated to differ materially in humans from estimates based on animal data in one third of cases. Candidate selection can also be performed in an iterative manner. In this paradigm changes to the structure of the molecule are made based on initial phase 0 study results. The new molecule is then tested in another phase 0 study until acceptable parameters are obtained. These phase 0 studies are also informative for determining the first dose for the subsequent phase I study.

Once a drug is being introduced into humans, an early readout regarding standard bioavailability, pharmacokinetics, and pharmacodynamics is desirable. While bioavailability, pharmacokinetics, and basic safety parameters can often be obtained in healthy volunteers, pharmacodynamic parameters may only be informative in individuals with the targeted disease. This favors the inclusion of patients for whom the new therapy is targeted to be included early during phase I. For some therapeutic areas such as oncology, the risk to benefit ratio is such that only individuals with the specific-targeted tumor may be included. Imaging provides a key pharmacodynamic measure in early phase trials within oncology and neurosciences.

Currently a biomarker plan which is a consideration of the key anticipated pharmacodynamic effects of the drug is part of the clinical development plan. Individuals charged with developing this plan may have little or no familiarity with imaging and may restrict their plan to evaluation of parameters that can be assayed from serum samples. For several indications such as prostatic hypertrophy, osteoporosis, and oncology, imaging early in development provides information that will increase the quality of subsequent decision making.

Often there is an argument that the incorporation of imaging parameters into phase I trials will exceed the planned budget for a specific phase I study. This is more common in organizations where the phase I unit is organized into a distinct group with a limited operating budget. As stated previously, when viewed as an integrated development effort, if imaging assessments can provide scientifically valuable information regarding efficacy or safety that will impact subsequent development decisions (including project advancement versus termination), then they will be highly cost-effective. Of course the inclusion of imaging parameters when their outcome will not be used in the decision-making process is in essence for academic interest only. In this situation, they are simply a cost with no preplanned value.

Phase II and Beyond: Scientific Considerations (Strategic and Technical)

Imaging endpoints differ in many respects from patient-reported outcomes or binary clinical outcomes such as the occurrence or nonoccurrence of a myocardial infarction. Therefore assuming that well-trained clinicians can implement imaging parameters into clinical trials within their area of therapeutic expertise can lead to unanticipated outcomes. The implementation of imaging endpoints requires a much greater attention to operational detail than occurs at most clinical visits during a research study. Since clinicians responsible for study design and conduct may not be sufficiently experienced in imaging principles, it is not surprising that the most common criticism from imaging authorities or experts regarding the design and implementation of imaging endpoints in clinical trials is that they are poorly conceived from a scientific perspective. Studies which are flawed scientifically can be well executed but will still not result in regulatory approval and will not recoup the

initial investment. Therefore a firmly grounded scientific basis is the foundation for the successful use of imaging endpoints.

Many of the key questions involved in developing a design for imaging parameters are listed in Table 6.2. Conceptually, one needs to determine what information related to the physiology of the disease process will be obtained through the use of imaging. The question that follows is whether this information will be of practical use either in the clinical development process or in clinical practice.

For example, say you are evaluating a drug that is intended to slow the rate of decline in disease for patients with emphysema. You can use pulmonary function tests to follow the severity of emphysema, so what additional information, if any, would be gained from the addition of imaging endpoints to the trial? You perform some investigative work and determine that there is evidence that CT findings correlate with the presence and severity of morphologic emphysema better than do results of pulmonary function tests [20]. Your initial assessment is that incorporation of pulmonary CT into your phase II trial will improve clinical decision making. Therefore you wish to employ an imaging endpoint in your trials.

Now that you have decided from a strategic perspective to pursue the use of imaging, technical considerations arise that need to be worked through. What are the appropriate imaging modality and appropriate technique to use? Assuming that high-resolution CT is selected as the imaging modality of choice, additional questions that require the involvement of technical experts remain. Should the CT scans be obtained using 10 and 1.5 mm collimation, or should 5 and 1.0 mm collimation be selected? Should software programs be utilized to highlight areas of abnormally low attenuation? If so, which model scanners and which software programs will provide reproducible data? What are the advantages and disadvantages of proceeding with one approach versus another? The need to involve individuals with expertise not only on the clinical side but also with technical expertise related to image acquisition early in the process of study design becomes apparent.

You have been diligent in your research and are now presenting your protocol to the protocol review committee. This committee includes individuals with highly variable skills and knowledge regarding the therapeutic area. How do you increase the likelihood that sound scientific decisions will be made in a timely manner? Table 6.3 provides a framework for making decisions on whether to incorporate imaging endpoints into clinical trials.

Rather than attempting to quantify the value added by the imaging data from low to high, we have categorized the value into essential, supportive, and nice to have.

Essential Studies

Imaging endpoints may be performed to assess efficacy or safety endpoints. In 1997 fenfluramine which was used for weight loss was withdrawn from the market due to evidence that its use caused a thickening of the leaflet and chordae tendineae. Fenfluramine and its active metabolite norfenfluramine are agonists of 5-HT_{2B}

Table 6.3 Strategic consideration for incorporation of imaging endpoints

Category	Scientific standards	Example (s)
Essential	Scientific standard of care	Echocardiography for the evaluation of valvular function
	Regulatory standard	Radiographs for rheumatoid arthritis
	Required to demonstrate proof of principle	PET imaging to demonstrate presence or absence of a new chemical entity to specific areas in the CNS
	Proof of concept	Bone mineral density as a surrogate for fracture risk
	Mechanism of action	Demonstrates the intended MOA of the compound in either preclinical or clinical studies
	Increased sensitivity for detection of change relative to regulatory standard	MRI for changes in the joint in rheumatoid arthritis
	Defining a target patient population prospectively through imaging	Fracture reduction based on bone mineral density at time of study initiation
Supportive	Adds information regarding treatment induced changes in physiological parameters	CIMT, pulse wave velocity
	Data are anticipated to be informative relative to subsequent development decisions	Imaging in phase II studies for rheumatoid arthritis, oncology
Nice to have	Studies that provide evidence of additional benefit	Body composition for diabetes
	Characterization of specific populations that benefit	Subpopulations not identified in the registration studies
	Use of the pharmacological agent in clinical settings	Demonstrating how use of imaging in practice can improve clinical outcomes
Potential for future development	Enhancement to currently labeled treatment regimens	Lifecycle management
	Mechanistic studies	Improved understanding of disease pathophysiology

receptors, which are postulated to have led to a pathological increase in cell division in the heart valves. Supporting evidence for this mechanism is the finding that other drugs acting on 5-HT_{2B} receptors are associated with similar findings [21]. Therefore in the development of lorcaserin which is a selective 5-HT_{2C} receptor agonist, with a 100:1 relative binding affinity for 5-HT_{2C} relative to other receptors, it was necessary to evaluate the impact on cardiac valvular function. The performance of echocardiograms prior to and during treatment is the current scientific standard for evaluation of cardiac valvulopathy.

Regulatory guidance is available online from FDA and EMA regarding specific indications. These guidances are built upon historical precedent. Most

health authorities are risk averse consistent with their mandate to protect public health. They will often insist upon maintaining the current imaging modality within the phase III registration trials (e.g., radiographs for rheumatoid arthritis) and will enable the incorporation of additional imaging data into the product label if a case can be made that these new data are clinically pertinent. If the imaging modality (e.g., MRI for rheumatoid arthritis) is used in clinical practice beyond research purposes, then this will usually meet the criteria for clinically pertinent. It is essential to meet with global health authorities and to provide your scientific rationale for the use of specific imaging modalities within the development program. If the health authorities can follow the scientific rationale, they are more likely to support its inclusion in the prescribing information upon approval.

While some use the terms proof of principle and proof of concept interchangeably, we will use them distinctively for the purpose of drug development. Proof of principle involves the interaction between the drug and its intended target in the species of interest which is usually the human. For a drug intended for depression, localization to specific anatomical regions in the CNS by PET scan, in conjunction with in vitro receptor-binding studies and in vivo animal studies together, may demonstrate proof of principle. The principle being that the drug binds selectively to a particular receptor that is localized in its anticipated area in the human brain. Proof of concept is the demonstration that this drug through binding to this receptor will translate into a clinical improvement in depression. This proof of clinical concept will occur by evaluating specific dosing regimens in patients with depression.

In drug development for osteoporosis, the demonstration of increased bone mineral density by DXA (Dual energy X-ray absorptiometry) was sufficient for securing a marketing license until fluoride was marketed. Fluoride administration resulted in marked increases in bone mineral density but was associated with an increased fracture risk. The reasons for this were subsequently elucidated through evaluation of bone biopsies. Currently demonstration of an increase in bone mineral density in conjunction with bone biopsy data demonstrating good bone quality represent proof of concept for osteoporosis. A phase III trial is still required in order to demonstrate a reduction in fracture risk [22].

Advances in imaging technology may include the development of new modalities, novel applications of existing technology and most commonly improved precision and detection limits with a new generation of hardware. If the technology has been validated, which is a requirement for commercializing a new generation of scanners, and presents some advantages over existing methods, then it should be considered for use up through phase II. One example is the use of a new generation of high-resolution CT to determine if a product [23] for emphysema can favorably modify lung structure or delay disease progression. Another is the use of MRI in rheumatic diseases. The imaging modality used for phase III will require discussion with key global health authorities.

Supportive Studies

Imaging may at times serve as both an efficacy and a safety endpoint. In oncology, imaging of tumors is standardized. The primary endpoint in most clinical oncology studies is patient survival or progression free survival. Reduction in tumor size is often a secondary endpoint. From a drug development perspective, tumor size is also a safety parameter and most dosing regimens that document increases in tumor volume will not be progressed further.

For purposes of discussion, we classify supportive indications for imaging as circumstances where imaging data have a high probability of adding value to the clinical development decision making, but they are not expected to be pivotal in driving decision making. In cardiovascular development clinical outcome trials are commonly required. Surrogate parameters such as lipoprotein changes are not sufficiently robust in predicting the result of outcome studies. Therefore imaging studies such as carotid intima-medial thickness (CIMT), pulse wave velocity, or other assessments can be used to determine if there is an additional effect beyond lipid changes. These studies may assist in dose selection or in determining whether to invest in phase III. However, these studies are not considered essential.

Many circumstances occur during drug development when decisions regarding the timing of specific assessments must be made. In the previously discussed example for rheumatoid arthritis, management accelerated the imaging data into phase II in order to have higher quality data in planning for phase III. Strictly speaking, these data are not required and a dosing decision could have been made based on traditional biomarkers such as C-reactive protein. As is the case for much of the data that fall into the supportive category, if they have meaningful economic value, they will merit consideration.

For assessment of multiple sclerosis examples of imaging parameters include optic nerve magnetization transfer ratio, retinal nerve fiber layer thickness (by optical coherence tomography), brain lesion magnetization transfer ratio, MRI brain T1 hypointensity load, or new T2 lesions, the latter of which is the regulatory standard. PET scanning is being used more commonly in CNS disorders. In summary, this therapeutic area will involve imaging studies that have a high probability of yielding scientific information that will be of value during the clinical development program.

Nice to Have Studies

Imaging studies that are essential in phase IIIb studies to gain approval for additional indications are considered essential. This category of "nice to have" is defined as imaging studies that will not affect decisions regarding whether or not to continue development of a compound for its primary indication and will not affect a decision by health authorities regarding marketing authorization for that indication. These studies are

commonly performed to provide additional evidence of clinical benefit. Examples include body composition studies for diabetes drugs which have been used to highlight differences between agents, QCT in osteoporosis, and MRI in osteoarthritis [24].

Imaging studies to better understand the pharmacological effects of a new chemical entity may be performed as a nested sub-study in a phase III program or in a phase IV trial. In general the rationale for their conduct is primary based on marketing considerations. Phase IV studies that can have a considerable public health impact are those which evaluate the use of an imaging assessment on patient care and clinical outcomes. Examples include the role of bone mineral density measurements in the management of osteopenia or the role of radiographs in the management of rheumatic disorders.

Potential for Future Development

This category refers to studies designed to test a hypothesis for which there is no immediate return on investment for the current compound. They may be performed with the intention of a return on investment that is beyond the time horizon for the current compound. These may include the development of new imaging biomarkers for specific diseases. For example, one may wish to validate MRI endpoints for disease progression such that these endpoints may be discussed with health authorities. If these new imaging endpoints are more sensitive in determining disease progression and are accepted by health authorities, then phase III clinical trials may be able to be performed with fewer subjects. Similar paradigms hold for other therapeutic areas such as osteoporosis.

Many of these studies are carried out in partnership with academic institutions. They may seek to improve upon the clinical outcomes achieved in the registration trials by modifying the treatment regimen according to the data from imaging endpoints. For instance, disease progression may not be associated with clinical symptoms until late in the disorder. A demonstration of disease progression through use of imaging endpoints (e.g., in rheumatoid arthritis, osteoporosis, atherosclerotic heart disease) may result in more aggressive therapy and/or improved patient compliance that will yield improved outcomes. Healthcare practitioners will compare different treatments for different populations in order to prioritize amongst available therapeutic options. This may include a comparison of medical to surgical options. Researchers interested in better defining the pathophysiology of the disease may utilize a pharmacologic agent as a probe to define disease subtypes. It may also be used as a proof of concept for a new indication. The potential scope of imaging studies outside of industry related clinical trials is expansive and beyond our intended scope.

Protocol Development

When one is primarily mimicking a predecessor's clinical development strategy, protocol development is straightforward. For first in class compounds and for novel therapeutic indications with high unmet need, the potential for a huge success is

apparent, but so is the risk of failure. In the end, the one individual who has the greatest impact on the success of a clinical trial is the person charged with protocol development. Success is not predicted on a specific IQ score, but rather on the wisdom of seeking and interpreting seemingly disparate information and most importantly being diligent in working through all of the scientific issues. If the imaging component of a protocol is written with statements along the line of "high-resolution pulmonary CT will be obtained at baseline and at the end of study visit," this must be accompanied by a detailed explanation of what is meant by high-resolution pulmonary CT. This information is best suited to the imaging charter which can be referenced in the body of the protocol. Since the protocol should describe all of the study procedures, the imaging charter is considered part of the protocol and needs to be included in health authority communications regarding scientific guidance.

No matter how expert you feel that you are in a certain area it is important to listen to others both internally and externally. Engaging staff at prospective clinical sites can provide a good reality check that is pertinent. It is important to understand prior to study initiation, what will really happen at clinical sites and how they will manage specific protocol instructions. Engage a number of external consultants, but rapidly determine which ones provide value to you and forge ongoing relationships with these individuals or organizations. Keep internal and external stakeholders informed regarding your progress and decision making. Stay focused; thousands or even millions of patients may be eagerly awaiting the outcome of the trial you are designing.

Regulatory Considerations

The pharmaceutical and medical device industries are highly regulated due to the potential for adverse events as a consequence of their products. Earlier in the development process, the risk of adverse events relative to any clinical benefit that may be derived is higher. The benefit to risk ratio continues to increase throughout the development process such that at the time of marketing authorization the benefit to the patient significantly outweighs the risk. Health authorities are charged with protecting patient safety throughout the development process. Interaction with global health authorities is required at key points during development. These include but are not limited to the investigational new drug application which is required to administer the product to humans.

The IND application must contain information in three broad areas:

- Animal Pharmacology and Toxicology Studies – Preclinical data to permit an assessment as to whether the product is reasonably safe for initial testing in humans. Also included are any previous experiences with the drug in humans (often foreign use).
- Manufacturing Information – Information pertaining to the composition, manufacturer, stability, and controls used for manufacturing the drug substance and the drug product. This information is assessed to ensure that the company can adequately produce and supply consistent batches of the drug.

- Clinical Protocols and Investigator Information – Detailed protocols for proposed clinical studies to assess whether the initial phase trials will expose subjects to unnecessary risks; also information on the qualifications of clinical investigators – professionals (generally physicians) who oversee the administration of the experimental compound – to assess whether they are qualified to fulfill their clinical trial duties; and finally, commitments to obtain informed consent from the research subjects, to obtain review of the study by an institutional review board (IRB), and to adhere to the investigational new drug regulations.

Once the IND is submitted, the sponsor must wait 30 calendar days before initiating any clinical trials. During this time, FDA has an opportunity to review the IND for safety to assure that research subjects will not be subjected to unreasonable risk (FDA.gov IND Application).

One should not assume that if no response is forthcoming within the specified time interval that the regulatory agency is in full agreement with the sponsor's plan. Regulators may find themselves in situations where there are more documents to review than is possible within a particular time. If there are key areas that should be resolved prior to initiating first in man studies, it is best to be proactive and to indicate to the agency that communication on a particular topic is sought. Depending on the complexity of the topic and work tendencies within a health authority, communication may be in writing, by teleconference, or at a face-to-face meeting. Requesting a pre-IND meeting is highly recommended as this is an excellent opportunity to discuss issues where there is any level of doubt.

Given that regulators are very busy, it is important to provide them with high-quality documents that clearly state the key questions which need to be agreed upon. The scientific considerations need to be stated in a logical and easy to follow manner. Regulators need to balance patient safety risks with the potential benefits of the product under investigation. Their objective is aligned with industry in that they want safe and effective products to be brought to market. It is critical to listen very carefully to the guidance provided, to clarify the scientific advice, and also to challenge it based on either scientific evidence or regulatory precedent.

Global health authorities are available to meet with sponsors throughout the drug development process to resolve issues that arise. It is important to realize that health authorities are responsible for protecting patient safety during the development process, but are not responsible for deciding how to develop the drug. While formal scientific advice is available from several health authorities, it is "advice" and does not mean that following the advice is mandated nor does it guarantee that if the advice is followed and the primary endpoint is met, that approval will be granted. On several occasions, sponsors have asked health authorities their opinion on the best way to proceed with a specific drug candidate in development. These questions are out of scope for the regulators. It is up to the sponsor to propose a development plan. The regulators will review the plan and provide concerns, objections, or endorsement of its components. Be respectful of the agencies time and communicate professionally. Frequent communication on trivial matters will be more likely to cause a health authority to view the sponsor as incompetent rather than fostering a positive relationship.

When meeting with health authorities, "there should be free, full, and open communication about any scientific or medical question that may arise during the clinical investigation" (CFR title 21). This directive means that full disclosure is required. Therefore, all potentially pertinent data need to be presented. Suppression of potentially unfavorable data is unacceptable and will lead to difficulties in the future. Potential safety signals should be clearly identified, and the sponsor should present a plan for their evaluation during the ongoing clinical development.

Considerations for Trial Design and Conduct

Your imaging partner should be versed in the regulatory standards for the therapeutic area and also be up to date and in compliance with good clinical practice (GCP) standards and with CFR 21 part 11. Most contract research organizations with good imaging services and expertise will be able to provide consulting services during study design for nominal fees in comparison to the total study costs. Getting your imaging partner or an imaging consultant involved early in protocol design will expedite the process and will enhance the quality of the trial.

Table 6.4 outlines different situations that may be encountered as part of the clinical development process. Similar principles apply whether we consider a laboratory parameter such as glycosylated hemoglobin, a patient-reported questionnaire, or an imaging parameter. The first and most important principle is that the sponsor is responsible for and accountable for the development program and its outcomes. The health authority will provide guidance and may mandate certain procedures to maintain patient safety including placing a program on clinical hold, but the sponsor is ultimately responsible for their program. When guidances exist that are current, the task is straightforward. Regulatory guidances for imaging parameters are behind those of other clinical endpoints such that interaction with imaging professionals who have been involved in health authority interactions with successful programs are currently recommended. As published guidelines are written by FDA and others, followed by accumulation of experience with these guidelines, the need for such interaction may be reduced. When changes in the imaging endpoint are sought due to emerging endpoints which may have enhanced predictability for disease progression, discussion with health authorities regarding methods used to validate these endpoints will be required. Differences in image acquisition, read methodologies, or other parameters should be detailed in the imaging charter and posed as specific questions in briefing documents.

Advocating for a modification of the traditional regulatory pathways in the absence of scientific information that clearly justifies modification of the existing approach is likely to be futile. If new scientific information is available that is compelling for say a change in endpoints despite no change in the standard of care, then a change can be effected. In order to successfully modify existing regulatory precedent, a sponsor will need to be very well organized with the support of the appropriate professional organizations and key individuals therein. The rationale and

Table 6.4 Regulatory paradigms in clinical development

Regulatory standard	Available guidance or precedent	Sponsors' objective
Established	A guidance exists which is current with the standard of care	Sponsor wants to follow this regulatory path
	Sponsor wants to modify existing regulatory standards to optimally position their new product	Sponsors' target product profile involves a modification to current product labels in the existing class (will need compelling scientific arguments supported by well-respected scientists)
Evolving	A guidance exists which the sponsor does not consider to reflect current or emerging standards of clinical practice	Sponsor proposes an alternative path to approval
Absent or rudimentary	No guidance exists	Opportunity to set standards
	The new chemical entity is partially addressed by some existing guidances but with conflicting direction	Challenging as the health authority may prefer to adopt a single related guidance rather than create a new one to accommodate this therapeutic agent

benefits to the public will need to be apparent to all involved. A modification that benefits one sponsor over another is less likely to be ratified.

In situations where the guidance is ten or more years old and practice standards have evolved considerably, it will be helpful to meet with the health authorities early in the development process to propose your anticipated development plan leading to approval. When given sufficient time and faced with obviously outdated guidances, the health authorities will usually update the guidances during the conduct of your development program. Risks are best managed by working closely with the health authorities such that your scientific rationale that is driving the need for updating the guidance is reflected in the final document. In this situation, it is prudent to thoroughly map out all of the options and to consider opinions external to one's organization including providers across different geographic regions. Since one prefers to have a single global standard, the sponsor will need to engage team members with effective communication skills who can develop that single global standard in a series of interactions with various health authorities.

When no regulatory guidance exists, this may mean that you are in the process of solving a significant unmet medical need. Health authority staff want to participate in bringing novel and safe treatments to market. They will be energized at the prospect of satisfying an unmet medical need and will usually prioritize your meeting over others especially if the treatment under investigation has promising preliminary data. When moving along this path, try to maintain maximum flexibility and try to avoid committing to a final strategy until you have fully interpreted the end of phase II data. The reasoning for this waffling is that when you are going into

unchartered territory, the potential for unanticipated situations is increased. The health authorities also do not want to err or retract their position, so the delay in commitment should be mutual.

The most complex situation is when your product does not fit neatly into any of the existing guidances. Say you have a product which counteracts some of the cytokines that are thought to be responsible for the increased cardiovascular morbidity and mortality associated with rheumatoid arthritis. If one follows the cardiovascular guidances, a clinical outcomes trial is recommended. However, patients with painful arthritis will not agree to be randomized to a placebo, so a placebo-controlled trial is impractical. There are no proven agents with this capability so an active comparator trial is not scientifically valid. In these situations modifications to the existing guidances should be made in order to provide the opportunity to bring such a therapeutic entity to market. Negotiating this path will involve considerable challenges.

End of Phase II Meeting

The sponsor has the right to request a meeting at the end of phase II. At this juncture, the product has demonstrated a positive proof of concept for efficacy, and safety appears to be acceptable. In general all sponsors should take advantage of this legislated opportunity. A briefing book should be submitted in advance that reviews the key efficacy and safety data to date. Proposals for the phase III program and the specific indications that are sought should be clearly described. Be thorough and include the imaging component and all key aspects of the proposed program. Agreements from this meeting are put into official minutes that will be used when evaluating the phase III program for product approval. Preparation for this meeting is a crucial step in the development process. A face-to-face meeting is preferred in most cases. External consultants should be utilized as needed and can be brought to the health authority meeting. One expert may include an external imaging consultant. They may attend in person or via teleconference even for a face-to-face meeting between the sponsor and health authority.

At times different scientific advice will be obtained from different health authorities. If the sponsor takes the advice of the health authority that recommends the most comprehensive phase III program, then no further interactions are required. Often there are differences in the recommendations that warrant further interactions with specific health authorities during the implementation of the phase III program. A special protocol assessment will be performed by the FDA at a sponsor's request. EMA will provide scientific advice. Many health authorities will not be current with imaging standards; therefore briefing documents must be well written and should not assume any specialized knowledge. It is good practice to have your regulatory documents pertinent to imaging endpoints drafted and reviewed in conjunction with your imaging partner.

Pre-submission Meetings

While optional, this meeting should be considered essential for several reasons. A phase III program takes several years to conduct during which new information becomes available. This meeting provides the health authority the opportunity to notify the sponsor of any potential deficiencies in the overall development program to date. This may include safety issues that the agency is aware of that have been observed with other products that are approved or are under investigation. It can also include updated regulations or changes in policies regarding toxicology, manufacturing standards, or other aspects of the pending application. Secondly, it provides the agency with a summary of the key issues involved in review of the sponsor's application. It enables them to more efficiently allocate their resources. It also speeds up the review process for the primary reviewers and provides for a scientific exchange between the reviewing division and the sponsor.

Advisory Board Hearing

Presentation of a drug's clinical research program either in a closed session to regulators from member states in the European Union or in a public forum in the United States is becoming more common. These are typical for drugs with a new mechanism of action and for drugs with clinically relevant safety concerns or potential safety concerns. Preparation for these meetings is extensive and should include one or more imaging experts who were involved in the reading and interpretation of the imaging results.

Financial Considerations

Within the pharmaceutical industry, a small number of projects provide exceptional financial returns which provide the financing for overall R & D. A product's revenues drop precipitously upon patent expiration. This has the consequence of needing to factor in the remaining patent exclusivity into clinical development decision making. Different organizations adopt varying approaches in managing R & D budgets. It is intriguing that the adoption of innovative methodologies such as phase 0 studies is more common in biotechnology companies than in large pharmaceutical organizations. Perhaps it is a reflection of the types of individuals who are drawn to the smaller biotechnology companies, or it may be that limited availability of funds drives more innovative solutions.

Invariably, the use of imaging technologies can increase the cost especially in early phase studies. In organizations where decision making is compartmentalized (e.g., a fixed budget for all phase I or early development studies), or where

the goal is simply to advance the compound to the next phase of development, one may face challenges in the incorporation of imaging parameters. The smaller biotech companies generally have fewer assets and are focused on the value of these assets over their entire life cycle. These organizations are also pushed harder to demonstrate results early so that they can attract future funding. While these factors will influence decision making, a thorough analysis regarding the potential advantages of imaging early in the development program relative to later stages should be performed. This should then be compared with overall program costs. It may be useful to consult with individuals experienced in conducting these imaging studies who will be able to assist in flushing out the advantages and limitations of various approaches.

Imaging is being increasingly utilized to evaluate potential safety signals. In many of these situations, it is prudent to initiate these studies during phase II for several reasons. First, a dose-dependent change may be observed which supports a pharmacologically mediated effect. This can then be factored into the decision of whether to proceed to phase III and if so with what doses. Secondly, if no evidence of the potential safety issue is observed, in addition to being reassuring, the experience in phase II will be helpful in the design and implementation of the larger phase III program.

In phase III when imaging endpoints such as fracture are the primary endpoints required for product registration, the trials are powered accordingly. For secondary imaging endpoints in large phase III programs, it is often cost-effective to perform these sub-studies in a limited number of centers where historical performance has been good. Decisions will also need to be made regarding the incorporation of imaging variables that are not essential for registration but have financial value post approval for commercialization. These can be placed into the phase III program but for regulatory and financial reasons are often better served as standalone studies conducted independent of the phase III program.

Estimating Costs for Imaging Endpoints

The cost for image acquisition represents a minority of the overall imaging costs for a phase III clinical trial. Therefore, while the cost for an MRI may be many times that of radiographs, the overall cost between imaging modalities will not be as large. The largest driver of costs in phase III will be the overall number of subjects enrolled. Patient retention will impact the number of evaluable subjects. Since the analysis is usually performed as intent to treat with last observation carried forward, patients who discontinue participation early are more likely to demonstrate a reduction in pharmacological benefit relative to those completing a full course of treatment. Therefore, patient retention programs need to be incorporated not just for the overall trial but also specific to the imaging component since this component is frequently managed at a location distinct from the clinical study site by different study personnel. Site performance will also affect the trial outcome as increased

variability in image acquisition will diminish the ability to detect a true difference with treatment even when one exists. Therefore selection of an imaging partner with the ability to effectively manage the study sites with a focus on minimizing variability at the sites is essential for success. The central imaging lab that is selected is the single most important investment decision and should be made early on as the clinical protocol is being developed. It is also important that the imaging partner selected be independent of the clinical investigators. Submissions have been rejected on the basis of a potential conflict of interest when the cooperative oncology group enrolling patients in the trial also controlled the imaging data and its assessment.

Key facets that are the cornerstone of a successful imaging laboratory include their operational focus on quality control. This starts with site selection, training, and maintaining active dialog with the sites. Challenges arise with all trials. The skills and experience of the imaging team to manage through these issues are relevant. Other pertinent aspects include the setting up and maintaining of an imaging database, the read methodology, and operational aspects of conducting the reads, and aligning the imaging assessments with the other trial activities.

Value of Imaging Partners

Whether your imaging partners are internal or external to your organization, the degree of success achieved within the project will be driven by the people involved. Red flags should go up when an external organization espouses their new technology which moves images around electronically with remarkable efficiency such that this can all be conducted flawlessly without human interaction.

This is the antithesis of the requirements for the successful conduct and management of imaging in clinical trials, which is still a people-based system. It is the technologist interacting with study subjects who acquires the images. It is the technologist or study coordinator who will transmit the images to a central location. Following the development of an acquisition protocol, it is the study management team who will train the sites, provide ongoing supervision, detect anomalies, and reeducate the site staff that is pivotal.

Demand from your imaging partner professionals both technical knowledge and superior communication skills. Consider the tradeoff between low-budget proposals that have insufficient human resources versus those who have ample personnel and quality controls in place. The technology platform should be proven, should be compliant from a regulatory perspective, and should provide operational efficiencies. However, the technology should not be the only variable considered. Also keep in mind that nothing works flawlessly. Issues will be identified during the conduct of the trial. If your partner is good, these will be identified promptly. This requires significant human interaction. Excellent communication which is dependent on the individuals on the imaging and study teams is the most important variable for both the study outcome and workplace satisfaction.

Operational Considerations

The process by which clinical trials are managed in many large pharmaceutical companies can place the imaging component at risk. A common industry practice is to develop a study protocol for late development which is then provided to an operations group that conducts a feasibility study. This feasibility study will ask a potential study site regarding the anticipated number of patients that they can enroll, whether they have equipment available for the imaging modality, and if they have participated in other clinical trials with similar requirements. The study protocol design will be finalized. Clinical trial research organizations will be asked to bid on the project. The CRO may bid for the imaging component of the trial, or specialized imaging central labs may be invited to bid. Finally an imaging "vendor" is brought on board to execute the agreed upon "scope of work." The group of individuals responsible for the imaging is charged with an operational task. There is minimal or no opportunity to contribute to the strategic imaging elements of the trial. Imaging strategy that is significantly flawed often will not be detected until well into the trial conduct. The imaging vendor as they are commonly referred to may be treated as subservient to the sponsor's personnel. Since they are not true partners, the management of the imaging vendor may be reluctant to communicate inadequacies in the imaging component to the sponsor in a timely manner with the result that some trials may not yield the intended imaging results.

Skilled professionals dedicated to conducting imaging trials often are able to contribute significant value. It is highly recommended that imaging core labs be interviewed early and usually more than once in a consulting role as the trial design is being developed. By engaging potential imaging "partners" early on, it will be apparent which organization is a better match for the particular project. Your imaging partner will be able to provide guidance regarding variables that can impact variability in image acquisition (e.g., hardware, software, patient positioning). They will also likely have experience with some study sites under consideration and will also be able to share characteristics of reliable sites as well as early warning signs for sites with poor quality control.

They will draft and discuss a trial-specific imaging charter. The imaging charter is a detailed protocol specific to the imaging component of the trial. It should be completed at the same time as the overall study protocol and should be submitted with the study protocol for a special protocol assessment (FDA) or scientific advice (EMA). In addition to the imaging charter, detailed training materials for the clinical sites specific to the trial should be prepared by the imaging partner. Formal imaging training should be a key component of the investigator meeting. Depending on the situation, study sites may be required to qualify for participation by demonstrating proficiency within predefined standards. Usually study site monitoring is performed in a timely manner following enrollment of the first subject and is usually more intense early on until the site becomes more familiar with the study expectations and is more self sufficient. A similar practice should be used for monitoring of the imaging data. Instruction of site personnel regarding imaging should

be viewed as a study long endeavor. A standard regarding acceptable image acquisition and a system for notifying the site of substandard images requiring repeat scans are required.

The statistical analysis plan for the imaging data, especially if it is the primary endpoint, will be discussed in the imaging charter. It is worthwhile to put together a comprehensive analysis plan prior to the final study protocol. The reasons for this are not only regulatory, but also scientific and operational. Unintended events will almost always occur in the conduct of clinical trials. In the writing of a statistical analysis plan, items such as visit windows, handling of out of window, missing and duplicate data, and unscheduled visits all need to be addressed. Events that may influence the imaging endpoints such as surgical procedures, specific concomitant medications, or development of specific comorbidities need to be discussed. Decisions will need to be made regarding whether to control for these effects in the statistical analysis. If a decision is made that these are clinically relevant, it will require accurate capturing of these events. This will impact design of the case report forms as well as study monitoring. Standards for acceptable images, time windows for repeat imaging studies, and other considerations may influence the imaging charter and final protocol. It is best to work through these considerations up front rather than engaging in protocol amendments.

When assessing an imaging core lab, it is important to request a dedicated study team. The experience and leadership capabilities of the team leader and whether team members will be dedicated to your study or will be juggling multiple responsibilities are relevant and should be captured up front and if deemed appropriate included in the scope of work or other relevant document. There will be turnover of personnel at the study sites during the conduct of the clinical study. Therefore, additional training of site staff will be required as this occurs. Equipment changes will also occur especially if the study duration is longer than 1–2 years. Planning for this should also be performed, and rules for managing the situation should be part of the initial study documentation.

In the previously mentioned example, where high-resolution CT scans are being used to evaluate lung density in COPD patients, it is critical to know in advance of study initiation, whether the clinical sites can provide high-quality data using the specified protocol. Variability both within sites and between sites is a critical factor in determining the success or failure of the study. From a statistical perspective, for a fixed number of subjects, the higher the variability of a particular study parameter, the higher the p-value and therefore the less likely one is to demonstrate a statistically significant treatment effect. If the variability increases, more patients will be required to achieve a similar p-value to that which could be achieved with fewer patients and lower variability. Therefore minimizing variability is a key operational objective within clinical trials. Large multinational phase III clinical trials involve differences in imaging equipment (hardware and software) as well as differences in patient positioning and related imaging procedures. Discussion will be required in order to decide on appropriate tradeoffs between minimizing variability and conducting the trial within a reasonable time frame. Usually, more industrialized regions will have more recent equipment compared to other regions, although many exceptions exist. Different manufacturers will tend to have dominant market share

in different regions. Individuals familiar with the hardware and software across manufacturers are needed in order to estimate the difference in measurements that will occur between products for the endpoints under investigation.

If clinical assessments are made prior to and following an intervention, some individuals will conclude that the selection of equipment will not be relevant since in essence the delta should be similar between all equipment. In practice post hoc assessment of data from clinical trials commonly reveals differences between patients assessed with different scanners. These differences may be greater in specific patient subgroups such as the obese. Therefore, specification of the equipment that is acceptable for the clinical study is required for all trials. All efforts should be made to ensure that equipment does not change at the site level during the conduct of the clinical trial.

It follows that minimizing variability at the site level can be best achieved by employing consistent procedures for image acquisition. This can be best managed through training, maintaining a consistent staff, and providing ongoing feedback from the central imaging laboratory to the sites. Excellent and ongoing communication between the sites and central lab is essential to achieve this objective.

Compliance

All clinical development programs which result in a health authority submission will be reviewed by health authority personnel prior to approval. It is critical that good clinical practice and ICH standards be adhered to throughout the clinical program. Documentation of all actions taken during the trial with a full audit trail including all entries clearly identifying the study personnel, time, and date is mandatory. One must be able to reconstruct all activities from an audit trail. Ideally study site monitors should be familiar with the imaging component. When it is the primary endpoint, 100 % source verification is appropriate. If there is any concern regarding the quality control procedures of an imaging provider, do not retain them until you are satisfied. When quality issues arise, try to work through them with your imaging partner. If this is not possible, a second provider may need to be brought in. There is a precedent for non-approval of imaging submissions due to compliance issues. Imaging standards are evolving rapidly. To the extent possible, management of anticipated issues should be prespecified in the imaging charter. An open and transparent relationship between the sponsor, CRO, and imaging provider in conjunction with well-defined responsibilities and a detailed scope of work is the best recipe for success.

Summary

In summary, medical imaging continues to evolve rapidly. We are beginning the process of applying consistent scientific principles to the design, implementation, analysis, and interpretation of imaging parameters. Health authority guidances

regarding imaging are emerging. The use of imaging to assess disease pathophysiology should be considered for all development programs. Once a decision is made to proceed with imaging parameters, experienced professionals should work together to minimize variability so that pharmacological effects can be demonstrated most efficiently. The use of imaging within development programs has been and will continue to increase over time. The acquisition of skills pertinent to the design and implementation of imaging parameters within clinical trials will be an asset to most biopharmaceutical organizations.

References

1. DiMasi JA, Hansen RW, Grabowski HG. The price of innovation: new estimates of drug development costs. J Health Econ. 2003;22(2):151–85.
2. Adams CP, Brantner VV. Spending on new drug development 1. Health Econ. 2010; 19(2):130–41.
3. Critical path initiative 2004 report. Updated 2010. Available at: http://www.fda.gov/downloads/ScienceResearch/SpecialTopics/CriticalPathInitiative/CriticalPathOpportunitiesReports/ucm113411pdf2011.
4. Woodcock J, Woosley R. The FDA critical path initiative and its influence on new drug development. Annu Rev Med. 2008;59:1–12.
5. Goodwin R, Giaccone G, Calvert H, Lobbezoo M, Eisenhauer EA. Targeted agents: how to select the winners in preclinical and early clinical studies? Eur J Cancer. 2012;48(2):170–8.
6. von Haehling S. Trial of lorcaserin for weight management. N Engl J Med. 2010; 363(25):2468–70.
7. McKenney JM, Davidson MH, Shear CL, Revkin JH. Efficacy and safety of torcetrapib, a novel cholesteryl ester transfer protein inhibitor, in individuals with below-average high-density lipoprotein cholesterol levels on a background of atorvastatin. J Am Coll Cardiol. 2006;48(9):1782–90.
8. Barter PJ, Caulfield M, Eriksson M, Grundy SM, Kastelein JJ, Komajda M, Lopez-Sendon J, Mosca L, Tardif JC, Waters DD, Shear CL, Revkin JH, Buhr KA, Fisher MR, Tall AR, Brewer B. Effects of torcetrapib in patients at high risk for coronary events. N Engl J Med. 2007; 357(21):2109–22.
9. Kontush A, Guerin M, Chapman MJ. Spotlight on HDL-raising therapies: insights from the torcetrapib trials. Nat Clin Pract Cardiovasc Med. 2008;5(6):329–36.
10. Biomarkers Definitions Working Group. Biomarkers and surrogate endpoints: preferred definitions and conceptual framework. Clin Pharmacol Ther. 2001;69(3):89–95. www.bioclinica.com/assets/uploads/medical-imaging-core-lab-act.pdf
11. Behrman RE, Woodcock J. FDA regulations for drug development. Science. 2010;329(5987):33.
12. DiMasi JA, Feldman L, Seckler A, Wilson A. Trends in risks associated with new drug development: success rates for investigational drugs. Clin Pharmacol Ther. 2010;87(3):272–7.
13. DiMasi JA. The value of improving the productivity of the drug development process: faster times and better decisions. Pharmacoeconomics. 2002;20 Suppl 3:1–10.
14. Suleiman OH, Gorovets A. FDA presentation on imaging standards. Available at: http://www2.rsna.org/re/TwoTopicImagingWorkshopPresentations/Index%20Files/Suleiman%20Gorovets/pdf2011.
15. Kremer JM, Blanco R, Brzosko M, Burgos-Vargas R, Halland AM, Vernon E, Ambs P, Fleischmann R. Tocilizumab inhibits structural joint damage in rheumatoid arthritis patients with inadequate responses to methotrexate: results from the double-blind treatment phase of a randomized placebo-controlled trial of tocilizumab safety and prevention of structural joint damage at one year. Arthritis Rheum. 2011;63(3):609–21.

16. McEwan AS, Graham MM, Conti PS. Site qualification in multicenter clinical trials. J Nucl Med. 2009;50(8):30N.
17. Wagner CC, Langer O. Approaches using molecular imaging technology – use of PET in clinical microdose studies. Adv Drug Deliv Rev. 2011;63(7):539–46.
18. Buchan P. Smarter candidate selection–utilizing microdosing in exploratory clinical studies. Ernst Schering Res Found Workshop. 2007;59:7–27.
19. Jekunen AP, Pauwels EK, Kairemo KJ. Microdosing in early lead discovery. Bioanalysis. 2010;2(3):421–8.
20. Stolk J, Putter H, Bakker EM, Shaker SB, Parr DG, Piitulainen E, Russi EW, Grebski E, Dirksen A, Stockley RA, Reiber JH, Stoel BC. Progression parameters for emphysema: a clinical investigation. Respir Med. 2007;101(9):1924–30.
21. Roth BL. Drugs and valvular heart disease. N Engl J Med. 2007;356(1):6–9.
22. Lundy MW, Stauffer M, Wergedal JE, Baylink DJ, Featherstone JD, Hodgson SF, Riggs BL. Histomorphometric analysis of iliac crest bone biopsies in placebo-treated versus fluoride-treated subjects. Osteoporos Int. 1995;5(2):115–29.
23. Stolk J, Ng WH, Bakker ME, Reiber JH, Rabe KF, Putter H, Stoel BC. Correlation between annual change in health status and computer tomography derived lung density in subjects with alpha1-antitrypsin deficiency. Thorax. 2003;58(12):1027–30.
24. Basu A, Jensen MD, McCann F, Mukhopadhyay D, Joyner MJ, Rizza RA. Effects of pioglitazone versus glipizide on body fat distribution, body water content, and hemodynamics in type 2 diabetes. Diabetes Care. 2006;29(3):510–4.

Chapter 7
Evaluating and Working with an Imaging Core Laboratory

Eric S. Forsthoffer and Joel Krasnow

Abstract This chapter will take you through all of the key components to evaluate selecting an imaging core laboratory for a clinical trial from the sponsor perspective. There is a corresponding checklist that will ensure none of the key components are overlooked during the selection process.

Keywords Corporate infrastructure • Regulatory experience • MCC metrics • Strategic partnership

Introduction

This chapter takes the perspective of the sponsor who is retaining an imaging core lab as part of a specific clinical trial or a clinical trial program. We make the assumption that the clinical development program is targeting regulatory approval at some point in the future and that conduct of the clinical trial will be performed consistent with good clinical practice and in a manner that will satisfy reviewing regulatory authorities. In discussing the attributes and behaviors of the imaging core lab, we will follow the traditional sequence of events that occurs during the clinical trial process and how to utilize metrics to effectively monitor trial progress and guide interventions as needed. The traditional sequence of events starts with establishing a partnership between the sponsor and the imaging core lab. This includes

E.S. Forsthoffer
Department of Business Development, BioClinica Inc,
826 Newtown-Yardley Road, Newtown, PA 18940, USA
e-mail: eric.forsthoffer@bioclinica.com

J. Krasnow, MD, MBA (✉)
Department of Strategy and Regulatory Affairs, TherapeuticsMD,
Boca Raton, FL, USA
e-mail: jkras@optonline.net

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_7, © Springer-Verlag London 2014

aspects related to general corporate characteristics such as number of employees, financial viability, experience, and degree of success with respect to regulatory approvals, as well as previous findings on health authority audits. Next we go through study design, study start-up, study conduct, data management and documentation, data analysis and reporting, regulatory support for health authority interactions, and project-specific attributes. In addition to these primarily technical factors, we will briefly discuss the culture of the organization that you are considering partnering with as this can become an important facet that can impact on performance between the sponsor and imaging core lab.

Corporate Infrastructure

The corporate infrastructure represents the demographics of the imaging core lab. This includes the number of employees, physical locations, financial status, and years in the imaging core lab business. Is the organization new or established? Is it growing, stable, or consolidating? What are the reasons for this? Is imaging an established core competency or a business expansion opportunity? From the sponsor perspective, one wants to ensure that the imaging core lab that we consider retaining is able to provide the necessary resources, both human and financial to the project for the entire life of the project including the period of health authority review. Therefore, the firm must be financially stable and ideally making a profit since companies losing money will for their own survival need to either remove resources or come to the sponsor for more funding to adequately resource the project. The ideal situation is a company that is stable or growing slowly. Rapidly growing organizations are associated with higher turnover at the project level which means that the sponsor team will be reorienting new team members at an above industry average over the duration of the project.

The nature of the sponsors' objective will also influence the selection of an imaging partner. Consider the example where there is a preclinical finding of heart failure in toxicology studies at doses greater than 50-fold the maximal predicted human dose. The team would like to incorporate an echocardiogram assessment at baseline and on the last day of drug administration in a phase IB study. They feel that echocardiograms performed locally in the radiology department are sufficient for their purpose. In this case having the CRO leading the study collect the echocardiogram reports may be sufficient. If there is greater concern regarding the potential for an adverse event of heart failure due to the mechanism of action of the drug or the calculated therapeutic window is only 3–5 times the predicted human dose, then a greater focus on the echo findings is warranted and an experienced imaging core lab would be desired.

The corporate structure of your imaging partner needs to mesh with that of the sponsors. The imaging partner should have a project team structure that aligns with the structure of the sponsors' team. Ready access to senior management within the imaging organization is essential for efficiently managing challenges that arise.

The imaging partner must be able to support the clinical trial in the geographic locations where the sponsor intends to recruit. This often means a requirement for a global infrastructure. It is good practice to drill down into these global requirements including the need for adequate resources to cover the volume from specific time zones.

The processes and procedures of your imaging partner will impact on the amount of sponsor resources that will be needed to successfully manage the study. Organizations that have well-established standard operating procedures that are reviewed and updated at regular intervals will benefit the sponsor. Inquire about the internal quality control measures employed by the imaging core lab. A high-level assessment can be easily obtained by requesting data from recent health authority audits that are routinely performed as part of submission reviews. For projects that merit increased scrutiny of the imaging core lab, the sponsor should request an on-site visit where they can assess the capabilities and can make a determination regarding the robustness of the imaging processes. Two additional aspects which are markers of successful partners are the current investment in R&D and their track record of successful health authority approvals. Imaging companies need to stay on top of new developments in their field. The rate of change in imaging technology is quite rapid. As a sponsor, it is imperative to know whether the guidance that will be provided is up to date from both a technical and regulatory perspective. Look for ongoing projects and relationships with the imaging hardware manufacturers and with leading institutions or companies developing new imaging standards. From a regulatory perspective, look for the presence of relationships with key imaging leaders within FDA and EMA. Look for an ongoing and consistent record of product approvals where imaging was a key component of the submission from the major health authorities.

One common mistake within study teams is that they focus on managing a specific trial and do not focus on the overall objective which is product registration. There are significant differences between successfully completing a trial and gaining timely regulatory approval. When imaging endpoints are key efficacy or safety parameters for regulatory approval, an imaging partner who has successfully navigated the approval process is a very valuable partner. This factor should be heavily weighted when deciding between imaging partners as within the pharmaceutical industry the cost of a non-approval or a deferred approval will usually be several multiples of the entire imaging contract.

Study Design

The most important milestone that should be achieved during the study design phase is to get an imaging partner on board. When the sponsor views the imaging core lab as solely an operational vendor, the imaging core lab is retained after the protocol is finalized. To date, I have yet to see a phase II/III protocol where some improvements to the protocol were not recommended by the imaging partner. I have

also observed many instances where protocol amendments were required or when decisions were made to compromise on some imaging aspects due to the desire not to amend the protocol or imaging charter due to the late engagement of an imaging partner.

Clinical trial protocols have many facets. Protocol development involves not only a thorough literature review but also dialog with those who are at the forefront of the field under investigation. Very few clinicians have had exposure or training in the principles of diagnostic imaging. Therefore, inclusion of imaging experts internal to the organization in conjunction with your imaging partner is preferred.

The powering of the study will be dependent upon several factors. These include the clinically meaningful change, the detection limit of the imaging technology, and the scan-to-scan variability within an individual subject. The scan-to-scan variability will be impacted by standardization of the acquisition procedure and by training of the investigational sites.

It is wise to involve the imaging partner in the protocol design since you may want to include in the investigational site feasibility assessment the availability of specific imaging hardware and software. The available hardware will impact the imaging sensitivity and may also impact reproducibility. Software updates are also common and may also impact key variables that will impact the power calculation. Since most-experienced imaging partners survey the investigational sites for technical and personnel information during start-up activities, early collaboration with the imaging core lab could gain efficiencies and remove duplicate efforts from the overall site feasibility and site survey processes conducted by the sponsors' team or their delegates.

Study Start-Up Activities

Following completion of the feasibility assessment and discussions on protocol design, a final protocol has been agreed upon. Due to differences in radiation exposure standards, clinical practice, hardware availability, and other considerations, specific geographic regions have been selected for participation. A single or a series of investigator meetings are being planned. Investigational sites are being identified, and study contract negotiations are underway. Investigational site qualifications need to be done together for both the imaging and other clinical aspects of the trial. Thus, the systems and personnel from both the sponsor and imaging partner need to be aligned. Informed consents need to be drafted, and the imaging partner will be asked to respond to questions or requested changes to the imaging component of the informed consent by the respective institutional review boards.

In many organizations there is a 6-month time period from final protocol to the investigator meeting. Prior to or immediately following "final" study protocol, it may be of value to the sponsor to conduct a reproducibility study. This study will involve taking patients who would be eligible for the study and imaging them

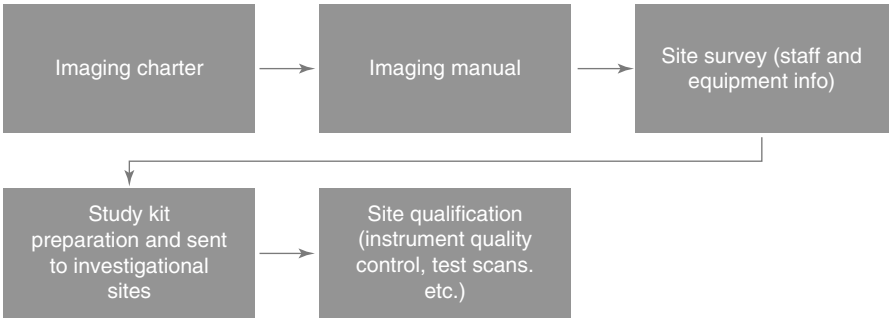


Fig. 7.1 Critical start-up points of the imaging workflow

according to proposed study guidelines. After acquisition of the image, the patient would get off of the imaging apparatus. Several minutes later the same process would be repeated. This type of study can be used both to determine the within subject variability between scans and to identify key factors involved in determining the final image read out. By identifying key variables that impact upon the final image acquisition, instructions to the study sites can be generated that specifically focuses on these areas. This will enhance the imaging investigational site training. A reduction in scan-to-scan variability has the effect of increasing the power of the study to demonstrate a statistical difference between the investigational and comparator groups.

Several tasks led by the imaging core lab with assistance from the sponsor and CRO need to occur before the first patient can be enrolled in the clinical trial. As mentioned previously, selecting and engaging your imaging core lab vendor early on in the process is key, as it will enable sufficient time to complete these start-up activities with the required quality for a successful study. A high-level flow diagram in Fig. 7.1 highlights the start-up activities from the imaging core lab perspective. All of these activities are dependent upon and driven by the clinical protocol. Therefore, it is essential that a well-developed protocol has been completed prior to study start-up activities.

The imaging core lab should be able to take the lead in generating the imaging charter. The imaging charter may be included as part of a special protocol assessment or scientific advice which will impact the timelines of development and finalization. It will include detailed information regarding image acquisition, read methodology, and data management. The importance of an imaging charter has increased with the draft FDA guidance document released in August 2011, focusing on the content and importance of the imaging charter which is explained in great detail in Chap. 4. The final protocol is imperative to ensure there is no delay in finalizing the imaging charter or the need to produce several amendments.

Imaging manual refers to the detailed instructions regarding image acquisition that is contained within the training manual developed for the investigational sites. This document must fully describe the imaging time points, the imaging modalities,

de-identification procedures, image submission procedures, source data storage regulations, query resolution process, and imaging protocol to be followed by the investigational site technologists. The imaging protocol that needs to be followed for a clinical research study differs markedly in comparison to everyday clinical practice. There is much more attention to detail and more documentation involved in performing research studies. Therefore, the technologists need to review the imaging manual document in full before scanning any subjects. We will touch upon this further when we discuss the importance of investigational site training and qualification.

The site survey will capture the investigational sites' contact information and equipment information which is necessary for ensuring the site is capable of participating in the trial as well as identifying the need for site training when personnel changes occur. Any issues identified with the investigational sites' equipment capabilities must be flagged to the sponsor and CRO to discuss options and associated risks with that investigational sites participation. In order for the imaging core lab to send the site surveys, they will need to receive a site list from the CRO containing the following required information: investigational site number, investigational site name, study coordinator name, and email address. If this required information is not included in the site list, the imaging core lab will be unable to survey the sites thus possibly causing a delay in start-up. Prioritizing investigational sites for this activity by the study initiation visit dates will be more effective.

Study kits are prepared by the imaging core lab and sent to all the participating investigational sites. A typical study kit will include an imaging binder and media (CDs, films, etc.) and mailers to submit the image data to the imaging core lab. If the imaging core lab has the ability for the investigational sites' to submit image data electronically and the investigational sites' have the capability to do so less materials/forms will have to be generated and sent to the sites via courier saving on shipping costs. Just like the site survey, the study kit must be sent to the site at the appropriate time to avoid duplicate work and unnecessary follow-up. This requires clear communication between the imaging core lab and CRO to ensure these activities take place when IRB approval is complete and the SIV is scheduled for the best response from investigational site.

Investigational site qualification refers to the process where the imaging lab certifies that the investigational site is able to successfully conduct all of the procedures required for the clinical trial as detailed in the imaging manual. While this process can increase the time required for having the investigational sites ready to acquire and submit image data, it directly improves the quality of the image data being submitted to the imaging core lab. Qualification can include test scans being submitted for review and approval, phantom scans and instrument quality control. This needs to be highlighted in the risk management plan to ensure the study team takes the appropriate actions with investigational site qualification in respect of time and the imaging modality or modalities involved in the clinical trial. Poor quality scans can have a major effect on the outcome of a trial. Therefore it is imperative that the investigational sites demonstrate proficiency not only at study initiation but throughout the study. This requires ongoing monitoring by the imaging lab.

Investigational Site Training and Qualification

The imaging manual document will need to be generated by the imaging partner. The key elements in maintaining consistency in image acquisition should be highlighted. Investigational sites should identify a primary and a backup technician who will be performing the image acquisitions. Their credentials should be reviewed by the core lab. The technicians should attend the investigator meeting, and special sessions should be devoted to review of the protocol and imaging guideline contents that are relevant. A formal assessment should be performed at the investigator meeting to determine whether the content was understood and is able to be acted upon according to the needs of the trial. Similar to the way the clinical monitor reviews the patient data from the first few subjects in detail with the study investigational sites, the imaging core lab should review the first few images being acquired in detail to ensure that they are consistent with the image standards set up for the trial. Should the image quality not meet prespecified standards, for trials where the imaging assessment is the primary endpoint there is no value in randomizing the subject as without a valid baseline assessment there is no way to generate data on change from baseline. Should there be minor issues with the investigational site these may be managed remotely. However, whenever there are significant issues, trained individuals from the imaging core lab should go to the investigational site, ideally when a patient is scheduled for imaging to assess and remedy the situation. In certain situations such as pivotal phase III trials the sponsor may wish to qualify individual investigational sites prior to permitting randomization of any subjects. This usually involves acquiring images from several patients and sending the images to the core lab for verification of image quality. Once the investigational site has demonstrated proficiency, then they are qualified to begin randomizing subjects.

Investigational site training should not be viewed as a onetime event at the investigator meeting. There will be some imaging technicians who are unable to attend the group training. There will be loss of recall regarding specific procedures over time especially at investigational sites less experienced in conducting these assessments and at slow enrolling investigational sites. A training plan should be requested from the imaging core lab that outlines all activities including investigational site remediation activities that may be required spanning the entire study interval. The training plan should detail how a need for training will be identified proactively via various quality gates established by the imaging core lab.

Study Conduct

Study conduct involves the collection and communication of data between the investigational site, imaging core lab, and study sponsor. The core lab should provide the investigational site with a secure process for transmitting the images together with subject number and core information required for interpreting the images. If the acquired image needs to meet certain criteria for study enrollment,

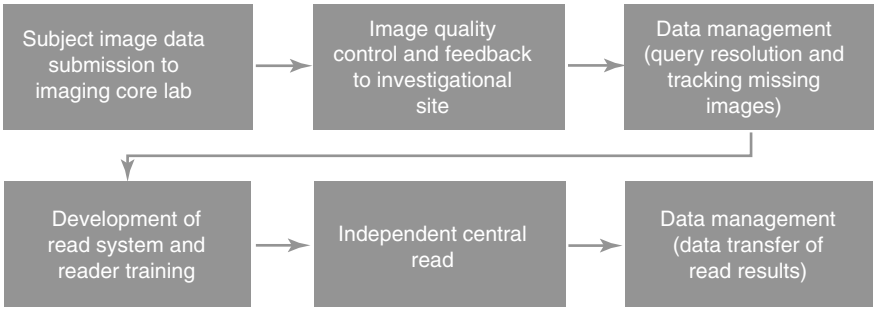


Fig. 7.2 Critical study conduct points of the imaging workflow

such as a bone mineral density for inclusion in an osteoporosis trial, then turnaround should be sufficiently rapid so as to permit good operational flow at the investigational site level. The same system that securely transmits images between the investigational site and core lab should enable transfer of the images to the blinded readers where assessments can be recorded. Queries pertaining to the data will originate within the core lab to the investigational sites. In the current era of electronic data capture, the imaging core lab should have the ability to do this electronically. The core lab should review with the sponsor the systems, procedures, and data standards that they have in place. They should be able to provide real-time reports regarding the number and type of queries and be able to drill down to the investigational site and subject level upon request. Prior to engaging a core lab, one may want to inquire regarding metrics for similar trials they have performed in the past.

Similar to the conduct of the nonimaging components of the trial, a monitoring plan should be in place for the imaging component. The core lab systems should provide a full audit trail with date and time stamped entries that identify the individual entering the data that are CFR part 11 compliant. In essence the documentation system should allow any auditor to be readily able to reconstruct the events that occurred during the trial. Successfully conducting a clinical trial requires not only technical skills but also good interpersonal communication skills and good attention to detail. As a sponsor one should insist on meeting the team members that the core lab plans to dedicate to your study. You should also inquire as to whether these team members have other significant responsibilities or are dedicated primarily to your project. You should feel comfortable that the team has sufficient experience to solve the problems that will invariably be encountered during the conduct of the trial. Finally, you should agree on a plan for project oversight from both the sponsor and imaging core lab perspective including when certain milestones are achieved such that data will be transferred to the sponsor for assessment of data integrity and analysis.

As we did with study start-up, we are going to now discuss the key tasks and associated best practices for study conduct following the flow diagram in Fig. 7.2.

The investigational site image data submission to the imaging core lab is an extremely important area that needs to be focused on. Image data must be submitted

to the imaging core lab within 3 days of acquisition to maintain high quality. The 3-day window is commonly missed. It is important to understand the process at the investigational sites in order to provide the best solution via electronic submission or simple process improvements at the investigational sites.

If the investigational sites are submitting the image data within 3 days of acquisition, the imaging core lab will have the ability to identify issues early and implement corrective actions with the investigational sites via the image quality control and feedback process. The imaging core lab will need to have highly qualified and certified modality-specific imaging technologists for the required image review per the study protocol and imaging charter. The experience of the imaging technologists is important to ensure imaging-related queries are being generated when required. Some imaging core labs will generate imaging-related queries when they are not necessary or not generate them when they are needed, which will be reflected in the independent central read.

Managing the resolution of queries and tracking down any missing image data from the investigational sites must be ongoing with close collaboration between the CRO and imaging core lab. The process for following up with the investigational sites needs to be clearly stated in the communication plan as well as the appropriate escalation paths when issues need to be escalated. Re-occurring meetings with the imaging core lab and CRO will drive the necessary communication to monitor investigational site, as well as aligning the CRO and the imaging core lab's activities. Responsiveness from investigational sites is historically poor which is understandable and should be anticipated. Sites commonly do not follow instructions regarding submitting image data as soon as possible after each time point is acquired for each subject. Instead, they send multiple time points together which is referred to as batching. The CRO and imaging core lab must be open in acknowledging these issues and should work together to develop effective solutions.

The development of the read system and reader training has to be completed in order for the central read to commence. There are numerous tasks that have to occur for these activities to be completed. Therefore, the imaging core lab must set the appropriate expectations, roles, and responsibilities to ensure each task is completed and nothing gets overlooked. If one of these tasks is overlooked, it may very well impact the ability to deliver the read results when required. Flexibility in designing a read system can easily improve the power of your data by being able to perform additional analysis. Experienced imaging core labs will bring this to your attention and involve the relevant experts when necessary. Reader training is best accomplished in a face-to-face meeting with imaging core lab, readers (radiologists, oncologists, cardiologists, etc.), sponsor and/or CRO. The ability of the imaging core lab to calibrate the readers through the initial reader training will be reflected in the adjudication rate. Adjudication rates vary per indication, and an experienced imaging core lab will be able to advise you on what is expected and what is abnormal before the independent central read begins.

Once the independent central read has begun, the imaging core lab is responsible for communicating status updates and loading all available subjects into the read systems. The rule for when a subject is to be read needs to be established well in

advance. The imaging core lab will use the read plan in order to monitor expected vs. actuals. A bell-shaped enrollment curve is desired by the imaging core lab as it prevents the need to shorten timelines or add resources for the interim and/or final data analysis like when there is a bolus of subjects enrolled at the end of the enrollment period. The time from final read to interim or final analysis needs to be looked at closely as the time necessary to complete this task will vary per read methodology, the required image data being available to be read, the selected readers' availability, and time needed to send the data transfer containing the read results.

Data transfer of read results should be a smooth process as it is the final critical point in the process when tension is at its highest point. An experienced imaging core lab will start discussions of the data transfer early and finalize the required specification document shortly after the design of the read system has been finalized. The sooner this can be done and a test transfer can be generated and approved by the recipient (sponsor, CRO, or third party), the better as it allows flexibility to review the read results earlier than expected if necessary.

Once the final data transfer is completed many imaging core labs feel that their work is done for the most part which is incorrect. An experienced imaging core lab will assist with the interpretation of the data from the images with the health authority submission. Presentation can make or break any deal in real life and this also rings true with submission. You have to put the results in context of both a clinical and therapeutic response. An imaging core lab advising the sponsor about the data significance will be a great asset to the health authority submission.

Risk Mitigation Plan

In addition to the monitoring plan, a risk mitigation plan for the imaging component is another document that will benefit the clinical trial immensely. The imaging partner should lead this process by going through deviations from the intended imaging process and should gain consensus on how to manage these deviations prior to study initiation. This should all be clearly detailed in a risk management document that focuses on the foreseeable risks specific to the clinical trial. Risks associated with investigational site start-up, investigational site training, missing images, resolution of queries, independent review progress, data transfers are all crucial to discuss at the start of the clinical trial, and this open dialog needs to continue throughout the trial. Transparency between all parties is critical to success. The experience of an imaging core lab feeds into this document and is a good test to determine if you selected an experienced imaging core lab or not.

Study Closeout, Analysis, and Communication

Study closeout is an intense time for the study team as they are under pressure to close out all queries in order to lock the clinical trial database. The imaging core lab must finalize the reads, perform internal quality checks regarding within reader and

between reader variability, as well as resolve differences between readers through the prespecified adjudication process. Since it is relatively easy for the imaging core lab to be on the critical path towards database lock, it is imperative that the sponsor work proactively with them and the investigational sites to collect all needed data and to resolve discrepancies. The amount of work to be done at this critical time is inversely related to the ongoing efforts of the study team during the trial. This is when imaging core labs with substandard processes and reporting systems or those who do not communicate openly with the sponsor will "suddenly" become the rate-limiting step for database lock. From the sponsor side, periodic data transfers followed by sufficiently in-depth analyses should alert the sponsor ahead of time to any issues that require resolution. Sufficient resources both on the sponsor and imaging core lab side should be applied to the project well in advance of the last patient completing the study.

The analysis plan should be outlined in the protocol and imaging charter. More detailed documentation of the analysis plan must be finalized in the statistical analysis plan and associated documents in advance of database lock. At times some anatomical structures are not evaluable due to previous surgery or other circumstances. This may necessitate manual coding for some subjects. Interaction between sponsor statisticians and the imaging core lab may be required. Following prespecified statistical analyses, the data must be interpreted and communicated in clinical study reports as well as submission documents to health authorities. Invariably, there will be questions that arise regarding the imaging data such as differences in study drug efficacy according to scanner type, geography, or specific patient demographic or disease variables. Individuals from the imaging core lab who have had experience with other studies can assist in the interpretation of this data. Their expertise may also be very helpful in responding to imaging-related questions from the health authorities. Certain health authorities will also want to audit the actual images from the clinical studies. However, rather than travel to the core lab, they will request that the images be sent to them. In such cases which are becoming more common, it is important that the core lab has experience with the required specification of the viewing system used by the health authority reviewers.

Metrics Champion Consortium (MCC)

When evaluating the operational effectiveness of an imaging partner, there are standardized measurements which are utilized within the industry that can serve as a useful assessment tool. The Metrics Champion Consortium (MCC) was established in 2006, focusing on improving clinical trial processes via standardized performance metrics. In January 2009, the MCC published the standardized imaging metrics for clinical trials. The metrics focus on four key elements: quality, timeliness, efficiency/cost, and cycle time. There are a total of twenty (20) standardized MCC imaging metrics, version 1.1 issued November 2011. We have compiled two tables which separate the quality metrics from the metrics that focus on timeliness, efficiency/cost, and cycle time as often different individuals are responsible for these different metrics within the sponsor's organization.

Table 7.1 MCC version 1.1 quality metrics. One additional metric which is not part of the MCC is included in this table as it is a great indicator of quality at one of the final steps in the process

MCC metric #	Metric category	Area targeted	Metric definition
10	Quality	Image QC	Percentage of suboptimal (but evaluable) images
11	Quality	Image QC	Percentage of non-evaluable images vs. total received
12	Quality	Image QC	Percentage of non-evaluable baseline images vs. total received
13	Quality	Data management	Percentages of missing imaging visits
14	Quality	Data management	Percentage of investigational site queries
19	Quality	Protocol	Number of image acquisition technique-related amendments per modality per protocol
n/a	Quality	Data transfer	Percentage of the data transfers meeting the data transfer specification document

As a sponsor it is critical to clearly communicate to both the imaging partner and CRO what quality metrics are targeted for the study. Quality can be defined as the percentage of non-evaluable images during the central read, the amount of missing imaging visits, the number of queries that are generated for images not acquired by the imaging protocol, and/or data transfers meeting the expectations of the data transfer specifications document. If quality is not discussed at the beginning of the study, it will most likely not be discussed throughout the clinical trial. By discussing quality as a team, quality will stay in the forefront of everyone's mind and result in a successful study. Once quality is defined, the team can focus on how to monitor and develop standard practices for addressing any challenges the team may face. Table 7.1 represents the MCC quality metrics as well as an additional suggested quality metric.

All of the quality metrics focusing on image quality have a direct impact on the imaging endpoint. Non-evaluable images at baseline mean that a change from baseline cannot be calculated rendering the patient as noninformative for the imaging endpoint. Non-evaluable images post-baseline will need to be imputed according to methodology agreed upon by the health authorities. These methods are deliberately conservative, meaning that the missing data will be treated in a manner that usually will reduce any treatment effect that an evaluable image would have provided. Suboptimal but evaluable images will diminish the accuracy of the reading and therefore serves to increase the scan-to-scan variability which also diminishes the ability to demonstrate a treatment effect. Missing imaging visits will also need to be imputed, thus every effort should be made to obtain the scheduled scans.

Amendments to the image acquisition technique reflect a failure to fully anticipate events occurring during the trial, may result in data collected using multiple techniques, and have several undesirable operational consequences. The number of investigational site queries involves several factors that can have opposite effects. A lack of queries may indicate that the imaging core lab is not being thorough in their review. Excess queries may indicate poor investigational site performance, or poor communication between the CRO, imaging core lab, and study site. Data transfers

Table 7.2 MCC version 1.1 timeliness, efficiency/cost, and cycle time metrics

MCC metric #	Metric category	Area targeted	Metric definition
1	Efficiency/cost	Business development	Average percentage of variance in the imaging budget
2	Cycle time	Business development	Average number of calendar days from imaging study award to contract signature
3	Timeliness	Project start-up	Percentage of investigational sites qualified vs. actual
4	Cycle time	Project start-up	Average number of calendar days from investigational site designated ready to first date of image receipt
5	Cycle time	Image acquisition and submission	Average number of calendar days from image acquisition to image receipt
6	Cycle time	Image QC	Average number of calendar days from image receipt to initial feedback to investigational site
7	Cycle time	Image processing	Average number of calendar days from image QC complete to reporting of eligibility results
8	Cycle time	Image processing	Average number of calendar days from image receipt to ready for independent review
9	Cycle time	Image processing	Average number of calendar days from when the image is designated for review to completion of the review
15	Cycle time	Data management	Average number of calendar days an imaging query is outstanding
16	Cycle time	Data transfers	Average number of calendar days from last patient reviewed to delivery of dataset
17	Timeliness	Data transfers	Average number of calendar days from original estimate to actual for export submission
18	Cycle time	Project start-up	Number of weeks to develop and write independent review charter
20	Timeliness	Protocol	Percentage of images acquired at investigational sites within agreed-upon timeframe for imaging time point (as defined by protocol)

also involve communication between the imaging core lab and the receiving organization as well as adherence to the transfer specifications. It is important to perform several transfers over the course of the trial as information technology systems may be updated, and the historical data transfer specifications may no longer function as they did previously. Close communication between all parties is the most important aspect to a successful trial. Also, close monitoring and prompt attention to these quality metrics often impacts the outcome of the trial.

Table 7.2 focuses on the timeliness, efficiency/cost, and cycle time metrics.

Timeliness, efficiency/cost, and cycle time are a direct reflection of the experience, specifically within project management, at the imaging core lab, CRO, and sponsor as well as the overall performance level of the imaging core lab. Setting the appropriate expectations and having strong communication between all parties involved greatly improve these metrics stated previously. The key to these types of metrics is to monitor on a continuous basis and have action plans established. The most difficult metrics to achieve prespecified targets typically include investigational site involvement. An agreed-upon action plan for identifying investigational sites with trending issues will lead to having predetermined corrective actions depending on the level of severity. These corrective actions should include putting the investigational site on hold to provide the proper retraining or possibly closing the investigational site to enrolling new subjects or participating in the trial. In a lot of cases, the investigational sites are not addressed appropriately and continue to impede progress of achieving the targets for these metrics and the overall quality and timeliness of the trial.

An imaging core lab should have the capabilities to track MCC metrics or a variation of the MCC metrics. Depending on the need and goal of a clinical trial, not all of the metrics listed previously may apply, but the majority usually does. The implementation of operational metrics is useful in focusing team activities towards prespecified goals. By capturing metrics on a monthly interval, it is very easy to see the areas that require additional focus or process improvements. The relationship between the imaging core lab, CRAs, and investigational sites is crucial.

The impact of not meeting the desired target for these metrics will vary depending on the indication and central read design, but the imaging core lab must be proactive and should develop solutions on how to tackle these challenges via training, communication, setting expectations correctly and early, and identify issues immediately when appropriate corrective actions can be taken. Training is a key component to ensuring that key metrics are met within the agreed-upon target. In general, the more high-quality training that can be applied at the onset of a clinical trial, the fewer issues and corrective actions will have to be implemented during the trial. The better performing imaging core labs through their experience know how to mitigate commonly experienced issues and demonstrate proficiency in rapidly identifying and successfully managing deviations from the operational plan. One recently published peer-reviewed paper published in conjunction with the MCC details the advantages of using metrics for imaging in clinical trials with case examples [1].

All of the MCC metrics are important, but there are key metrics that require additional attention as they feed directly into the quality delivered by the investigational sites and imaging core lab. I have never worked on a clinical trial that has no investigational site issues. There will always be at least a few sites that require intervention and these metrics will help identify them early. Metric 5 is a common metric that does not meet its target and prevents the imaging core lab from providing the best quality control as possible for the image data received. When investigational sites batch the image data as stated previously, it does not allow the imaging

core lab to proactively manage the image quality from the investigational sites. It is important for CRAs assisting in monitoring this activity via accessible reports from the imaging core lab to remind the investigational sites to send the images promptly following acquisition. This will greatly improve this metric of time from image acquisition to submission. Metric 6 measures the imaging core lab's ability to provide feedback to the investigational sites in a timely manner. When a protocol allows a window for having repeat images performed, it is based off of when the image data was acquired thus requiring a short cycle time for metrics 5 and 6. Metric 15 measures the time an imaging query is outstanding which closes the loop ensuring that the corrective action is taken by the investigational site as quickly as possible in order to avoid repeated imaging acquisition issues. This is another metric which requires cooperation from the CRAs and investigational sites. We personally feel that metric 11 is the most important metric to determine the level of quality applied to the clinical trial as all of the other 5 key metrics are contributing factors to metric 11, percent of non-evaluable images. You could have a high percentage of queries across the study, but if you have a low percent of non-evaluable images, it tells you that the CRO and imaging core lab took the appropriate steps for maintaining quality. Metric 13 is the best indicator of the level of communication between the imaging core lab and CRO. If the percentage of missing image data is high, it means that you will effectively be losing study subjects from inclusion in the independent central read for either the primary endpoint of the study or secondary time points. Last but not least is metric 14 which allows you to measure if the investigational site training applied was appropriate or not. This metric also will tell you if problem sites were identified and the proper corrective actions mentioned previously in this chapter were taken.

Culture and Financial Strategy

The culture of an organization is heavily influenced by its leadership. Corporate leaders hire, retain, and advance individuals based on performance characteristics that are valued. Look for organizations where the team is striving towards success. Be wary of teams that do not delve deeply into the project details who seek primarily to reassure you of their capabilities. For complex projects one may consider retaining one or more individuals from the imaging core lab as consultants during study design to ensure that the technical and communication skills are up to expectations. Finally, ensure that the goals and objectives are aligned between the sponsor and imaging core lab teams. The imaging core lab should benefit from the delivery of high-quality imaging data. The challenge is that the quality as reflected in within subject variability will not be evident until well into the trial. The culture of the imaging core lab, the thoroughness with which the sponsors' project proposal is worked through and the willingness to tackle ongoing challenges can be assessed. A face-to-face meeting with the prospective project team leaders is highly instructive.

Different imaging core labs utilize different contracting strategies. For discussion purposes, these are divided into 3 categories. The first is the low-bid strategy where the high-level project description provided by the sponsor is covered with a focus on the imaging parameters. Many details are not specified and several tasks that can reasonably be anticipated are not present. The bid is the lowest because it covers the least with respect to contracted services. The risk here is that as the trial progresses and additional services are needed, the sponsor is in fact hostage to the imaging vendor such that the initial low bid may turn out to be the high-cost selection. In addition it is much more difficult to implement new processes midway in a trial. This may lead to regulatory complications if some patients are managed differently from others.

The second strategy which can also be influenced by the sponsor team is the take-no-risk strategy. A good core lab will discuss sources of variability within an imaging program. A somewhat inexperienced but highly motivated study leader may opt to provide the same imaging hardware and software for all investigational sites. Similarly, training and investigational site monitoring may be performed at intervals that are more than usual and have not been demonstrated to improve results. Similar within and between reader variability assessments may be performed many more times than required. The adjudication process may be overly complex. While there are times when certain elements of this approach may be prudent, teams should be able to make reasonable tradeoffs in structuring their imaging program.

The third strategy is one where the imaging core lab has the experience to outline in sufficient depth and with contingencies for anticipated issues such as retraining of a percentage of investigational sites a complete study proposal. Their proposal should explain the rationale behind key decisions. When meeting face-to-face to discuss the proposal, the imaging core lab representatives should be able to explain the available options for each component in the proposal, along with their recommendation and rationale. Based on the nature of the project, the imaging core lab should be able to guide the sponsor regarding where investments have historically had a positive return. This works best when the imaging core lab is transparent and sufficiently experienced. The best situation is predicated on having experienced personnel working on the project from both the sponsor and imaging core lab sides. This will enable generation of a fair and comprehensive scope of work contract that enables sound project planning with few if any events that occur beyond those specified in the contract.

In summary, evaluating and deciding on which imaging core lab to use for one's clinical trials is a very important decision for the sponsor. A critical aspect is to get the imaging partner on board early in the process when the protocol design is still being developed. Developing a strategic partnership with an imaging core lab automates this critical aspect. Since this selection can make the difference between a successful program and one that is not, appropriate time and attention should be made in this process. Key considerations include the corporate metrics, the imaging core lab culture and work approach, as well as their systems and track record of success. A checklist incorporating all of the points we discussed can be found in Appendix 7.1 at the end of this chapter as an easy-to-use tool to assist you when evaluating and working with an imaging core laboratory.

Appendix 7.1: Checklist for Selecting an Imaging Core Lab

Imaging core lab capabilities	Yes	No	Not required	Comments
Is there global infrastructure?	☐	☐	☐	
Is there a sufficient amount of employees?	☐	☐	☐	
Is there a high turnover rate?	☐	☐	☐	
Is this an established organization (How many years have they been in business)?	☐	☐	☐	
Is the organization financially stable?	☐	☐	☐	
Is imaging an established core competency of the organization?	☐	☐	☐	
Is there ready access to senior management within the organization?	☐	☐	☐	
Does the organization have well-established SOPs that are reviewed and updated on a regular basis?	☐	☐	☐	
Does the organization have current investment in R&D?	☐	☐	☐	
Does the organization have a track record with successful health authority approvals?	☐	☐	☐	
Does the organization have relationships with key imaging leaders within the FDA and EMA?	☐	☐	☐	
Does the organization have experienced and sufficient medical and scientific staff?	☐	☐	☐	
Does the organization have the capabilities to capture, track, analyze, and take appropriate actions from MCC metrics or a variation of MCC metrics?	☐	☐	☐	
Does the organization have sufficient training methods for the investigational sites, CRAs, and sponsor?	☐	☐	☐	
Does the organization have an in-house electronic solution for transmitting image data?	☐	☐	☐	
Is the database tracking system and independent read system 21 CFR part 11 compliant?	☐	☐	☐	
Does the organization's study team at the organization have sufficient experience?	☐	☐	☐	
Does the organization have a dedicated experienced team developing imaging charters?	☐	☐	☐	
Does the organization's study team develop a risk mitigation plan as a standard practice?	☐	☐	☐	
Does the organization have experience with the required specification of the view system used by the health authority reviewers?	☐	☐	☐	
Does the imaging core lab have the ability to apply a governance structure via a relationship/alliance director?	☐	☐	☐	

Reference

1. Yu HJ, Miller CG, Flitcraft D. Metrics in medical imaging: changing the picture. ACRP: The Monitor. Aug 2012. p. 36–40.

Part II
Therapeutic Specifics of Medical Imaging

Chapter 8
Monitoring Responses to Therapy in Oncology

Binsheng Zhao and Lawrence H. Schwartz

Abstract Image-based, quantitative response assessment methods play an increasingly important role in monitoring tumor response to therapy in oncology clinical trials. The fate of a new drug should not be wrongfully determined because of imprecise assessment methods used in the clinical trial. Therefore, it is crucial to select the most appropriate response assessment method, which considers the mechanisms of interactions between drugs and cancers as well as drug-induced tumor changes as they are captured by the imaging modalities and associated response assessment methodologies. This chapter addresses the role and progress of the widely accessible imaging modality of computed tomography (CT) and advanced image analysis techniques in monitoring tumor response to therapy in oncology clinical trials and clinical care.

Keywords Oncology clinical trials • Solid tumors • Response assessment • Quantitative methods • Computed tomography (CT)

Introduction

Radiologic images have been used for decades to gauge the effectiveness of therapeutic interventions [1]. Increasingly, novel quantitative imaging techniques are being incorporated into oncology clinical trials, where they serve as surrogate biomarkers for various aspects of tumorigenesis or as indicators that facilitate evaluation of the efficacy of experimental therapies. Moreover, properly designed imaging studies can significantly affect the size, duration, cost, and success of clinical trials and ultimately affect patient care. Indeed, in this modern era, which has

B. Zhao, DSc (✉) • L.H. Schwartz, MD
Department of Radiology, Columbia University Medical Center,
New York, NY 10032, USA
e-mail: bz2166@cumc.columbia.edu; lhs2120@cumc.columbia.edu

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_8,

witnessed the development of targeted therapies and personalized medicine, development of a tissue biomarker that predicts sensitivity to a targeted therapy has become an essential step in the clinical success of a novel anticancer agent. Imaging can help identify such predictive tissue biomarkers, which can allow us to subdivide tumors into sensitive and resistant populations. In day-to-day oncology practices, imaging has been widely used to assist radiologists in the early detection of metastasis and in identification of ineffective and toxic therapies so that patients can be promptly switched to an alternative treatment option.

But, classic imaging approaches may not be as appropriate for many new cancer therapies that are being developed. For 30 years, the standard way to assess a patient's response to treatment in both clinical trials and clinical practice has been to monitor tumor changes measured bidimensionally per World Health Organization (WHO) criteria [2, 3] or, since 2000, unidimensionally using the Response Evaluation Criteria in Solid Tumors (RECIST) guidelines [4, 5]. However, many of the new classes of anticancer therapies are cytostatic drugs that may not cause as rapid tumor shrinkage or may cause less tumor shrinkage than previous generations of cytotoxic chemotherapies. Instead of size reduction, therapy-induced tumor changes may be associated with development of central necrosis or other complex changes. Such new patterns of change seen on radiographic images are challenging traditional response assessment methods, which are based on measuring tumor diameters, predominantly on longitudinal computed tomography (CT) or magnetic resonance imaging (MRI) scans.

While functional and molecular imaging techniques, e.g., positron emission tomography (PET) and dynamic contrast-enhanced (DCE) MRI, hold great promise, they are immature, expensive, rarely accessible, and prone to measurement variability. In contrast, CT is well developed and globally ubiquitous and is the standard clinical imaging modality for monitoring the growth of solid tumors. Furthermore, in clinical trials, there are many quantitative endpoints that depend upon CT findings and may correlate with overall survival. These endpoints include (but are not limited to) objective response rate (ORR), progression-free survival (PFS), disease-free survival (DFS), and time to progression (TTP). Ultimately, in monitoring therapies, the goal of imaging is for the images to serve as a successful surrogate endpoint for a patient's response to therapy. Thus, tumor shrinkage that is quantified by ORR, PFS, etc., as measured on CT should ideally correlate with prolonged survival. CT has also facilitated identification of target lesions in baseline examinations, detection of new lesions in follow-up studies, and confirmation of tumor responses after completion of therapy.

In this chapter, we will take CT as an example to address the role and progress of medical imaging and image analysis techniques in monitoring tumor responses to therapies in oncology clinical trials and clinical care. We will start with a brief overview of conventional response assessment methods and then address limitations of these standard response criteria, especially in the era of therapies targeting specific molecules. We will then introduce revised and modified RECIST criteria for lymphoma, mesothelioma, hepatocellular carcinoma (HCC), and gastrointestinal stromal tumor (GIST). Last but not least, we will discuss the use of volumetric CT to improve tissue biomarker discovery for novel therapies in non-small cell lung cancer (NSCLC).

Conventional Response Assessment Methods

Tumor change with therapy plays a fundamental role in medical oncologic care. A reduction in tumor size, termed a "response" to therapy, indicates that the patient is gaining some degree of benefit from treatment. In contrast, an increase in tumor size, termed "disease progression," suggests a tumor that is refractory to therapy and that a change of treatment is needed. In clinical trials and also in clinical care, tumor sizes are measured mainly on CT, and based on size changes, tumor responses to therapies are generally assessed by WHO and by the currently recommended RECIST guidelines.

WHO Criteria

The first guideline, known as the WHO criteria, that attempted to use objective metrics to assess tumor responses to therapy and to standardize reports of clinical outcomes from cancer treatment trials was codified by the WHO and published in 1981 [2, 3]. The WHO criteria utilize the cross product (i.e., bidimensional measurement) of the greatest diameter of the tumor and its greatest perpendicular diameter in a transverse plane to approximate tumor burden. Based on the change in the sum of these cross products of tumors, the WHO criteria recommend reporting results of cancer treatment using the following four categories: complete response (CR), partial response (PR), stable disease (SD), and disease progression (PD) (Table 8.1). A size reduction of 50 % or more from the baseline study was considered to be a PR, whereas a size increase of 25 % or more was deemed to be PD. The presence of any new lesion would be considered PD, and any "substantial" enlargement in tumor size that was not easily measured would also be considered PD.

Table 8.1 The WHO and the RECIST response criteria

Response category	WHO definition	RECIST definition
Complete response (CR)	Disappearance of all disease, as confirmed at 4-week follow-up	Disappearance of all disease, as confirmed at 4-week follow-up
Partial response (PR)	50 % decrease in the sum of the cross-products of measurable disease, as confirmed at 4 weeks	30 % decrease in the sum of the maximal diameters of measurable disease, as confirmed at 4 weeks
Stable disease (SD)	Neither PR nor PD	Neither PR nor PD
Disease progression (PD)	25 % increase in the sum of the cross-products of measurable disease or the presence of new disease	20 % increase in the sum of the maximal diameters of measurable disease or the presence of new disease

RECIST

In the middle of the 1990s, the European Organization for Research and Treatment of Cancer (EORTC), the National Cancer Institute (NCI) of the USA, and the NCI of Canada trials group set up a task force to review existing response assessment criteria. Based upon a retrospective review of clinical trials involving approximate 4,000 patients, and considering advances that had been made, particularly in medical imaging technologies, a new set of guidelines for assessing the response of solid tumors to anticancer therapies was released in 2000 [4]. Known as RECIST, these recommendations included the adoption of a simplified approach to measuring tumors utilizing only the greatest diameter (unidimensional measurement) and the sum of the greatest diameters of the target lesions. RECIST defines the rules to select target lesions on baseline scans including the number (up to 10 per patient and 5 lesions per organ) and the size of target lesions ($\geq$10 mm). The establishment of these new criteria was guided by a number of important principles: (1) the need to maintain the standard four-response category system (i.e., CR, PR, SD, PD) (Table 8.1), with a size reduction of 30 % or more for PR and a size increase of 20 % or more for PD; (2) the goal of maintaining consistency of results such that no major discrepancy in the meaning of PR would exist between the older WHO criteria and the new RECIST criteria; (3) the recognition of both the arbitrary nature of the cutoff value for PR and the need to maintain this cutoff until other potentially more reliable or powerful surrogates could be developed; (4) concern about categorizing patients as PD too easily; and (5) recognition that cytostatic agents may not have the same measurement "activity" and that other serum markers and specific tumors may present unique challenges.

Assuming that a tumor is spherical and changes size symmetrically, a size reduction of 30 % defined by the unidimensional RECIST method corresponds to a size decrease of 50 % by the bidimensional WHO criteria. Considering calling for PD too soon, a size increase of 20 % proposed by the RECIST method is compatible with a size increase of 44 % by the WHO method (Table 8.2). When comparing clinical trials evaluated by RECIST with old studies, tumor progression may be detected later because of the increased threshold for PD. Since its establishment in 2000, unidimensional RECIST guideline has been widely accepted as the standard method for assessing tumor responses to systemic therapies.

RECIST 1.1

Continuous evaluating and updating of RECIST guidelines was suggested by the RECIST Working Group at the time the criteria were published. Based on an intensive analysis of data collected for more than 6,000 clinical trial patients [6] and the

Table 8.2 Relationship between changes in diameter, product, and volume

	Diameter, $2r$ (%)	Product, $(2r)^2$ (%)	Volume, $4/3\pi r^3$ (%)
Response	Decrease	Decrease	Decrease
	30	**50**	65
	50	75	87
Disease progression	Increase	Increase	Increase
	12	**25**	40
	20	44	73
	25	56	95
	30	69	120

Used with permission from Therasse et al. [4]
Numbers in bold font represent the RECIST (diameter) and WHO (product) criteria for change in tumor size to meet response and disease progression definitions

reported inadequacies of using the RECIST criteria in prospective clinical trials [7], a revised set of RECIST (version 1.1) was published in 2009 [5]. Major modifications in this new release included (1) a reduced number of target lesions from 10 to 5 per patient and from 5 to 2 per organ, (2) the need for response confirmation only in nonrandomized trials and only where response is the primary endpoint, (3) use of the short axis to measure malignant lymph nodes, and (4) requirements for a 20 % increase and for a minimum absolute increase of 5 mm in the sum of all target lesions' diameters for PD. The Working Group believed that it is not yet time to adopt volumetric and functional assessments (e.g., DCE MRI, DCE CT, or 18F-fluorodeoxyglucose (FDG) PET) because these techniques have not been standardized, are not widely available, and have not, through studies, received thorough clinical validation.

Limitations of Conventional Response Assessment Methods

Limitations of RECIST have been well described [4, 5]. First, changes in tumor maximal diameters measured on an axial plane between longitudinal imaging scans cannot fully capture changes in total tumor burden, especially along the z-axis (Fig. 8.1a, b). Second, the response cutoffs (e.g., 50 % or more reduction in the sum of tumor bidimensional measurements is considered to be a PR by the WHO) were developed by evaluating the measurement error of antiquated response assessment modalities used during the 1970s and early 1980s (i.e., physical palpation or plain X-ray measurements) [8, 9]. These cutoff values probably do not reflect variability in measuring tumor diameters using today's tumor measurement tools (e.g., an electronic ruler on a diagnostic workstation) [10, 11] on modern CT scans. Third, conventional response assessment methods disregard changes in tumor component as seen with tumor necrosis or tumor density decreases, a potential new dimension allowing evaluation of anticancer effects of antiangiogenic agents with anatomical imaging.

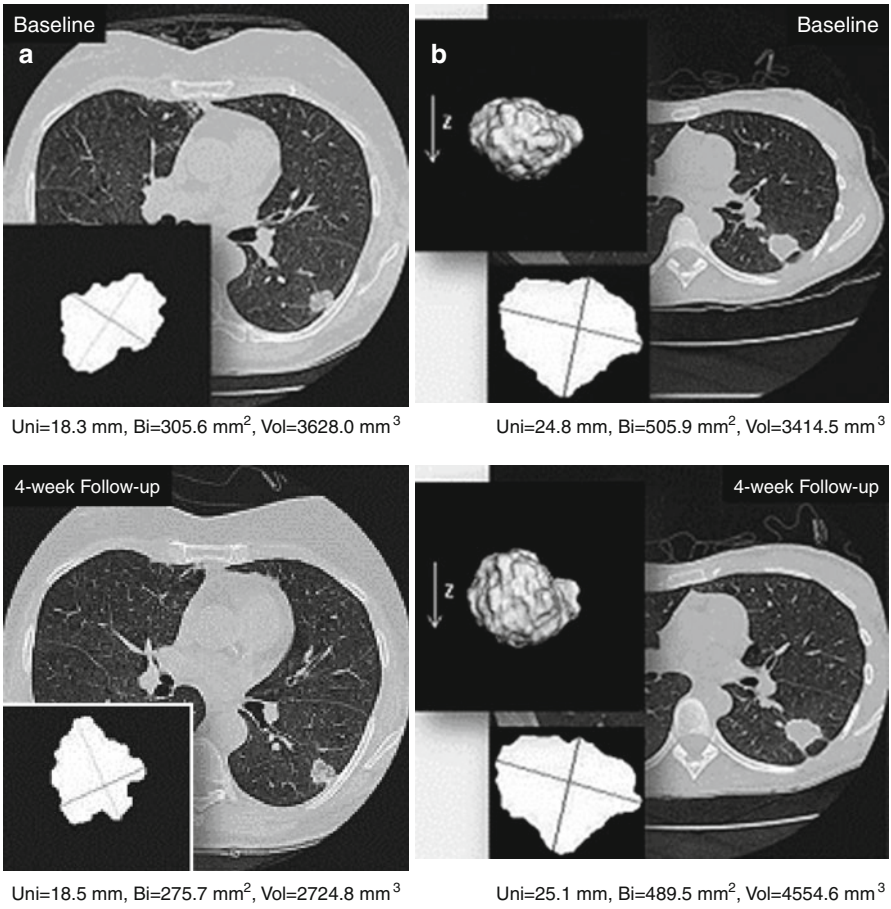


Fig. 8.1 Asymmetric growth of lung cancer. Two examples of NSCLC tumors (**a**, **b**) taken from a clinical trial testing gefitinib. (**a)** Line direction along which the greatest tumor diameter was measured on baseline and follow-up scan images changed. Percentage changes in unidimensional, bidimensional, and volume measurements were 1.1, 9.8, and 24.9 %, respectively. (**b**) Unidimensional and bidimensional measurements did not detect tumor change, but the volumetric technique did. Percentage changes in unidimensional, bidimensional, and volume measurements were 1.2, 3.2, and 33.4 %, respectively (**b**: Used with permission from Zhao et al. [43])

Revised and Modified RECIST and Beyond

Over the past decade, revised and modified RECIST guidelines have been suggested for certain types of tumors that do not lend themselves to unidimensional or bidimensional measurements due to their origin, extent, posttreatment residue, and density changes with targeted therapies. In the remaining sections of this chapter, we will discuss, for certain types of cancers, limitations of RECIST and its revisions and modifications that have been proposed and that are or will be evaluated.

Lymphoma

Lymphoma usually resides in the normal structure of the lymph nodes. Variations in the size of normal nodes can make selection of target lesions at baseline scans and determination of new lesions at follow-up scans both difficult and inconsistent. Furthermore, posttreatment residual masses often consist of non-tumor components such as fibrosis, necrosis, or inflammation that can be indistinguishable from tumors on CT and affect classification of CR and PR rates in clinical trials [12].

To resolve inconsistencies that can arise in lymphoma clinical trials, in 1999 an international working group (IWG) of lymphoma experts published a set of guidelines, based on their consensus, for the standardization of response assessment in adult patients with indolent and aggressive non-Hodgkin's lymphomas (NHL) [13]. The IWG criteria specifically defined a posttreatment size range of normal lymph nodes by taking into account the baseline size. For lymph nodes greater than 1.5 cm at baseline, a CR should be declared if these nodes have regressed to less than 1.5 cm after therapy. For nodes less than 1.5 cm but greater than 1.1 at baseline, the size of the normal nodes after therapy should be no larger than 1.0 cm in order to qualify as a CR. The IWG guidelines have provided clinicians with uniform criteria to interpret and assess outcomes of lymphoma clinical trials. However, these criteria cannot differentiate viable tumor components from necrosis or fibrosis [14]. With the increased availability of PET with ^{18}F-FDG radiotracers and the use of immunohistochemistry and flow cytometry, an International Harmonization Project (IHP) significantly revised the IWG criteria for lymphoma clinical trials [15]. The new IHP criteria evaluate all types of lymphomas and tumor responses to therapy by jointly considering tumor changes measured on both FDG PET and CT (Table 8.3). Additionally, the IHP criteria suggest post-therapy time intervals at which response should be assessed (i.e., after 3 weeks or more to evaluate the effects of chemotherapy and between 6 and 12 weeks to evaluate chemoimmunotherapy and radiation therapy).

Mesothelioma

Malignant pleural mesothelioma (MPM) typically grows as a rind of tumor encasing the lung in an irregular pattern (Fig. 8.2a). Measuring tumor size by replacing the "longest in-plane diameter" per RECIST can be problematic and will unlikely capture the change in tumor burden. Lack of reproducibility due to the circumferential and axial growing patterns of MPM is another major problem when using RECIST to assess tumor changes [16]. Because of these limitations and recent reports on the inadequacy of the RECIST criteria in the response assessment of MPM [17–19], current practice is to modify the RECIST criteria so that they can better capture the unique growth pattern of MPMs [20].

Table 8.3 IHP response criteria for lymphoma

Response	Definition	Nodal masses	Spleen, liver	Bone marrow
CR	Disappearance of all evidence of disease	(a) FDG-avid for PET positive prior to therapy; mass of any size permitted if PET negative (b) Variably FDG-avid or PET negative; regression to normal size on CT	Not palpable, nodules disappeared	Infiltrate cleared on repeat biopsy; if indeterminate by morphology, immunohisto-chemistry should be negative
PR	Regression of measurable disease and no new sites	≥50 % decrease in SPD of up to 6 largest dominant masses; no increase in size of other nodes (a) FDG-avid or PET positive prior to therapy; one or more PET positive at previously involved site (b) Variably FDG-avid or PET negative; regression on CT	≥50 % decrease in SPD of nodules (for single nodule in greatest transverse diameter); no increase in size of liver or spleen	Irrelevant if positive prior to therapy; cell type should be specified
SD	Failure to attain CR/PR or PD	(a) FDG-avid or PET positive prior to therapy; PET positive at prior sites of disease and no new sites on CT or PET (b) Variably FDG-avid or PET negative; no change in size of previous lesions on CT		
Relapsed disease or PD	Any new lesion or increase by ≥50 % of previously involved sites from nadir	Appearance of a new lesion(s) >1.5 cm in any axis, ≥50 % increase in longest diameter of a previously identified node >1 cm in short axis Lesions PET positive if FDG-avid lymphoma or PET positive prior to therapy	>50 % increase from nadir in the SPD of any previous lesions	New or recurrent involvement

Used with permission from Cheson et al. [15]

Abbreviations: *CR* complete remission, *FDG* [^{18}F]fluorodeoxglucose, *PET* positron emission tomography, *CT* computed tomography, *PR* partial remission, *SPD* sum of the product of the diameters, *SD* stable disease, *PD* progressive disease

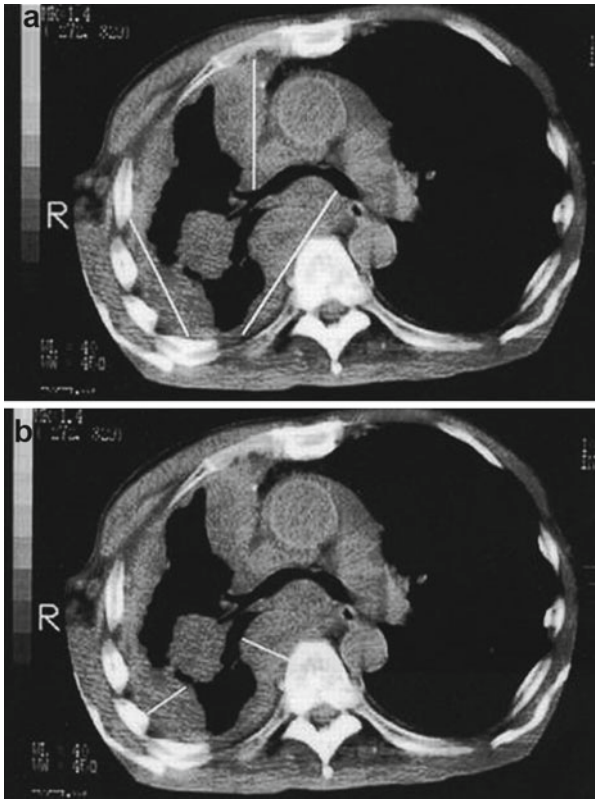


Fig. 8.2 Measuring pleural mesothelioma using conventional RECIST and modified RECIST. (**a**) *Lines* represent possible interpretations of "greatest tumor diameter" per conventional RECIST. (**b**) *Lines* represent suggested measurement sites that are perpendicular to fixed structures such as chest wall and vertebral column, according to the modified RECIST guidelines (Used with permission from Byrne and Nowak [20])

The modified RECIST criteria for MPM measure tumor thickness perpendicular to fixed anatomical structures such as the chest wall, mediastinum, or vertebral column at two sites for each of three separate levels on transverse CT planes (Fig. 8.2b). The unidimensional measurement is defined as the sum of the six tumor thicknesses, and the response evaluation follows the RECIST guidelines. In the modified RECIST criteria, the anatomical landmarks where measurements should be taken have been defined to improve measurement reproducibility on serial scans. Using the modified RECIST method to reevaluate two clinical trials, Byrne et al. found no change in response rate as assessed originally by the WHO criteria. However, the median survival for responding patients was significantly higher than that for nonresponding patients (15.1 month vs 8.9 month, $p = 0.03$) [20].

In the late 1990s, Pass and colleagues published a study showing that the preoperative tumor volume was representative of tumor T status in MPM and predictive of overall and progression-free survival as well as postoperative stage [21]. However, lack of automated or semiautomated volume quantification tools prevented further validation of these important findings. Recently, with the help of a computer algorithm, Fan et al. found a strong association between MPM patient survival and change in tumor volumes measured at two cycles after the onset of induction chemotherapy [22].

HCC and GIST

HCC is one of the most common malignancies worldwide and is the fastest growing cancer in the USA. GIST is a classic tumor model for the development, in modern drug discovery, of anticancer therapies that target specific molecules such as enzymes or receptors rather than killing cells. Many such therapeutic agents have been proposed for these two cancers. The sized-based RECIST method, especially with an arbitrary 30 % cutoff value defining a response (i.e., a 30 % or greater reduction in tumor size), has been shown to be misleading in the evaluation of tumor responses. Indeed, responding tumors may only minimally decrease in size or even slightly increase in size, but they may undergo internal necrosis and hemorrhage, hyalinization, and fibrosis [23–25]. Accurate and sensitive response assessment methods are thus imperative for the success of these clinical trials as well as for continued discovery of novel, target-specific, anticancer agents.

HCC: mRECIST for Locoregional Treatments

A variety of locoregional treatments have been developed for HCC in the past decade. However, such therapies are hard to evaluate by WHO and RECIST criteria because of the development of central necrosis, an outcome of all effective locoregional therapies.

In 2000, an expert panel on HCC organized by the European Association for the Study of the Liver (EASL) revised the response assessment for HCC by taking into account therapy-induced tumor hypodense areas and necrosis [26]. The concept of viable tumor, i.e., an enhanced tumor component in the arterial phase of dynamic CT or MRI, was then used to assess HCC responses to therapies and was soon accepted by the American Association for the Study of Liver Diseases (AASLD) [27]. To standardize a growing number of complex HCC clinical trials, a formal guideline, called the modified RECIST assessment (mRECIST), was subsequently established for HCC by the AASLD panel and was published in 2008 [28]. The mRECIST method emphasizes the importance of standardization of dynamic contrast-enhanced imaging techniques because the viable tumors are best depicted and measured on arterial-phase images. Instead of measuring an entire tumor to assess treatment response, the mRECIST method suggests that one should measure the longest diameter of only the viable tumor component in each tumor area (Fig. 8.3a, b) [29]. The four categories and the corresponding cutoff values for tumor response and progression used by the RECIST guidelines, however, remain unchanged.

GIST: Choi's Criteria for Targeted Therapies

A number of research groups reported significant underestimation of tumor responses by the RECIST method while monitoring GISTs treated with imatinib mesylate, a targeted therapy [30–32]. Although PET scanning has proven useful

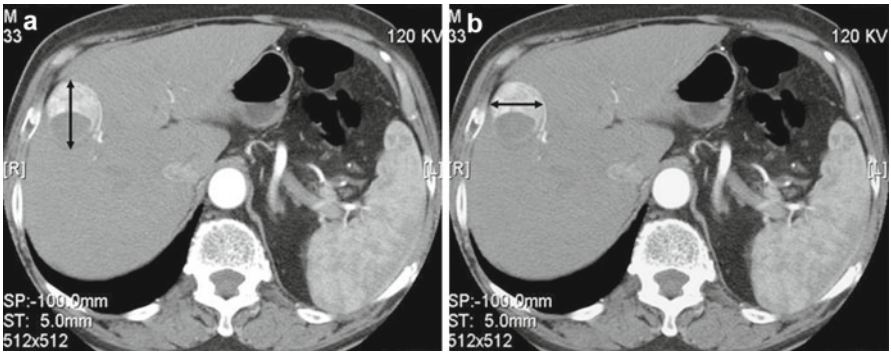


Fig. 8.3 Use of mRECIST criteria in the assessment of HCC responses to therapy. Target tumor response measurements on arterial-phase CT scans. (**a**) Measurement of greatest overall tumor diameter according to conventional RECIST criteria and (**b**) measurement of greatest viable tumor diameter according to the mRECIST method for HCC (Used with permission from Lencioni and LIovet [29])

for monitoring tumor responses in GIST patients, the high cost of PET, its lack of worldwide availability, and its lack of standardization led Choi and her colleagues to develop alternative CT criteria in evaluating responses for patients with GIST [32].

Choi's criteria define a 10 % decrease or more in a unidimensional measurement or a 15 % decrease or more in density (as measured by Hounsfield units) on a selected image plane as a partial response [32]. In a study of metastatic GIST patients treated with imatinib, the group reported that Choi's criteria reached a sensitivity of 97 % and a specificity of 100 % in identifying the responders assessed by PET, whereas the RECIST method only had a sensitivity of 52 %, though specificity was also 100 % [32].

HCC: Necrosis to Tumor Ratio

Necrotic lesions frequently develop a cyst-like appearance without a significant change in anatomic dimensions, when evaluated by CT or MRI. In a pivotal clinical trial of sorafenib in HCC, a partial response according to the WHO criteria occurred in only 2 % of subjects [33]. However, there was clear clinical benefit as 33.6 % of patients had stable disease (SD) for ≥16 weeks, and central tumor necrosis in response to sorafenib was common. Using baseline and follow-up triphasic CT scans, Abou-Alfa et al. then calculated lesion and necrosis volumes with the help of computer software. They found that the necrosis to tumor ratio (N/T) was significantly associated with responses, with responders having greater increases (in the ratio between necrosis volume and tumor volume) relative to baseline, as compared to nonresponders (Fig. 8.4a–d). The study did not show an association between the N/T ratio and overall survival [34].

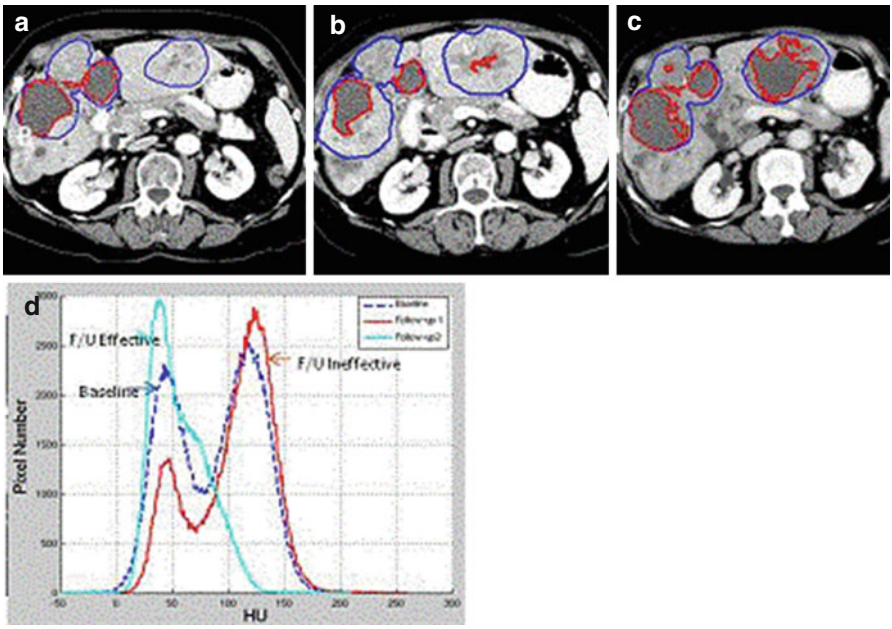


Fig. 8.4 Contrast-enhanced CT scans of an HCC patient enrolled in a growth factor inhibitor trial where target lesion tumors are outlined in *blue* and necrosis is outlined in *red*. (**a**) At baseline, the ratio of necrosis volume to tumor volume (N/T) was 34 %; (**b**) at 6 weeks post-therapy there was less necrosis and a decrease in the N/T ratio to 20 %. It was considered to be an ineffective treatment; (**c**) the patient was then switched to sorafenib, with an increase in the N/T ratio up to 47 %; (**d**) density histogram of the tumor at baseline (*blue*), at the first follow-up (*red*), and at the final time point (*turquoise*)

HCC: Choi's Criteria for Targeted Therapies

Studies on HCC treated with sunitinib reported considerably low response rates (<10 %) when evaluated by RECIST criteria, even though improved patient survival was observed [33, 35, 36]. Increases in hypodense regions rather than reductions in tumor size were often seen in the sunitinib-treated tumors. However, antitumor effects of antiangiogenic agents in HCC could not be captured by the RECIST criteria. For a recently published phase II clinical study of HCC patients treated with sunitinib, Faivre and colleagues reported a drastically increased response rate from 3.8 % (by RECIST criteria) to 65.4 % (by Choi's criteria). They also found that, using Choi's criteria, responding patients had a significantly longer time to progression than nonresponders (7.5 month vs 4.8 month, $p=0.0182$). However, no significant difference in median overall survival between the two groups was observed [37].

NSCLC

In 2004, somatic activating mutations in the tyrosine kinase (TK) domains of the epidermal growth factor receptor (*EGFR*) gene were discovered in a subset of non-small cell lung cancer (NSCLC) patients who showed remarkable responses to an

inhibitor of the TK within the *EGFR* [38–40]. This has opened the door to genotype-directed therapies for NSCLC. That is, treatment options can be selected based on an individual patient's clinical characteristics and tumor biology, and novel therapies are developed to target key genetic mutations.

The rapid increase in the number of systemic agents in NSCLC has demanded novel biomarkers that better determine drug-induced tumor changes. Such biomarkers should be indicative of underlying biologic processes in the tumor and be of high precision so that they can serve as a valuable means for screening of promising anticancer agents. One potential technique of particular interest is volumetric CT, a technology with the potential to more accurately capture tumor growth dynamics. In a phase II neoadjuvant NSCLC trial [41], Zhao and colleagues incorporated an analysis of thin-section volumetric CT scans to determine the value of early radiographic changes (i.e., 3 weeks post-therapy) in predicting the biologic activity of gefitinib therapy in a subset of *EGFR*-enriched NSCLC patients (Fig. 8.5a–c). The authors found that, compared to unidimensional measurements, volume measurements allow significantly better dichotomization of these molecular subtypes, indicating that volume change has promise as an investigational method for early detection of the biologic activity of a systemic therapy in NSCLC [42].

Summary

Insensitive methods in evaluating patient responses to cancer treatments can delay drug discovery and mislead those doing patient management. The conventional RECIST guideline has its pitfalls when applied to the solid tumors whose origin, morphology, and extent are not suitable for linear measurements (e.g., lymphoma and mesothelioma) or when used to assess treatments that may not necessarily reduce tumor size (e.g., interventional and targeted therapies in HCC). In the past decade, revisions and modifications of the RECIST criteria were proposed for certain types of cancers treated with these novel therapies, aiming at better assessing drug efficacy. Yet, the full potential of modern CT and computerized image analysis in accurately assessing tumor changes over time, and thus optimally interpreting tumor responses to therapies, has not been well explored.

Even though the RECIST criteria are considered as the standard guidance for evaluation of modern clinical trials, its response cutoff values have not been validated against state-of-the-art CT and advanced tumor measurement tools. In a recent contemporary study, Zhao and her colleagues revealed, for the first time, the magnitudes of the variability in tumor unidimensional, bidimensional, and volumetric measurements made on two repeat CT scans performed within 15 min in patients with non-small cell lung cancer (NSCLC) [10]. The high reproducibility of both radiologists' and computer-aided tumor size measurements suggests that a thorough reevaluation of conventional RECIST criteria should be done for assessing novel targeted therapies.

There is no doubt that volume measurements are more accurate in quantifying changes in tumor burden than current diameter measurements. Changes in necrosis to tumor ratio and/or in tumor density can be quantified by computer-aided algorithms. However, to be accepted by the oncology community as better imaging

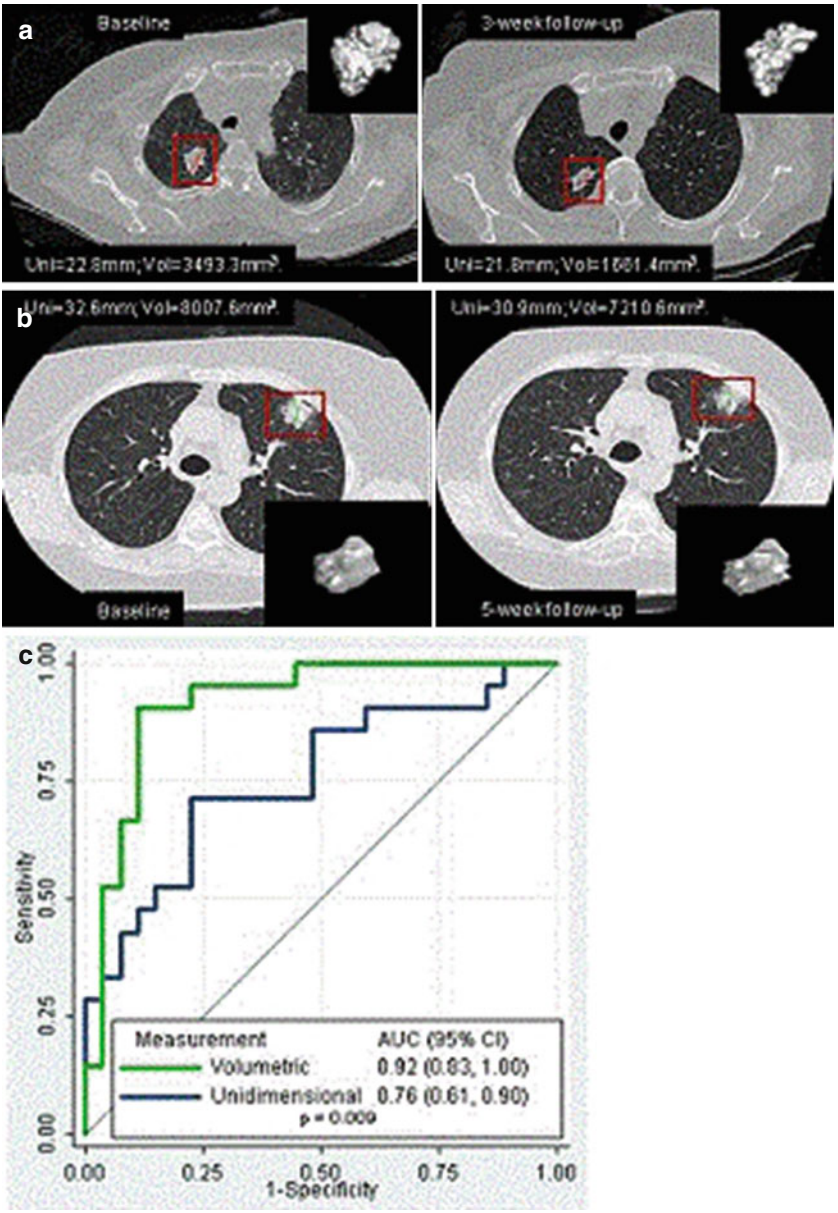


Fig. 8.5 Volumetric CT technique in tissue biomarker discovery. (**a**) Shows an *EGFR* mutant NSCLC (likely response to gefitinib therapy). Changes in unidimensional and volume measurements were −4.4 and −52.4 %, respectively. The volumetric technique detected tumor change at 3 weeks post-therapy, whereas the diameter technique did not. (**b**) Shows a KRAS mutant NSCLC (resistant to gefitinib therapy). Changes in unidimensional and volume measurements were −5.2 and −10.0 %, respectively. Both measurement techniques showed no real tumor change at 3 weeks post-therapy. (**c**) Receiver operating characteristic curve (ROC) shows a significantly higher area under the curve (AUC) of the volumetric measurement than unidimensional measurement, indicating that early volumetric change is a better metric for predicting *EGFR* mutation status than is early diameter change (**a**, **c**: Used with permission from Zhao et al. [42])

biomarkers of tumor responses, these new quantitative metrics along with the proposed (or to be proposed) response criteria need to be validated in prospective multicenter clinical trials. The new metrics should be reproducible, and tumor responses to therapies assessed by any new criteria should be correlated with clinical outcomes (e.g., survival).

There is an ever-improving understanding of tumor biology and of the underlying mechanisms of interactions between drugs and cancers. There are also advances in medical imaging technologies and computerized image analysis methods. Therefore, optimal strategies to monitor tumor responses to novel therapies at anatomical, functional, or molecular levels should be developed by jointly considering the best possible imaging modalities, by standardizing imaging acquisition techniques, and by developing advanced response assessment methods.

References

1. Schwartz LH and Zhao B. Response assessment in Oncology. In: imaging in oncology, 3rd ed. Ed(s) Husband JE, Reznek RH. Informa Healthcare; 2010.
2. World Health Organization. WHO handbook for reporting results for cancer treatment. Geneva: World Health Organization; 1979.
3. Miller AB, Hoogstraten B, Staquet M, Winkler A. Reporting results of cancer treatment. Cancer. 1981;47:207–14.
4. Therasse P, Arbuck SG, Eisenhauer EA, et al. New guidelines to evaluate the response to treatment in solid tumors. J Natl Cancer Inst. 2000;92:205–16.
5. Eisenhauer EA, Therasse P, Bogaerts J, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45:228–47.
6. Bogaerts J, Ford R, Sargent D, et al. Individual patient data analysis to assess modifications to the RECIST criteria. Eur J Cancer. 2009;45:248–60.
7. Therasse P, Eisenhauer EA, Verweij J. RECIST revisited: a review of validation studies on tumour assessment. Eur J Cancer. 2006;42:1031–9.
8. Morel CG, Hanley JA. The effect of measuring error on the results of therapeutic trials in advanced cancer. Cancer. 1976;38:388–94.
9. Warr D, Mckinney S, Tannock I. Influence of measurement error on assessment of response to anticancer chemotherapy: proposal for new criteria of tumor response. J Clin Oncol. 1984; 2:1040–6.
10. Zhao B, James LP, Moskowitz C, et al. Evaluating variability in tumor measurements from same-day repeat CT scans in patients with non-small cell lung cancer. Radiology. 2009; 252:263–72.
11. Oxnard GR, Zhao B, Sima CS, et al. Variability of lung tumor measurements on repeat computed tomography (CT) scans taken within 15 minutes: implications for care and clinical research. J Clin Oncol. 2011;10:3114–9.
12. Grillo-Lopez AJ, Cheson BD, Horning SJ, et al. Response criteria for NHL: importance of "normal" lymph node size and correlations with response rates. Ann Oncol. 2000;11:399–408.
13. Cheson BD, Horning SJ, Coiffier B, et al. Report of an international workshop to standardize response criteria for non-Hodgkin's lymphomas. J Clin Oncol. 1999;17:1244–53.
14. Juweid ME, Wiseman GA, Vose JM, et al. Response assessment of aggressive non-Hodgkin's lymphoma by integrated international workshop criteria and fluorine-18-fluorodeoxyglucose positron emission tomography. J Clin Oncol. 2005;23:4652–61.
15. Cheson BD, Pfistner B, Juweid ME, et al. Revised response criteria for malignant lymphoma. J Clin Oncol. 2007;25:579–86.
16. Oxnard GR, Armato Iii SG, Kindler HL. Modeling of mesothelioma growth demonstrates weaknesses of current response criteria. Lung Cancer. 2006;52:141–8.

17. Monetti F, Casanova S, Grasso A, et al. Inadequacy of the new response evaluation criteria in solid tumors (RECIST) in patients with malignant pleural mesothelioma: report of four cases. Lung Cancer. 2004;43:71–4.
18. Ceresoli GL, Chiti A, Zucali PA, et al. Assessment of tumor response in malignant pleural mesothelioma. Cancer Treat Rev. 2007;33:533–41.
19. van Klaveren RJ, Aerts JGJV, de Bruin H, et al. Inadequacy of the RECIST criteria for response evaluation in patients with malignant pleural mesothelioma. Lung Cancer. 2004;43:63–9.
20. Byrne MJ, Nowak AK. Modified RECIST criteria for assessment of response in malignant pleural mesothelioma. Ann Oncol. 2004;15:257–60.
21. Pass HI, Temeck BK, Kranda K, Steinberg SM, Feurestein IR. Preoperative tumor volume is associated with outcome in malignant pleural mesothelioma. J Thorac Cardiovasc Surg. 1998;115:310–8.
22. Liu F, Zhao B, Krug LM, et al. Assessment of therapy responses and prediction of survival in malignant pleural mesothelioma through computer-aided volumetric measurement on CT scans. J Thorac Oncol. 2010;5:879–84.
23. Wang J, Chen LT, Tsang YM, et al. Dynamic contrast-enhanced MRI analysis of perfusion changes in advanced hepatocellular carcinoma treated with an antiangiogenic agent: a preliminary study. Am J Roentgenol. 2004;183:713–9.
24. Choi H. Critical issues in response evaluation on computed tomography:Lessions from the gastrointestinal stromal tumor model. Curr Oncol Rep. 2005;7:307–11.
25. Hong X, Choi H, Loyer EM, et al. Gastrointestinal stromal tumor: role of CT in diagnosis and in response evaluation and surveillance after treatment with imatinib. Radiographics. 2006;26:481–95.
26. Bruix J, Sherman M, Llovet JM, et al. EASL panel of experts on HCC European association for the study of the liver. Clinical management of hepatocellular carcinoma. Conclusions of the barcelona-2000 EASL conference. J Hepatol. 2001;35:421–30.
27. Bruix J, Sherman M, Practice Guidelines Committee, American Association for the Study of Liver Diseases. Management of hepatocellular carcinoma. Hepatology. 2005;42:1208–36.
28. Llovet JM, Di Bisceglie AM, Bruix J, et al. Panel of experts in HCC-design clinical trials. Design and endpoints of clinical trials in hepatocellular carcinoma. J Natl Cancer Inst. 2008;100:698–711.
29. Lencioni R, LIovet JM. Modified RECIST (mRECIST) assessment for hepatocellular carcinoma. Semin Liver Dis. 2010;30:52–60.
30. Stroobants S, Goeminne J, Seeger M, et al. 18FDG-positron emission tomography for the early prediction of response in advanced soft tissue sarcoma treated with imatinib mesylate. Eur J Cancer. 2003;39:2012–120.
31. Antoch G, Kanja J, Bauer S, et al. Comparison of PET, CT, and dual-modality PET/CT imaging for monitoring of imatinib (STI-571) therapy in patients with gastrointestinal stromal tumors. J Nucl Med. 2004;45:357–65.
32. Choi HA, Charnsangavej C, Faria SC, et al. Correlation of computed tomography and positron emission tomography in patients with metastatic gastrointestinal stromal tumors treated at a single institution with imatinib mesylate: proposal of new computed tomography response criteria. J Clin Oncol. 2007;25:1753–9.
33. Abou-Alf GK, Schwartz LS, Ricci S, et al. Phase II study of sorafenib in patients with advanced hepatocellular carcinoma. J Clin Orthod. 2006;24:4293–300.
34. Abou-Alfa GK, Zhao B, Capanu M, et al. Tumor necrosis as a correlate for response in subgroup of patients with advanced hepatocellular carcinoma (HCC) treated with sorafenib. Stockholm: ESMO; 2008.
35. Cheng AL, Kang YK, Chen Z, et al. Efficacy and safety of sorafenib in patients in the Asia-Pacific region with advanced hepatocellular carcinoma: a phase III randomized, doube-blind, placebo-controlled trial. Lancet Oncol. 2009;10:25–34.
36. Llovet JM, Ricci S, Mazzaferro V, et al. Sorafenib in advanced hepatocellular carcinoma. N Engl J Med. 2008;359:378–90.

37. Faivre S, Zappa M, Vilgrain V, et al. Changes in tumor density in patients with advanced hepatocellular carcinoma treated with sunitinib. Clin Cancer Res. 2011;17:4506–12.
38. Lynch TJ, Bell DW, Sordella R, et al. Activating mutations in the epidermal growth factor receptor underlying responsiveness of non–small-cell lung cancer to gefitinib. N Engl J Med. 2004;350:2129–39.
39. Paez JG, Jänne PA, Lee JC, et al. EGFR mutations in lung cancer: correlation with clinical response to gefitinib therapy. Science. 2004;304:1497–500.
40. Pao W, Miller V, Zakowski M, et al. EGF receptor gene mutations are common in lung cancers from "never smokers" and are associated with sensitivity of tumors to gefitinib and erlotinib. Proc Natl Acad Sci U S A. 2004;101:13306–11.
41. Kris MG, Pao W, Zakowski M, et al. Prospective trial with preoperative gefitinib to correlate lung cancer response with EGFR exon 19 and 21 mutations and to select patients for adjuvant therapy. Proc Am Soc Clin Oncol. 2006;24:369s.
42. Zhao B, Oxnard GR, Moskowitz CS, et al. A pilot study of volume measurement as a method of tumor response evaluation to aid biomarker development. Clin Cancer Res. 2010;16:4647–53.
43. Zhao B, Schwartz LH, Moskowitz CS, et al. Lung cancer: computerized quantification of tumor response—initial results. Radiology. 2006;241(3):892–8.

Chapter 9
Cardiac Imaging in Clinical Trials

George P. Heyrich and Joel Krasnow

Abstract Cardiovascular disease is highly prevalent in developed countries. The development of pharmaceuticals and medical devices for cardiovascular disease requires knowledge of the relevant imaging technologies and their relative strengths and weaknesses. Cardiovascular imaging can be approached from either an anatomic or a physiological perspective. This chapter is organized by disease process with discussion of the current and evolving imaging standards.

Keywords SPECT • Cardiac PET • CT angiography • Echocardiography • Intravascular ultrasound

Introduction

Cardiovascular disease is highly prevalent in developed countries. It is associated with the rapidly growing epidemics of obesity, tobacco use, and diabetes mellitus and is a major cause of morbidity, mortality, and healthcare expenditures. Cardiac imaging modalities play an essential role in the accurate diagnosis of cardiovascular disease. Their role is also essential in monitoring the effectiveness of medical treatment and disease progression. Cardiovascular imaging is responsible for a large percentage of healthcare costs of overall imaging budgets.

G.P. Heyrich, MD, FACC, FSCAI (✉)
Mercer Bucks Cardiology, 104 Pheasant Run, Suite 128,
Newtown, PA 18940, USA
e-mail: meptca@aol.com

J. Krasnow, MD, MBA
Department of Strategy and Regulatory Affairs, TherapeuticsMD,
Boca Raton, FL, USA
e-mail: jkras@optonline.net

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_9, © Springer-Verlag London 2014

Table 9.1 Key attributes of cardiovascular imaging modalities

Imaging modality	Radiation dose	Key measurements	Common applications
Echocardiography (2D, 3D, exercise stress, dobutamine stress, transesophageal)	None	Wall thickness, cardiac chamber size in systole or diastole, valvular anatomy and function, intracardiac hemodynamics	Serial assessments in heart disease. Assessment of pharmacologic interventions
Carotid ultrasound	None	Intima-media thickness	Carotid atherosclerosis
Intravascular ultrasound	None	Plaque volume and characteristics within the vessel wall and degree of luminal stenosis	Coronary arteries and peripheral vascular arteries
Optical coherence tomography	None	Can differentiate between lipid, calcium, and fibrosis	Plaque characterization, vessel sizing
Cardiac CT angiography	Moderate	Still images of the heart and blood vessels, coronary calcium content	Coronary artery disease
Cardiac MRI	None	Structure and function of the myocardium	Cardiac function
Coronary angiography	Mild	Anatomy of the coronary vessels	Coronary artery disease, ventricular function, valvular assessment
SPECT	Moderate	Perfusion defects	Cardiac ischemia and function

There are several imaging modalities that are commonly used in the diagnosis and management of cardiovascular disorders. These are listed in Table 9.1 along with emerging technologies. Cardiovascular trials may have either anatomic or physiological endpoints. Anatomic endpoints include vessel patency or degree of stenosis, carotid intimal media thickness, amount of valvular stenosis, and quantitative/qualitative assessment of left ventricular function. Physiological endpoints include stroke volume, cardiac output, the degree of regurgitation/stenosis severity of the mitral or aortic valve, and cardiovascular hemodynamics.

Ischemic Heart Disease

Athlerosclerotic plaques begin to appear in the vasculature during childhood and progressively increase over time. A diagnosis of ischemic heart disease (in the absence of congenital cardiac anomalies) is made when the atherosclerotic process has advanced to the point of causing ischemic symptoms either at rest or under stress conditions. The diagnosis of coronary artery disease may also be made anatomically by demonstrating vascular stenosis >50 %. Thus, the significance of an individual's heart disease is defined by its quantity and location.

The diagnosis of ischemic heart disease or coronary artery disease may involve several imaging modalities. The initial approach to patients with suspected ischemic heart disease is to perform myocardial stress testing. Therefore, we will begin with a discussion of imaging techniques commonly employed in conjunction with stress testing.

Nuclear Medicine

Nuclear medicine encompasses both diagnostic imaging and treatment of disease and may also be referred to as molecular medicine or molecular imaging and therapeutics. Nuclear medicine uses certain properties of isotopes and the energetic particles emitted from radioactive material to diagnose or treat various pathology. Different from the typical concept of anatomic radiology, nuclear medicine enables assessment of physiology. This function-based approach to medical evaluation has useful applications in most subspecialties, notably oncology, neurology, and cardiology. Gamma cameras are used in, for example, scintigraphy, SPECT, and PET to detect regions of biologic activity that may be associated with disease. A relatively short-lived isotope, such as sestamibi, is administered to the patient. Isotopes are often preferentially absorbed by biologically active tissue in the body and can be used to identify myocardial perfusion and function. Images are acquired after collimated photons are detected by a crystal that gives off a light signal, which is in turn amplified and converted into count data.

Scintigraphy (*scint*) is a form of diagnostic test wherein radioisotopes are taken internally, for example, intravenously or orally. Then, gamma camera captures and forms two-dimensional images from the radiation emitted by the radiopharmaceuticals.

SPECT is a 3D tomographic technique that uses gamma camera data from many projections and can be reconstructed in different planes. A dual-detector head gamma camera combined with a CT scanner, which provides localization of functional SPECT data, is termed a SPECT/CT camera and has shown utility in advancing the field of molecular imaging.

Positron emission tomography (*PET*) uses coincidence detection to image functional processes. Short-lived positron emitting isotope, such as ^{18}F, is incorporated with an organic substance such as glucose, creating F18-fluorodeoxyglucose, which can be used as a marker of metabolic utilization. This is more fully described in Chap. 1. Utilization of PET imaging allows a comparison of myocardial functional status with perfusion imaging. Areas of underperfused but functionally viable myocardium can be noninvasively identified to determine patient eligibility for invasive percutaneous or surgical revascularization strategies.

PET images can be viewed in comparison to computed tomography scans to determine an anatomic correlate. Modern scanners combine PET with a CT, or even MRI, to optimize the image reconstruction involved with positron imaging. This is

performed on the same equipment without physically moving the patient off of the gantry. The resultant hybrid of functional and anatomic imaging information is a useful tool in noninvasive diagnosis and patient management.

Myocardial Stress Testing

Most exercise stress testing involves the BRUCE protocol where an individual starts exercising on a treadmill and is progressed through increasing levels of difficulty that elicit increases in heart rate due to the increased oxygen demands. In patients who are unable to exercise, pharmacologic stress can be induced with intravenous coronary vasodilating agents including adenosine and Lexiscan. When these oxygen demands are not met, there may be associated ST segment changes on the ECG, patient symptoms, or areas of underperfusion that can be detected through imaging. Monovalent cations are known to be transported into myocardial cells. Therefore, cations such as potassium, thallium, rubidium, and ammonia have been used with single-photon emission computed tomography (SPECT) imaging. Superior imaging quality has been established with technetium-99m-labeled isonitriles (sestamibi and tetrofosmin) which are the current standard of tracer for SPECT. These isotopes are injected intravenously at baseline and again at the time of peak exercise. See Fig. 9.1 for an example image using technetium-99m. They are rapidly taken up into the myocardium. Areas where the tracer is not taken up represent either abnormal regional blood flow or previous infarction. By comparing the baseline scan to the subsequent scan at peak load, one can distinguish between the two. Stenotic arteries are not able to vasodilate to the same degree as normal vessels which lead to relative differences in myocardial perfusion.

A normal exercise SPECT is associated with a rate of MI or cardiac death of 0.65 % annually, while an abnormal test is associated with a 4.30 % annual rate. Normal pharmacologic nuclear myocardial perfusion scans are associated with a 1.78 % annual risk of MI or cardiac death, while an abnormal test yields a 9.98 % annual rate [1]. These differences are attributed to greater comorbidities and the associated increased atherosclerotic disease burden in the pharmacologically diagnosed group. While the rates of MI/cardiac death with a negative SPECT are relatively low, due to the large numbers of these tests being performed with approximately ten million being performed annually in the United States, the event rate per million negative procedures is 6,500 for exercise SPECT and 17,800 for pharmacologic SPECT. One reason for this is that nonobstructive plaques vulnerable to rupture will not be detected.

If one is interested in performing a clinical trial for patients at risk for cardiac events due to coronary artery disease, is a positive exercise or pharmacologic SPECT a useful inclusion criterion? A meta-analysis comparing SPECT to coronary angiography where coronary artery stenosis >50 % was used to define CAD, the sensitivity, specificity, and diagnostic accuracy for SPECT was 82, 76, and 83 %, respectively [2]. PET yielded the most favorable results with sensitivity of 91 %, specificity of 89 %, and diagnostic accuracy of 89 % making it a preferred noninvasive test for the diagnosis of coronary artery disease.

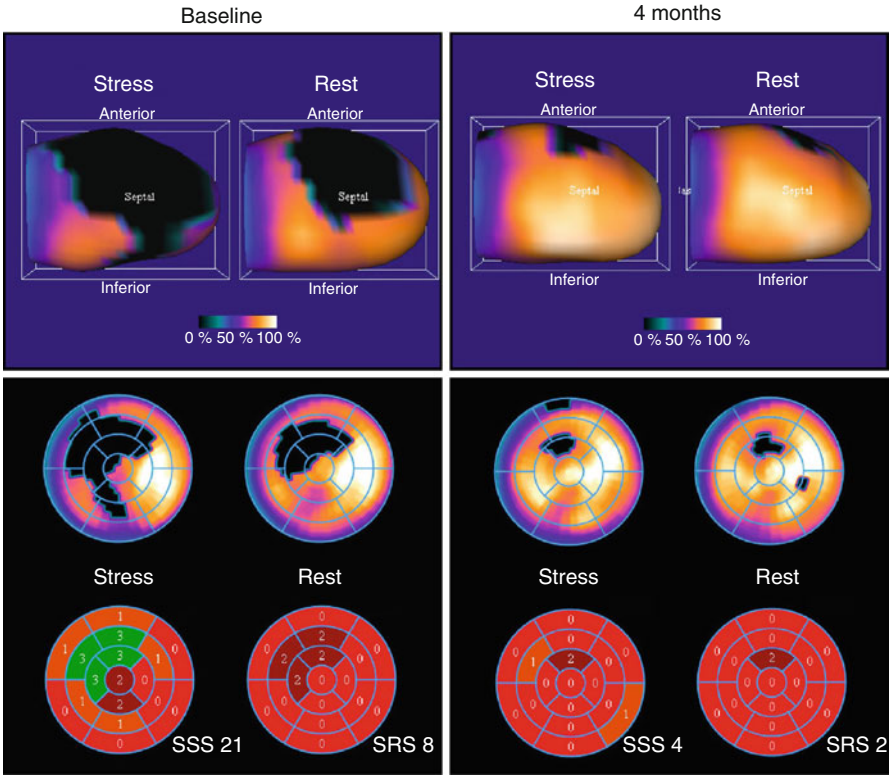


Fig. 9.1 Cardiac SPECT imaging. Technetium-99m sestamibi single-photon emission computed tomography (SPECT) scintigraphy looking at myocardial perfusion. Patient with acute anterior MI presented an important perfusion defect of the anteroseptal and apical territories (*top left*). The bull's-eye image (*bottom left*) combining the 17 segments revealed an important ischemic territory with summed stress score (*SSS*) of 21 and a summed rest score (*SRS*) of 8. The summed difference (SDS = SSS–SRS) was 13 (not shown), indicating a significant redistribution (reversible perfusion defect). Four months later (*rights*), perfusion defect and ischemia were significantly reduced as revealed by stress and rest SPECT (Used with kind permission of Springer Science + Business Media from Mansour et al. [6])

Cardiac Positron Emission Tomography (PET)

Cardiac positron emission tomography (PET) is a form of diagnostic imaging in which patients are evaluated using a PET scanner after intravenous injection with a radioisotope. Although several isotopes have been used for cardiac PET imaging, the most widely employed in clinical practice are rubidium-82 and nitrogen-13 ammonia.

The requirements to perform cardiac PET imaging include:

- Facility: taking into consideration clinical workflow, as well as regulatory requirements such as requisite shielding from radiation exposure.
- Capital equipment: PET or PET/CT scanner.

- Radiopharmaceutical: rubidium-82 generator system or close access to cyclotron produced isotopes such as nitrogen-13 ammonia.
- Personnel: including specialty trained physician, radiation safety, physics, nursing, and technologist support.
- Operations: stress test monitoring, as well as emergency response equipment, processing and review workstations, and administrative and support personnel are additional considerations.

This form of diagnostic imaging has traditionally been perceived as cost prohibitive in comparison to general nuclear medicine cardiac stress testing using single-photon emission computed tomography (SPECT). However, due to significant gains in access to scanners, related to the widely accepted role of PET/CT in clinical oncology, cardiac PET is likely to become more widely available, particularly given various clinical and technical advantages that might make this a potential test of choice in the diagnosis of coronary artery/heart disease. An example set of images are shown in Fig. 9.2.

Computed Tomography Angiography (CTA)

CT combines the use of X-rays with computerized analysis of the images. Beams of X-rays are passed from a rotating device through the area of interest in the patient's body from several different angles to obtain projection images, which then are assembled by computer into a three-dimensional picture of the area being studied. See Chap. 1 for more details. Contrast material is injected in order to better visualize the coronary vessels. With the advent of subsecond rotation combined with multislice CT (up to 128-slice), high resolution and high speed can be obtained at the same time, allowing excellent imaging of the coronary arteries (cardiac CT angiography). Images with an even higher temporal resolution can be formed using retrospective ECG gating. In this technique, each portion of the heart is imaged more than once while an ECG trace is recorded. The ECG is then used to correlate the CT data with their corresponding phases of cardiac contraction. Once this correlation is complete, all data that were recorded while the heart was in motion (systole) can be ignored, and images can be made from the remaining data that happened to be acquired while the heart was at rest (diastole). In this way, individual frames in a cardiac CT investigation have a better temporal resolution than the shortest tube rotation time.

Because the heart is effectively imaged more than once (as described previously), cardiac CT angiography results in a relatively high radiation exposure around 12 millisieverts (mSV). Newer prospective gating techniques and imaging protocols in combination with the development of larger imaging scanner 128-slice and higher are allowing significant reduction in radiation exposure. Methods are available to decrease this exposure, however, such as prospectively decreasing radiation output based on the concurrently acquired ECG (also known as tube current modulation). This can result in a significant decrease in radiation exposure, at the risk

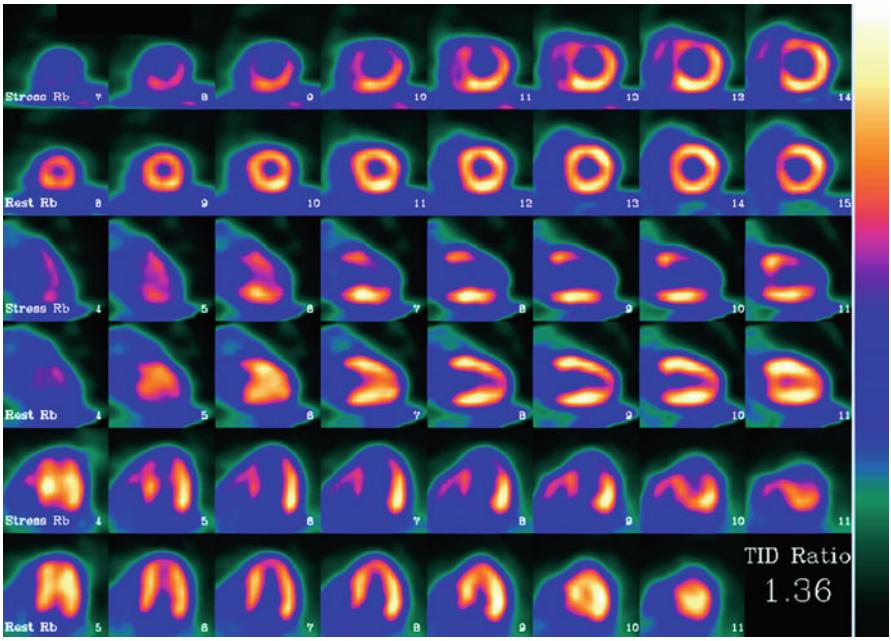


Fig. 9.2 Cardiac positron emission tomography. Abnormal rubidium-82 PET myocardial perfusion scan at stress and rest. Short axis (*top rows*), vertical long axis (*middle rows*), and horizontal long axis (*bottom rows*) slices demonstrate severe and extensive ischemia involving the anterior, apical, and septal walls (Used with kind permission of Springer Science + Business Media from Fox and Strauss [7])

of compromising image quality if there is any arrhythmia during the acquisition. The significance of radiation doses in the diagnostic imaging range has not been proven, although the possibility of inducing an increased cancer risk across a population is a source of significant concern. This potential risk must be weighed against the competing risk of not performing a test and potentially not diagnosing a significant health problem such as coronary artery disease.

Currently, it appears that the greatest utility of cardiac CT lies in excluding significant coronary artery disease rather than confirming its presence. This is because the test has a high sensitivity (greater than 90 %) and thus a negative test result means that a patient is very unlikely to have coronary artery disease and can be worked up for other causes of their chest symptoms. This is termed a high negative predictive value. A positive result is less conclusive and often will be confirmed (and possibly treated) with subsequent invasive angiography. The positive predictive value of cardiac CTA is estimated at approximately 82 % and the negative predictive value is around 93 %. A meta-analysis at the patient level comparing computed tomography angiography to invasive coronary angiography in the diagnosis of arterial stenosis between years 2006 and 2009 determined a sensitivity of 98.2 % and

specificity of 81.6 %. The median positive predictive value was 90.5 % and negative predicted value was 99.0 % [3]. A retrospective study compared males and females at low (defined as risk <30 %) and intermediate (defined as risk of 30–90 %) risk of CAD assessed CTA against invasive coronary angiography. Women at low and intermediate risk had a sensitivity of 97 and 99 %, a specificity of 79 and 72 %, a positive predictive value of 80 and 83 %, and a negative predictive value of 97 and 98 %, respectively. Values for men were comparable with sensitivity of 100 and 99 % and negative predictive values of 100 and 99 %, respectively [4]. Thus, a negative CTA is a very good noninvasive manner to detected coronary artery disease. The diagnostic accuracy of the modality is now accepted and the field has shifted towards reducing the radiation dose associated with the procedure.

Dual-source CT scanners, introduced in 2005, allow higher temporal resolution by acquiring a full CT slice in only half a rotation, thus reducing motion blurring at high heart rates and potentially allowing for shorter breath-hold time. This is particularly useful for ill patients who have difficulty holding their breath or who are unable to take heart-rate-lowering medication.

The speed advantages of 64-slice MSCT have rapidly established it as the minimum standard for newly installed CT scanners intended for cardiac scanning. Manufacturers are now actively developing 256-slice and true "volumetric" scanners, primarily for their improved cardiac scanning performance.

The latest MSCT scanners acquire images only at 70–80 % of the R-R interval (late diastole). This prospective gating can reduce effective dose from 10 to 15 mSv to as little as 1.2 mSv in follow-up patients acquiring at 75 % of the R-R interval. Effective doses at a center with well-trained staff doing coronary imaging can average less than the doses for conventional coronary angiography.

CT Angiography of Vascular Structures

CT angiography remains the mainstay in imaging of the thoracic and abdominal aorta. This technique provides an accurate measurement of segmental regions of the aorta. Three-dimensional reconstruction can also be performed with automated software. Image analysis allows precise measurement for procedure planning, device measurement for interventional placement, and endovascular graft placement for thoracic and abdominal aneurysm.

Serial CT angiograms of the aorta allow an effective method for monitoring progression of disease targeting timing of interventional therapies. In addition, serial studies are included in protocols to monitor the short- and long-term effectiveness of therapies and post-procedural complications. Examples include stent graft endoleaks and graft migration.

In addition, quantitative analysis of vascular plaque presence and its extent have been utilized as surrogate endpoints in the effectiveness of pharmacotherapy in clinical trials.

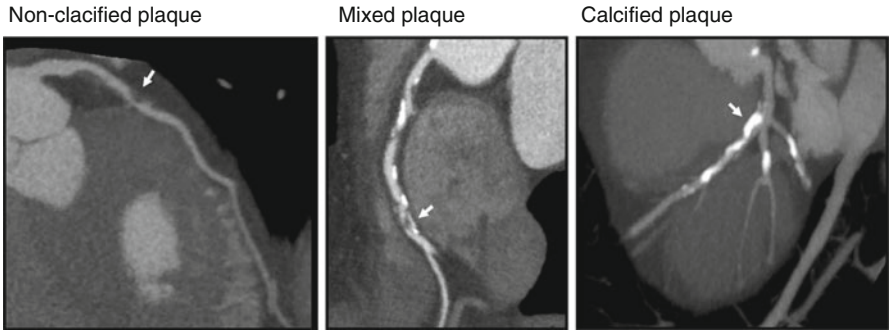


Fig. 9.3 Computed tomography angiography. Plaque composition assessed with CTA. Curved multi-planar reconstructions showing three distinct plaque characteristics observed on CTA with noncalcified plaque (*arrow, left panel*), mixed plaque (*arrow, middle panel*), and calcified plaque (*arrow, right panel*) (Used with kind permission of Springer Science + Business Media from van Werkhoven et al. [8])

Prognostic Implications of an Abnormal Coronary CT Angiogram (CCTA)

The clinical outcomes published to date for coronary CT angiogram (CCTA) are comparable in trial size, patient population, and median follow-up to those with SPECT. A CCTA indicating no coronary artery disease is associated with a 0.15 % annual risk of overall mortality. In the same study there were no deaths related to coronary artery disease during the follow-up period [5] (The annualized mortality rate for nonobstructive CAD was approximately 1 % and for obstructive CAD was approximately 4 %.) Thus, CCTA enables one to predict three distinct levels of risk and importantly identifies a population that is at very low risk of a cardiac event. Thus, for the purpose of identifying individuals at risk for a first MI, CCTA is the imaging modality to identify the burden of coronary artery disease. See Fig. 9.3 for plaque identification with CCTA.

Coronary Artery Calcium (CAC) Score

Coronary artery calcium (CAC) scoring provides a non-contrast CT scan which quantifies the extent of calcification present in the coronary circulation. Calcification is a surrogate marker for atherosclerotic plaque deposition in the coronary artery. The extent and severity of coronary calcification can be localized to an individual coronary artery distribution. The summary score of coronary calcification can then be age adjusted to the overall population to further risk stratify the individual's risk of having significant obstructive coronary artery disease. CAC does not identify the degree of stenosis present in a coronary artery, but rather provides a method of

screening for its presence. Identification of high-risk patients using CAC can provide an effective screening mechanism for either including or excluding patients from clinical trials where preexisting coronary artery disease would need to be considered.

Management of Coronary Artery Stenosis

Historically cardiac ischemia was managed exclusively with coronary artery bypass grafting. With improvement in diagnosis and percutaneous angioplasty techniques, the majority of patients can be managed in the cardiac catheterization lab.

Cardiac Catheterization

Cardiac catheterization remains the "gold standard" for determining the presence, extent, localization, and severity of coronary artery disease. Coronary angiography also allows the visualization of coronary artery collaterals, intracoronary thrombus, dynamic coronary artery spasm, and congenital anomalies of the coronary arteries. Cardiac catheterization allows detection of the severity of coronary artery stenosis and lesion location, determination of coronary artery diameter, and evaluation of severity of coronary distribution at risk subtended by the coronary stenosis. Upon identification of obstructive coronary artery disease, medical decision making is performed to identify the best method of treatment. These modalities include medical therapy, percutaneous intervention (PTCA/stent), or coronary artery bypass surgery. Assessment of left ventricular function is also realized with performance of left ventriculography. The overall measurement of left ventricular ejection fraction and regional wall motion analysis is obtained. Determining the severity of valvular stenosis and regurgitation is also performed by integration of hemodynamic assessment of individual cardiac chambers and visualization with contrast imaging.

Cardiac catheterization is an invasive imaging modality. Local anesthetic is injected into the skin to numb the area. A puncture is then made with a needle in either the femoral artery in the groin or the radial artery in the wrist (Seldinger technique), before a guidewire is inserted into the arterial puncture. A plastic sheath (with a stiffer plastic introducer inside it) is then threaded over the wire and pushed into the artery. The wire is then removed and the side port of the sheath is aspirated to ensure arterial blood flows back. It is then flushed with saline. This arterial sheath, with a bleedback prevention valve, acts as a conduit into the artery for the duration of the procedure.

Catheter placement and movement are monitored on specialized X-ray angiography equipment allowing multiple projections for imaging with hemodynamic monitoring. Catheters are inserted using a guidewire and moved towards the heart. Once in position above the aortic valve, the guidewire is then removed. The catheter is then engaged with the origin of the coronary artery (either left coronary artery or

right coronary artery), and X-ray opaque iodine-based contrast is injected to visualize the coronary vessels on the X-ray fluoroscopy image. A catheter is placed in the left ventricular chamber and contrast injected to visualize the left ventricular function, regional wall motion analysis, and valvular integrity.

Quantitative Coronary Angiography (QCA)

Quantitative coronary angiography provides a computerized methodology of evaluating obstructive coronary artery disease. The greatest advantage of quantitative coronary angiography is its theoretical freedom from observer influences and bias, thereby minimizing significant potential intraobserver and interobserver variability. Many techniques are available for computer applications that permit quantification of coronary stenosis. A quantitative analysis of the angiogram requires some form of optical magnification of the cineangiographic image which, in turn, permits computer-assisted definition and quantitation of disease severity. With the off-line techniques, image acquisition proceeds in the conventional manner, with the generation of a digital cineangiogram displayed on the image processor. Digital quantitation of selected image frames can be made with or without electronic magnification, easily accomplished with modern digital quantitative software and analysis systems. Online digital systems and computer application packages have become widely available through commercial distribution. These systems are designed to facilitate accurate clinical analysis of thrombolysis, stent deployment, and other endovascular interventions. See Fig. 9.4 for an example image of QCA.

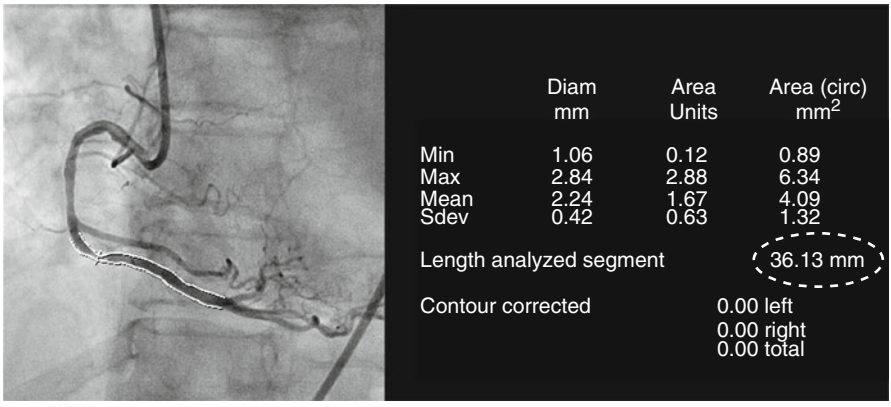


Fig. 9.4 Quantitative coronary angiography assessment. An example of the measurements by two-dimensional quantitative coronary analysis. The angiographic length of a segment was determined by measuring the distance between the proximal and distal origin sites of the branches. The minimal luminal diameters were measured from the center of the stenosed lesion to the outline of the vessel wall (Used with kind permission of Springer Science + Business Media from Lee et al. [9])

QCA provides a methodology for determining absolute measurement of the minimal and reference diameters, percent area, and diameter of area stenosis and of extent and asymmetry of atherosclerotic plaque. In addition, reference segments of adjoining nonobjective segments are cataloged. Serial comparisons of coronary artery dimensions post coronary intervention have provided the mainstay in demonstrating effectiveness of coronary stenting and mechanical interventional techniques. QCA has also been utilized in evaluating the impact on coronary artery regression utilizing pharmacologic and biologic therapies. Recent clinical trials have utilized coronary angiography in concert with intravascular ultrasound/OCT to demonstrate the effectiveness and response to mechanical coronary intervention and pharmacologic treatments.

Coronary Artery Stent Placement

Traditional ("bare metal") coronary stents provide a mechanical framework that holds the artery wall open, preventing stenosis, or narrowing, of coronary arteries. PTCA with stenting has been shown to be superior to angioplasty alone in patient outcome by keeping arteries patent for a longer period of time.

Newer drug-eluting stents (DES) are traditional stents that are coated with a polymer and bioactive pharmacologic agent, which, when placed in the artery, release certain drugs over time. It has been shown that these types of stents help prevent restenosis of the artery through several different physiological mechanisms. These locally eluted drugs act upon the suppression of tissue growth at the stent site and local modulation of the body's inflammatory and immune responses.

Intravascular coronary ultrasound and quantitative coronary angiography (QCA) have been widely used in clinical trials in quantifying the initial treatment impact in acute gain in coronary artery lumen post-intervention and subsequent angiographic and IVUS follow-up to monitor the degree of late loss occurring at the stent treatment site and correlation with clinical outcomes and endpoints.

Magnetic Resonance Imaging (MRI)

A magnetic resonance imaging instrument (MRI scanner) uses powerful magnets to polarize and excite hydrogen nuclei (single proton) in water molecules in human tissue, producing a detectable signal which is spatially encoded, resulting in images of the body. MRI uses three electromagnetic fields: a very strong (on the order of units of teslas) static magnetic field to polarize the hydrogen nuclei, called the static field; a weaker time-varying (on the order of 1 kHz) field(s) for spatial encoding, called the gradient field(s); and a weak radio frequency (RF) field for manipulation

of the hydrogen nuclei to produce measurable signals, collected through an RF antenna. This is described in more detail in Chap. 1.

Like CT, MRI traditionally creates a two-dimensional image of a thin "slice" of the body and is therefore considered a tomographic imaging technique. Modern MRI instruments are capable of producing images in the form of 3D blocks, which may be considered a generalization of the single-slice, tomographic, concept. Unlike CT, MRI does not involve the use of ionizing radiation and is therefore not associated with the same health hazards. However, there is well-identified health risks associated with tissue heating from exposure to the RF field and the presence of implanted devices in the body, such as pacemakers. These risks are strictly controlled as part of the design of the instrument and the scanning protocols used.

Because CT and MRI are sensitive to different tissue properties, the appearance of the images obtained with the two techniques differs markedly. In CT, X-rays must be blocked by some form of dense tissue to create an image, so the image quality when looking at soft tissues will be poor. In MRI, while any nucleus with a net nuclear spin can be used, the proton of the hydrogen atom remains the most widely used, especially in the clinical setting, because it is so ubiquitous and returns a large signal. This nucleus, present in water molecules, allows the excellent soft tissue contrast achievable with MRI.

Limitations of cardiac MRI in trial use include patient acceptance of study time and space constraints and increased study cost compared to computerized tomography. The ability of cardiac MRI to measure ventricular function, mass, and volumes accurately, reproducibly, and on serial studies without significant risk and without radiation exposure to the patient is a major strength for cardiac MRI in clinical trials. The measurements have been utilized to demonstrate the effectiveness of pharmacologic therapies in hypertensive management.

Cardiac MRI provides an accurate method of identifying cardiac function and myocardial perfusion. These data are extremely helpful in research involving myocardial infarction. It provides an accurate means of quantifying the size, transmural extent, and anatomic location of myocardial infarction. Perfusion imaging is performed with addition of intravenous gadolinium enhancement. In addition, the use of T1- and T2-weighted imaging can be utilized to myocardial tissue characteristics allowing quantification of myocardial infarct size.

MRI Vascular Imaging

MRI of the vasculature is an accurate method of determining the vessel dimensions as well as the components of the vessel wall (Fig. 9.5a–d). MR angiography may require the use of supplemental IV gadolinium contrast. Serial measurement of vascular dimensions allows monitoring the effectiveness of therapies and procedural planning for endovascular and surgical intervention if required. These serial

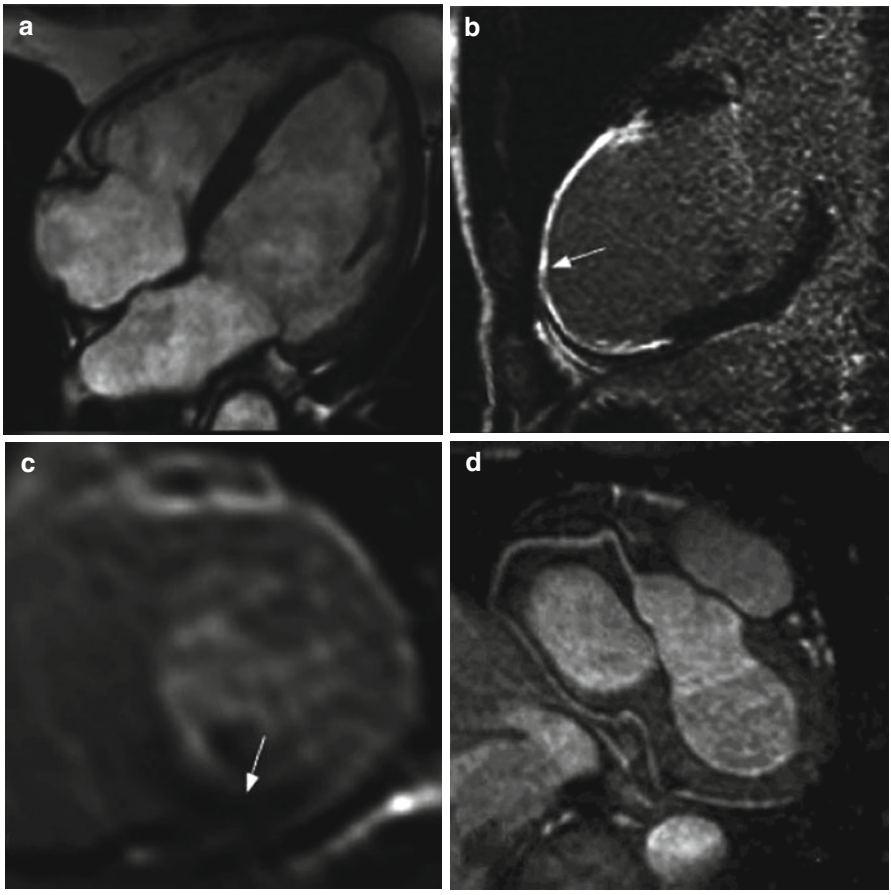


Fig. 9.5 Cardiac magnetic resonance imaging. Sample cardiac MR images. (**a**) Standard four-chamber SSFP view of the heart LV showing both atria and ventricles. (**b**) Vertical long-axis late gadolinium enhancement view of the LV showing infarction of the anteroapical wall (*arrow*). (**c**) Short-axis view of the LV showing a perfusion defect in the inferior wall (*arrow*). (**d**) Typical MR angiogram of a right coronary artery (From Greenwood et al. [10]; © 2009; licensee BioMed Central Ltd.http://www.trialsjournal.com/content/10/1/62)

measurements also provide quantitative and qualitative analysis of the effectiveness of pharmacologic and procedural techniques.

Vascular wall imaging of large arteries such as the aorta and carotid artery can be utilized to gain insight into the atherosclerotic process. Plaque severity including wall area/volume, plaque eccentricity, and minimal lumen area can be measured. In addition, atherosclerotic plaque composition can be identified including fibrous cap classification, identification of intraplaque hemorrhage, and plaque components of fibrosis and lipid-rich elements. The unique noninvasive abilities of MRI to identify various components affecting the vascular wall allow it to remain a valuable imaging tool in cardiovascular imaging trials.

Two-Dimensional Echocardiography

Two-dimensional (2D) echocardiography is the backbone of echocardiography. A background to ultrasound technology can be found in Chap. 1. By displaying anatomic structures in real-time tomographic images, comprehensive visualization of the components of the beating heart is achieved. The distance of ultrasound echoes along the vertical axis represents the depth of echo-producing structures, with brightness indicating the intensity of the returning echo.

M-mode echocardiography is useful for quantitating single dimensions of walls and chambers, which can be used to estimate chamber volumes and left ventricular (LV) mass when those structures are geometrically uniform. The Doppler technique uses reflections from moving red blood cells to characterize blood flow in the central and peripheral circulation. Doppler echocardiography complements M-mode and 2D echocardiography by providing functional information regarding intracardiac hemodynamics, including systolic and diastolic flow, blood velocities and volumes, severity of valvular lesions, location and severity of intracardiac shunts, and assessment of diastolic function. See Fig. 9.6a–d for an example of 2D Doppler echocardiography.

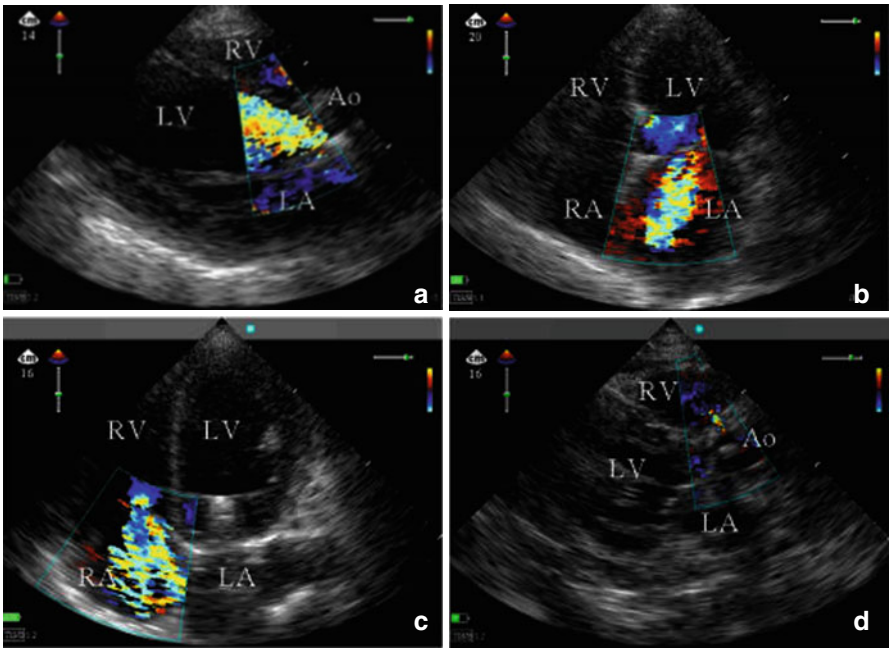


Fig. 9.6 2D Doppler echocardiography. Examples of color Doppler mapping still frames obtained from handheld echocardiographic examinations of four distinct patients. (**a**) Parasternal long-axis view obtained from patient with massive aortic insufficiency. (**b**) Apical four-chamber view obtained from a patient ventilated for an acute pulmonary edema and associated systolic murmur. (**c**) Apical four-chamber view obtained from a ventilated patient with shock. (**d**) Parasternal long axis view obtained from a patient presenting with septic shock and new onset systolic murmur (From Vignon et al. [11]. © 2003; licensee BioMed Central Ltd; http://ccforum.com/content/7/5/R84)

There are four types of Doppler: pulsed wave, continuous wave, color flow mapping, and tissue Doppler. Pulsed-wave Doppler is useful for localizing and timing flow that is moving within the physiological range of velocities. Continuous-wave Doppler, which lacks spatial resolution, is useful for accurately measuring the gradients that drive pathological flow jets. Color flow mapping, by measuring velocity along each sector line of the 2D image and displaying the information as color-coded pixels, provides a composite picture of flow over a larger area. It is most useful for screening the valves for regurgitation and stenosis, imaging systolic and diastolic flow, detecting the presence of intracardiac shunts, and detecting coronary flow. Tissue Doppler detects the amplitude and phases of the relatively slow motion of the LV myocardium. The supplemental use of intravenous echo-contrast agents can also improve endocardial visualization allowing more accurate assessment of regional wall motion assessment and perfusion.

Echocardiographic parameters are helpful in obtaining serial measurements of cardiac chamber size and left ventricular wall thickness/function to evaluate responses to pharmacologic interventions and disease states. Examples include evaluating change in ejection fraction in response to biventricular synchronized pacemaker implantation and changes in LV wall thickness/mass to antihypertensive therapies.

Digital acquisition and storage has many advantages in clinical trials, including high image quality, reproduction of images without loss of information, and long-term storage and transportability. Another important advantage is the ability to link study sites to core laboratories via the Internet or secure servers, with virtually no concern about geographic distance.

Carotid Ultrasound

Carotid atherosclerosis has proved to be a useful surrogate for coronary atherosclerosis in epidemiological and prospective interventional trials of anti-atherosclerotic agents. Since similar pathophysiological mechanisms are present in both the coronary and carotid arteries, imaging of the carotid arteries has been used as a gauge of pharmacologic activity on atherosclerotic progression. Ultrasonography permits noninvasive detection and quantification of abnormalities of carotid arterial structure, including wall thickening, plaque formation, and lumen enlargement. High-resolution B-mode ultrasound permits accurate and reproducible identification and measurement of the combined thickness of the intimal and medial layers of the carotid artery (Fig. 9.7). Several large epidemiological studies have shown significant associations between carotid intima-media thickness (CIMT) and both prevalent and incident coronary and cerebrovascular disease. Accordingly, measurement of CIMT has been a mainstay of cardiovascular epidemiological research for more than two decades.

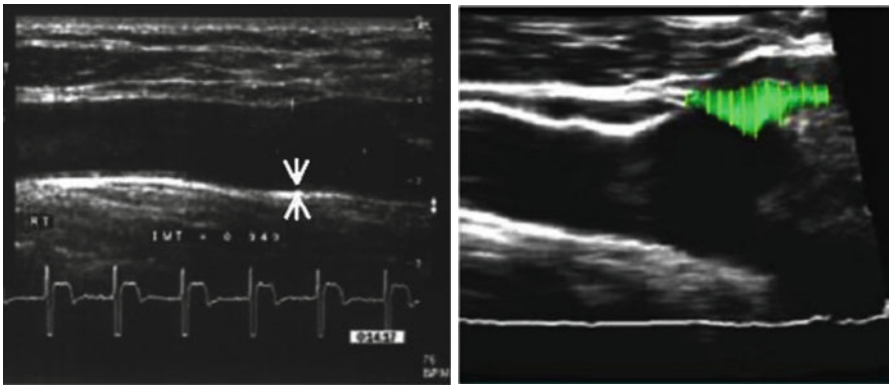


Fig. 9.7 Ultrasound in carotid anatomy. Ultrasound images used for the determination of carotid anatomy in a subject. The panel on the *left* shows an image of the right carotid artery used to determine intima-media thickness (*IMT*), with the *arrows* at the far carotid wall showing where IMT was determined. The panel on the *right* shows an image used to determine total plaque volume (*TPV*), with the encircled colored region defining one of the plaques identified (From Pollex et al. [12]. © 2006 ; licensee BioMed Central Ltd; http://www.cardiovascularultrasound.com/content/4/1/28)

When scanning and reading are performed carefully, the reproducibility and reliability of CIMT measurement can be excellent. Cross-sectional analyses suggest that age-related increases in mean CIMT average approximately 0.010 mm/year for women and 0.014 mm/year for men in the internal carotid artery and 0.010 mm/year for both genders in the common carotid artery. Similar values have been observed in prospective studies. Because the magnitude of clinically relevant differences in percentiles of CIMT and the progression rates are close to the resolution of vascular ultrasound transducers, highly standardized protocols are needed for performing and interpreting studies, which underscores the importance of high-quality, detailed image-acquisition protocols and highly skilled and trained ultrasonographers.

Imaging of the Vascular Anatomy: Intravascular Ultrasound (IVUS)

Intravascular ultrasound (IVUS) is a medical imaging methodology using a specially designed catheter with a miniaturized ultrasound probe attached to the distal end of the catheter. The proximal end of the catheter is attached to computerized ultrasound equipment. It allows the application of ultrasound technology to see from inside blood vessels out through the surrounding blood column, visualizing the endothelium (inner wall) of blood vessels in patients.

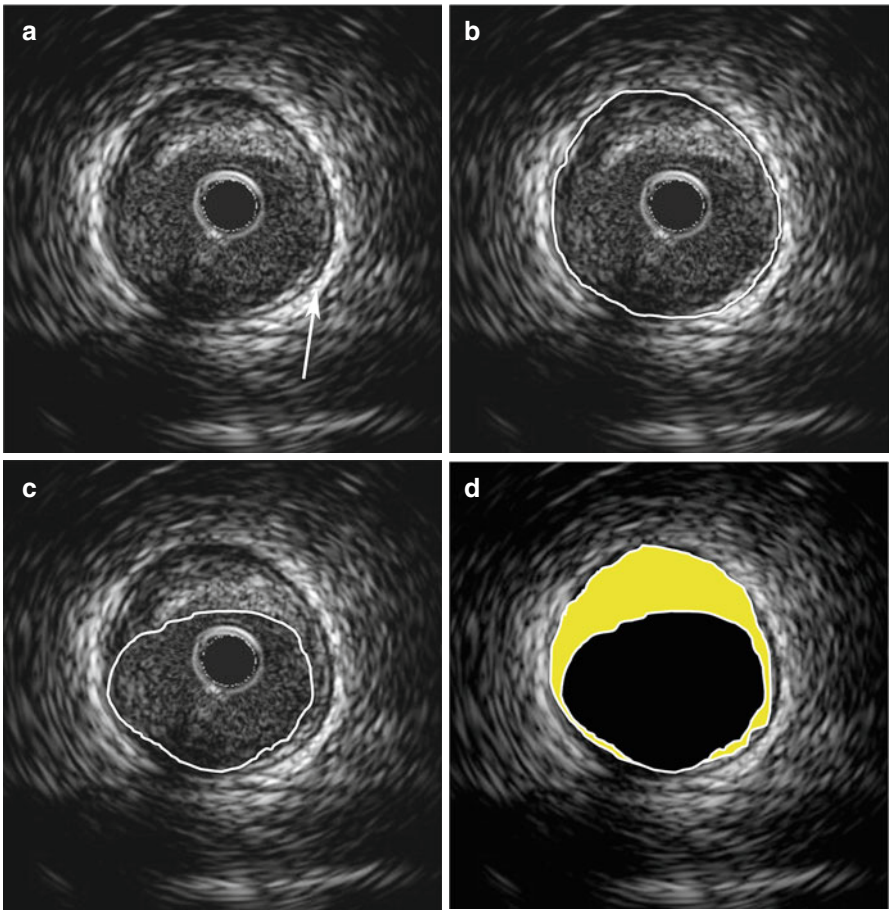


Fig. 9.8 Coronary artery intravascular ultrasound. An example of single cross section of a segment of the left anterior descending coronary artery obtained by intravascular ultrasound (IVUS) showing a fibro-fatty plaque (**a**). The external elastic lamina is shown by the *arrow*. The method of measuring plaque area is demonstrated. First, the external elastic lamina is identified and the area is traced (**b**); similarly, the lumen area is traced (**c**). The plaque area for this cross section is defined as the (EEL area) − (lumen area) (**d**). Plaque volume is determined by the summation of the individual plaque areas (Used with kind permission of Springer Science + Business Media from Ragosta [13])

The coronary arteries are the most frequent imaging target for IVUS, although IVUS can also be utilized in imaging the peripheral arterial system and aorta. IVUS is used in the coronary arteries to determine the amount of atheromatous plaque built up at any particular point in the epicardial coronary artery as shown in Fig. 9.8a–d. The progressive accumulation of plaque within the artery wall over decades is the foundation for vulnerable plaque. Sudden disruption or plaque rupture is a leading cause of acute myocardial infarction. The coronary plaques most likely to be predisposed to disruption tend to have higher concentrations of soluble

lipid components. Stable atherosclerosis or fibrotic plaque deposition tends to cause more progressive coronary artery stenosis limiting coronary blood flow with stenosis or at rest with severe disease. IVUS is of use to determine both plaque volume within the wall of the artery and the degree of stenosis of the artery lumen. It can be especially useful in situations in which angiographic imaging is considered unreliable, such as for the lumen of ostial lesions or where angiographic images do not visualize lumen segments adequately, such as regions with multiple overlapping arterial segments. It is also used to assess the effects of treatments of stenosis such as with hydraulic angioplasty expansion of the artery, with or without stents, and the results of medical therapy over time.

Arguably the most valuable use of IVUS is to visualize plaque, which cannot be seen by angiography. It has been increasingly used in research to better understand the behavior of the atherosclerosis process.

IVUS enables accurately visualizing not only the lumen of the coronary arteries but also the atheroma (membrane-/cholesterol-loaded white blood cells) "hidden" within the wall. IVUS has thus enabled advances in clinical research providing a more thorough perspective and better understanding.

In the early 1990s, IVUS research on the restenosis problem after angioplasty led to recognition that most of the restenosis problem (as visualized by an angiography examination) was not true restenosis. Instead it was simply a remodeling of the atheromatous plaque, which was still protruding into the lumen of the artery after completion of angioplasty, the stenosis only appearing to be reduced because blood and contrast could now flow around and through some of the plaque. The angiographic dye column appeared widened adequately; yet considerable plaque was within the newly widened lumen and the lumen remained partially obstructed. This recognition promoted more frequent use of stents to scaffold the plaque outward against the inner artery walls, out of the lumen.

IVUS measurements are regularly used in coronary stent clinical trials to quantify regions of scar formation (instent restenosis) in relation to stent distribution as well as to quantify the amount of therapeutic endothelialization that occurs following stent placement. Quantification of the thickness of the healing response following stent implantation (late loss) can be related to the incidence of restenosis and differential magnitudes between bare-metal coronary stents and various drug-eluting stents. IVUS measurements are regularly correlated with quantitative coronary angiographic measurements in coronary stent trials.

Additionally, IVUS examinations, as they were done more frequently, served to reveal and confirm the autopsy research findings of the late 1980s, showing that atheromatous plaque tends to cause expansion of the internal elastic lamina, causing the degree of plaque burden to be greatly underestimated by angiography. Angiography only reveals the edge of the atheroma that protrudes into the lumen.

Perhaps the greatest contribution to understanding, so far, was achieved by clinical research trials completed in the United States in the late 1990s, using combined angiography and IVUS examination, to study which coronary lesions most commonly result in a myocardial infarction. The studies revealed that most myocardial infarctions occur at areas with extensive atheroma within the artery wall, however

very little stenosis of the artery opening. The range of lumen stenosis locations at which myocardial infarctions occurred ranged from areas of mild non-flowing-limiting stenosis to lesions of greater than 95 % stenosis.

Current clinical uses of IVUS technology include checking how to treat complex lesions prior to coronary intervention and allowing appropriate coronary device selection – atherectomy versus balloon angioplasty. An additional application of IVUS is to evaluate appropriate stent sizing and strut apposition to the coronary artery endothelium vital to effective drug delivery with drug-eluting stent technology. If a stent is not adequately expanded to allow symmetric strut deployment to the endothelial wall, turbulent flow may occur between the stent and the wall of the vessel, a potential risk marker for a devastating complication of stent thrombosis.

The primary disadvantages of IVUS being used routinely in a cardiac catheterization laboratory are its expense, the increase in the time of the procedure, and the potential complication rate secondary to its invasive nature.

Methodology for IVUS

To visualize a coronary artery, angiographic techniques are used and the physician positions the tip of a guidewire, usually 0.36 mm (0.014″) diameter with a very soft and pliable tip and about 200 cm long. The physician steers the guidewire from outside the body, through angiography catheters and into the blood vessel branch to be imaged.

The ultrasound catheter tip is slid in over the guidewire and positioned, using angiography techniques so that the tip is at the farthest-away position to be imaged. The sound waves are emitted from the catheter tip and are usually in the 10–20 MHz range, and the catheter also receives and conducts the return echo information out to the external computerized ultrasound equipment which constructs and displays a real-time ultrasound image of a thin section of the blood vessel currently surrounding the catheter tip, usually displayed at 30 frames/s image. The guidewire is kept stationary and the ultrasound catheter tip is slid backwards, usually under motorized control at a pullback speed of 0.5 mm/s (The motorized pullback tends to be smoother than hand movement by the physician.)

The (a) blood vessel wall inner lining, (b) atheromatous disease within the wall, and (c) connective tissues covering the outer surface of the blood vessel are echogenic, i.e., they return echoes making them visible on the ultrasound display. By contrast, the blood itself and the healthy muscular tissue portion of the blood vessel wall are relatively echolucent, just black circular spaces, in the images.

Heavy calcium deposits in the blood vessel wall both heavily reflect sound, i.e., are very echogenic, but are also distinguishable by shadowing. Heavy calcification blocks sound transmission beyond and so, in the echo images, is seen as both very bright areas but with black shadows behind (from the vantage point of the catheter tip emitting the ultrasound waves).

Applications for IVUS

IVUS, as outlined previously, has been the best technology, so far, to demonstrate the anatomy of the artery wall in living animals and humans. It has led to an explosion of better understanding and research on both (a) the behavior of the atherosclerosis process and (b) the effects of different treatment strategies for changing the evolution of the atherosclerosis disease process.

Validating the Efficacy of New Treatments

Because IVUS is widely available in coronary catheterization labs worldwide and can accurately quantify arterial plaque, especially within the coronary arteries, it is increasingly being used to evaluate newer and evolving strategies for the treatment of coronary artery disease, including the statins and other medical therapies.

Optical Coherence Tomography (OCT) in Cardiology

Intravascular optical coherence tomography (OCT) has recently been proposed as a high-resolution imaging method for plaque characterization. Optical coherence tomography is an optical analog of ultrasound imaging because it measures the amplitude of backscattered light (optical echoes) returning from a sample as a function of delay. In vitro studies have shown that the resolution of OCT can resolve the thin fibrous caps thought to be responsible for plaque vulnerability. Additionally, the intrinsic optical properties of typical plaque constituents have provided sufficient contrast in these studies for OCT to differentiate between lipid, calcium, and fibrous tissue.

Identification of vulnerable plaques might lead to a therapeutic strategy specifically designed for a given patient, such as balloon angioplasty, stenting, or local delivery therapy to prevent acute coronary syndromes.

Summary and Conclusion

The cardiovascular system can be visualized through a wide variety of noninvasive and invasive methodologies and procedures. Imaging modalities allow an objective means of reproducibly identifying and quantifying the cardiovascular system's response to disease state and therapeutic agents.

Cardiovascular imaging in clinical trials remains an important modality in providing endpoints for effectiveness of therapies and monitoring for potential toxicities.

Imaging findings have been utilized as a primary endpoint for an investigation as well as secondary endpoints. The broadness of imaging modalities and their applicability continue to reinforce their value in measuring clinical effectiveness and in obtaining regulatory approval for new and innovative treatments.

Future Directions

Future directions in clinical imaging of the cardiovascular system include the development of new hardware and software to minimize ionizing radiation exposure and improving imaging resolution while decreasing amount of time required to perform the study.

The integration of cardiovascular imaging with biologically specific targeted pharmacologic or radioactive therapies will provide new insight into evaluating the effectiveness of atherosclerotic treatments and better understanding in basic mechanisms of disease. Identification of new pathways of disease modification will allow new targets for therapeutic agents.

Development of next generation contrast agents will allow lower toxicity profiles for imaging studies and potential multimodality same setting imaging.

References

1. Navare SM, Mather JF, Shaw IJ, et al. Comparison of risk stratification with pharmacologic and exercise stress myocardial perfusion imaging: a meta-analysis. J Nucl Cardiol. 2004;11: 551–61.
2. Al Moudi M, Sun Z, Lenzo N. Diagnostic value of SPECT, PET and PET/CT in the diagnosis of coronary artery disease: a systematic review. Biomed Imaging Interv J. 2011;7(2):e9.
3. Paech DC, Weston AR. A systematic review of the clinical effectiveness of 64-slice or higher computed tomography angiography as an alternative to invasive coronary angiography in the investigation of suspected coronary artery disease. BMC Cardiovasc Disord. 2011;11:32.
4. Dharampal AS, Papadopoulou SL, Rossi A, Weustink AC, Mollet NR, Meijboom WB, Neefjes LA, Nieman K, Boersma E, de Feigter PJ, Krestin GP. Computed tomography coronary angiography accuracy in women and men at low to intermediate risk of coronary artery disease. Eur Radiol. 2012;22(11):2415–23.
5. Hulten EA, Carbonaro S, Petrillo SP, et al. Prognostic value of cardiac computed tomography angiography. A systematic review and meta-analysis. J Am Coll Cardiol. 2011;57:1237–47.
6. Mansour S, Roy DC, Bouchard V, et al. COMPARE-AMI trial: comparison of intracoronary injection of $CD133^{+}$ bone marrow stem cells to placebo in patients after acute myocardial infarction and left ventricular dysfunction: study rationale and design. J Cardiovasc Transl Res. 2010;3(2):153–9.
7. Fox JJ, Strauss HW. Imaging the heart in the cancer patient. In: Strauss HW, Mariani G, Volterrani D, Larson SM, editors. Nuclear oncology. New York: Springer Science + Business Media; 2013.
8. van Werkhoven JM, Bax JJ, Nucifora G, Jukema JW, Kroft LJ, de Roos A, Schuijf JD. The value of multi-slice-computed tomography coronary angiography for risk stratification. J Nucl Cardiol. 2009;16(6):970–80.

9. Lee JB, Chang SG, Kim SY, Lee YS, Ryu JK, Choi JY, Kim KS, Park JS. Assessment of three dimensional quantitative coronary analysis by using rotational angiography for measurement of vessel length and diameter. Int J Cardiovasc Imaging. 2012;28(7):1627–34.
10. Greenwood JP, Maredia N, Radjenovic A, Brown JM, Nixon J, Farrin AJ, Dickinson C, Younger JF, Ridgway JP, Sculpher M, Ball SG, Plein S. Clinical evaluation of magnetic resonance imaging in coronary heart disease: the CE-MARC study. Trials. 2009;10(1):62. doi:10.1186/1745-6215-10-62#_blank.
11. Vignon P, Chastagner C, François B, Martaillé J-F, Normand S, Bonnivard M, Gastinne H. Diagnostic ability of hand-held echocardiography in ventilated critically ill patients. Crit Care. 2003;7(5):R84–91.
12. Pollex RL, Al-Shali KZ, House AA, Spence DJ, Fenster A, Mamakeesick M, Zinman B, Harris SB, Hanley AJG, Hegele RA. Relationship of the metabolic syndrome to carotid ultrasound traits. Cardiovasc Ultrasound. 2006;4(1):28. doi:10.1186/1476-7120-4-28#_blank.
13. Ragosta M. Techniques for phenotyping coronary artery disease in the cardiac catheterization laboratory for applications in translational research. J Cardiovasc Transl Res. 2011;4(4):385–92.

Chapter 10
Neuroimaging in Clinical Trials

Sumit N. Niogi and Apostolos J. Tsiouris

Abstract Neuroimaging plays a vital role when designing a clinical trial to study tumors, degenerative diseases, stroke, and autoimmune disorders of the central nervous system. Standardization of diagnostic criteria and analysis techniques to report severity of disease serve as objective criteria to judge outcome in these studies. This chapter will review the main imaging tools used when designing a neuroimaging clinical trial and the diagnostic and measurement criteria used to evaluate investigations in neuro-oncology, stroke, multiple sclerosis, and Alzheimer's disease. This chapter will also provide a brief review of these pathologies to provide a better understanding of the topics most commonly investigated with neuroimaging.

Keywords Cancer • Brain tumors • Multiple sclerosis • Stroke • Alzheimer's disease • MRI • Medical devices • RECIST

Introduction

Over the past several years, neuroimaging has become the cornerstone of medical care in diseases and disorders that affect the brain. Numerous medical imaging techniques are currently used for diagnosis, tracking treatment responses and disease progression, as well as providing prognoses. Similarly, neuroimaging plays a vital

S.N. Niogi, MD, PhD
Department of Radiology, Weill Cornell Med College,
525 East 68th Street, New York, NY 10065, USA
e-mail: sun2003@med.cornell.edu

A.J. Tsiouris, MD (✉)
Department of Radiology, Neuroradiology Section,
New York-Presbyterian Hospital/Weill Cornell Medical Center,
525 East 68th Street, Starr 630C, New York, NY 10065, USA
e-mail: apt9001@med.cornell.edu

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_10,

role in the design of a clinical trial. Beyond qualitative descriptors that aid diagnosis, with the standardization of diagnostic criteria and analysis techniques, neuroimaging methods yield quantitative metrics that can be used to objectively classify severity of disease and serve as a statistic to judge outcome in these studies.

Traditional tools for structural assessment such as computerized tomography (CT) and magnetic resonance imaging (MRI) are readily used in clinical trials to detect abnormalities such as tumors, infarctions, demyelination, blood products, and degenerative disease. Functional magnetic resonance imaging (fMRI), positron emission tomography (PET), magnetic resonance spectroscopic imaging (MRSI), and single-photon emission computed tomography (SPECT) are additional imaging modalities that assess hemodynamic or metabolic parameters to yield functional information. More recently applied methods such as diffusion tensor magnetic resonance imaging (DTI) are used in the research community to identify and quantify microstructural changes in white matter for diseases such as multiple sclerosis and Alzheimer's disease. Susceptibility-weighted imaging (SWI), a recently commercialized MRI sequence, is even more sensitive to blood products than the traditional T2*-weighted gradient-echo MRI scans.

With such a wide array of neuroimaging tools available, it is critical to understand the utility of each methodology, the limitations, and how to interpret the findings in a standardized fashion. Presented in this chapter is an overview of the major neuroimaging techniques and how they are used in clinical trials.

Structural Neuroimaging Techniques

The two most common tools for in vivo structural assessment of the brain are CT and MRI. CT stands out as a valuable and superior tool to detect skull fractures, acute brain hemorrhage, brain herniation and hydrocephalus. Conversely, MRI outperforms CT for the detection of most brain pathology including tumors, infections, neurodegenerative diseases, microhemorrhages, and white matter diseases. At subacute time points and beyond, MRI becomes significantly more sensitive than CT for subarachnoid hemorrhages due to changes in the composition of blood. As a result, MRI is typically the methodology of choice for clinical trials in neuroimaging due to the increased sensitivity and detailed information it can provide about the pathology [1]. However, due to several contraindications to MRI, claustrophobic patients and longer imaging times, CT retains tremendous value for emergent assessment, in patients with implanted devices deemed unsafe for MRI and in patients that require continuous monitoring.

Computed Tomography (CT)

Head CT scanners became available in 1972 when introduced by Sir Godfrey Hounsfield. This prototype scanner produced 80×80 images with 3 mm pixel resolution and required about 5 min to obtain each axial image [2]. Since the early

1990s, newer "helical" CT scanners produce x-rays that form a spiral path, hence the name [2]. CT scanners now use multiple detector arrays that allow the acquisition of several images simultaneously. Today, CT scanners obtain pictures of the brain with submillimeter resolution in seconds. The newest CT scanners have dual x-ray sources that acquire image data in about half the time, making patient motion less of an issue for CT because of its shorter image acquisition time than MRI [3].

Contrast in CT images is due to the attenuation of the x-ray beam through the patient. The attenuation coefficient in each pixel is compared to the attenuation through water, calculating the dimensionless Hounsfield unit (HU) [2]. Each HU is equivalent to 0.1 % of the attenuation of water. This normalized scale ranges from about −1,000 to +3,000 where −1,000 corresponds to air, soft tissues are in the range of −100 to +100, water is 0, and dense bones, metal, or contrast agents register as high as +3,000. Most importantly, HU are quantitative and can be used for accurate diagnosis in many clinical settings using objective criteria.

Various contrast agents are used to create or increase differences in attenuation between different anatomic structures for better visualization. Iodinated contrast agents were first established in the 1950s and continue to be commonly used today [4]. High-density iodinated contrast agents increase x-ray attenuation. Despite very low toxicity, the high osmolality of these iodinated agents can cause certain side effects resembling allergic reactions ranging from urticaria to an anaphylactoid process. While a prior reaction remains the best predictor of a future adverse event following the use of a contrast agent, recurrent reactions only range from 8 to 25 % [5]. A second serious complication includes intravenous (IV) contrast agent-induced nephropathy. While the pathogenesis remains unclear, it occurs only in patients with compromised renal function before contrast agent injection. Since IV contrast is eliminated by the kidneys, calculated creatinine clearance is the best way to determine risk and follow this potential complication. In patients with renal impairment, alternative procedures such as MRI should be considered. If it cannot be avoided, ample hydration before and after the scan is recommended in addition to the administration of acetylcysteine [4].

Newer nonionic contrast agents have been formulated that have lower osmolalities and a lower incidence of side effects. Advantages of nonionic contrast agents include reduced incidence of life-threatening anaphylactoid reactions and improved tolerance secondary to the blood-isotonic character of these substances compared to iodinated contrast [4].

There are numerous applications for the use of contrast in CT. Applications include CT myelography for mechanical filling of cavity structures (intrathecal administration), detection of arterial occlusions and perfusion deficits in stroke (intravenous administration), and functional assessment of the blood-brain barrier (intravenous administration). During CT myelography, contrast agents make the luminal area of the spinal canal radiopaque after mixing with the cerebral spinal fluid. This can be used to appreciate intraspinal diseases such as spinal canal stenoses, masses, or cavities. With angiographic and perfusion imaging, contrast agents can reveal the occluded vessel and affected areas of the brain which lack contrast enhancement. Contrast agents will also opacify the lumens of cerebral aneurysms, making them very conspicuous. Contrast leaking through the blood-brain barrier

into the brain parenchyma after intravenous administration indicates functional disruption of the blood-brain barrier that may be caused by a tumor, inflammation, or infection [6].

Availability, cost-effectiveness, short imaging times, minimal contraindications, and the ability to detect critical lesions are all reasons why CT is the initial imaging modality of choice for acute assessment. In cases of acute injury from trauma or stroke, early diagnosis with aggressive management can prevent secondary complications. Furthermore, acute CT imaging can be used to establish a baseline to track the progression of recovery or injury.

Magnetic Resonance Imaging (MRI)

MRI provides tremendous insight into subtle structural abnormalities unappreciated by CT. MRI outperforms CT in visualization of numerous pathologies including most subtle intracranial abnormalities and axonal injury. Beyond 48–72 h post-injury, MRI becomes more sensitive to subarachnoid hemorrhages than CT. Additionally, MRI is far superior to CT at detecting nonhemorrhagic posttraumatic lesions [7].

Unlike CT, there are numerous variables that can be adjusted during the acquisition of MRI that produces unique types of images. Certain clinical MRI sequences may be insensitive to certain types of pathology. A brief review of the different MRI pulse sequences will reinforce the notion that multiple sequences should be used to maximize the likelihood of detecting abnormalities. Some standard image sequences for head examination using MRI include T1-weighted and T2-weighted FLAIR (fluid-attenuated inversion recovery), diffusion-weighted imaging, and T2*-weighted GRE (gradient-recalled echo).

T1-weighted imaging (T1WI) is created using short echo times (TE) and short repetition times (TR). Fat, which has a long longitudinal relaxation time, has high signal and will appear bright on T1WI. Water has less longitudinal magnetization and will appear dark on T1WI. T1WI has excellent anatomic visualization and can be useful for detecting the presence of focal atrophy. Because most lesions have increased water content relative to normal brain, lesions frequently appear relatively hypointense on T1WI. Hyperintense lesions on T1WI indicate fat-containing structures, proteinaceous fluid, or subacute hemorrhage.

T1WI is also used for contrast-enhanced MRI scans yielding a hyperintense signal and making pathology more conspicuous. Like with CT, IV contrast in MRI can be used to opacify blood vessels and detect defects in the blood-brain barrier. The most common contrast used is gadolinium-based, usually chelated to DTPA to prevent heavy metal toxicity. Gadolinium-based MRI contrast agents can be excreted by the kidneys or by the liver into the biliary system. In cases where the blood-brain barrier is absent or damaged, contrast accumulates and causes a hyperintensity on T1WI. The use of contrast T1-weighted MRI improves detection and characterization of pathologic processes such as neoplasm and inflammation in the brain [8].

T2-weighted imaging (T2WI) is routine and extremely helpful in detecting pathologic brain lesions. Tissues with long T2 relaxation times appear bright on T2WI. Pure water/CSF appears brightest since it has a very long T2 relaxation time. Fat has a shorter T2 relaxation time when compared to water; therefore on T2WI, fat appears less hyperintense than water. To distinguish between CSF and adjacent pathology, FLAIR technique can be applied to suppress the signal from CSF by using an initial inversion radiofrequency pulse. This causes increased contrast in brain adjacent to CSF, allowing better visualization of pathology. T2-weighted FLAIR imaging is the most sensitive technique used to detect white matter pathology and is superior to other imaging modalities and MR sequences at appreciating cortical surface contusions and periventricular lesions. Additionally, hemorrhagic contusions after a few hours can be visualized on T2-weighted imaging, the sensitivity to which increases over the next several days as the deoxyhemoglobin oxidizes to a paramagnetic methemoglobin [9]. Lastly, it is generally accepted that CT is superior to conventional MRI in detecting acute subarachnoid hemorrhages; however, there are cases when T2-weighted FLAIR sequences can appreciate minimal or subacute hemorrhages overlooked by CT [10].

Gradient-echo (GRE) imaging has high sensitivity to magnetic susceptibility. As a result, T2*-weighted GRE pulse sequences are primarily used to detect traumatic microhemorrhages and calcium deposits in the brain. Blood breakdown products such as deoxyhemoglobin, methemoglobin, and hemosiderin induce local inhomogeneities in the magnetic field of the brain that cause a loss of T2*-weighted signal intensity. As such, T2* GRE imaging is very sensitive for detection of cerebral microhemorrhages. While parenchymal hemorrhage appears hypointense, flowing blood is hyperintense, a property of GRE sequences that is exploited for MR angiography [11].

A similar type of sequence to T2* GRE imaging is susceptibility-weighted images (SWI) that utilizes the paramagnetic properties of hemorrhagic blood products but has greater spatial resolution than a typical T2*-weighted GRE sequence and takes advantage of susceptibility-induced phase shifts of the local magnetic field to enhance sensitivity to blood products. It has been shown that SWI is more sensitive to traumatic microhemorrhages than other MR sequences [12]. Lastly, SWI can differentiate calcium from hemorrhage in the brain, two substances that are not differentiable on T2* GRE imaging.

Diffusion-weighted imaging (DWI) is a functional sequence that reflects the motion of water, but because it is commonly produced using a spin-echo echoplanar (EPI) sequence with diffusion gradients applied after the 180° pulse, the contrast in the image is also due in part to T2 weighting. The signal intensity is inversely proportional to the amount of water molecule diffusion. Thus, CSF appears very dark, white matter appears dark gray, and gray matter appears light gray. The EPI acquisition is remarkably fast such that whole brain diffusion imaging can take less than 1 min. DWI is an extremely valuable sequence because it is extremely sensitive to acute infarction (approximately 95 % sensitivity in the first 3 h). This is because acute infarctions produce cytotoxic edema restricting water motion and leading to DWI hyperintensity. Similarly, abscesses and encephalitis

are associated with cytotoxic edema and produce DWI hyperintensity; this property makes them very conspicuous and improves differentiation from other pathologies. On the other hand, vasogenic edema and fluid collections cause increased diffusion that leads to a hypointense signal. This type of signal is frequently seen in reactive edema around masses and necrotic tumors as well as within cysts. Quantitative measurements can be made on DWI by calculating the apparent diffusion coefficient (ADC). DWI is clinically useful for the evaluation of cerebral ischemia, tumors, and infection [13].

Diffusion tensor imaging (DTI) is performed using a series of DWI scans, each measuring the diffusion along a specific direction, to measure the 3D profile of water diffusion. The degree of diffusion along a single direction, the diffusion anisotropy, can be quantified and serve as a proxy for white matter integrity. Moreover, since water tends to flow along white matter pathways instead of orthogonal to it, tracing the anisotropic diffusion can produce 3D representations of white matter pathways, also known as DTI tractography. Tractography can be used for quantitative assessment of the relative number of neurons, as 3D regions of interest, or for preoperative planning to visualize the location of viable white matter pathways prior to a neurosurgical procedure [7].

Magnetization transfer imaging (MTI) was first described by Wolff and Balaban in 1989 [14]. MTI is based on the phenomena that protons bound in macromolecular structures exhibit T1 relaxation coupling with protons in the aqueous phase [14]. Applying an off-resonance saturation pulse can selectively saturate bound protons. The result is a transfer of magnetization from bound water to free water protons, leading to a reduction in signal intensity detected by MRI from these free protons. A decrease in the magnetization transfer ratio, a quantitative comparison of the signal intensity with and without magnetization transfer, is indicative of neuropathology. As a result, the MTI can provide a quantitative measure of the structural integrity of tissues. Initial studies suggest that MTI may be more sensitive to structural integrity than conventional MRI scans, but there are limited data available and correlations with clinical outcome have been weak [15].

MRI is useful in late follow-up stages of brain injury to detect cerebral atrophy. MRI is sensitive to changes in edema and can be used for patient monitoring. Unlike CT, it has utility in monitoring the changing status of a hemorrhage. Additionally, MRI is of great benefit to pediatric populations. Unlike CT, MRI uses no ionizing radiation. Thus, if sequential imaging is needed, MRI can be used to minimize exposure to radiation that is a concern in pediatric TBI cases.

In the future, indications for MRI may change. Once technical problems preventing the use of MRI in the acute phase are overcome (e.g., availability of MRI scanners in emergency settings and use on patients with life support or medical monitoring devices), the use of this modality in the acute setting may increase. Considering that the cost of MR exams is decreasing, image acquisition times are getting shorter, and newer MRI sequences have greater sensitivity to the neuropathology, there is great potential that MRI may be used in acute settings in the future.

Neuroimaging Methods to Assess Function

Functional Magnetic Resonance Imaging (fMRI)

Functional MRI (fMRI) is a widely used research technique to study brain function. Furthermore, fMRI is a promising tool because it has better temporal and spatial resolution when compared to other functional imaging modalities such as PET and SPECT. fMRI can easily be combined with structural MRI to provide better mapping of function to structure. Unlike other functional imaging modalities, fMRI does not require a radioactive isotope or ionizing radiation.

The blood-oxygen-level-dependent (BOLD) signal which fMRI utilizes to measure brain function is based on the assumption that an increase in neuronal activity will result in an increase in local blood flow to subserve the increased metabolic demand of neuronal cells in the region. The increase in blood flow brings with it a transient decrease in deoxyhemoglobin. The result is a relative increase in oxygenated blood that causes a local magnetic field gradient that produces a signal detectable by the MR scanner. In short, fMRI effectively measures changes in local cerebral blood as a proxy for changes in regional brain activity. To perform an fMRI experiment, the patient performs a simple task to isolate a specific domain of cognitive function while the MR scanner measures the BOLD signal [13].

fMRI studies show that brain activity can be reduced possibly due to damaged catecholaminergic systems or increased to facilitate task performance. There are a few limitations to consider in using fMRI in a clinical setting. First, the same contraindications that exist for structural MRI also exist for fMRI. Secondly, there is limited use in patients that are not alert or cooperative because the patients must be able to adequately perform multiple cognitive tasks. The tasks used for evaluation must also be standardized before fMRI is widely used in a clinical setting. Finally, interpretation of fMRI data requires statistical software and also technical expertise to perform the data analysis. Patient data must be compared to normal data; a premorbid scan or a database of appropriately matched controls is required.

Positron Emission Tomography (PET)

Positron emission tomography (PET) is another functional imaging modality that shows promise for examining functional deficits due to neuropathology. While PET does not approach the same spatial resolution of fMRI, current PET scanners achieve resolutions around 4 mm, exceeding SPECT resolution. Additionally, PET can be coregistered to CT or MRI images to provide additional anatomical detail. Hybrid PET/CT systems facilitate this objective.

PET can be used to detect blood flow, oxygen extraction, cell death, and glucose metabolism among other functions depending on the tracer or radioligand used. This radiotracer is injected into the patient intravenously, subsequently undergoing

beta decay and releasing a positron. Soon after positron release, it collides with and annihilates an electron. The annihilation produces two gamma photons that travel away from each other at 180°. The PET scanner is designed to detect the photon pair. Ultimately, the data are transformed with a computer to produce a tomographic image [16].

The function measured depends on the type of radiotracer or radioligand used; the most common tracer is fluoride 18 (F^{18}) fluorodeoxyglucose (FDG) which is a derivative of glucose and is used to study glucose metabolism. Cobalt 55 (Co^{55}) can be used to identify areas where cell death is occurring. Blood flow and oxygen extraction fraction can be determined using an oxygen 15 tracer (O^{15}). Because some tracers have a short half-life, the radiotracer must be made on-site. This limits the use of PET scanners at many institutions without a cyclotron facility on-site to produce the necessary radionuclides [16].

Single-Photon Emission Computed Tomography (SPECT)

Single-photon emission computed tomography (SPECT) is a less expensive alternative to PET and fMRI. Its low cost and accessibility at most institutions make it an attractive imaging modality to study brain activity. The most common tracer used is technetium-99m-hexamethylpropyleneamine oxime (^{99m}Tc-HMPAO). As this radiotracer decays, it emits a photon that is detected by the SPECT gamma camera. The tracer accumulates in endothelial cell membranes within a few minutes of injection and lasts for about 24 h. The long half-life allows for multiple scans with one injection. After data collection, a computer reconstructs the data to form a tomographic image of activity throughout the brain. The newer triple head cameras (collimators) achieve a resolution around 1 cm. Pinhole and multi-pinhole collimators can potentially achieve submillimeter resolution, but these are used primarily in the context of small animal imaging [17]. Fortunately, SPECT can be coregistered with CT or MRI scans to provide greater anatomical precision [1].

Magnetic Resonance Spectroscopic Imaging (MRSI)

Magnetic resonance spectroscopic imaging (MRSI) is a noninvasive in vivo technique that can be used to obtain several molecular images of endogenous metabolites. In addition to spatial information, MRSI can be used to measure the metabolite profile of a region or voxel. Unlike PET or SPECT, MRSI does not require the use of an exogenous contrast agent and uses the same hardware as MRI. MRSI has a unique value in the diagnosis of brain tumors and other brain pathology; it is currently the only noninvasive technique that can measure metabolite profiles and obtain molecular images. MRSI can be used to measure N-acetylaspartate (NAA) which is a marker of neuronal loss [18]. It can also measure myoinositol, which

serves as the putative marker for glia. In general, disease processes in the brain cause decreases in NAA and increases in myoinositol. MRSI can also measure lipids and lactate which are observed in necrosis and numerous other pathological processes in the brain. Furthermore, MRSI can measure choline, a membrane turnover marker, and creatine that relates to the energy states of the cell. Beyond the measure of metabolic profiles, MRSI provides the spatial metabolite distribution which is extremely helpful in delineating the borders of a lesion, measuring the amount of tumor cell infiltration within a region (which can be used to differentiate different types of tumors), and can be used for monitoring therapy or progression of a disease state. Unfortunately, MRSI suffers from long acquisition times, partial volume effects, and the same contraindications as MRI.

Neuroimaging of Neuro-oncology

MRI Score for Tumor Grading

In clinical trials, one can distinguish low- and high-grade gliomas using nine MRI criteria. These include heterogeneity, cyst formation, necrosis, hemorrhage, contralateral hemispheric extension, edema and/or mass effect, border definition, internal flow voids, and degree of contrast enhancement. Each criterion is scored through visual assessment (score for each criteria in Table 10.1) and average score is calculated. Typically, a score of 9 or greater signifies a high-grade glioma, while a lesion with a score less than 9 is characterized as a low-grade glioma [19].

Response Assessment Criteria

The best evaluation for the treatment of tumors is based on patient survival. In those with recurrent disease, progression-free survival (PFS) through radiologic assessment serves as a reliable marker. Any recorded change in tumor size as a result of treatment typically corresponds with duration of survival. With the emergence of anticancer therapies, standardization of radiographic assessment became essential to reliably monitor therapeutic response.

In 1979, the World Health Organization (WHO) published the first major set of criteria to measure the response rate that was subsequently followed by the Response Evaluation Criteria in Solid Tumors (RECIST) in 2000 and RECIST 1.1 in 2009 [20]. In general, the response rate is based on how the tumor shrinks anatomically. However, this requires serial measurements of tumor burden that can be both costly and tedious for both patient and investigator. As found in both WHO and RECIST, this change is defined by either having a complete response, partial response, stable disease, or progressive disease. Unfortunately, interpretation of these findings

Table 10.1 MRI score for tumor grading

Criteria	Measurement	Score
Heterogeneity (HET)	None	0
	Mild	1
	Moderate	2
	Marked	3
Cyst formation/necrosis (CN)	None	0
	Equivocal	1
	Yes	2
Hemorrhage (HEM)	None	0
	Equivocal	1
	Yes	2
Crossing midline (CM)	None	0
	Equivocal	1
	Yes	2
Edema and/or mass effect (EM)	None	0
	Mild	1
	Moderate	2
	Severe	3
Definition of border (BD)	Well defined	0
	Poorly defined	1
Flow void (FV)	None	0
	Equivocal	1
	Yes	2
Degree of contrast enhancement (CE-D)	None	0
	Slight	1
	Moderate	2
	Marked	3
Heterogeneity of contrast enhancement (CE-HET)	None	0
	Homogenous	1
	Heterogeneous	2

remains challenging because certain therapies may not shrink tumor size, but may be successful in preventing progression associated with good outcome (i.e., patient survival).

World Health Organization (WHO) Criteria

In 1979, the WHO attempted to standardize treatment response assessment [21]. The WHO criteria entail measuring the tumor size based on the product of the bidimensional measurement of tumors (i.e., greatest perpendicular dimensions; Fig. 10.1a–d), summing these dimensions over all tumors, and then categorizing changes in these summed products as follows:

- *Complete response*: Tumor has disappeared for at least 4 weeks.
- *Partial response:* 50 % or greater reduction in sum of tumor size products from baseline, confirmed at 4 weeks.

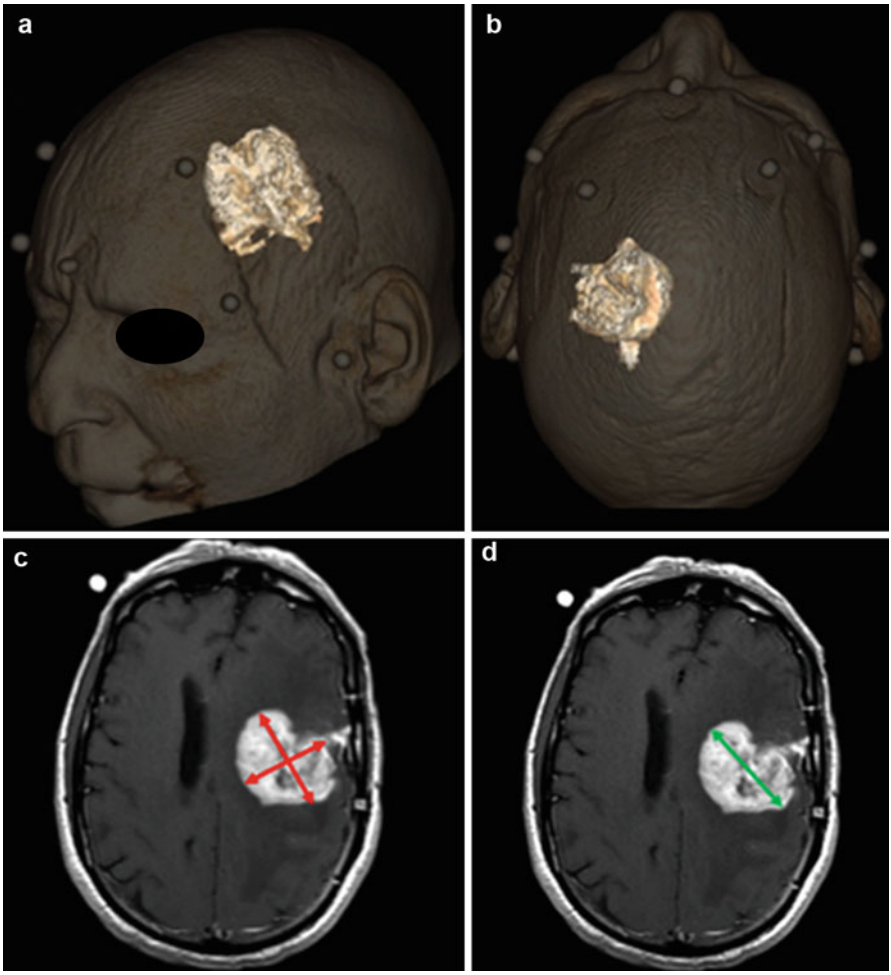


Fig. 10.1 Measurement of tumor size in a patient with a glioblastoma. (**a**, **b**) 3D volume rendered measurement of tumor volume. Measurement can be done manually by selecting or "painting" the tumor on multiplanar sequential images or by semiautomated methods by manually placing a seed within the tumor and using a seed growing algorithm to determine tumor volume. (**c**) WHO criteria uses a 2-dimensional measurement, i.e., the maximum cross product of perpendicular measurements (*red lines*). (**d**) RECIST use a unidirectional measurement, i.e., the longest axis (*green line*)

- *Stable disease*: Neither partial response nor complete response nor progressive disease.
- *Progressive disease*: At least a 25 % increase in tumor size in one or more lesions, with no complete response, partial response, or stable disease documented before increase in size, or development of new tumor sites.

With the WHO criteria there is a slight bias toward categorizing the response as "progressive disease" because small (11 %) increases in the perpendicular

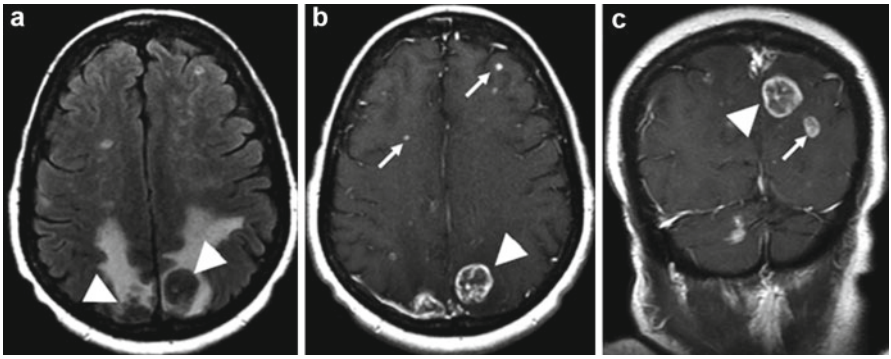


Fig. 10.2 Example of metastatic disease. (**a**) T2 FLAIR; extensive vasogenic edema surrounds two posterior parietal metastases (*arrowheads*). (**b**, **c**) Post-gadolinium T1WI; examples of target lesions that are measurable (*arrowheads*) and nontarget lesions (*arrows*) that are not measurable, but should be followed

dimensions will yield a 25 % increase in the product. Additionally, the WHO criteria are ambiguous regarding how many tumor foci should be measured, how small a lesion could be measured, and how progression should be defined. Still, the WHO criteria are used currently in some trials and were used to define clinical response rates for two decades. While the criteria is not commonly used today, understanding the WHO response criteria is essential for comparing current studies with ones from the past.

RECIST 1.0

The RECIST 1.0 were originally published in 2000 in recognition of the limitations of the WHO criteria [21]. The criteria were designed to serve as an end point for clinical trials that assess tumor treatments. RECIST provides guidelines regarding how many lesions to assess. The most significant difference is the use of a unidimensional measurement as opposed to the bidirectional cross product stated in the WHO criteria. RECIST 1.0 specifies assessment of up to ten target lesions (no more than five per organ) by measuring the single longest dimension of the tumor (Fig. 10.1a–d). The sum of these unidimensional measurements is used as the metric for determining response [21].

RECIST also distinguishes “measurable” and “nonmeasurable” tumors. Using spiral CT with 5 mm or thinner slices, the minimum “measurable” lesion to be assessed is 1 cm (using conventional CT the minimum is 2 cm). Lesions of adequate size for measurement are described as “measurable.” RECIST also use the designations “target” and “nontarget” lesions. All “target” lesions are measurable while only some nontarget lesions are measurable (Fig. 10.2a–c). However, both types of lesions can contribute to disease progression and to complete response. Target lesions are to be selected based on their size and suitability for accurate

repeated measurements up to a total of ten lesions, but no more than five per organ. All other lesions or sites of disease should be identified as "nontarget" lesions and recorded at baseline. For RECIST 1.0, measurements of these "nontarget" lesions are not required, but the presence of each should be noted throughout the follow-up.

The RECIST categories for response include:

- *Complete response*: Disappearance of all tumor foci for at least 4 weeks
- *Partial response*: A decline of at least 30 % in tumor diameters for at least 4 weeks
- *Stable disease*: Neither partial response nor progressive disease
- *Progressive disease*: At least a 20 % increase in the sum of all tumor diameters from the lowest tumor size

The RECIST are much more stringent than the WHO criteria for progressive disease. As a result, the time to disease progression can be shorter with the WHO criteria than with RECIST using the same data. In other instances, such as when new lesions are noted, the WHO and RECIST 1.0 yield the same conclusions.

RECIST 1.1

In 2009, the RECIST were modified based on data and outcomes from a variety of clinical trials [22]. In RECIST 1.1, fewer lesion assessments were found to yield better concordance of response classifications. In randomized studies for tumor progression, RECIST 1.1 suggests that as few as three lesions may be used. This potentially reduces the number of measurements by as much as much as 70 % compared to RECIST 1.0.

Notable revisions in RECIST 1.1 include [22]:

- Measurability: Longest dimension in the plane of measurement with a minimum size of 10 mm by CT and MRI (no less than double the slice thickness).
- Lymph node assessment: Normal if short axis <10 mm, "measurable" if short axis ≥15 mm, and "nonmeasurable" if short axis is 10–15 mm. (RECIST 1.0 made no specific recommendations for lymph nodes.)
- Cystic lesions: Simple cysts are not considered malignant. Complex cystic lesions are "nonmeasurable" lesions. Cystic metastases can be considered "measurable" lesions. (In RECIST 1.0 all cystic lesions were considered non measurable.)
- Measurable disease at baseline: In studies where the primary end point is tumor progression, the protocol must specify if entry is restricted to those with measurable disease or whether patients having nonmeasurable disease only are also eligible.
- Baseline: All lesions up to a maximum of five lesions (2 per organ) are target lesions. (RECIST 1.0 states a maximum of 10 lesions, 5 per organ.)

- Complete response: Pathologic lymph nodes must reduce to <10 mm. (RECIST 1.0 did not address lymph nodes.)
- Progressive disease: There must be a 20 % increase in size and a minimum of 5 mm increase over the smallest recorded size since the treatment started. (RECIST 1.0 did not require an absolute increase in size.)
- Too small to measure: Target lesions that become difficult to assess on CT are reported as "too small to measure" and a value of 5 mm should be assigned (not addressed in RECIST 1.0).
- Lesions that split: The longest diameters of the fragmented portions should be added together to calculate the target lesion sum (not addressed in RECIST 1.0).
- PET: FDG-PET can complement CT scanning to assess new lesions (not addressed in RECIST 1.0).
- Lesion reappearance: In the setting of complete response, reappearance of a lesion would be considered progressive disease (not addressed in RECIST 1.0).

Currently, RECIST 1.1 will be used in virtually every clinical trial of new solid tumor therapeutics because regulatory agencies have accepted RECIST as the standard in response assessment for clinical trials [20, 22]. It is important to note that response measurements using the WHO, RECIST 1.0, and RECIST 1.1 are not identical and therefore do not produce identical results. However, there exists good concordance among one-dimensional measurements of RECIST, two-dimensional measurements (e.g., WHO), and volumetric measurements (Fig. 10.1a–d). Three-dimensional measurements appear to be inferior to these other criteria.

Macdonald Criteria

While RECIST has become an almost universal standard for solid tumor response, an exception exists for high-grade gliomas. Currently, the Macdonald criteria remain the most widely used standard to evaluate tumor response in clinical studies of high-grade gliomas [23].

In 1990, the Macdonald criteria were published for the response assessment of high-grade gliomas [24]. The Macdonald criteria was developed using the same two-dimensional WHO measurements of tumor size (i.e., the product of the maximal cross-sectional enhancing diameters of the tumor). Although the criteria were originally developed for CT, they have been extended for use in MRI that is the current methodology of choice for tumor assessment.

The response assessment in neuro-oncology (RANO) working group has recently proposed a 2010 updated response criteria to enhance the interpretation of clinical trials involving novel agents [23]. The RANO working group recognized that there is a great deal of heterogeneity of high-grade gliomas and difficulty in measuring some lesions. As a result, the proposed update recommends that between two and five lesions be measured with emphasis put on selecting the largest lesions that allow for reproducible and repeated measurements. Readers should refer to the review by Wen [23] for a detailed description of the proposed updated response criteria. The following is a brief summary of the proposed changes:

Updated definitions of radiographic response [23]:

- *Complete response*: Requires all of the following:
  - Complete disappearance of all enhancing measurable and nonmeasurable disease sustained for at least 4 weeks
  - No new lesions
  - No corticosteroids
  - Stable or improved clinically
- *Partial response:* Requires all of the following:
  - 50 % decrease compared with baseline in the sum of products of perpendicular diameters of all measurable enhancing lesions sustained for at least 4 weeks
  - No new lesions
  - Stable or reduced corticosteroid dose
  - Stable or improved clinically
- *Stable disease:* Requires all of the following:
  - Does not qualify for complete response, partial response, or progression
  - Stable clinically
- *Progression:* Defined by any of the following:
  - 25 % increase in sum of the products of perpendicular diameters of enhancing lesions
  - Any new lesion
  - Clinical deterioration

Updated criteria for entry into clinical trials for recurrent high-grade glioma [23]:

- Patients are required to have a 25 % increase in the sum of the products of perpendicular diameters of contrast-enhancing lesions while on stable or increasing doses of corticosteroids.
- New contrast-enhancing nonmeasurable disease in trials for which progression-free survival is the primary end point.

Criteria for determining first progression [23]:

- Patients should be excluded from clinical trials for recurrent disease if they received radiotherapy within 12 weeks unless the progression is outside the radiation field (i.e., beyond the high-dose region or 80 % isodose line) or the disease is confirmed by pathology.
- 12 weeks after chemoradiotherapy completion, progressive disease is defined by:
  - New contrast-enhancing lesion outside the radiation field
  - Increase $\geq$25 % in the sum of the products of the perpendicular diameters between scans 12 weeks apart on stable or increasing doses of corticosteroids

- Clinical deterioration not attributable to medication or comorbid conditions
    - Sufficient to declare progression
    - Not valid for entry into clinical trial
- Significant increase in T2/FLAIR non-enhancing lesions for those receiving antiangiogenic therapies on stable or increasing doses of corticosteroids. Cannot be the result of comorbid events

Limitations of Response Criteria

Each of the response criteria suffers from limitations. For example, irregularly shaped tumors are difficult to measure. Furthermore, there are no guidelines on how to select lesions or assess multifocal tumors which greatly affects inter-rater reliability. Each criteria published does not assess non-enhancing tumors. There is also a bias to classify disease as progressive because any significant increase in size of the contrast-enhancing component signifies a tumor progression, which will alter therapy. However, contrast enhancement is nonspecific and can be influenced by medications such as corticosteroids and antiangiogenic agents. Additionally, postsurgical changes, radiation effects, and inflammation all increase enhancement but do not signify an increase in tumor burden. The most commonly experienced limitations for standardization of response criteria are discussed in the following sections.

Pseudoprogression

Radiotherapy with concurrent and adjuvant chemotherapy such as temozolomide is currently part of the standard treatment for GBMs [25]. In 20–30 % of patients, MRI following the first radiotherapy session will show increased contrast enhancement that eventually subsides without any change in therapy [25]. This phenomenon is termed pseudoprogression (Fig. 10.3a–f). It is thought to arise from a transient increase in the permeability of tumor vasculature following radiotherapy. Identifying pseudoprogression is critical to prevent the premature discontinuation of effective therapy. In clinical trials of patients with progressive disease, failure to identify and exclude patients with pseudoprogression will result in a falsely high response rate. That is, the therapy being investigated will be considered incorrectly to be active. PET utilizing ^{18}F-FDG can be used to assist in differentiating tumor from radiation necrosis by determining relative glucose metabolism [26]. Potentially, MR perfusion showing elevated rCBV values may be indicative of recurrent tumor.

Enhancement as a Result of Therapy

Surgery often causes marginal enhancement 48–72 h after tumor resection making it difficult to distinguish postoperative changes from residual enhancing disease. Given the time window of postoperative enhancement, a baseline MRI scan should

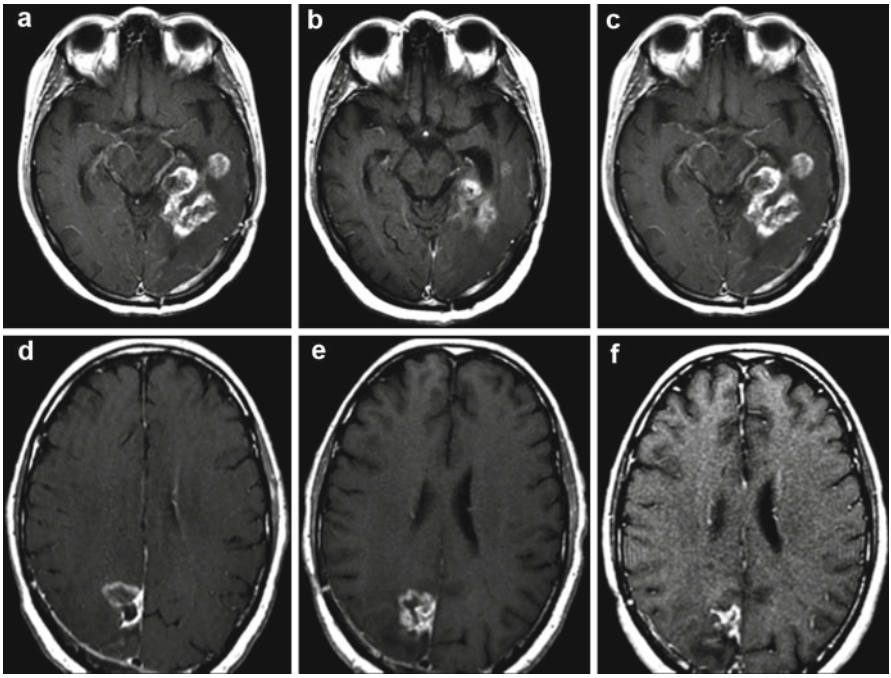


Fig. 10.3 Post-gadolinium T1WI images demonstrate a pseudoresponse (**a**, **b**, **c**) and pseudoprogression (**d**, **e**, **f**). (**a**) GBM prior to bevacizumab therapy; (**b**) after bevacizumab therapy demonstrating a pseudoresponse; (**c**) follow-up 3 months later demonstrating biopsy-proven true progression of disease; (**d**) GBM prior to temozolomide therapy; (**e**) after temozolomide therapy demonstrating pseudoprogression; (**f**) follow-up 3 months later showing regression of tumor off therapy

ideally be obtained within 24–48 h after surgery. This MRI scan should include diffusion-weighted imaging to determine whether new enhancement found in subsequent scans is the result of ischemia or tumor recurrence. Similar transient enhancements can be caused by chemotherapy wafers, immunotoxins delivered by convection-enhanced delivery, regionally administered gene and viral therapies, immunotherapies, focal irradiation with brachytherapy, and stereotactic radiosurgery. Unfortunately, no imaging modality, even diffusion-weighted MRI, has the specificity to conclusively differentiate recurrent tumor from treatment effects. If making this distinction were critical to clinical trials, surgical sampling would be necessary to obtain a definitive diagnosis.

Pseudoresponse/Pseudoregression

Pseudoresponse typically occurs after administration of antiangiogenic agents, such as those targeting vascular endothelial growth factor (VEGF) or VEGF receptor. These agents can produce a decrease in contrast enhancement between 24 and 48 h after administration. The decrease in enhancement is currently thought to be related to the normalization of permeable tumor vessels and it should not be mistaken to be a true

anti-glioma effect. As a result, falsely interpreted high response rates due to pseudoprogression may yield only modest survival benefits. Treatments with therapies such as bevacizumab produce a dramatic improvement in tumor enhancement on MRI [27]. The effect is mostly on the blood-brain barrier as opposed to a true tumor response. The rapid speed of the response is suggestive of a change in vascular permeability as opposed to a reduction in tumor size. Studies demonstrate that agents that target VEGF or VEGF receptor may initially produce a reduction in tumor contrast enhancement, but subsequently these tumors have increasing non-enhancing signals on T2 FLAIR [23, 28]. Even though enlarging areas of non-enhancing tumor is indicative of infiltrative disease, current response criteria fail to measure non-enhancing components. Still, there are emerging data suggesting that the degree of initial response due to vascular normalization may also correlate with survival [28] (Fig. 10.3a–f).

Failure to Measure Non-enhancing Tumor

Because high-grade gliomas are infiltrative, they often contain regions without disruption of the blood-brain barrier and thus no enhancement on MRI. Unfortunately, most criteria do not account for the non-enhancing component of the tumor, which is particularly challenging for WHO grade II and grade III gliomas that frequently contain non-enhancing components. In higher-grade gliomas, patients that have an initial reduction in tumor enhancement may develop an increase in non-enhancing infiltrative tumor as visualized on T2-weighted scans. Abnormal T2 signal in the brain parenchyma is difficult to interpret because it may be caused by a host of factors including radiation effects, demyelination, changes in medication (e.g., decreased corticosteroids), ischemic injury, infection, postoperative changes, or other treatment effects. T2 signal changes that are suggestive of infiltrating tumor include associated mass effect (as determined by sulcal effacement, ventricular compression, and thickening of the corpus callosum), infiltration of the cortical ribbon, and location outside of the radiation field [23].

Neuroimaging of High-Grade Gliomas

Gliomas account for about half of all primary brain tumors in adults with an annual incidence of about five per 100,000 people [29]. They typically occur supratentorially in adults and infratentorially in children. There exists a simple histological grading system for gliomas based on nuclear atypia, mitoses, endothelial proliferation, and necrosis. High-grade gliomas have at least two of these features. High-grade gliomas are expansive and frequently extend beyond the tumor margins. Additionally, all high-grade gliomas show some amount of anaplasia and are subdivided into WHO grade III anaplastic astrocytomas (AA) and WHO grade IV glioblastoma multiforme (GBM) [29]. While CT can be used to differentiate tumor grade in some cases, MRI is the modality of choice due to better tissue contrast that aids in the assessment and grading of all brain neoplasms.

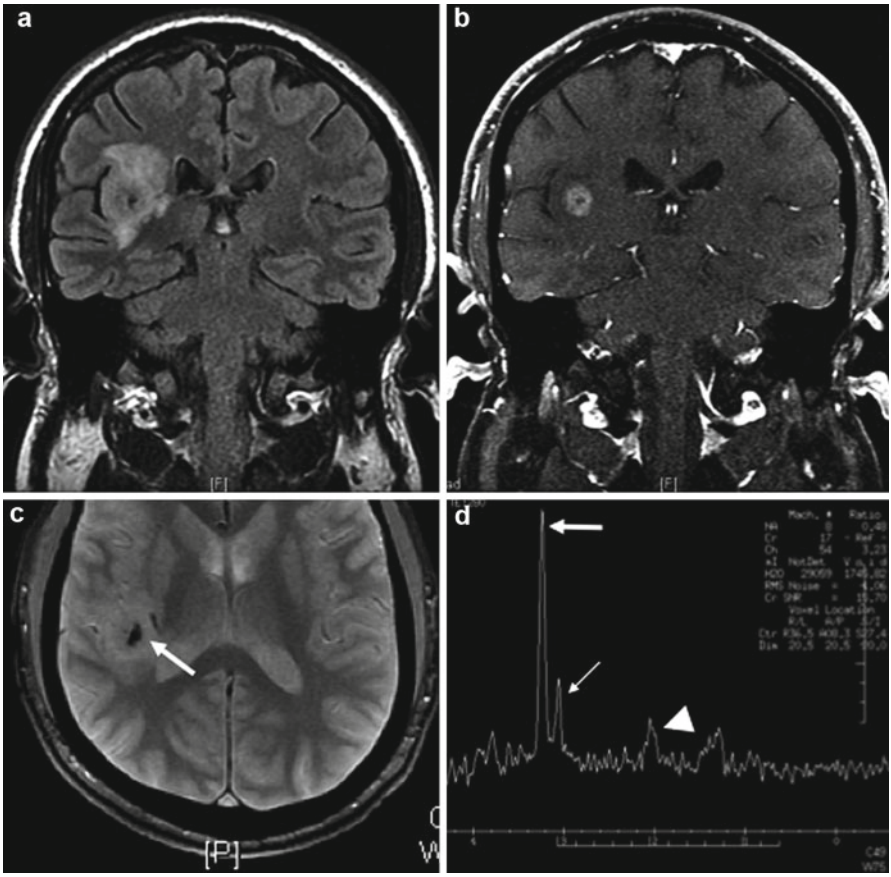


Fig. 10.4 Example of a grade III astrocytoma. (**a**) Coronal T2 FLAIR showing mass-like infiltrative T2 signal abnormality involving the insular cortex and subcortical white matter. (**b**) Post-gadolinium T1WI demonstrates central enhancement. (**c**) Foci of internal calcification are noted on the gradient-echo image (*arrow*). (**d**) MRS demonstrates an elevated choline to creatine (Cho:Cr) ratio of 3:1 characteristic of malignancy and a marked reduction of NAA, a normal neuronal marker (Cho, *thick arrow*; Cr, *thin arrow*; NAA, *arrowhead*)

Anaplastic astrocytomas (AA) are infiltrating lesions in which about 75 % arise from low-grade gliomas while the remainder forms as a new primary neoplasm. Unlike GBMs, AAs lack necrosis and vascular proliferation. Instead, the presence of pleomorphic astrocytes and mitosis are characteristic of most AAs [29]. On imaging, AAs present as poorly defined heterogeneous lesions. On CT scans, typically there exists a hypodense area of peritumoral edema. When enhancement is present, it is nodular or heterogenous. On MRI, AAs can present with a range of findings better reflecting the variability in histology of AAs. T1WI shows a heterogenous isointense to hypointense lesion. High-signal areas on T2WI represent the tumor itself and a peritumoral vasogenic edema. On T1-weighted contrast scans there is usually patchy enhancement [30] (Fig. 10.4a–d).

GBMs are primary neoplasms in 60 % of cases, while the remainder arises from AAs. These tumors usually occur in subcortical white matter of the temporal (31 %), parietal (24 %), frontal (23 %), and occipital (16 %) lobes [31]. Brainstem gliomas can also occur, but they are rare and usually found in children secondary to transformation from a low-grade glioma. On imaging, necrosis is the hallmark of GBMs. On CT, GBMs appear as low-density lesions that are heterogenous due to necrosis, hemorrhage, and increasing cellularity. Additionally, edema can be found surrounding the tumor and extending along adjacent white matter tracts causing a mass effect. On contrast-enhanced CT, GBMs appear with heterogeneous ringlike enhancement and an irregular and nodular wall. On T1-weighted MRI, GBMs appear heterogenous and characteristically have a necrotic center, peritumoral edema, and a thick irregular margin. As with CT, post-contrast T1-weighted imaging shows irregular infiltrative enhancement. On T1WI, the necrotic center will appear hypointense, while an enhanced irregular border surrounds the mass. Any peripheral non-enhancing areas most likely represent vasogenic edema. It is important to note that the enhanced ringlike structure is not the true border of the tumor since malignant glioma cells can be found beyond a 2-cm margin. On T2WI, GBMs also appear heterogeneous due to necrosis, hemorrhage, or tumor vascularity. The rim of vasogenic edema is better appreciated on T2WI. Unfortunately, distinguishing tumor from edema is challenging due to similar signal characteristics on T2 [30] (Fig. 10.5a–d).

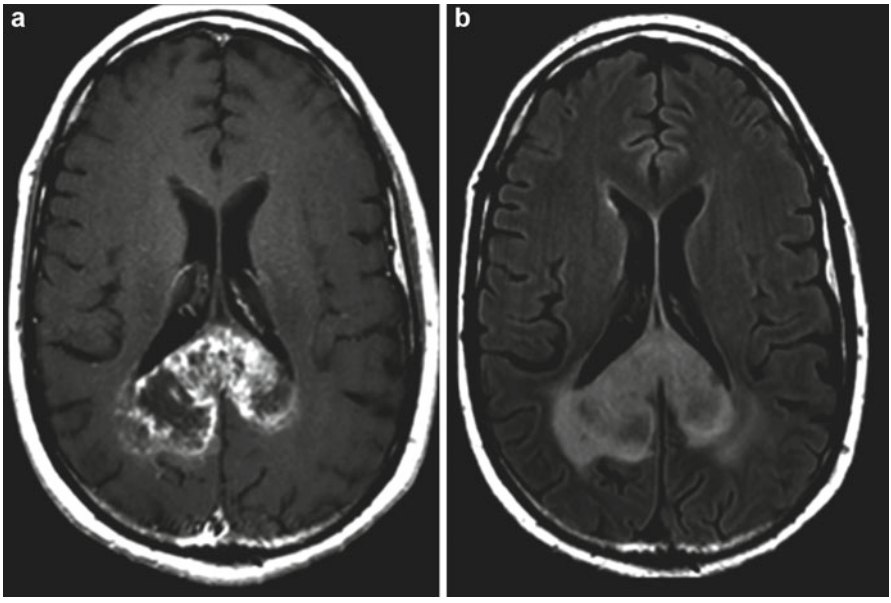


Fig. 10.5 Example of a grade IV glioblastoma. (**a**) Post-gadolinium T1WI shows a heterogeneously enhancing mass centered on the splenium of the corpus callosum; (**b**) T2W FLAIR imaging of the same patient demonstrates the corresponding mass-like T2 signal abnormality; (**c**) MR perfusion rCBV function map clearly shows elevated rCBV within the neoplasm (*black arrow*) indicating marked tumor neovascularity; (**d**) post-gadolinium T1WI of a different GBM patient demonstrates secondary leptomeningeal dissemination along the septum pellucidum (*white arrow*)

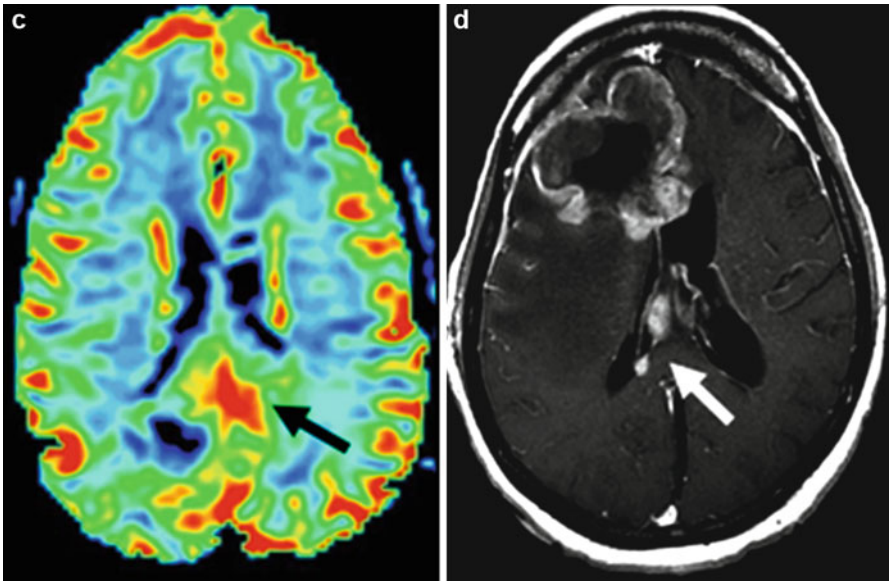


Fig. 10.5 (continued)

On DWI, ADC values have been shown to correlate with tumor cellularity and tumor grade. GBMs have high cellularity compared to lower-grade gliomas. The increased cellular density causes restricted diffusion in the extracellular space, which produces low ADC values. As a result, high-grade gliomas present as hyperintense lesions on DWI.

Anatomically, GBMs can spread along white matter tracts and cross hemispheres via the corpus callous (Fig. 10.5a–d). Other common pathways for dissemination include the optic radiations, the internal capsule, the anterior commissure, and posterior commissure. GBMs remain solitary lesions 99 % of the time. Rare multicentric tumors commonly spread via the meningeal-subarachnoid space and appear widespread throughout the brain without evidence of intracranial connections (those with some continuity are multifocal) [32]. They can also disseminate throughout the brain via subependymal, intraventricular, or direct penetration routes through white matter. To support its own tumor growth, GBMs undergo angiogenesis, which is a critical parameter to classify tumor grade.

Determination of the regional vascularity is an important parameter to grade gliomas. While contrast enhancement can indicate a disruption in the blood-brain barrier, it cannot reveal the vascularity of the tumor. Furthermore, as many as 33 % of high-grade gliomas demonstrate no significant contrast enhancement and 20 % of low-grade gliomas demonstrate contrast enhancement. Perfusion-weighted imaging can be used to assess the underlying vascularity of the tumor. Regional cerebral blood volume (rCBV) calculated from perfusion-weighted imaging is a sensitive marker of the microvasculature of tumor (Fig. 10.5a–d) [33]. Numerous studies show that rCBV correlates positively with tumor grade [33, 34]. Low-grade gliomas

typically have rCBV value that range between 1.11 and 2.14, while high-grade gliomas range between 3.54 and 7.32 [35]. CBV mapping also has the added benefit of directing stereotactic biopsy sites to highly vascular areas so as to reduce sampling errors in histopathologic diagnosis of gliomas [33].

Studies suggest that MRSI through the calculation of metabolite ratios and levels can be helpful in grading tumors (Fig. 10.5a–d). Choline (Cho) is a cell membrane marker that increases during high rates of cell membrane turnover or proliferation; elevations in Cho commonly represent high-grade tumors. It is important to note that while low levels of Cho are typically associated with low-grade tumors, in GBMs with extensive necrosis, Cho levels may be low. Creatine (Cr), a marker for energy metabolism, is commonly low in high-grade tumors that consume energy due to increased metabolic activity. N-acetylaspartate (NAA), a marker for neuronal integrity, is also commonly low when high-grade tumors are present. In high-grade gliomas, lactate levels are increased after the tumor outgrows the blood supply and begins to utilize anaerobic glycolysis. Myoinositol, a glial cell marker, can also be used to distinguish high-grade and low-grade gliomas because GBMs tend to have low levels of this metabolite. Ratios of these metabolites, namely, Cho/Cr, NAA/Cr, and NAA/Cho, can aid in the grading of tumors [12, 18]. Table 10.2 summarizes MRSI findings useful in grading tumors.

The differential diagnosis of GBM includes abscess, metastasis, lymphoma, and tumefactive multiple sclerosis. Brain abscesses, in particular, remain a diagnostic challenge because the presenting clinical manifestations and neuroradiologic appearances often overall with GBMs. Imaging findings of abscesses that may be helpful in differentiating include a thin wall with ring-like enhancement that is often thinner along the medial margin, daughter rings, and a hypointense rim on T2-weighted images.

Metastatic Tumors

Metastatic brain tumors can be difficult to differentiate from GBMs, which can be a challenge in many clinical trials [29, 31]. Metastatic brain tumors originate from a

Table 10.2 MRSI in tumor grading

MRSI marker	Low-grade tumor (I, II)	High-grade tumor (III, IV)
Choline (Cho)	↔ or ↓	↑↑
Creatine (Cr)	↑	↓↓
N-Acetylaspartate (NAA)	↓	↓↓
Cho/Cr	↓↓	↓↓↓↓
NAA/Cr	↓↓	↓↓↓↓
NAA/Cho	↓	↓↓↓↓
Lactate	↔ or ↓	↓↓↓
Myoinositol	↑↑↑↑	↑
Lipids	↔ or ↑	↑↑↑

primary systemic malignancies and lymphoma. Current therapy for MRI documented solitary metastasis is surgical removal, but this constitutes only 25–35 % of all patients because many are not appropriate candidates for surgery (non-solitary lesions, inaccessible tumor location, etc.).

On non-contrast head CT, metastatic brain tumors appear hypodense to isodense and are frequently well circumscribed with some having low-attenuation vasogenic edema or a necrotic low-attenuation center. Rarely, metastases may appear hyperdense indicating hemorrhagic origin from lung, breasts, thyroid, melanoma, or renal cell cancers. Hyperdensity can also indicate high cellularity as seen in lymphoma or small-cell lung cancer. Like in GBMs, most metastases will enhance on contrast-enhanced CT [34].

On MRI, metastatic brain tumors appear on T2WI as isointense to hyperintense compared to gray matter with edema that is even more hyperintense and predominately infiltrates the white matter (Fig. 10.2a–c) [30, 34]. On T1WI, these lesions appear hypointense to isointense, with surrounding hypointense edema. Central hyperintensity on T2WI is indicative of necrotic centers or cystic tumors. In contradistinction, internal T2 hypointensity is indicative of a high nuclear-to-cytoplasmic ratio found in cellular metastases such as lymphoma [34]. Like GBMs, almost all metastases enhance on post-gadolinium-enhanced MRI. However, this enhancement can be greatly altered by steroid therapy, radiation, or chemotherapy [30, 34].

Unfortunately, primary tumors and metastases frequently have similar enhancement patterns. Extension of tumor and edema into the corpus callosum or gray matter is more likely seen in GBMs. On the other hand, multiple lesions are more likely found in metastases. MRS also demonstrates elevated choline to creatine ratios with a decrease in NAA as it does with gliomas [18]. Also similar to GBMs, rCBV in perfusion-weighted imaging is often increased; however, in the peritumoral regions the rCBV is almost always higher in gliomas than metastatic lesions, a finding that can help in differentiations.

WHO Grade II Astrocytomas

Low-grade astrocytomas typically lack mitotic figures, necrosis, and vascular proliferation, which are all hallmarks of higher-grade gliomas. Well-differentiated diffuse infiltrative astrocytomas are classified as WHO grade II and comprise approximately 25 % of all gliomas. There are numerous subtypes, but the most frequently encountered is the fibrillary astrocytoma [29–31].

Grade II astrocytomas typically occur in the cerebral hemispheres in adults or in the brainstem for children [31] (Fig. 10.6a–c). Cerebral astrocytomas are often seen in the frontal and temporal lobes and have a far better prognosis than brainstem astrocytomas, partially due to challenges in surgical resection at this location. Low-grade astrocytomas can be non-infiltrative or they can infiltrate surrounding brain; all subtypes can progress to higher grades.

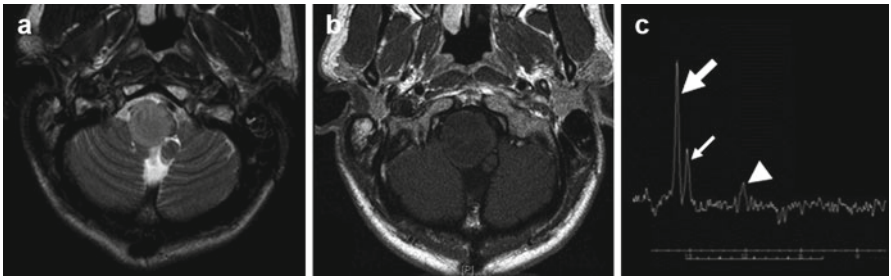


Fig. 10.6 Example of a grade II astrocytoma. (**a**) T2WI demonstrates an expansile T2 hyperintense brain stem mass. (**b**) Post-gadolinium T1WI demonstrates no internal enhancement, typical of most grade II astrocytomas. (**c**) MRS demonstrates a choline to creatine (Cho:Cr) ratio of 2:1 characteristic of a lower-grade neoplasm and a reduction of NAA, the normal neuronal marker (Cho, *thick arrow*; Cr, *thin arrow*; NAA, *arrowhead*)

Diffuse Infiltrative Low-Grade Astrocytomas

These tumors typically appear homogeneous and rarely enhance on CT, making them difficult to detect. As with all brain tumors, MRI is markedly more sensitive than CT, although the tumor's borders can extend well beyond the margins identified on MR. On T1WI, they are hypointense while they appear hyperintense on T2WI with little or no enhancement following contrast administration. Still, contrast enhancement is not a reliable marker for high histopathologic grade since even low-grade tumors can enhance on occasion [30]. These tumors may have cystic foci, but typically lack the peritumoral edema and mass effect seen in higher grades.

Numerous imaging signs and findings can be used as prognostic indicators. For example, large tumor volume and involvement in many lobes is associated with progression to higher grades and worse outcome. Similarly, high cerebral blood volumes are predictive of a worse prognosis, as these are associated with microvascularity and angiogenesis.

Non-infiltrative Low-Grade Astrocytomas

Unlike infiltrative or higher-grade astrocytomas, these tumors have well-circumscribed borders. There are numerous subtypes with varying imaging characteristics, but most have good prognosis. In general, these low-grade tumors are treated by surgical resection and are not included in many clinical trials. Common subtypes include pilocytic, pleomorphic xanthoastrocytomas, and subependymal giant cell astrocytomas which are reviewed in the following subsections [36].

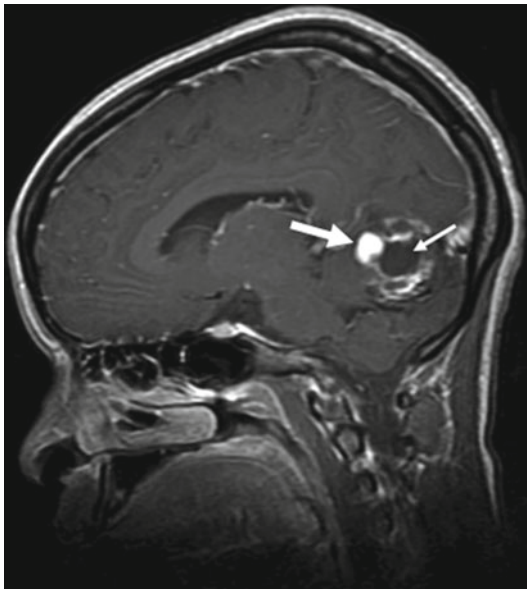


Fig. 10.7 Example of a grade I pilocytic astrocytoma. A hypervascular enhancing mural nodule (*thick arrow*) with non-enhancing cystic component (*thin arrow*) is characteristic of a pilocytic astrocytoma on this post-gadolinium sagittal T1WI

Pilocytic Astrocytoma

Pilocytic astrocytomas are classified as WHO grade I neoplasms characterized by spongiform tissue containing microcysts. These slowly growing tumors sometimes spontaneously regress. On imaging they appear classically as cystic tumors with an enhancing mural nodule. On CT, the solid component can be hypodense (43 %), isodense (51 %), or hyperdense (6 %). On MRI these tumors have pronounced contrast enhancement 94 % of the time due to increased vascularity; however, increased vascularity is not indicative of higher histological grade when present in pilocytic astrocytomas. On T1WI, the tumor is hypointense compared to gray matter and hyperintense compared to gray matter on T2-weighted scans [30, 36] (Fig. 10.7).

Pleomorphic Xanthoastrocytomas

Pleomorphic xanthoastrocytomas are classified as WHO grade II tumors and unlike pilocytic astrocytomas are more prone to malignant transformation. Typically, they are found peripherally along the temporal lobe leptomeninges. They are difficult to distinguish from gangliomas due to cysts and mural nodules. On CT, these tumors are cystic without calcifications and are hypervascular, receiving their blood supply from the meningeal arteries. On T1WI, these tumors are isointense to gray matter, but hyperintense on T2WI. Meningiomas, pilocytic astrocytomas, gangliogliomas, and oligodendrogliomas should be on the differential when evaluating pleomorphic xanthoastrocytomas due to the similar appearance on imaging [36].

Designing a Neuro-oncology Clinical Trial

In phase II studies that have the goal of evaluating the effectiveness of new therapeutic agents rely heavily on radiographic assessments. Phase I and phase III studies, on the other hand, have different end points (toxicity and survival) and therefore do not rely as greatly on neuroimaging assessments. In phase II trials, the radiographic response serves as the major end point. In designing a phase II trial, the subjects enrolled are usually those with progressive tumors. Both enrollment (i.e., defining progressive disease) and the end point of the trial are based on definitions from the relevant response criteria (e.g., RECIST or MacDonald criteria). Most clinical trials occur in patients with malignant gliomas where complete response is rare. In these cases, a better end point is duration of stable disease or PFS. The length of these trials is usually based on PFS after 6 or 12 months. Time to progression is another possible end point in phase II trials. It is important to note that actual tumor size reduction is a better marker of effective therapeutic response (which PFS does not require) than duration of stable disease.

In designing the clinical trial, a baseline imaging study is recommended, ideally within 2 weeks of treatment initiation. Subsequent scan acquired serially based on length of time (e.g., 2 months) or treatment cycles (e.g., every two treatments) are necessary based on the treatment and research protocol. The investigator should commit to one imaging modality for the entirety of the study, almost always preferably MRI.

There are several newer imaging techniques beyond CT and conventional MRI reviewed earlier in this chapter that investigators might want to include in the imaging protocol. These techniques which include MRSI, dynamic MR perfusion (rCBV), and diffusion tensor imaging can help distinguish pseudoresponse or postsurgical effects, yield additional quantitative parameters, and may be more sensitive to the detection or pathologic processes of tumors.

Multiple Sclerosis (MS)

Over the past several decades, our understanding of MS has evolved; currently it is thought to be a complex disease process with an unpredictable course. MS can be relatively benign to overwhelmingly disabling. A simplified description of MS is that it is a demyelinating disease. An accurate understanding of the disease process remains elusive, but many investigators believe it to be autoimmune in nature and that MS attacks may be linked to an unknown environmental trigger such as a virus [37].

While there is currently no cure for MS, there are several therapies in existence and several more currently being investigated [37]. MRI is the primary imaging technique used to evaluate the efficacy of clinical trials for novel neurotherapeutics. MRI is essential for initial evaluation to rule in or rule out disease, as a prognostic

tool at first presentation, and serves as the primary outcome measure in clinical trials. MS therapeutics aim at reducing symptoms, delaying relapses, and delaying long-term disability [37]. As such, a combination of imaging and clinical markers is necessary to evaluate the efficacy of novel therapies. The following sections will summarize the clinical markers used to assess disease activity in addition to how MRI is used in clinical trials of MS.

Clinical Markers for Multiple Sclerosis

Standard measures of MS severity include the Expanded Disability Status Scale (EDSS) and the Multiple Sclerosis Functional Composite (MSFC), both of which focus on physical and cognitive function. Clinical markers to assess MS activity include time to relapse, annual relapse rate, and perhaps more importantly, quality of life. EDSS or the newer MSFC are typically employed in most clinical trials.

EDSS, Expanded Disability Status Scale

EDSS serves as a commonly used scale to measure disability in MS and is currently approved by drug agencies such as the FDA and EMEA as a clinical marker. EDSS is a 20-step (scored from 0 to 10 in half steps) overall disability rating scale in MS [38]. EDSS assesses seven functional categories including visual, brainstem, pyramidal, cerebellar, sensory, bowel/bladder, and cerebral. EDSS is particularly insensitive to clinical change not associated with impaired gait and in those with milder disease. Most clinical trials in the past use time to change by one step in EDSS in a survival analysis as the main outcome measure. While common, this metric does not take into consideration that progression is frequently reversible in many cases. Still, EDSS remains the gold standard for clinical assessment [38, 39].

MSFC, Multiple Sclerosis Functional Composite

In 1997, the National Multiple Sclerosis Society Clinical Assessment Task Force presented the MSFC in response to the weaknesses in EDSS. MSFC is a metric scale (raw scores are converted to z-scores based on a reference population) [38]. The MSFC is a three-part quantitative assessment that includes the Timed 25-FootWalk, the 9-Hole Peg Test (for arm function), and the Paced Auditory Serial Addition Test (for cognitive function). MSFC changes were found to reflect severity of disease and in the IMPACT study was shown to be more sensitive than EDSS [40]. Unfortunately, the three domains of function assessed by MSFC are infrequently affected in those with early stage or mild disease. Furthermore, visual and

sensory deficits are often found in isolation, although these deficits have a significant impact on the MSFC overall [38, 39].

Imaging of Multiple Sclerosis

MRI serves as the central marker for outcome in many MS clinical trials. Furthermore, several criteria depend on MRI findings to diagnose MS. A typical imaging protocol for MS includes [41]:

- T2W FLAIR (sagittal and axial and/or 3D volumetric) or dual-echo proton density (axial)
- T2W fast spin echo (axial)
- Gadolinium-enhanced T1WI (axial and sagittal T1W fast spin echo and/or 3D volumetric T1W spoiled gradient echo [SPGR])

More advanced sequences sometimes acquired include:

- Magnetization transfer (MT) imaging
- Dual inversion-recovery imaging
- Diffusion tensor imaging (DTI)
- 2D and 3D chemical shift imaging
- Magnetic resonance spectroscopy (MRSI)

On conventional MRI, new lesions are visualized on gadolinium-enhanced T1WI (Fig. 10.8a–d) as nodular enhancement in previously normal-appearing white matter (NAWM). Lesions are commonly seen as hyperintense on T2W scans. In about two-thirds of cases, these lesions appear hypointense on non-contrast T1WI [42]. As a lesion evolves the associated enhancement typically fades by 2 months and it can become T1 isointense with white matter when remyelination occurs. Aggressive lesions may have ring-like enhancement and central spherical hypointensity on T1-weighted scans. An “open ring sign” is characteristic of MS, which occurs when the lesion lies alongside gray matter (referred to as juxtacortical lesions). Complete rings occur when lesions are completely within white matter. Lesion T2-hyperintensity also reduces over time as edema resolves and the lesion remyelinates. Brainstem and spinal cord lesions often resolve completely on MRI whereas T2 lesions in other areas frequently remain somewhat evident. Hyperintensities on T2WI are nonspecific and can be indicative of Wallerian degeneration, gliosis, inflammation, edema, demyelination, or axonal loss [37]. These lesions frequently appear in the periventricular region spreading along the periventricular collecting veins. On sagittal T2W FLAIR, these characteristic lesions appear hyperintense, forming the so-called Dawson’s fingers (Fig. 10.8a–d).

Unfortunately, lesion evolution is quite complex, as is the course of MS. T1-hypointense resolution often coincides with fading of enhancement, but this is not always the case. It appears that enhancement is more common in older patients where as hypointense lesions are more common in those with progressive disease [37]. Persistent T1- hypointense lesions are indicative of irreversible axonal loss.

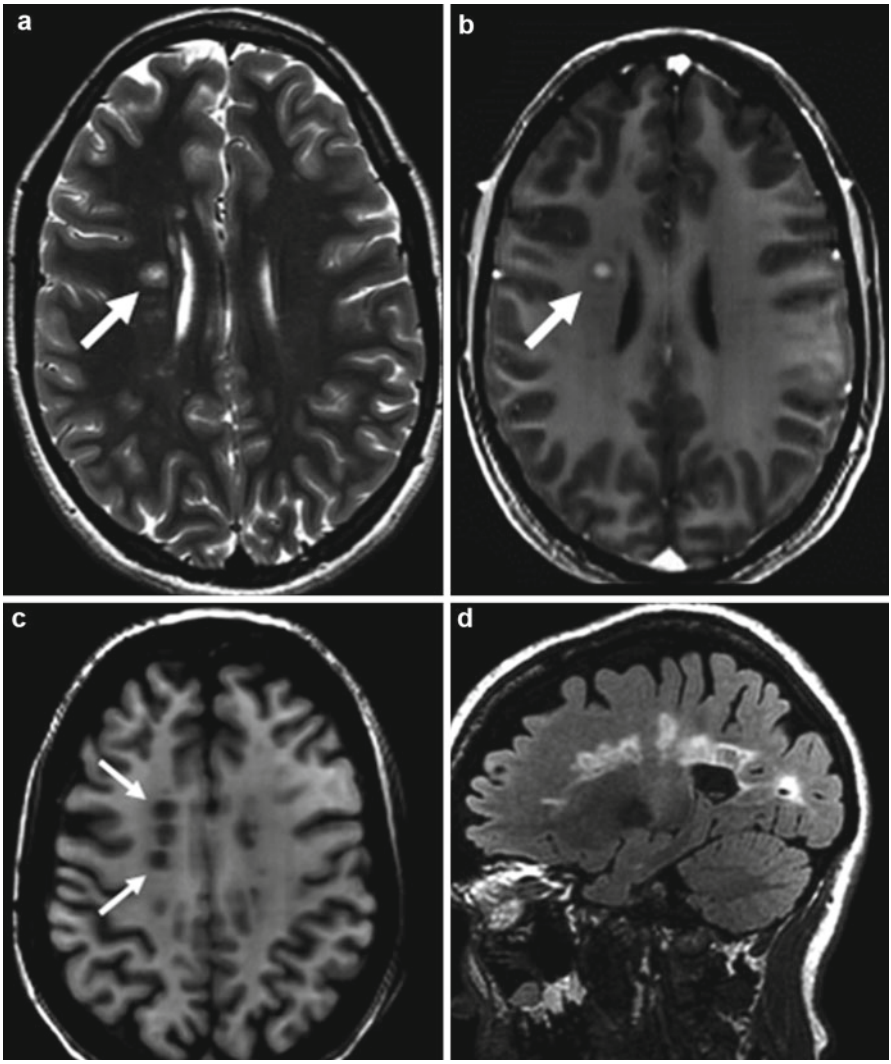


Fig. 10.8 Multiple sclerosis. (**a**) Hyperintense lesion on T2WI that enhances on post-gadolinium T1WI (**b**). Enhancing lesions are indicative of active disease. (**c**) T1WI demonstrates multiple hypointense lesions ("*black holes*") on a separate patient. T1-hypointense lesions persistent over time are indicative of irreversible axonal loss. (**d**) Sagittal T2 FLAIR demonstrates the characteristic "Dawson's fingers"

Clinical Correlations and Role in Clinical Trials of MRI

As mentioned previously, T2-hyperintensity is nonspecific so it is not surprising that correlation of T2 burden of disease (T2 BOD) with clinical impairment and outcome is poor. Additionally, T2 lesions may occur in areas of the brain that are clinically silent. Still, T2 BOD may play a role in clinically isolated syndrome in terms of

prognosis. The Early Treatment of MS study (ETOMS) reported that CIS patients with at least 9 T2-hyperintense lesions and 1 Gd-enhancing lesion progressed to MS in 41 % of cases versus only 11 % in those without either finding [40]. Alternatively, CIS patients with more than 1.2 cubic cm of T2 BOD have over a 50 % risk of developing MS [40]. In clinical trials, new or enlarging T2-hyperintensities are used to monitor treatment. T2 BOD decrease or subsequent scans without new T2-hyperintense lesions can be used as end points in clinical trials [41, 42].

Enhancing T1 lesions in MS are commonly caused by increased permeability of the blood-brain barrier due to inflammation. Clinically, the appearance of enhancing lesions is highly correlated with clinical symptoms when the lesions occur in a location that is not clinically silent. For example, Bruck and colleagues [43] showed that enhancing lesions over the optic nerve in 94 % of patients had acute optic neuritis. When used in clinical trials as a measure of disease activity, it is important to note that drugs that suppress T1 enhancement do not necessarily have clinical benefit. On the other hand, all disease-modifying agents currently used for MS are known to affect T1 enhancement. Thus, trials using enhancement as a primary outcome measure are designed to test whether the therapeutic has potential before proceeding to a large-scale clinical trial [41, 42].

New T1-hypointense lesions can be due to a combination of inflammation, edema, gliosis, and demyelination (or early remyelination). T1-hypointense lesions with minimal enhancement on gadolinium-enhanced T1WI are likely due to edema or partial remyelination that commonly resolves and becomes isointense within 6 months. T1-hypointense lesions, or black holes, have strong clinical correlation with disability measures such as EDSS. Total T1-hypointense lesion volume and its change over time correlate significantly with impairment. Those with T2-hyperintense lesions that are also T1-hypointense are more likely to be progressive. In clinical trials, longitudinal change of T1-hypointense lesion volume can serve as an outcome measure. Additionally, measurements of changes in normalized brain volume and brain parenchymal fraction can also serve as biomarkers for quantification of neurodegeneration secondary to MS [41].

MS Criteria Involving Neuroimaging

Poser Criteria

The Poser criteria was developed in 1983 to incorporate technological advances (i.e., MRI) in the diagnosis of MS [44]. The Poser criteria are:

- Clinically definite MS
  - 2 attacks and clinical evidence of 2 separate lesions
  - 2 attacks, clinical evidence of one, and paraclinical evidence of another separate lesion

- Laboratory-supported definite MS
  - 2 attacks, either clinical or paraclinical evidence of 1 lesion, and cerebrospinal fluid (CSF) immunologic abnormalities
  - 1 attack, clinical evidence of 2 separate lesions, and CSF abnormalities
  - 1 attack, clinical evidence of 1 and paraclinical evidence of another separate lesion, and CSF abnormalities
- Clinically probable MS
  - 2 attacks and clinical evidence of 1 lesion
  - 1 attack and clinical evidence of 2 separate lesions
  - 1 attack, clinical evidence of 1 lesion, and paraclinical evidence of another separate lesion
- Laboratory-supported probable MS
  - 2 attacks and CSF abnormalities

McDonald Criteria

The McDonald criteria was developed in 2001 with the intention to replace the older Poser criteria making use of advances in MRI for the diagnosis of MS. The McDonald criteria is shown in Table 10.3 [45].

An attack is defined as:

- Neurological disturbance of kind seen in MS.
- Subjective report or objective observation.
- 24 h duration, minimum.
- Excludes pseudo-attacks, single paroxysmal episodes.
- 2 separate attacks must occur at least 30 days between the onset of two events.

An "abnormality" is defined on MRI as having three of the following four:

- 1 gadolinium-enhancing or 9 T2-hyperintense lesions if no gadolinium-enhancing lesion
- 1 or more infratentorial lesions
- 1 or more juxtacortical lesions
- 3 or more periventricular lesions
- (1 spinal cord lesion = 1 brain lesion)

MRI evidence of dissemination in time is defined as having one of the following:

- A Gd-enhancing lesion demonstrated in a scan done at least 3 months following onset of clinical attack at a site different from attack
- In the absence of Gd-enhancing lesions at 3-month scan, follow-up scan after an additional 3 months showing Gd lesion or new T2 lesion

Table 10.3 MacDonald criteria for use of MRI in diagnosis in MS

Clinical presentation	Additional data needed
2 or more attacks (relapses) 2 or more objective clinical lesions	None; clinical evidence will suffice (additional evidence desirable but must be consistent with MS)
2 or more attacks 1 objective clinical lesion	Dissemination in space, demonstrated by: MRI Or a positive CSF and 2 or more MRI lesions consistent with MS Or further clinical attack involving different site
1 attack 2 or more objective clinical lesions	Dissemination in time, demonstrated by: MRI Or second clinical attack
1 attack 1 objective clinical lesion (monosymptomatic presentation)	Dissemination in space by demonstrated by: MRI Or positive CSF and 2 or more MRI lesions consistent with MS *and* Dissemination in time demonstrated by: MRI Or second clinical attack
Insidious neurological progression suggestive of MS (primary progressive MS)	Positive CSF *and* Dissemination in space demonstrated by: MRI evidence of 9 or more T2 brain lesions Or 2 or more spinal cord lesions Or 4–8 brain and 1 spinal cord lesion Or positive VEP with 4–8 MRI lesions Or positive VEP with <4 brain lesions plus 1 spinal cord lesion *and* Dissemination in time demonstrated by: MRI Or continued progression for 1 year

An "abnormality" on cerebrospinal fluid (CSF) is defined as:

- Oligoclonal IgG bands in CSF (and not serum)
- Elevated IgG index

An "abnormality" on evoked potentials (EP) is defined as having a delayed but well-preserved wave form.

Barkhof Criteria

The Barkhof criteria has been shown to have high specificity for predicting conversion from clinically isolated syndrome to MS or, in other words, in predicting a second relapse after a first episode suggestive of MS [41, 44].

Having at least three out of the following four criteria is suggestive of MS:

- At least nine lesions on the T2-weighted images
- Presence of at least three periventricular lesions
- Presence of at least one juxtacortical lesion
- Presence of at least one infratentorial lesion

Stroke

Neuroimaging has become an integral tool for the diagnosis and management of stroke. Acutely, imaging is used distinguish ischemic stroke from hemorrhagic stroke. Currently, the standard of care to rule out hemorrhagic stroke is a non-contrast CT scan. If there is no evidence of hemorrhage within a stroke, then intravenous thrombolysis with tissue plasminogen activator (tPA) can be safely administered within 3 (and more recently 4.5) hours of stroke onset given no other contraindications exist. While it is possible to define ischemic stroke with higher sensitivity using MRI, emergent availability, contraindications, and difficulty interpreting results within a clinically meaningful time frame prevent more routine use of MRI [46].

Parenchymal changes on CT in acute ischemia include hypo-attenuation and sulcal effacement. The hypo-attenuation likely represents an increase in water content due to fluid shifts from the interstitial to intracellular space. At hyperacute periods, the hypo-attenuation may be due to a decrease in cerebral blood volume (CBV). In these areas, the hypo-attenuation is characteristic of irreversibly injured tissue representing the infarct. The sulcal effacement is likely secondary to vasodilatation and a resultant increase in CBV. Sulcal effacement is characteristic of the penumbra that may be salvaged with reperfusion of the area [46]. On CT, one may use the Alberta Stroke Program Early CT Score (ASPECTS) to assess the extent of ischemia within the distribution of the middle cerebral artery. ASPECTS is a 10-point scale in which 10 is normal and 0 reflects ischemia along the entire distribution [47].

With MRI, T1WI, T2WI, and T2* GRE or SWI can be used to evaluate the age of blood products in the parenchyma. Compared to CT, MRI is superior for the detection of subarachnoid blood, particularly at later time points or when the bleed is small. T2W FLAIR imaging is the sequence of choice for the detection of subarachnoid blood. Table 10.4 summarizes the appearance of blood on MRI at varying time points [46].

On MRI, the most powerful sequence to assess for stroke is DWI; strokes can be visualized within minutes on DWI and will remain positive for up to 2 weeks (Fig. 10.9a–d). DWI has been shown to be greater than 90 % sensitive to hyperacute infarction, vastly superior to CT (40–60 % sensitivity) [46].

Both CT and MRI can be used to image the cerebral vasculature (Fig. 10.9a–d). Magnetic resonance angiography (MRA) can be performed with or without contrast agents. When using contrast agents, the signal-to-noise ratio is improved and thus smaller vessels can be evaluated. Furthermore, with contrast agents, MRA can

Table 10.4 Appearance of blood on MRI at varying time points

	Blood constituents	T1	T2	SWI/T2*
Hyperacute <24 h	Oxyhemoglobin Deoxyhemoglobin Plasma	↔	↓ (rim) ↑ (core)	↓
Acute (1–3 days)	Deoxyhemoglobin Plasma	↔	↓	↓
Early subacute 3–7 days	Intracellular Methemoglobin Plasma	↑	↓	↓
Late subacute 7–14 days	Extracellular Methemoglobin	↑	↑	↓
Chronic >14 days	Hemosiderin	↓	↓	↓

better visualize stenotic areas. CT angiography (CTA) always requires administration of a contrast agent. CTA is a much faster technique and allows for rapid identification of arterial stenoses, dissections, and occlusions [46].

Both CT and MRI can perform perfusion (or for MRI, perfusion-weighted) imaging using bolused contrast agents. On MRI, the changes in intensity are not linear so absolute values of CBF are not possible. However, it possible to obtain quantitative flow estimates by normalizing the perfusion maps to a reference value. CT perfusion, on the other hand, can be used to measure absolute perfusion values (CBF, CBV, TTP, MTT, Tmax). These techniques can be used to distinguish the infracted area from the penumbra (Fig. 10.9a–d) [13, 46].

Alzheimer's Disease (AD)

Alzheimer's disease (AD) is the most common cause of dementia in elderly people, and given the aging population, this condition is growing to vast proportions. There is a recognized need for effective treatments, particularly disease-modifying agents. To test these therapies, reliable biomarkers are required. In addition to clinical makers and certain chemical markers, neuroimaging has become the primary technique to gauge response in clinical trials. Furthermore, neuroimaging techniques may potentially be the modality of choice for diagnosis and prognosis. In fact, four imaging modalities have been used as secondary end points for AD clinical trials. These include MRI, fMRI, MRS, and PET [48–50].

The Alzheimer's Disease Neuroimaging Initiative (ADNI) examined how brain imaging can be used to measure progression of mild cognitive impairment and AD with the purpose of providing a standard assessment tools to aid future clinical trials [49]. The ADNI MRI core recommends the following two sequences as a minimum requirement for future AD neuroimaging studies: 3D T1-weighted sequenced (Siemens MPRAGE or GE's IR-FSPGR) and 2D T2W FLAIR sequence.

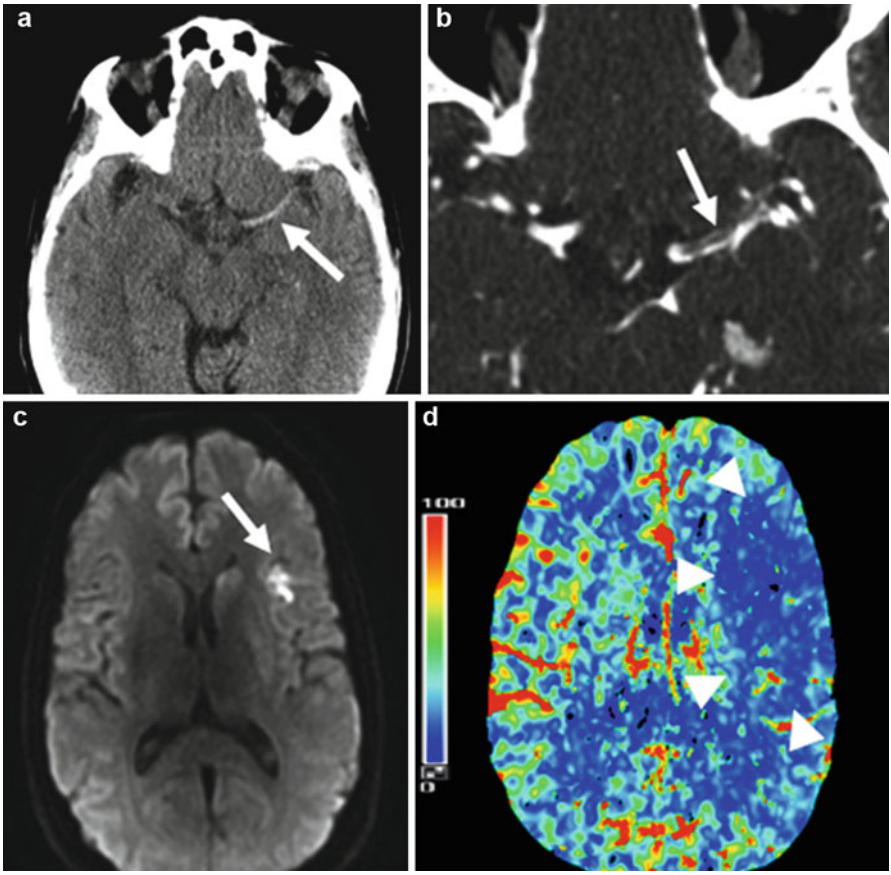


Fig. 10.9 Acute ischemic stroke. (**a**) Non-contrast CT demonstrates asymmetric hyperdensity within the left middle cerebral artery (MCA) also known as the "hyperdense MCA sign"; this finding is indicative of acute thrombosis, an early sign of ischemic stroke. (**b**) Contrast-enhanced CT angiography demonstrates the thrombus within the left MCA (*arrow*) resulting in occlusion. (**c**) DWI shows a small core infarction (*arrow*) within the anterior insula that appears hyperintense due to restricted diffusion secondary to cytotoxic edema. (**d**) CT perfusion cerebral blood flow (CBF) function map demonstrates a large penumbra of decreased blood flow involving the entire left MCA territory (*arrowheads*) indicative of tissue at imminent risk for infarction

Additionally, a gradient echo (GRE) or susceptibility-weighted (SWI) sequence should be acquired for micro-hemorrhage detection. Given the increased SNR, imaging was recommended exclusively at ultrahigh-field 3-T magnets. Additional advanced sequences that can be obtained for AD studies include arterial spin-labeling perfusion imaging (ASL) and diffusion tensor imaging (DTI) [49].

In AD, MRI typically shows decreased gray matter in the parahippocampus, the hippocampus, the amygdala, the posterior association cortex, and the subcortical nuclei (Fig. 10.10a, b). Using MRI, potential biomarkers for AD include

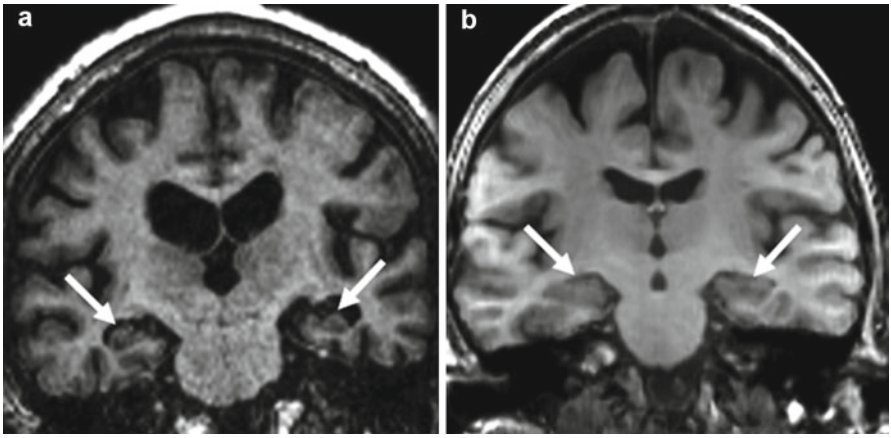


Fig. 10.10 (**a**) Coronal T1-FSPGR MRI of an Alzheimer's patient shows atrophy of the hippocampi (*arrows*). (**b**) Normal hippocampi are included for comparison

hippocampus volume, entorhinal cortex volume, whole brain volume, cortical thickness, and patterns of regional gray matter atrophy (using voxel-based morphometry). If longitudinal scans are collected, the rate of atrophy can be calculated as a potential biomarker [49].

MRS can also be used as a modality to assess AD. NAA levels have been found to be decreased in AD, independent of atrophy. Other compounds such as choline, creatine, myoinositol, and glutamine may also be potentially affected by AD [48].

Studies indicate that fMRI activation patterns in AD patients are altered in the temporal and parietal lobes. Specific treatments that alter attention networks in AD may produce a measurable effect in the fMRI BOLD signal in these areas. Given the wide availability of fMRI, its use as an end point in multicenter clinical trials is reasonable [48].

As part of the ADNI study, a PET core was devised to test whether PET could serve as an outcome measure to track drug effects [50]. Initial studies examined the efficacy of FDG-PET in AD since FDG uptake is reduced secondary to neuronal activity impairment. Changes in the FDG-PET signal correlated with cognitive impairment and can be used as a secondary outcome measure in AD clinical trials. More recently, several studies have shown that the Pittsburgh compound B (^{11}C-PIB), a ^{11}C-labelled thioflavin analogue that binds selectively to amyloid beta, can be visualized in PET scans of patient with AD. Amyloid PET can be used to measure the amyloid-beta plaque burden in AD patients, independent of changes in brain atrophy. While it is unclear if amyloid PET measures serve as a useful end point, it may be used as a prognostic or diagnostic indicator [50].

A combination of MRI, fMRI, MRS, and PET is likely to hold the most value in clinical trials to monitor drug effects. Additionally, these biomarkers can be used as clinical diagnostic and prognostic indicators [48].

Medical Devices: Contraindications

There is an ongoing challenge to maintain a safe environment when performing medical imaging. While medical devices are generally safe for CT, many devices such as stents, neurostimulators, and pacemakers are potential contraindications to MRI. Since new devices are released every day, there are websites available that endeavor to maintain updated information on this topic (e.g., www.mrisafety.com). To complicate matters further, some devices may be safe at 1.5 T or lower fields but unsafe at 3.0 T or higher fields. Moreover, devices that may be safe for imaging can produce artifacts on both CT and MRI that may render portions of the scan unreadable [51].

Thousands of objects have been evaluated for their use in MRI [51]. Implanted devices may become heated during an MRI or its function may be altered. Other devices may be displaced or torqued in the presence of a magnetic field. The MR Task Group of the American Society for Testing and Materials (ASTM) put forth a new set of terms defining an objects relative safety in an MR scanner [51]. A "MR safe" item is one that poses no known hazard in all MRI environments. These include nonconducting, nonmetallic, and nonmagnetic items. "MR conditional" items are those that pose no known hazards in specific MRI environments. These environments are defined using field conditions such as static field strength, spatial gradient, time-varying magnetic field, radiofrequency fields, and specific absorption rate. Additional conditions (such as the routing of leads for neurostimulators) may be specified as well. Finally, "MR unsafe" items are those with known hazards in all MRI environments [51].

References

1. Kumar RA, Khandelwal N, et al. Comparison between contrast-enhanced magnetic resonance imaging and technetium 99m glucohepatonic acid single photon emission computed tomography with histopathologic correlation in gliomas. J Comput Assist Tomogr. 2006; 30(5):723–33.
2. Brody AS. New perspectives in CT and MR imaging. Neurol Clin. 1991;9(2):273–86.
3. Rogalla P, Kloeters C, et al. CT technology overview: 64-slice and beyond. Radiol Clin North Am. 2009;47(1):1–11.
4. Rutten A, Prokop M. Contrast agents in X-ray computed tomography and its applications in oncology. Anticancer Agents Med Chem. 2007;7(3):307–16.
5. Bettmann MA, Heeren T, et al. Adverse events with radiographic contrast agents: results of the SCVIR Contrast Agent Registry. Radiology. 1997;203(3):611–20.
6. Runge VM. A review of contrast media research in 1999–2000. Invest Radiol. 2001;36(2):123–30.
7. Niogi SN, Mukherjee P. Diffusion tensor imaging of mild traumatic brain injury. J Head Trauma Rehabil. 2010;25(4):241–55.
8. Burtea C, Laurent S et al. Contrast agents: magnetic resonance. Handb Exp Pharmacol. 2008;(185 Pt 1):135–65.
9. Lee B, Newberg A. Neuroimaging in traumatic brain imaging. NeuroRx. 2005;2(2):372–83.

10. Bakshi R, Kamran S, et al. Fluid-attenuated inversion-recovery MR imaging in acute and subacute cerebral intraventricular hemorrhage. AJNR Am J Neuroradiol. 1999;20(4):629–36.
11. Reichenbach JR, Venkatesan R, et al. Theory and application of static field inhomogeneity effects in gradient-echo imaging. J Magn Reson Imaging. 1997;7(2):266–79.
12. Ashwal S, Babikian T, et al. Susceptibility-weighted imaging and proton magnetic resonance spectroscopy in assessment of outcome after pediatric traumatic brain injury. Arch Phys Med Rehabil. 2006;87(12 Suppl 2):S50–8.
13. Holdsworth SJ, Bammer R. Magnetic resonance imaging techniques: fMRI, DWI, and PWI. Semin Neurol. 2008;28(4):395–406.
14. Wolff SD, Balaban RS. Magnetization transfer contrast (MTC) and tissue water proton relaxation in vivo. Magn Reson Med. 1989;10(1):135–44.
15. Bagley LJ, Grossman RI, et al. Magnetization transfer contrast: its utility as a technique and its application to central nervous system pathology. Neurology. 1999;53(5 Suppl 3):S49–51.
16. Weber WA. Positron emission tomography as an imaging biomarker. J Clin Oncol. 2006;24(20):3282–92.
17. Rahmim A, Zaidi H. PET versus SPECT: strengths, limitations and challenges. Nucl Med Commun. 2008;29(3):193–207.
18. Gruber S, Stadlbauer A, et al. Proton magnetic resonance spectroscopic imaging in brain tumor diagnosis. Neurosurg Clin N Am. 2005;16(1):101–14, vi.
19. Asari S, Makabe T, et al. Assessment of the pathological grade of astrocytic gliomas using an MRI score. Neuroradiology. 1994;36(4):308–10.
20. Wahl RL, Jacene H, et al. From RECIST to PERCIST: Evolving Considerations for PET response criteria in solid tumors. J Nucl Med. 2009;50 Suppl 1:122S–50.
21. Miller AB, Hoogstraten B, et al. Reporting results of cancer treatment. Cancer. 1981;47(1):207–14.
22. Eisenhaue EA, Therasse P, et al. New response evaluation criteria in solid tumours: revised RECIST guideline (version 1.1). Eur J Cancer. 2009;45(2):228–47.
23. Wen PY, Macdonald DR, et al. Updated response assessment criteria for high-grade gliomas: response assessment in neuro-oncology working group. J Clin Oncol. 2010;28(11):1963–72.
24. Macdonald DR, Cascino TL, et al. Response criteria for phase II studies of supratentorial malignant glioma. J Clin Oncol. 1990;8(7):1277–80.
25. Gerstner ER, McNamara MB, et al. Effect of adding temozolomide to radiation therapy on the incidence of pseudo-progression. J Neurooncol. 2009;94(1):97–101.
26. Chao ST, Suh JH, et al. The sensitivity and specificity of FDG PET in distinguishing recurrent brain tumor from radionecrosis in patients treated with stereotactic radiosurgery. Int J Cancer. 2001;96(3):191–7.
27. Zuniga RM, Torcuator R, et al. Efficacy, safety and patterns of response and recurrence in patients with recurrent high-grade gliomas treated with bevacizumab plus irinotecan. J Neurooncol. 2009;91(3):329–36.
28. Sorensen AG, Batchelor TT, et al. A "vascular normalization index" as potential mechanistic biomarker to predict survival after a single dose of cediranib in recurrent glioblastoma patients. Cancer Res. 2009;69(13):5296–300.
29. Chandana SR, Movva S, et al. Primary brain tumors in adults. Am Fam Physician. 2008;77(10):1423–30.
30. Jacobs AH, Kracht LW, et al. Imaging in neurooncology. NeuroRx. 2005;2(2):333–47.
31. Kleihues P, Sobin LH. World Health Organization classification of tumors. Cancer. 2000;88(12):2887.
32. Jellinger K. Glioblastoma multiforme: morphology and biology. Acta Neurochir (Wien). 1978;42(1–2):5–32.
33. Arvinda HR, Kesavadas C, et al. Glioma grading: sensitivity, specificity, positive and negative predictive values of diffusion and perfusion imaging. J Neurooncol. 2009;94(1):87–96.
34. Mechtler L. Neuroimaging in neuro-oncology. Neurol Clin. 2009;27(1):171–201, ix.
35. McMillan KM, Rogers BP, et al. Physiologic characterisation of glioblastoma multiforme using MRI-based hypoxia mapping, chemical shift imaging, perfusion and diffusion maps. J Clin Neurosci. 2006;13(8):811–7.

36. Sievert AJ, Fisher MJ. Pediatric low-grade gliomas. J Child Neurol. 2009;24(11):1397–408.
37. Bakshi R, Minagar A, et al. Imaging of multiple sclerosis: role in neurotherapeutics. NeuroRx. 2005;2(2):277–303.
38. D'Souza M, Kappos L, et al. Reconsidering clinical outcomes in Multiple Sclerosis: relapses, impairment, disability and beyond. J Neurol Sci. 2008;274(1–2):76–9.
39. Foley JF, Brandes DW. Redefining functionality and treatment efficacy in multiple sclerosis. Neurology. 2009;72(23 Suppl 5):S1–11.
40. Comi G, Filippi M, et al. Effect of early interferon treatment on conversion to definite multiple sclerosis: a randomised study. Lancet. 2001;357(9268):1576–82.
41. Barkhof F, Calabresi PA, et al. Imaging outcomes for neuroprotection and repair in multiple sclerosis trials. Nat Rev Neurol. 2009;5(5):256–66.
42. Bar-Zohar D, Agosta F, et al. Magnetic resonance imaging metrics and their correlation with clinical outcomes in multiple sclerosis: a review of the literature and future perspectives. Mult Scler. 2008;14(6):719–27.
43. Bruck W, Bitsch A, Kolenda H, Bruck Y, Stiefel M, Lassmann H. Inflammatory central nervous system demyelination: correlation of magnetic resonance imaging findings with lesion pathology. Ann Neurol. 1997;42:783–93.
44. Poser CM, Brinar VV. Diagnostic criteria for multiple sclerosis. Clin Neurol Neurosurg. 2001;103(1):1–11.
45. Polman CH, Wolinsky JS, et al. Multiple sclerosis diagnostic criteria: three years later. Mult Scler. 2005;11(1):5–12.
46. Butcher K, Emery D. Acute stroke imaging. Part I: fundamentals. Can J Neurol Sci. 2010;37(1):4–16.
47. Demchuk AM, Coutts SB. Alberta Stroke Program Early CT Score in acute stroke triage. Neuroimaging Clin N Am. 2005;15(2):409–19, xii.
48. Hampel H, Frank R, et al. Biomarkers for Alzheimer's disease: academic, industry and regulatory perspectives. Nat Rev Drug Discov. 2010;9(7):560–74.
49. Jack Jr CR, Bernstein MA, et al. Update on the magnetic resonance imaging core of the Alzheimer's disease neuroimaging initiative. Alzheimers Dement. 2010;6(3):212–20.
50. Jagust WJ, Bandy D, et al. The Alzheimer's disease neuroimaging initiative positron emission tomography core. Alzheimers Dement. 2010;6(3):221–9.
51. Shellock FG, Spinazzi A. MRI safety update 2008: part 2, screening patients for MRI. AJR Am J Roentgenol. 2008;191(4):1140–9.

Chapter 11
Imaging in Musculoskeletal, Metabolic, Endocrinological, and Pediatric Clinical Trials

Colin G. Miller, Hui Jing Yu, and Cornelis van Kuijk

Abstract Clinical trials in skeletal pathology are abundant and comprise predominantly of trials in osteoporosis, rheumatoid arthritis, osteoarthritis, fracture healing, and bone marrow disease, including genetic disorders of the skeletal system predominantly in pediatric populations. Furthermore, many metabolic and endocrinological syndromes also affect the musculoskeletal system. Radiological end points in clinical trials for the evaluation of the musculoskeletal system are numerous and have a unique set of challenges which are usually disease specific. The imaging modalities employed for these end points include conventional radiography, ultrasound, computed tomography, dual X-ray absorptiometry, magnetic resonance imaging, and bone scintigraphy. This chapter will present the key disease areas, the imaging requirements, the characteristics, including the challenge of quantitative versus qualitative assessment, and the use of imaging as a biomarker in these diseases.

Keywords Osteoporosis • Rheumatoid arthritis • Osteoarthritis • Fracture healing • Bone marrow disease • Pediatric bone diseases

C.G. Miller, BSc, PhD, FICR, CSci (✉) • H.J. Yu, PhD
Department of Medical Affairs, BioClinica, Inc.,
826 Newtown-Yardley Road, Newtown, PA 18940, USA
e-mail: colin.miller@bioclinica.com; huijing.yu@bioclinica.com

C. van Kuijk, MD, PhD
Department of Radiology, VU University Medical Center,
De Boelelaan 1118, Amsterdam 1081 HZ, The Netherlands
e-mail: c.vankuijk@vumc.nl

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_11,

Introduction

Clinical trials in skeletal pathology are abundant and comprise of mostly trials in osteoporosis, fracture healing, bone marrow disease, degenerative joint disease (arthritis), and rheumatoid arthritis (joint inflammation). There are also a number of genetic disorders of the skeletal system which are more recently being studied and by definition are usually in pediatric subjects. Many metabolic and endocrinological syndromes affect the musculoskeletal system and require imaging as efficacy or safety end points which have impact in the design and uses of imaging modalities. Several types of imaging are used in clinical practice for these diseases. Conventional radiography is still the first line of imaging, complemented by ultrasound, computed tomography (CT), dual X-ray absorptiometry (DXA), magnetic resonance imaging (MRI), and bone scintigraphy. More recently the hybrid technique of PET-CT is emerging where the metabolic information gathered by positron emission tomography (PET) is combined with anatomical details provided by high-resolution 3D CT. PET-MRI scanners are being developed, which have many more technical challenges, but there are now software techniques to co-register these kinds of images acquired on scanners of different technologies.

For clinical trials designed to prove therapeutic efficacy for registration of a biologic or drug, by the Food and Drug Administration (FDA) or other regulatory bodies, the imaging biomarkers used are usually conservative and have to the most part remained unchanged for 30 plus years. There are a limited number of "fully" validated imaging techniques that are accepted. Validation of imaging techniques requires extensive knowledge of several parameters of the techniques described in detail in Chap. 2. In the following paragraphs we will discuss some of these diseases, their imaging characteristics, and the use of imaging as a biomarker in these diseases.

Osteoporosis

Osteoporosis is a disease that originates from a disturbance in bone metabolism. In normal bone metabolism there is a balance in bone turnover: osteoclasts (bone accretion) are counter-balanced by osteoblasts (bone formation or deposition). In primary ("aging") or secondary osteoporosis, the balance is negative. Patients are losing bone and will eventually fracture because the bone is simply not strong enough. Minor trauma or even normal use will lead to debilitating fractures, especially in the spine, hip, shoulder, and wrist. Drugs have therefore been developed that "restore the balance" or even shift the balance to net bone formation. These drugs include bisphosphonates [1–4], selective estrogen receptor modulators (SERMs) [5], parathyroid-acting drugs [6], vitamin D, and minerals (calcium, strontium) [7]. To prove drug efficacy, regulatory agencies require pivotal clinical trials which are designed to show a reduction (prevention) of fractures in patients with osteoporosis and show a positive effect on bone density.

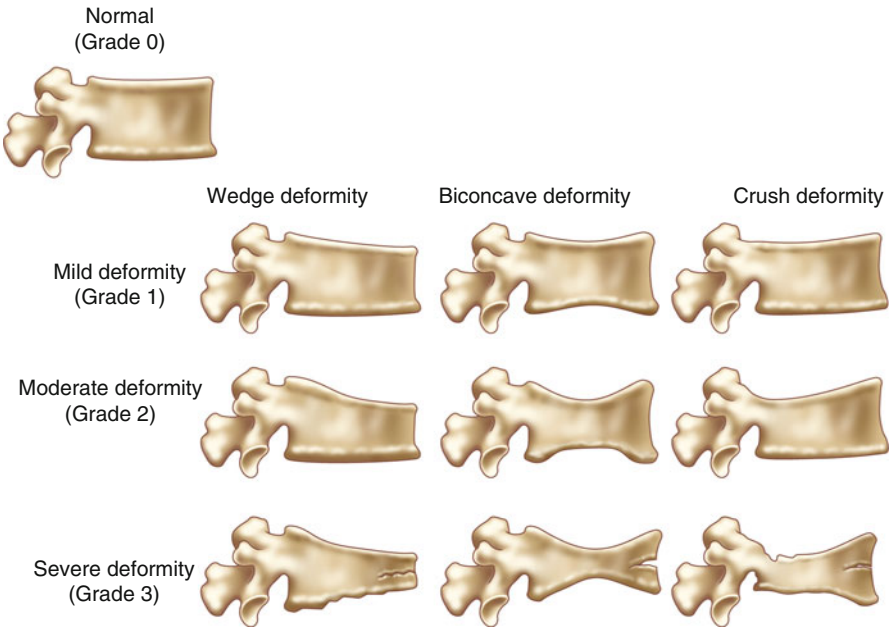


Fig. 11.1 Semiquantitative scoring system for vertebral deformity in osteoporosis, graphic representation (Adapted with permission from Genant et al. [8])

Fracture detection is generally evaluated by conventional radiography or X-ray. As vertebral fractures are common in osteoporosis, usually spine films are used to detect and grade prevalent and incident vertebral deformities/fractures. Bone density is usually measured with dual X-ray absorptiometry of the spine and hip, although quantitative computerized tomography is also used and is providing further insight into the bone biology and structural evaluation of the spine and femur

The detection and grading of vertebral fractures are usually evaluated by a semiquantitative grading method as described by Genant and colleagues [8]. An expert musculoskeletal radiologist will be required to read the spine films in a highly standardized and documented manner and assign grades of fracture (0, none; 1, mild; 2, moderate; 3, severe) to the distinct vertebral bodies of the thoracic and lumbar spine (from T4 to L4) (Fig. 11.1). The vertebrae inferior and superior to this area fracture rarely in osteoporosis and are difficult to evaluate due to overlying bony anatomy. In addition the vertebral bodies can be measured, using a 6-point measurement technique describing the posterior, mid-, and anterior height of the vertebral bodies. These measurements lead to different height ratios, such as the anterior-posterior ratio describing the wedge shape of the vertebral body. If the wedge exceeds a certain threshold (e.g., 25 %), the vertebral body is considered to be deformed/fractured. The literature discussing the semiquantitative technique and the quantitative morphometric technique is abundant [9–11]. As such it is considered a validated technique in clinical drug trials in osteoporosis.

Bone densitometry using DXA in osteoporosis has become a standard in clinical trials. In theory it is not the best technique for measuring bone density as it provides a two-dimensional outcome parameter (in gram per square cm or g/cm^2) while measuring a three-dimensional object. However, the regulatory agencies are acceptable of the data but not as a primary efficacy outcome in osteoporosis treatment. Once proven by a fracture study, DXA is an acceptable technique for both the assessment of prevention trials and more recently non-inferiority studies. However, DXA (or DEXA) has become the best validated technique just because of its accessibility, low radiation dose, and ease of use.

The challenge of using DXA for eligibility criteria has been described in more detail elsewhere [12]. However, briefly challenge is that usually for an osteoporosis study or similar, patients who are defined as osteoporotic have a so-called T-score (comparison against peak bone mass or Z-score which is age-matched control) of −2.5 (minus 2.5) or lower. This is gender, race, and anatomical area specific. Furthermore, the manufacturers have normative data bases which are not quite interchangeable, so some allowance has to be considered to ensure the population is uniform throughout the study [13]. To further reduce this variation, there are two manufacturers, GE Healthcare (Lunar) and Hologic Inc, that make 90–95 % of all the world's DXA instruments, so most studies are reduced to using just these types. There is a second challenge: the calibrations of the two instruments have a calibration difference of about 10–15 % (Chap. 1).

The second and ongoing challenge with using DXA in clinical trials is that it is a Type 1 Instrument (see Chap. 2). Therefore, there is need to monitor instrument performance or calibration. If there is a calibration shift or a change in the DXA instrument, then there has to be a process described that will evaluate the effect of the calibration shift to the subject data and then a second process to recalculate the subject BMD changes to compensate these calibration shifts. The end point should be that subjects' BMD results should be calculated as the percentage change from baseline and the results aggregated. This essentially removes inter-instrument variability. Therefore, at the start of the study, each site should measure phantom that covers a range of densities, such as the Bona Fide Phantom (BFP) (BioClinica Inc, Newtown, PA, USA), ten times without repositioning. If later during the course of the study a site changes instrument or has an instrument breakdown or change in the underlying calibration, the same BFP should measured again and the change in calibration evaluated using a regression analysis. If the measurement or calibration changes by more than twice the error of the BFP measurement (nominally 1 %), then a regression analysis can be applied to the subject BMD data acquired on that scanner, post-calibration change and the percentage change of the subject recalculated.

Quantitative computed tomography (QCT) provides a three-dimensional measure (in gram per cubic cm), thus true bone density, and has a better sensitivity to change as it measures specifically in the trabecular compartment (with high bone turnover) of the vertebral bodies in the spine. These are standard measurements with QCT that are used to report BMD in the lumbar spine, but it can also be applied to other skeletal parts. Peripheral QCT (pQCT) measurements are

performed on specially designed small-bore CT scanners. Like QCT in the spine, pQCT can provide separate measurement of the cortical and trabecular structure in peripheral regions such as the forearm, femur, and tibia. High-resolution QCT (HRQCT) is a further development in QCT measurements. HRQCT allows the analysis of trabecular structure with high-resolution thin slices. HRQCT is commonly used in research setting for microstructure analysis of bone specimens but can be extended to clinical settings.

QCT can be used to measure cortical and/or trabecular bone mineral density, and volumetric and cross-sectional areal bone geometry, allowing for additional assessments of bone quality and characteristics for osteoporosis. Cortical bone assessments have generally evaluated in the femur, but due to the thickness of the spine it has not been possible to accurately or precisely assess this bone compartment. Most femur assessments have evaluated the whole cortical shell [14, 15]. More advanced analysis techniques used include finite element analysis of the spine [16] and an analysis technique developed by Mindways Software Inc (Austin, TX, USA) [17, 18] which identifies and evaluates the four quadrants of the femoral neck cortical shell for both vBMD and thickness. Quadrant QCT analysis allows a noninvasive technique to elucidate anatomic distribution which may be critical in determining resistance to fracture, e.g., the superior cortex of the femoral neck is a stronger predictor for fracture than the inferior cortex [18]. The ability to segment out trabecular and cortical bone with QCT scans is particularly important for the evaluation of new therapeutic agents in each bone compartments. This has been recently shown by a new study using rosiglitazone where a negative therapeutic response was observed in 52 weeks [19]. If such a response was observable in a compound with relatively small therapeutic impact, as the authors state, it is highly likely that this end point may be of value in the treatment of osteoporosis.

Rheumatoid Arthritis

Rheumatoid arthritis is a progressive disease characterized by synovial joint inflammation, eventually leading to destruction of cartilage and underlying bone structures. For decades it was very difficult to treat. Drugs used were nonspecific like corticosteroids (against inflammation in general) and methotrexate (against tissue proliferation in general). Nowadays, disease-modifying antirheumatic drugs (DMARDs) like anti-tumor necrosis factor-alpha (or anti-TNFα) are used and being developed that are able to halt disease progression [20–27]. Furthermore, at the time of writing there are a slew of new DMARDs in development or in review with the regulatory agencies, such as the so-called JAK inhibitors [28], of which the first one has just been approved by the FDA, and a slew of interleukin (IL) compounds like IL-6 and IL-17. In imaging terms, rheumatoid arthritis is characterized by bone destruction and cartilage loss leading to joint destruction as assessed by bone erosions and decreased joint space narrowing, respectively. Disease progression is characterized by the joints being deformed and ultimately destroyed. Conventional radiography of the hands and

Table 11.1 The history of semiquantitative scoring systems in rheumatoid arthritis

Scoring system	Date of publication and reference
Steinbrocker Index	(1949) [30]
Kellgren's Method	(1957) [31]
Sharp Scoring Method	(1971) [32]
Larsen Scoring	(1977) [33]
Genant Scoring Method	(1983) [34]
Modified Sharp	(1985) [35]
The Sharp/van der Heijde Scoring Method	(1989) [36]
Modified Genant Scoring Method	(1998) [37]

feet are used to "grade" the disease. Very elaborate semiquantitative grading schemes have been developed over the years that encompass both joint space narrowing as well as bone erosions [29]. The historical timeline of these is shown in Table 11.1. The Sharp score is arguably the most documented, and its variation described by van der Heijde is the one most widely used in clinical drug trial to assess drug efficacy. It is now the scoring system of choice in the EMA guidelines for assessing DMARDs in clinical trials. As such these visual scoring systems are regarded fully validated. Standardized imaging protocols have been described for obtaining the radiographs of the hands and feet and are described fully elsewhere [38].

A new challenge is emerging in these trials: patients are being treated at a much earlier stage of the disease when there are no or minimal features of the disease visible on radiographs. Since the indication for DMARD requires the radiological demonstration of the decrease in the disease progression, many studies now require eligibility criteria that have to be centrally evaluated to show clear evidence of radiological disease. Furthermore, standard of care is being used as the comparator, and the trials are requiring many more subjects to show the new molecule has clinical and radiological benefit.

Magnetic resonance imaging (MRI) has been proposed as a new imaging biomarker for the assessment of rheumatoid arthritis. While at the time of writing it is still not accepted by the regulatory agencies as the primary end point for Phase III studies, it is being used very successfully in Phase II studies for "go/no-go" decisions for continuing drug development or dose-ranging studies [39]. It provides a visual interpretation of synovial inflammation, and in addition quantification of contrast uptake in the inflamed tissue has been investigated. As with radiographs there is a semiquantitative scoring system or the so-called RAMRIS (rheumatoid arthritis MRI scoring). This requires the evaluation by specialists in the field and is labor intensive. The MRI scans have also to be acquired in a very standardized manner with subjects lying prone in a scanner in the "superman" position or supine with their hands and wrist in a special coil. This can be very daunting and for those in pain, preventing motion during the 30–45 min, scan acquisition can be difficult. Also the preferred use of contrast agents further adds to the complexity of the study.

Novel inflammation-specific PET-tracers are being developed to try to assess disease activity, and more recently the evaluation of the pharmacologic

intervention is being investigated by the use of dynamic contrast-enhanced (DCE) MRI [40]. Ultrasound is having a role to play, particularly in Europe, and with the incentive to reduce radiation dose to patients, ultrasound of the joints has become a recognized end point for Phase IIb and Phase IV studies. Ultrasound, as discussed in Chap. 1, is very operator dependent, so this requires a high degree of site operator training if this modality is to be used in clinical trials. Furthermore, the site has to be very careful in labelling all the joints so the central readers can clearly identify the anatomy during the central read without access to the patient.

Osteoarthritis (Degenerative Joint Disease)

The classic description of osteoarthritis is cartilage lost due to wear and tear that eventually will lead to joint space narrowing and bone remodelling (osteophytes and sclerosis). However, more recently there are debates that it may be an inflammatory disease mediated by the so-called mechanokines or mechanical insult. Furthermore, there may be different pathophysiological pathways that are more clearly elucidated such as anterior cruciate ligament repair leading to knee osteoarthritis 20–30 years later, or a meniscal tear or meniscectomy versus a patient who has spent their life undergoing heavy labor and whose joints have undergone bony degeneration, remodelling, and cartilage destruction. Without going into the debate of the etiology, radiographically osteoarthritis is now recognized as a disease of the whole joint [41, 42]. Most clinical trials have focused on the knee due to the higher incidence although osteoarthritis occurs at the hip, shoulders and hand, with the latter two joints being non-weight bearing, so there is another argument as to whether this is truly primary osteoarthritis.

Osteoarthritis is usually detected on radiographs as joint space narrowing and specific features of bone remodelling that can be graded according to the severity of the disease. The Kellgren and Lawrence scale is the best known grading system originally being described in 1952 for knee and hips [43]. It is still the so-called gold standard for the eligibility criteria for clinical trials in osteoarthritis [44]. However, there are a number of different modifications to the original description with one paper citing ten different versions [45]. It is a scaling system that while it appears straight forward and simple is very difficult to obtain initial consensus between a group of radiologists due to the nuances in the disease and therefore requires “reader calibration” for use with a pool of readers in clinical or epidemiological clinical trials. Due to the slow rate of change in the characteristics of the joint assessed by the Kellgren and Lawrence scoring system, it is not used for efficacy. The regulatory authorities (FDA and EMA) still require joint space narrowing (JSN) as assessed by plain film radiographs to be the primary outcome in a disease-modifying anti- osteoarthritis drug (DMOAD) model. Joint space width (JSW) is a difficult end point to assess due to the reproducibility required to assess a change of 0.1 mm to 0.16 mm per year decrease in subjects with confirmed

osteoarthritis (Kellgren and Lawrence score 2 or 3). The acquisition protocol has to be very clearly defined, and the one arguably shown to be the most reliable is the modified Lyon-Schuss using a plexiglass positioning device [46]. With good quality acquisition the precise measurement of JSW can be obtained. Even then, there are several different methodologies that have been described [47, 48], but usually this is the medial aspect at a fixed anatomical point, but could be the narrowest within the predefined area, or even the mean of the tibial plateau/femoral condyle space.

The use of MRI for the assessment of OA has, as with RA, gained a place in clinical development especially in Phase II. However, at the time of writing, there is no one set of criteria or measurements that clearly provides the go/no-go signal that has been accepted by the FDA. MRI assessments can be broken down into quantitative and semiquantitative or scaling techniques. The former, at a minimum, evaluate cartilage thickness in different sub- anatomical areas of the medial and lateral cartilage [49–51]. They can also evaluate shape of the cartilage [52] using active shape modelling. There are a number of so-called "semiquantitative" scoring systems. The first one was arguably the Whole-Organ Magnetic Resonance Imaging Score of the osteoarthritis in the knee [53]. This has been superseded by the BLOKS (Boston Leeds Osteoarthritis Knee Score) [54], and a combination of the two has recently been developed, the so-called MOAKS (MRI Osteoarthritis Knee Score), by the same team [55].

The field of clinical trials in osteoarthritis is now littered with a number of failed drugs trying to prove DMOAD status. These include the risedronate study [56, 57] which failed the primary end point but provided significant insight in the field to improve future studies. The doxycyline study was one of the best conducted but was underpowered [58]. More recently, the calcitonin studies reached statistical significance with an MRI evaluation method but failed the primary end point of reduction in JSN by radiographs [59, 60]. Since this study had previously reported futility analysis failure, it can only be surmised that either the subjects were incorrectly enrolled or the quality control of the images was performed very poorly. In contrast the most recent program for an iNOS inhibitor, cindunistat, passed futility analysis and showed statistical significance at year 1 against placebo in those subjects with a modified Kellgren and Lawrence grade 2 (not grade 3). This is an important landmark study in which the results and methodology are both published as separate papers [44] led by the Hellio Le Graverand team [46], since it is the first time drug was shown to have statistically beneficial DMOAD properties with a radiographic end point. Unfortunately efficacy was lost at year 2 and the FDA requires statistical significance in radiographic joint space narrowing for 2 years.

Unlike joint space narrowing for osteoarthritis, the FDA has accepted MRI as the end point for focal cartilage defect healing using an implant [61]. For cartilage regeneration evaluation the so-called MOCART scale (magnetic resonance observation of cartilage repair tissue) [62] was developed. This has become a standard scoring system for focal cartilage repair and regeneration and is accepted by the FDA.

Fracture Healing

Radiographs as well as CT have been used to describe fracture healing. This is not trivial since the definition of fracture healing on radiographs is not quite clear. Usually bridging of cortical bone (which is usually circumferential) of at least 75 % of the fracture plane is used as a definition of successful fracture healing in tubular bones. This requires radiographs in at least two directions or a dedicated 3D CT scan. The RUST (Radiological Union Score for Tibial fractures) [63] has become the standard approach for this end point and evaluation, at least for fractures of the tibia.

Bone Marrow Disease

Bone marrow disorders can have different origins. Next to several types of leukemic disease and metastasis, there are more exotic diseases like Gaucher's disease. Radiographs depicting the skeletal status have been used to assess disease severity and disease progression. However, radiographs are sensitive to bone disease but less sensitive to bone marrow changes. MRI is the preferred technique to grade bone marrow burden. Only recently some imaging biomarkers have been validated for use in trials to study drug efficacy in Gaucher's disease [64].

Pediatric Bone Disease

The development of pediatric studies has lagged behind those of the adult, but in more recent years, mainly due to the emphasis by both the EMA and FDA to have new products developed in this specialized population and the so-called "pediatric exclusivity" program, there has been a larger number of studies of late. Further development in pediatric populations has occurred as there has been a focus in the pharmaceutical industry towards orphan drug indications and other unmet medical needs, of which many are genetic mutations and therefore present in children. Although the standard radiological techniques can be applied, there are challenges evaluating the growing skeleton. Plain radiographs have beam divergence, and therefore even measuring the length and hence growth velocity of the long bones is challenging, and radiopaque rulers have to be in position during the acquisition of radiographs.

For DXA the challenge is that 3-dimensional objects, the bones, are increasing over time but only displayed and measurements calculated in 2 dimensions, confounding longitudinal measurements. Z-score change is arguably the optimum method to achieve a meaningful end point, since this uses a normal reference data

set and hence growth changes in the evaluation of change in BMD seen in a pediatric population. The challenge is that many of the pediatric studies are in severely diseased children whose growth is already abnormal and whose level of pubertal on set and therefore growth patterns may be significantly distorted from the norm. So there have been a number of approaches of late to create a superior method and the development of height adjusted Z-score was developed [65, 66]. Essentially a subject's height is the taken from the standardized growth curves by comparing their height to the mean of the curve and giving them this age to then calculate the BMD Z-score. In other words, creating a bone age related to normal development. However, no one single methodology at the time of writing has come to the fore as the de facto standard.

Another approach with DXA has been the assessment of the distal femur [67]. This measurement was originally developed by the team at the Alfred I. duPont Hospital, Delaware, USA for assessment in children suffering from cerebral palsy. The side position for the patient, is comfortable and allows them to be relaxed and still for the measurement. This measurement has been further developed and expanded into other populations and has been successfully used in a number of clinical trials [68].

Peripheral quantitative computerized tomography (pQCT) has been used extensively in pediatric studies due to the ease of use, low radiation dosage, and a 3D evaluation of bone. These are dedicated systems of which there are two main manufacturers, Stratec and Scanco. Stratec is the most prevalent system and many studies have reported outcomes based on data collected by this instrumentation. As already stated, the challenge with DXA is the 2D evaluation of the growing bone. pQCT removes this challenge. More recently Mindways has developed a "pQCT" version of their software allowing a standard CT scanner to be used. The subject lies in the scanner in a "superman" position with arms outstretched so the forearms can be scanned avoiding radiation to the brain and torso. This will provide further investigator sites that can be employed in pediatric clinical trials without having to purchase expensive dedicated equipment.

Further to the forearm other anatomical sites can be easily measured using a full-body CT scanner. This has led to the development of a measurement by Leonard [69] at the Children's Hospital of Philadelphia, whereby the whole length of a bone such as the tibia or forearm can be measured. It is then possible to see the dynamic changes in bone growth and the lengthening from epiphysis to metaphysis and improvement in trabecular bone and/or cortical between time points at set anatomical locations.

The classic method of assessing bone age is using the atlas developed by Greulich and Pyle in 1958 [70] or Tanner and Whitehouse [71]. The assessment is made of the epiphyseal closures of the hand and wrist joints, usually in the left hand. It requires the evaluation of the radiographs by an experienced pediatric musculoskeletal radiologist. Due to the atlas being in annual chronological increments, except during the high growth times (puberty), where it is in 6-month increments, it is too imprecise to be used for efficacy assessments, except in

long-term (several year) studies. However, the FDA does require the evaluation of bone age at the start of a pediatric study.

Endocrinology and Safety Studies

Bone metabolism is under a highly complex endocrinological control. Therefore, many therapeutic agents have effect on this organ and calcium homeostasis. There are many studies which require evaluation of the bone density and fracture risk, usually by DXA in the population under study. This ranges from the use of isotretinoins for the treatment of acne [72] to the evaluation of BMD in patients being evaluated for the novel treatments in type II diabetes. The later is of particular note, since rosiglitazone was shown to increase the risk of hip fracture and further studies have shown loss of BMD [73]. Most new therapies being developed in this field will require monitoring of the bone mineral density due to the endocrinological interplay in this patient population.

Other endocrinological areas that require DXA assessments or BMD monitoring is where there is disease or therapeutic influence on the gonadal system. This includes growth hormone replacement, testosterone replacement and cessation (e.g., prostate cancer and benign prostate hyperplasia), endometriosis in women, and other estrogen replacement or intervention. In breast cancer this has become particularly critical, and arguably one of the longest running safety studies was conducted in women taking aromatase inhibitors. The so-called ATAC study had serial DXA measurements for 10 years [74].

Summary

Clinical trials evaluating the medical imaging of the skeletal system are numerous and have a unique set of challenges depending on the specific disease being studied. Although all imaging modalities are used depending on the imaging end point, plain film radiographs are the predominating imaging modality due to the ability to elegantly visualize bone. The challenge is that this only provides a two-dimensional view of the three dimensions, and so careful radiological interpretation is required or more views have to be obtained, and then the radiation to the patient increases. MRI evaluations of the skeletal system are becoming more prevalent, but cost and time in the scanner makes them prohibitive for most studies however CT scanners provide another 3D alternative, although radiation dose has to be considered carefully.

This chapter also encompasses a very wide range of metabolic disease areas, each with a different set of challenges, which means the contents provided here can only just provide a basic introduction to the topics. The reader is encouraged to read further texts on the specific areas, if more in depth knowledge is required [38, 12].

References

1. Cummings SR, et al. Effect of alendronate on risk of fracture in women with low bone density but without vertebral fractures: results from the Fracture Intervention Trial. JAMA. 1998;280(24):2077–82.
2. Harris ST, et al. Effects of risedronate treatment on vertebral and nonvertebral fractures in women with postmenopausal osteoporosis: a randomized controlled trial. Vertebral Efficacy With Risedronate Therapy (VERT) Study Group. JAMA. 1999;282(14):1344–52.
3. Black DM, et al. Once-yearly zoledronic acid for treatment of postmenopausal osteoporosis. N Engl J Med. 2007;356(18):1809–22.
4. Black DM, et al. Randomised trial of effect of alendronate on risk of fracture in women with existing vertebral fractures. Fracture Intervention Trial Research Group. Lancet. 1996;348(9041):1535–41.
5. Lufkin EG, et al. Antiresorptive treatment of postmenopausal osteoporosis: review of randomized clinical studies and rationale for the Evista alendronate comparison (EVA) trial. Curr Med Res Opin. 2004;20(3):351–7.
6. Rubin MR, et al. The anabolic effects of parathyroid hormone. Osteoporos Int. 2002; 13(4):267–77.
7. Binkley N, et al. A phase 3 trial of the efficacy and safety of oral recombinant calcitonin: the Oral Calcitonin in Postmenopausal Osteoporosis (ORACAL) trial. J Bone Miner Res. 2012;27(8):1821–9.
8. Genant HK, et al. Vertebral fracture assessment using a semiquantitative technique. J Bone Miner Res. 1993;8(9):1137–48.
9. Diacinti D, Guglielmi G. Vertebral morphometry. Radiol Clin North Am. 2010;48(3):561–75.
10. Guglielmi G, et al. Assessment of osteoporotic vertebral fractures using specialized workflow software for 6-point morphometry. Eur J Radiol. 2009;70(1):142–8.
11. Brett A, et al. Development of a clinical workflow tool to enhance the detection of vertebral fractures: accuracy and precision evaluation. Spine. 2009;34(22):2437–43.
12. Pearson D, Miller CG. Clinical trials in osteoporosis, vol. xiii. 2nd ed. Nottingham/Newtown: Springer Science + Business Media; 2007. p. 292.
13. Miller CG, Barden HS. Entrance criteria for clinical trials with DEXA. J Bone Miner Res. 1994;9(S1):S209.
14. Borggrefe J, et al. Quantitative computed tomographic assessment of the effects of 24 months of teriparatide treatment on 3D femoral neck bone distribution, geometry, and bone strength: results from the EUROFORS study. J Bone Miner Res. 2010;25(3):472–81.
15. Engelke K, et al. Regional distribution of spine and hip QCT BMD responses after one year of once-monthly ibandronate in postmenopausal osteoporosis. Bone. 2010;46(6):1626–32.
16. Niebur GL, et al. High-resolution finite element models with tissue strength asymmetry accurately predict failure of trabecular bone. J Biomech. 2000;33(12):1575–83.
17. Poole KE, et al. Changing structure of the femoral neck across the adult female lifespan. J Bone Miner Res. 2010;25(3):482–91.
18. Johannesdottir F, et al. Distribution of cortical bone in the femoral neck and hip fracture: a prospective case–control analysis of 143 incident hip fractures; the AGES-REYKJAVIK study. Bone. 2011;48(6):1268–76.
19. Miller CG, Nino AJ, Northcutt AR, Lewiecki EM, Paul G, Cobitz AR, Wooddell MJ, Bilezikian JP, Fitzpatrick LA. Evaluation of QCT cortical hip quadrant in a clinical trial with rosiglitazone: a potential new study endpoint. J Bone Miner Res. 2011:26(Suppl 1).
20. van der Heijde D, et al. Infliximab inhibits progression of radiographic damage in patients with active psoriatic arthritis through one year of treatment: results from the induction and maintenance psoriatic arthritis clinical trial 2. Arthritis Rheum. 2007;56(8):2698–707.
21. van der Heijde D, et al. Radiographic progression of ankylosing spondylitis after up to two years of treatment with etanercept. Arthritis Rheum. 2008;58(5):1324–31.

22. van der Heijde D, et al. Radiographic findings following two years of infliximab therapy in patients with ankylosing spondylitis. Arthritis Rheum. 2008;58(10):3063–70.
23. Gladman DD, et al. Adalimumab for long-term treatment of psoriatic arthritis: forty-eight week data from the adalimumab effectiveness in psoriatic arthritis trial. Arthritis Rheum. 2007;56(2):476–88.
24. Bathon JM, et al. A comparison of etanercept and methotrexate in patients with early rheumatoid arthritis. N Engl J Med. 2000;343(22):1586–93.
25. Kavanaugh A, et al. Golimumab in psoriatic arthritis: one-year clinical efficacy, radiographic, and safety results from a phase III, randomized, placebo-controlled trial. Arthritis Rheum. 2012;64(8):2504–17.
26. Tanaka Y, et al. Golimumab in combination with methotrexate in Japanese patients with active rheumatoid arthritis: results of the GO-FORTH study. Ann Rheum Dis. 2012;71(6):817–24.
27. Takeuchi T, et al. Golimumab monotherapy in Japanese patients with active rheumatoid arthritis despite prior treatment with disease-modifying antirheumatic drugs: results of the phase 2/3, multicentre, randomised, double-blind, placebo-controlled GO-MONO study through 24 weeks. Ann Rheum Dis 2012;72(9):1488–95.
28. LaBranche TP, et al. JAK inhibition with tofacitinib suppresses arthritic joint structural damage through decreased RANKL production. Arthritis Rheum. 2012;64(11):3531–42.
29. Ravindran V, Rachapalli S. An overview of commonly used radiographic scoring methods in rheumatoid arthritis clinical trials. Clin Rheumatol. 2011;30(1):1–6.
30. Steinbrocker O, Traeger CH, Batterman RC. Therapeutic criteria in rheumatoid arthritis. J Am Med Assoc. 1949;140(8):659–62.
31. Kellgren JH, Lawrence JS. Radiological assessment of rheumatoid arthritis. Ann Rheum Dis. 1957;16(4):485–93.
32. Sharp JT, et al. Methods of scoring the progression of radiologic changes in rheumatoid arthritis. Correlation of radiologic, clinical and laboratory abnormalities. Arthritis Rheum. 1971;14(6):706–20.
33. Larsen A, Dale K, Eek M. Radiographic evaluation of rheumatoid arthritis and related conditions by standard reference films. Acta Radiol Diagn (Stockh). 1977;18(4):481–91.
34. Genant HK. Methods of assessing radiographic change in rheumatoid arthritis. Am J Med. 1983;75(6A):35–47.
35. Sharp JT, et al. How many joints in the hands and wrists should be included in a score of radiologic abnormalities used to assess rheumatoid arthritis? Arthritis Rheum. 1985;28(12):1326–35.
36. van der Heijde DM, et al. Effects of hydroxychloroquine and sulphasalazine on progression of joint damage in rheumatoid arthritis. Lancet. 1989;1(8646):1036–8.
37. Genant HK, et al. Assessment of rheumatoid arthritis using a modified scoring method on digitized and original radiographs. Arthritis Rheum. 1998;41(9):1583–90.
38. Reid D, Miller CG. Clinical trials in rheumatoid arthritis and osteoarthritis, vol. viii. London: Springer; 2008. p. 325.
39. Ahmad HA. Rheumatoid arthritis: MRI as an efficacy endpoint in clinical trials. Int Clin Trials. Autumn 2010;82–4.
40. Hodgson R, et al. Dynamic contrast enhanced MRI of bone marrow oedema in rheumatoid arthritis. Ann Rheum Dis. 2008;67(2):270–2.
41. Conaghan PG, et al. Summary and recommendations of the OARSI FDA osteoarthritis Assessment of Structural Change Working Group. Osteoarthritis Cartilage. 2011; 19(5):606–10.
42. Hunter DJ, et al. Systematic review of the concurrent and predictive validity of MRI biomarkers in OA. Osteoarthritis Cartilage. 2011;19(5):557–88.
43. Kellgren JH, Lawrence JS. Rheumatism in miners. Part II: x-ray study. Br J Ind Med. 1952;9(3):197–207.
44. Hellio le Graverand M-P, et al. A 2-year randomised, double-blind, placebo-controlled, multicentre study of oral selective iNOS inhibitor, cindunistat (SD-6010), in patients with symptomatic osteoarthritis of the knee. Ann Rheum Dis. 2013;72(2):187–95.

45. Schiphof D, Boers M, Bierma-Zeinstra SM. Differences in descriptions of Kellgren and Lawrence grades of knee osteoarthritis. Ann Rheum Dis. 2008;67(7):1034–6.
46. Hellio Le Graverand M-P, et al. Considerations when designing a disease-modifying osteoarthritis drug (DMOAD) trial using radiography. Semin Arthritis Rheum. 2013; 43(1):1–8.
47. Nevitt MC, et al. Longitudinal performance evaluation and validation of fixed-flexion radiography of the knee for detection of joint space loss. Arthritis Rheum. 2007;56(5):1512–20.
48. Radiography Working Group of the OARSI-OMERACT Imaging Workshop, Le Graverand MP, Mazzuca S, Lassere M, Guermazi A, Pickering E, Brandt K, Peterfy C, Cline G, Nevitt M, Woodworth T, Conaghan P, Vignon E. Assessment of the radioanatomic positioning of the osteoarthritic knee in serial radiographs: comparison of three acquisition techniques. Osteoarthritis Cartilage. 2006;14(Suppl A):A37–43.
49. Eckstein F, et al. In vivo reproducibility of three-dimensional cartilage volume and thickness measurements with MR imaging. AJR Am J Roentgenol. 1998;170(3):593–7.
50. Eckstein F, et al. In vivo morphometry and functional analysis of human articular cartilage with quantitative magnetic resonance imaging–from image to data, from data to theory. Anat Embryol (Berl). 2001;203(3):147–73.
51. Eckstein F, et al. Long-term and resegmentation precision of quantitative cartilage MR imaging (qMRI). Osteoarthritis Cartilage. 2002;10(12):922–8.
52. Williams TG, et al. Measurement and visualisation of focal cartilage thickness change by MRI in a study of knee osteoarthritis using a novel image analysis tool. Br J Radiol. 2010;83(995):940–8.
53. Peterfy CG, et al. Whole-Organ Magnetic Resonance Imaging Score (WORMS) of the knee in osteoarthritis. Osteoarthritis Cartilage. 2004;12(3):177–90.
54. Hunter DJ, et al. The reliability of a new scoring system for knee osteoarthritis MRI and the validity of bone marrow lesion assessment: BLOKS (Boston Leeds Osteoarthritis Knee Score). Ann Rheum Dis. 2008;67(2):206–11.
55. Hunter DJ, et al. Evolution of semi-quantitative whole joint assessment of knee OA: MOAKS (MRI Osteoarthritis Knee Score). Osteoarthritis Cartilage. 2011;19(8):990–1002.
56. Spector TD, et al. Effect of risedronate on joint structure and symptoms of knee osteoarthritis: results of the BRISK randomized, controlled trial [ISRCTN01928173]. Arthritis Res Ther. 2005;7(3):R625–33.
57. Bingham CO, et al. Risedronate decreases biochemical markers of cartilage degradation but does not decrease symptoms or slow radiographic progression in patients with medial compartment osteoarthritis of the knee: results of the two-year multinational knee osteoarthritis structural arthritis study. Arthritis Rheum. 2006;54(11):3494–507.
58. Brandt KD, et al. Effects of doxycycline on progression of osteoarthritis: results of a randomized, placebo-controlled, double-blind trial. Arthritis Rheum. 2005;52(7):2015–25.
59. Manno RL, et al. OARSI–OMERACT initiative: defining thresholds for symptomatic severity and structural changes in disease modifying osteoarthritis drug (DMOAD) clinical trials. Osteoarthritis Cartilage. 2012;20(2):93–101.
60. Karsdal MA, et al. 64 oral calcitonin demonstrated symptom-modifying efficacy and increased cartilage volume: results from a 2-year phase 3 trial in patients with osteoarthritis of the knee. Osteoarthritis Cartilage. 2011;19:S35.
61. FDA. Guidance for industry preparation of IDEs and INDs for products intended to repair or replace knee cartilage additional. 2011. Available from: http://www.fda.gov/downloads/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/Guidances/CellularandGeneTherapy/UCM288011.pdf.
62. Marlovits S, et al. Magnetic resonance observation of cartilage repair tissue (MOCART) for the evaluation of autologous chondrocyte transplantation: determination of interobserver variability and correlation to clinical outcome after 2 years. Eur J Radiol. 2006;57(1):16–23.
63. Whelan DB, et al. Development of the radiographic union score for tibial fractures for the assessment of tibial fracture healing after intramedullary fixation. J Trauma. 2010;68(3):629–32.

64. Bracoud L, et al. Improving the accuracy of MRI spleen and liver volume measurements: a phase III Gaucher disease clinical trial setting as a model. Blood Cells Mol Dis. 2011;46(1):47–52.
65. Gordon CM, et al. Dual energy X-ray absorptiometry interpretation and reporting in children and adolescents: the 2007 ISCD pediatric official positions. J Clin Densitom. 2008;11(1):43–58.
66. Zemel BS, et al. Height adjustment in assessing dual energy x-ray absorptiometry measurements of bone mass and density in children. J Clin Endocrinol Metab. 2010;95(3):1265–73.
67. Henderson RC, et al. Pediatric reference data for dual X-ray absorptiometric measures of normal bone density in the distal femur. AJR Am J Roentgenol. 2002;178(2):439–43.
68. Henderson RC, et al. Bisphosphonates to treat osteopenia in children with quadriplegic cerebral palsy: a randomized, placebo-controlled clinical trial. J Pediatr. 2002;141(5):644–51.
69. Leonard MB. A structural approach to the assessment of fracture risk in children and adolescents with chronic kidney disease. Pediatr Nephrol. 2007;22(11):1815–24.
70. Greulich WW, Pyle SI. Radiographic atlas of skeletal development of the hand and wrist. Stanford: Stanford University Press; 1958.
71. Tanner JM, Hughes PC, Whitehouse RH. Radiographically determined widths of bone muscle and fat in the upper arm and calf from age 3–18 years. Ann Hum Biol. 1981;8(6):495–517.
72. DiGiovanna JJ, et al. Effect of a single course of isotretinoin therapy on bone mineral density in adolescent patients with severe, recalcitrant, nodular acne. J Am Acad Dermatol. 2004;51(5):709–17.
73. Fitzpatrick LA, et al. Mechanism of action study to evaluate the effect of rosiglitazone on bone in postmenopausal women with type 2 diabetes mellitus: rationale, study design and baseline characteristics. J Drug Assess. 2011;0:1–26.
74. Cuzick J, et al. Effect of anastrozole and tamoxifen as adjuvant treatment for early-stage breast cancer: 10-year analysis of the ATAC trial. Lancet Oncol. 2010;11(12):1135–41.

Chapter 12
Body Composition

John A. Shepherd

Abstract Human body composition is generally defined as the proportion of fat (adipose), muscle (parenchymal tissues), and bone (mineral) of the body. Body composition can be an indicator of disease or health. Besides fat, muscle, and bone, body composition can be described with varying levels of complexity. This chapter describes the models used to measure body composition and, because of their importance to clinical trials, goes into detail on the use of three imaging methods: dual-energy X-ray absorptiometry, quantitative computed tomography, and MRI.

Keywords Visceral adipose tissue • Subcutaneous adipose tissue • Percent body fat • Four-compartment model • Obesity

Introduction

Human body composition is generally defined as the proportion of fat (adipose), muscle (parenchymal tissues), and bone (mineral) of the body. Body composition can be an indicator of disease or health. Besides fat, muscle, and bone, body composition can be described with varying levels of complexity. This chapter describes the models used to measure body composition and, because of their importance to clinical trials, goes into detail on the use of three imaging methods: dual-energy X-ray absorptiometry (DXA), quantitative computed tomography (QCT), and MRI.

J.A. Shepherd, BS, PhD
Department of Radiology and Biomedical Imaging,
University of California at San Francisco,
1 Irving Street, Suite A-C108B, San Francisco, CA 94146, USA
e-mail: john.shepherd@ucsf.edu

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_12,

Modeling Body Composition

Wang and colleagues [1] summarized the various compartment models that are commonly used to describe human body composition. The simplest model is weight. It is in many cases hard to beat for precision, accuracy, and relationship to disease and health. If you divide by the participant's height squared, you have body mass index, BMI,

$$BMI = \frac{weight}{height^2} \frac{(kg)}{(\mathrm{m}^2)}$$

BMI is widely used in clinical trials because it is effectively free and proportional to whole-body percent fat by other methods. Flegal and colleagues [2] found that BMI, waste circumference, and waist–stature ratio were all highly correlated to percent body fat measured by dual-energy X-ray absorptiometry (DXA). The following are more complex models created by further subdividing body mass into compartments.

Two-Compartment Model

One of the simplest model used for body composition assessment is one that assumes that the body is divided in two compartments: fat mass (FM) and fat-free mass (FFM). Assuming that each of these compartments has a constant and uniform density (i.e., ~0.9 g/cm^3 for FM and 1.1 g/cm^3 for FFM), hydrodensitometry (underwater weighing) can be used to measure the whole-body density of an individual, and from this information the relative contribution of FM and FFM to total body weight can be calculated [3]. For example, the equation by Brozek and colleagues [4] has been widely used to convert body density into percent body fat:

$$\%\ Body\ \ fat = \left[\frac{4.57}{Total\ \ body\ \ density} - 4.142\right] \times 100$$

For the most part underwater weighing has been replaced by a technique called air displacement plethysmography (ADP) using a device called the BOD POD (Life Measurement Inc., Concord, CA, USA). Hydrodensitometry uses water displacement to measure body volume; the BOD POD uses air displacement within a sealed chamber. BOD POD is a broadly used clinical technique and is in use at research, clinical, and physical fitness facilities. Both BOD POD and hydrodensitometry assume a hydration constant of 0.723 to convert body density to % body fat. However, hydration can vary from person to person and by disease state.

The compartment models including a direct measure of water were developed to control for hydration status.

Three Compartment, Model 1

Initially, three compartment models were developed as extensions to the two-compartment models described previously. Body weight was divided into fat, water, and water-free lean (i.e., mineral and protein) [3]. The total body water measure is added to body density for this model. The advantages are that hydration status can change over time due to exercise, age, and disease state. The current gold standard method to measure total body water is the stable isotope dilution technique. A tracer, in these cases either deuterium (^{2}H) or oxygen-18 (^{18}O), is dissolved in the body by taking a drink of water that contains one of these stable isotopes. The tracer then distributes and reaches an equilibrium state throughout the body. A sample can then be taken of body water from saliva, urine, or blood, and by measuring the ratio of water containing the tracer to water, the total water volume can be determined. The limitations of this technique are that equilibrium typically takes 3–4 h and that protein is lumped with bone mineral. Thus, it is difficult to extract changes in protein from changes in bone and these changes are often coupled.

Three Compartment, Model 2 Using DXA

An alternative to the first three-compartment model described earlier is the use of DXA to measure three independent compartments. DXA defines the composition of the body as three materials having specific X-ray attenuation properties: bone mineral, lipid (triglycerides, phospholipid membranes, etc.), and lipid-free soft tissue. The term fat is commonly used to refer to adipose tissue. However, adipose tissue contains lipid-free mass such as water and proteins as well. Fat is best described chemically as the lipids in our body that are soluble in organic solvents and not in water, the largest category of body fat being triglycerides found in adipocytes. In this model, the non-lipid soft tissue mass is the sum of body water, protein, and soft tissue mineral mass. For each pixel in a DXA image, lipid, bone mineral, and soft tissue lean masses are being quantified. However, the distribution of the lipid, bone mineral, and non-lipid soft tissue within the volume projected onto the image pixel is not known. This model forces all tissue types into these three components. For example, the distinction between subcutaneous and visceral adipose tissue is lost for trunk measurements when both are projected in the same pixels. The same is true for skin, visceral non-adipose tissue, and muscle when all are projected in the same pixels. This limitation is true for most composition models that cannot represent the body as a true 3-dimensional volume. DXA systems are described in more detail in the later section.

Four-Compartment Model

The four-compartment model divides the body into the major molecular compartments of bone mineral, water, lipid, and protein:

$$Body\ \ mass = FM + TBW + BMC + total\ \ body\ \ protein$$

All but approximately 1 % of total body mass can be put into one of these molecular categories. The measures needed are body density using hydrodensitometry or air plethysmography, bone mineral content by DXA, and total body water. This model is considered by many to be the gold standard measure of total body composition, but it is rarely used in clinical trials because of the need for three measurement techniques and the length of time needed to acquire the water measure. In many studies, percent body fat using the three-compartment DXA model has been shown to be highly correlated to the full four-compartment model percent fat. Van Der Ploeg and colleagues [5] found the correlation to be $r^2 = 0.95$.

Tissue and Organ Compartment Models

If the composition or volume of an organ is needed independently of the surrounding tissue, a 3-dimensional imaging method, such as computed tomography (CT) or magnetic resonance imaging (MRI), is required. Three-dimensional (3D) imaging allows for organs to be isolated, slice by slice, through the volume from surrounding tissues. CT and MRI are the only methods that can isolate visceral fat from subcutaneous fat, muscle cross-sectional area from intramuscular adipose, and liver volume from other viscera as examples. The limitations to MRI and CT are that the access is limited, the scans are expensive in comparison to other methods, and, in the case of CT, significant dose is received by the participant.

Adipose Tissue

Adipose tissue is made up of lipid (85 %), proteins, minerals (3 %), and water (12 %) [6, 7] resulting in a physical density of approximately 0.92 kg/l. It is important to note that DXA, by definition, specifically measures the mass of total lipid, not adipose tissue. However, most researchers are interested in metabolic function. Adipose can be segregated anatomically to study its metabolic function. Adipose tissue is, in many cases, studied in terms of sub-compartments since many diseases are related to particular sub-compartments and not others. For example, visceral adipose is more strongly related to diabetes and cardiovascular disease than subcutaneous adipose. Table 12.1 is a summary of the subdivisions of adipose. Total AT can be measured by DXA, while the subdivisions of adipose can only be measured using CT or MRI. An example of a CT image showing superficial and deep subcutaneous adipose tissue is shown in Fig. 12.1a, b.

Table 12.1 Proposed classification of total body adipose tissue

AT component	Definition
Total AT	Sum of adipose tissue, usually excluding bone marrow and adipose tissue in the head, hands, and feet
Subcutaneous AT (SAT)	The layer found between the dermis and the aponeuroses and fasciae of the muscles. Includes mammary adipose tissue
Superficial subcutaneous AT	The layer found between the skin and a fascial plane in the lower trunk and gluteal–thigh area
Deep subcutaneous AT	The layer found between the muscle fascia and the fascial plane in the lower trunk and gluteal–thigh areas
Internal AT	Total adipose tissue minus subcutaneous adipose tissue
1. Visceral AT (VAT)	Adipose tissue within the chest, abdomen, and pelvis
2. Nonvisceral internal AT	Internal adipose tissue minus visceral adipose tissue
(a) Intramuscular AT	Adipose tissue within a muscle (between fascicles)
(b) Perimuscular AT	Adipose tissue inside the muscle fascia (deep fascia), excluding intramuscular adipose tissue
(i) Intermuscular AT	Adipose tissue between muscles
(ii) Paraosseal AT	Adipose tissue in the interface between muscle and bone (e.g., paravertebral)
(c) Other nonvisceral AT	Orbital adipose tissue; aberrant adipose tissue associated with pathological conditions (e.g., lipoma)

Used with permission from Shen et al. [8]

Visceral adipose fat can be further subdivided. Visceral or organ adipose tissue (VAT) is found in all three body cavities: intrathoracic (ITAT), intra-abdominal (IAAT), and intrapelvic (IPAT). However, most investigators report VAT as IAAT or the sum of IAAT and IPAT. The tree in Fig. 12.2 is a breakdown of the terminology proposed by Shen and colleagues for visceral adipose tissue components [8].

Subcutaneous Adipose Tissue (SAT)

SAT in the lower trunk and gluteal–thigh region can be subdivided into superficial and deep subcutaneous adipose tissue separated by a fascial plane (Table 12.1). Deep SAT is located primarily in the posterior half of the abdomen, while superficial SAT is more evenly distributed around the abdominal circumference [8, 9]. Differences have been reported between these two adipose tissue layers [8, 10–12]. Deep SAT has shown robust correlations to insulin resistance in both men and women, while superficial SAT showed little or no association [9]. Thus, the measure of SAT without distinction to these subcomponents dilutes the relationship. Furthermore, since whole-body DXA scans cannot discern the separating facial plane and the fact that the superficial and deep SAT overlap each other in DXA images, the ability of DXA to discern either component uniquely is unlikely.

SAT in the abdomen was examined in the Hill study outlined previously. The correlations of total abdominal SAT-defined CT area versus DXA were found to be $r=0.788$, while simple hip circumference [13] was $r=0.826$. Thus, there was no benefit by measuring SAT with DXA compared with the relatively simple hip circumference measure.

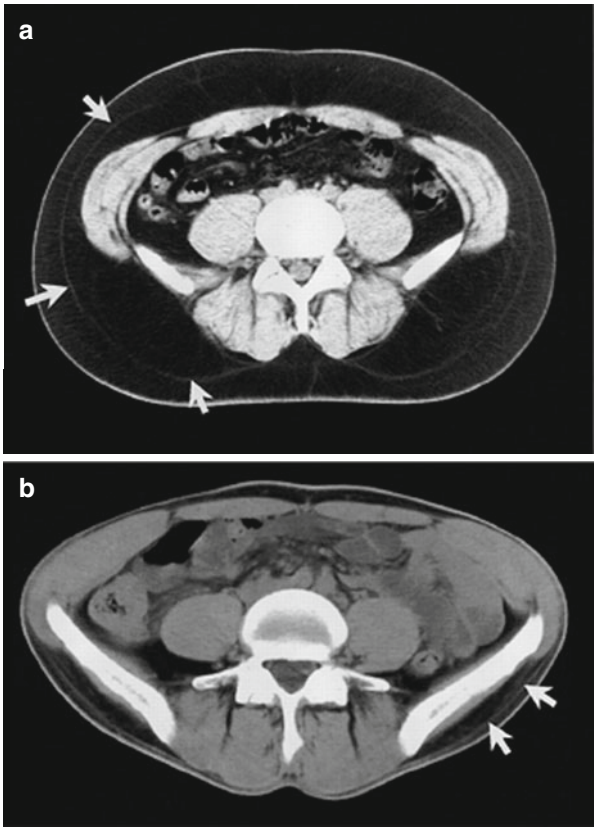


Fig. 12.1 (**a**, **b**). Abdominal axial CT scans of an obese (**a**) and a thin subject (**b**). Subcutaneous adipose tissue is divided into superficial and deep subcutaneous adipose tissue by a fascial plane, as indicated by the *white arrows*. It is unlikely that DXA could be used to discern superficial from deep SAT (Used with permission from Shen et al. [8])

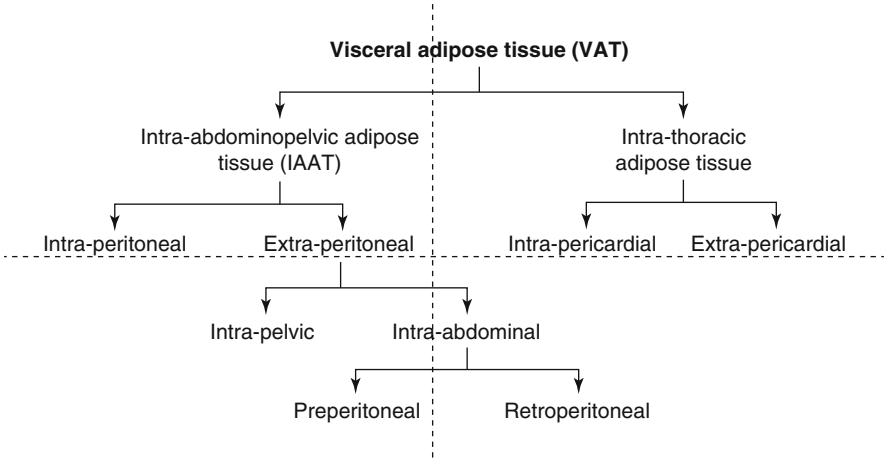


Fig. 12.2 Classification of visceral adipose tissue (*VAT*) as defined by Shen and colleagues [8] (Created from data in Shen et al. [8])

Body Composition Compartments and Measures

There are common terms used by many techniques to describe body composition. The truly independent values reported are relatively few. They are listed here.

Adipose Cross-Sectional Area (CSA)

The measure of adipose tissue is delineated within a CT or MRI slice using image segmentation. Subcutaneous and visceral adipose CSA are examples and are usually represented in cm^2.

Adipose or Organ Volume

Adjacent CT or MRI slices can be summed to create a volume of similar tissue. The volume of each slice is determined by the organ CSA and the thickness of the slice. Slice thicknesses typically range from 1 to 10 mm. Subcutaneous adipose or visceral can be summed in the abdomen region to create a total abdominal visceral and subcutaneous volume. The units of measure are typically cm^3.

Bone Mineral Content (BMC)

The mineral mass component of bone is called BMC. It is typically represented as hydroxyapatite, $Ca_{10}(PO_4)_6(OH)_2$. The units for BMC are typically grams. BMC is equivalent to the ashed weight of a bone sample and does not contain any mass of from the organic components of the bone (marrow, collagen, etc.). BMC can be defined using either DXA or CT but not MRI.

Fat Mass (FM)

Fat mass is generally defined as the mass that can be extracted using ether or other solvents from soft tissue. However, the definition can be confusing when looking between techniques. For example, DXA and CT (attenuation mode) cannot distinguish between the ether-extracted triglycerides and lipid membranes and connective tissue. Thus, DXA and CT will be influenced by the total lipid mass. FM is measured in either grams (g) or kilograms (kg).

Lean Soft Tissue Mass (LSTM)

Bone-free, fat-free soft tissue mass is the sum of all soft tissue lean, essentially water, protein, soft tissue mineral, and glycogen. It is measured in units of g or kg. This term is usually used for DXA whole-body scans.

Fat-Free Mass (FFM)

Fat-free mass is the sum of all the non-body lipid, such that

$$FFM = LSTM + BMC \ (\mathrm{g})$$

Soft Tissue Mass (STM)

Soft tissue mass is the sum of lean soft tissue and fat masses:

$$STM = FM + LSTM \ (\mathrm{g})$$

Total Body Mass (TBM)

Total body mass is the equivalent measure to scale weight, typically represented in units of kg or gm. In terms of the above,

$$TBM = FM + FFM = FM + BMC + LSTM \ (\mathrm{g})$$

TBM accuracy can be assessed against a calibrated scale. Studies that have investigated the agreement between scale mass and DXA total mass have found excellent agreement.

Percent Fat Mass (PCTFM)

Percent fat mass is a region's fat mass divided by the region's total mass times 100:

$$PCTFM = \frac{FM}{TBM} \times 100$$

Dual-Energy X-Ray Absorptiometry

For body composition, a DXA instrument scans the entire body using a dual-energy scanning protocol. The scan takes approximately 5 min and is considered lose dose. The test–retest precision is very high, 1.0 % or better standard deviation for repeat scans on the same individual for percent fat mass. In addition, DXA can report fat, lean soft tissue, and bone mineral mass compartments for anatomical subregions such as the arms, legs, and torso. For these reasons, DXA has been used in many clinical trials. The limitation to the DXA three-compartment model is that water is not explicitly solved for and changes in hydration can be interpreted as a change in the function lean tissue, such as protein. In addition, water is found in both adipose and parenchymal tissues. Monitoring change in muscle mass is confounded if adiposity also changes. DXA is a special imaging modality that is not typically available on general use X-ray systems because of the need for special beam filtering and near-perfect spatial registration of the two attenuations. Dedicated commercial DXA systems were first available in the late 1980s [14]. See Chap. 1.

The primary commercial application for DXA has been used to measure bone mineral density as an assessment of fracture risk and to diagnose osteoporosis, and the X-ray energies used are optimized for bone density assessment. For osteoporosis diagnosis, the lumbar spine, proximal hip, and sometimes the distal forearm are scanned. The regions of interest (ROIs) used and the diagnostic criteria are well defined. The whole body can also be scanned to measure whole-body bone mass and soft tissue body composition [15, 16]. For image areas that contain only soft tissue, lipid and lean tissue can be assessed [17], from which percent lipid mass can be calculated, while areas that contain bone use an estimated percent lipid from the surrounding tissue [18]. Reference populations have been scanned and defined by sex, ethnicity, and age. Diagnosis of disease is typically undertaken by comparing individuals to their peer group or to a young healthy population. Currently there are estimated to be over 50,000 whole-body DXA systems in use worldwide.

Why Use DXA for Clinical Trials Instead of Other Body Composition Methods?

The only imaging methods that can accurately estimate regional bone, fat, and lean soft tissue mass distributions are DXA, CT, and MRI. DXA is low dose in comparison to whole-body CT scanning and inexpensive compared to MRI. In addition, DXA %fat measures are more precise than is typical with CT and MRI. Lastly, DXA is easily tied to physical standards that are verifiable in the field, such as stearic acid and water, such that cross-calibration and pooling of data across clinical centers are possible. However, there are important limitations. First, DXA systems measure bone density in units of grams per unit area since it

does not have the ability to measure tissue thickness. Thus, DXA systems cannot tell the difference between thick low-density bone and thin high-density bone. Secondly, DXA can only solve for two materials simultaneously. This is a fundamental theorem of X-ray absorptiometry since the attenuation characteristics of any one material can be represented by combining two other materials together in the appropriate way [19]. Thus, soft tissue composition can only be solved in areas exclusive of bone, and bone mass can only be determined with an assumption of the soft tissue composition overlaying the bone. Since bone is contained in typically 40 % or more of the body image pixels, the soft tissue composition has to be estimated from surrounding tissue. In some cases, accurate estimates cannot be made, such as the head, hands, feet, and upper torso because there is no adequate soft tissue outside the bone projection and manufacturers turn to proprietary methods to reference the soft tissue. Third, there is a lack of standardization in the measured values. For example, the differences in BMC between the two largest manufacturers, Hologic (Hologic, Inc., Bedford, MA, USA) and GE Lunar (GE Healthcare, Madison, WI, USA) systems, are approximately 20 %. Equations have recently been published that provide a means of converting Hologic to GE measures and vice versa [20].

The cross-calibration relationships were derived in over 200 individuals ranging in age from 6 to 81 years. Care must be exercised though when using the equations in Table 12.2 since these equations were derived for specific software versions.

Special DXA Regions of Interest

Since the conception of DXA the subregions defined on whole-body scans have been the arms, legs, trunk, and head for soft tissue reporting. However, in recent years, there has been a progressive work to report special regions beyond these anatomical regions.

There are two regions of interest that measure abdominal fat: android and visceral adipose fat. Android fat and percent fat are defined by GE and Hologic within a region extending vertically from the top of the iliac crest to a height 1/5th the length between the crest and the chin and horizontally across the entire abdomen. There have been some reports on the relationship of this region to metabolic risk factors in both men and women [21]. The android region can also be paired with a gynoid special region to provide a densitometric definition of the waist-to-hip circumference ratio, the android-to-gynoid percent fat ratio [22].

DXA system can estimate visceral adipose fat as either a cross-sectional area (Hologic) or a tissue volume (GE) [23, 24]. These VAT estimates are made by subtracting off estimates of the overlaying subcutaneous fat from the total android fat. The correlation of VAT by DXA and VAT by CT is very high, greater than $r = 0.90$. It is too early to know if these DXA VAT estimates will completely replace CT measures in clinical trials and also be useful in the clinical evaluation of CVD risk. A report from a Hologic DXA system is shown in Fig. 12.3.

Table 12.2 Cross-calibration relationships for bone and soft tissue whole-body values. GE = "GE Healthcare Lunar"

Variables	Cross calibration equation (≤40 kg)	Cross calibration equation (>40 kg)	r^2 (pooled data)
BMD	$GE = 0.052 + 0.975 \times BMD_{Hologic} - (1.169E\text{-}6) \times LSTM_{Hologic} - (1.651E\text{-}6) \times TOT\_FAT_{Hologic}$	$GE = -0.148 + 0.986 \times BMD_{Hologic} + (2.462E\text{-}06) \times LSTM_{Hologic} + (2.906E\text{-}06) \times TOT\_FAT_{Hologic}$	0.96
	$Hologic = 0.024 + 0.826 \times BMD_{GE} + (3.47E\text{-}6) \times LSTM_{GE} + (1.17E\text{-}6) \times TOT\_FAT_{GE}$	$Hologic = 0.204 + 0.907 \times BMD_{GE} - (1.33E\text{-}6) \times LSTM_{GE} - (2.62E\text{-}6) \times TOT\_FAT_{GE}$	
BMC	$GE = -123.517 + 0.891 \times BMC_{Hologic} + 0.013 \times LSTM_{Hologic} + 0.007 \times TOT\_FAT_{Hologic}$	$GE = 5.9797 + 0.843 \times BMC_{Hologic} + 0.012 \times LSTM_{Hologic} - 0.001 \times TOT\_FAT_{Hologic}$	0.99
	$Hologic = 134.166 + 1.035 \times BMC_{GE} - 0.010 \times LSTM_{GE} - 0.009 \times TOT\_FAT_{GE}$	$Hologic = 3.741 + 1.085 \times BMC_{GE} - 0.009 \times LSTM_{GE} + 0.001 \times TOT\_FAT_{GE}$	
SubBMD	$GE = -0.034 + 1.183 \times SubBMD_{Hologic} - (4.395E\text{-}06) \times LSTM_{Hologic} - (7.820E\text{-}07) \times TOT\_FAT_{Hologic}$	$GE = -0.218 + 1.133 \times SubBMD_{Hologic} + 8.276E\text{-}07 \times LSTM_{Hologic} + 3.209E\text{-}06 \times TOT\_FAT_{Hologic}$	0.97
	$Hologic = 0.054 + 0.748 \times SubBMD_{GE} + (5.32E\text{-}06) \times LSTM_{GE} + (6.118E\text{-}07) \times TOT\_FAT_{GE}$	$Hologic = 0.219 + 0.810\ SubBMD_{GE} + (9.334E\text{-}08) \times LSTM_{GE} - (2.71E\text{-}06) \times TOT\_FAT_{GE}$	
SubBMC	$GE = -164.957 + 0.893 \times SubBMC_{Hologic} + 0.012 \times LSTM_{Holohic} + 0.007 \times TOT\_FAT_{Hologic}$	$GE = -21.31 + 0.883 \times SubBMC_{Hologic} + 0.009 \times LSTM_{Hologic} - 0.001 \times TOT\_FAT_{Hologic}$	0.99
	$Hologic = 164.144 + 1.038 \times SubBMC_{GE} - 0.009 \times LSTM_{GE} - 0.008 \times TOT\_FAT_{GE}$	$Hologic = 13.428 + 1.016 \times SubBMC_{GE} - 0.006 \times LSTM_{GE} + 0.001 \times TOT\_FAT_{GE}$	
PFAT	$GE = -2.173 + 1.119 \times PFAT_{Hologic}$		0.97
	$Hologic = 1.943 + 0.894 \times PFAT_{GE}$		
TM	$GE = 1.00 \times TM_{Hologic}$		0.99
ALSTM	$GE = -827.06 + 1.011 \times ALSTM_{Hologic}$		0.99
	$Hologic = 818.24 + 0.989 \times ALSTM_{GE}$		
TrunkPFAT	$GE = -3.275 + 1.248 \times TrunkPFAT_{Hologic}$		0.96
	$Hologic = 2.623 + 0.801 \times TrunkPFAT_{GE}$		
LegPFAT	$GE = 1.482 + 0.941 \times LegPFAT_{Hologic}$		0.94
	$Hologic = -1.576 + 1.063 \times LegPFAT_{GE}$		

(continued)

Table 12.2 (continued)

Variables	Cross calibration equation (≤40 kg)	Cross calibration equation (>40 kg)	r^2 (pooled data)
AndroidPFAT	$GE = -6.095 + 1.196 \times AndroidPFAT_{Hologic}$		0.95
	$Hologic = 5.095 + 0.836 \times AndroidPFAT_{GE}$		
GynoidPFAT	$GE = -10.917 + 1.406 \times GynoidPFAT_{Hologic}$		0.93
	$Hologic = 7.764 + 0.711 \times GynoidPFAT_{GE}$		
A/G Ratio	$GE = -0.401 + 1.449 \times A/G\ Ratio_{Hologic}$		0.86
	$Hologic = 0.2767 + 0.6899 \times A/G\ Ratio_{GE}$		

Used with permission from Shepherd et al. [20]

BMD total body bone mineral density, *BMC* total body mineral content, *LSTM* lean soft tissue mass, *FAT* fat mass, *TM* total mass, *PFAT* percent fat, *SubBMD* headless BMD, *SubBMC* headless BMC, *LegPFAT* percent fat of the legs, *TrunkPFAT* percent fat of the trunk, *ALSTM* appendicular lean soft tissue mass. *AndroidPFAT* percent fat of the android region, *GynoidPFAT* percent fat of the gynoid region, *A/G Ratio* (AndroidPFAT/GynoidPFAT)

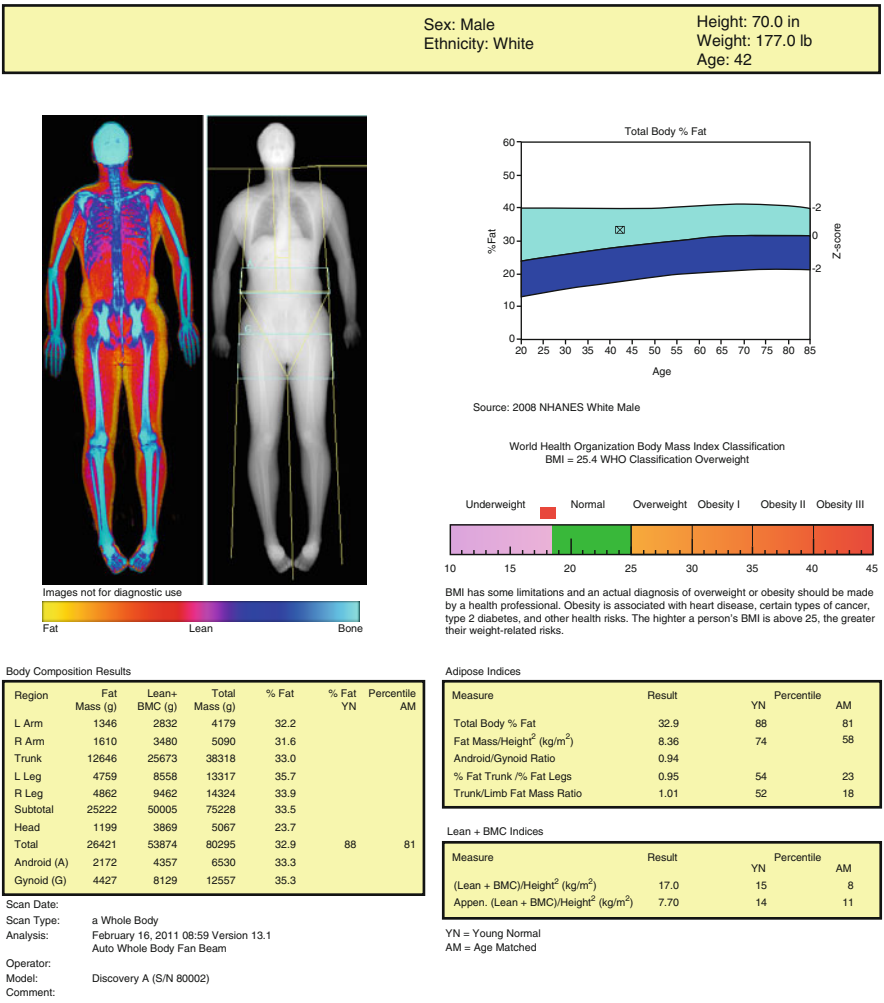


Fig. 12.3 DXA body composition report from a Hologic system. The individual regions used for the arms, legs, trunk, and whole body are divided by the *yellow cutlines*. Regions *A* and *G* are the android and gynoid regions. The rectangular subregion within region *A* is VAT region of interest

DXA Body Composition Reference Data

The largest study to date for body composition reference values is the National Health and Nutrition Examination Survey (NHANES). NHANES started collecting whole-body DXA scans in 1999 and continued through 2004. Over 22,000 Americans were selected to be representative of the USA and received a whole-body DXA scan. NHANES has reference values for percent fat, lean mass, and BMC

by sex, age (8–85 years), and ethnicity. Table representations can be found online at http://www.cdc.gov/nchs/nhanes/dxx/dxa.htm and summarized in a paper by Kelly and colleagues [25]. These reference values can be selected in the software of both Hologic and GE systems.

Scanning Obese Patients with DXA

There are several challenges in the scanning and analysis of heavy patients. First, the DXA systems have weight limits and table dimensions that restrict the size of the person to be scanned. DXA scanner table weight limits are generally 300 lb (136 kg), but some systems scan up to 450 lb (205 kg). Obese patients are also thicker and attenuate the X-ray more. Thus, some manufacturers provide special scan modes and analysis techniques for thicker patients. These scan modes in general are the same X-ray tube voltage settings but with either a higher mAs or slower scan time to increase the X-ray flux. GE systems will automatically alert the user to the need for the "thick" scan mode if the patient's weight exceeds a particular level. The dose is increased from 0.4 to 0.8 μSv. Hologic provides a "high-power whole-body" scan mode. This mode should be used if there is a noticeable increase in X-ray noise in the torso region. The dose is increased from 8.5 to 28.3 μSv. It is also sometimes difficult to fit an obese patient into the scan field. A "hemiscan" or "reflection" protocol can be used if the patient absolutely is too wide for the table. The patient is positioned off the center line of the scan table to ensure that one side, typically the right side, is completely included in the san field. See Fig. 12.4 (right). Tataranni and Ravussin [26] found that the accuracy of DXA body composition results of half-body scanning was not different from whole-body scanning ($r^2 \geq 0.98$).

DXA Quality Control

Scanner quality control (QC) procedures are used to monitor scanner performance throughout the course of a study or during general use. Longitudinal QC procedures consist of daily procedures used to monitor the performance of a single scanner over time. Cross-calibration procedures are used to monitor scanner variation between systems. There are few options for whole-body DXA quality control. Commercially available phantoms include the Hologic whole-body phantom (no longer manufactured), the BioClinica variable composition phantom and now the newer BioClinica body composition phantom (BBCP) (BioClinica, Inc., Newtown, PA, USA), and the OrthoMetrics whole-body phantom (OrthoMetrics, Inc., White Plains, NY, USA). All of these phantoms have been used for longitudinal calibration corrections and cross-calibration between similar systems. At the present time, none of the

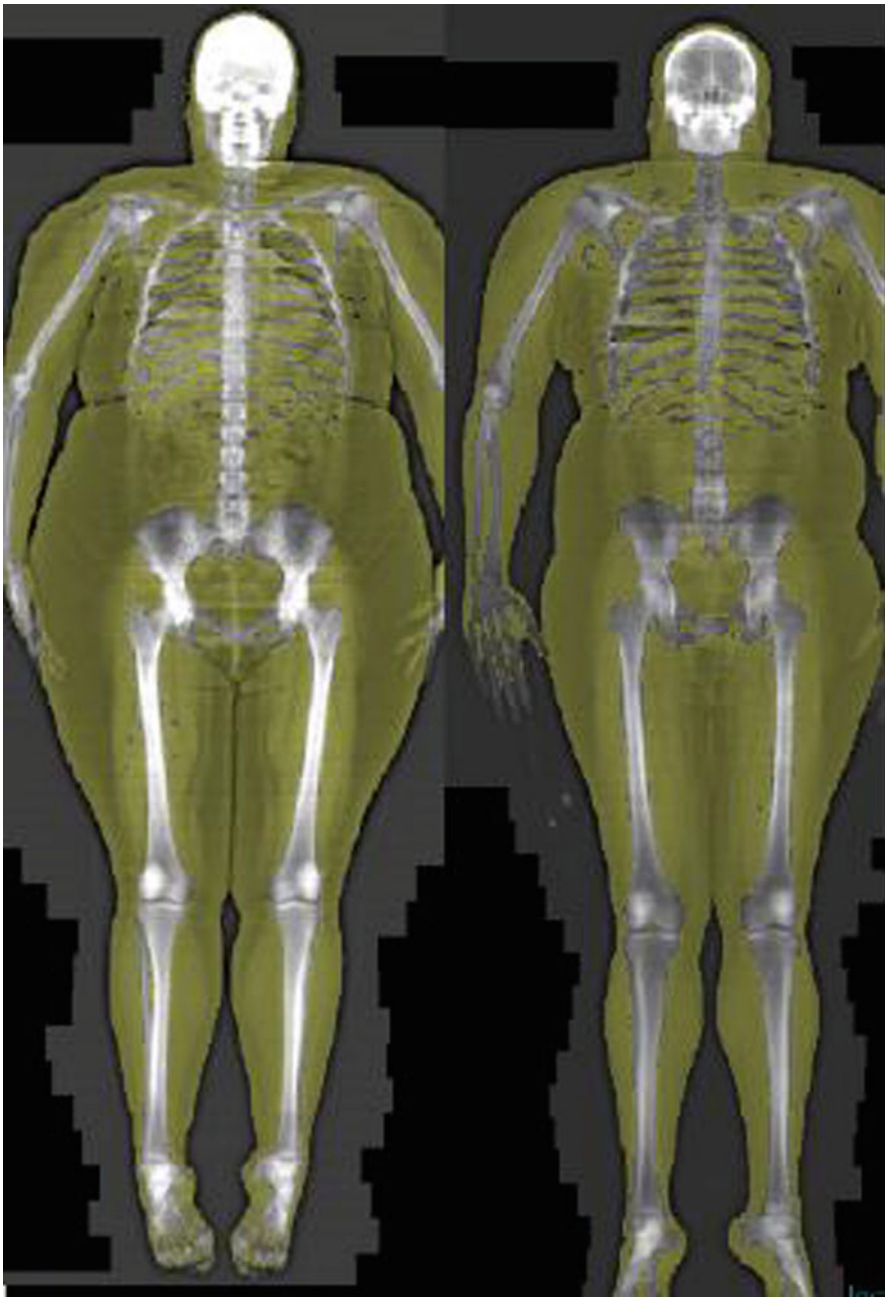


Fig. 12.4 Inaccurate scan since both arms are lost. Patient should be rescanned for hemiscan analysis (*left*). Shifted scan appropriate for imputation of the left side (*right*) (Courtesy of Mary K. Oates, CA, USA)

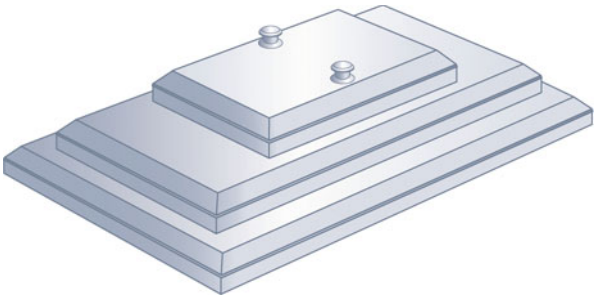


Fig. 12.5 Whole-body phantom assembly (Courtesy of Hologic, Inc.)

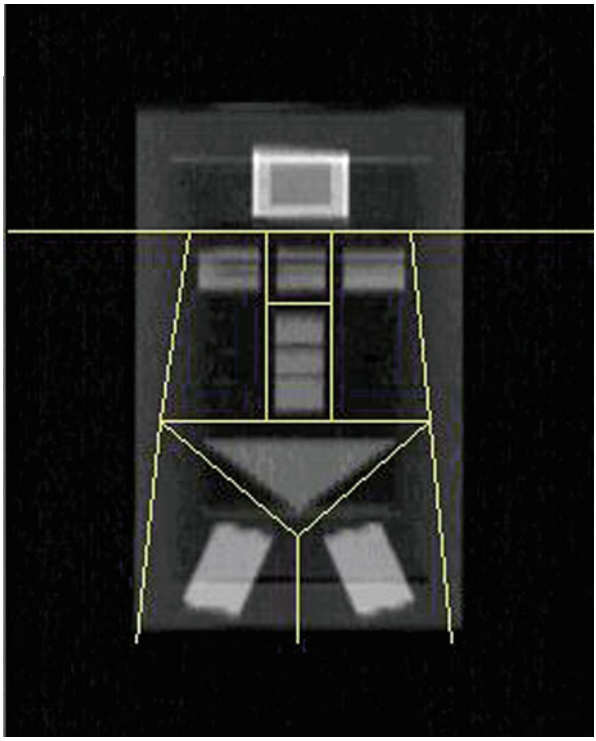


Fig. 12.6 DXA image of the whole-body phantom (Courtesy of Hologic, Inc.)

phantoms have been shown to be appropriate for cross-calibration between systems of different makes and models.

The assembled Hologic phantom is shown in Fig. 12.5. A DXA image of the phantom is shown in Fig. 12.6. The phantom is 40 kg assembled. Before lifting or transporting the phantom, break it down into its individual components. Do not attempt to lift the entire phantom assembly.

Changes in scanners, software, or location of scanners can have a large impact on the integrity of study data. For this reason, such changes are typically not allowed for the duration of a study or clinical trial without prior notification and approval of the study principal investigator.

These are the strategies that can be used to minimize precision error:

1. All operators should be formally trained in positioning and analysis for each scan mode used.
2. Patients should be scanned on the same densitometer, not a similar model in the same clinic. Scans from different make and model systems cannot be quantitatively compared.
3. The same operator should be used for the baseline and follow-up scans.
4. The patient should be positioned using the standardized procedure suggested by the manufacturer or study protocol.
5. The scan mode should not be changed between baseline and follow-up scans. A scanner may offer a quick, a normal, and a high-resolution option for a given skeletal site. Always use the scan mode for the follow-up as was used for the baseline.
6. Identical ROIs should be used on each scan and placed consistently. The "compare" or "copy" function should always be used if available.
7. Auto-analysis algorithms should be used and checked by operator and only modified when necessary and at a minimum.

Tissue Compartment Imaging Using CT and MRI

Computed tomography (CT) and magnetic resonance imaging (MRI) scans generate cross-sectional images of the body (slices). These slices can be processed to create 3-dimensional organ volumes. Full-bore clinical CT and MRI systems can scan any part of the body and accommodate a wide variety of body sizes. With a CT or MRI scan, estimation of fat tissue can be performed by either absorptiometry (CT), saturation mode (MRI), or segmentation (CT and MRI). In CT absorptiometry, the attenuation unit (*voxel*, the 3-dimensional volume element in a CT image) is decomposed into a fat and a lean tissue attenuation to derive a volume fraction for that voxel. An example of absorptiometry is the determination of liver fat content. In segmentation, the boundaries of well-defined regions of fat or lean tissues are segregated and quantified (quantitative CT). In MRI saturation modes, the fat or water is imaged independently of the other compartments and the voxel value is directly proportional to the amount of fat or water depending on the saturation mode. The segmentation approach is independent of attenuation, relying instead on the accurate delineation of boundaries of the regions of interest. If done by a skilled professional, the delineations can be very time consuming and thus usually limited to processing a few slices. Advanced image processing methods can automate the process and allow for the analysis of more slices [27]. CT is the leading technique for the study of regional fat content [28, 29]. Its most common research application for body composition is to quantify subcutaneous and visceral fat. There are many protocols used that vary in radiation dose and region of the abdomen images. A popular protocol uses a single transversal slice centered on the umbilicus. In using the

umbilicus as a marker, the scout scan can be eliminated to reduce dose. However, this protocol is only suitable for soft tissue measures since the slice falls arbitrarily onto the spine. Another variation is to use single-slice scans either centered in the intervertebral space or on a single vertebra. These require a short scout scan to position the slice. If the slice is centered in either T12, L1, L2, or L3, reference data can be used to evaluate that vertebrae's bone density. Arguably the slice for most optimum precision is the space between L2 and L3. CT volumetric bone density is a marker for osteoporosis. Attenuation values in Hounsfield units are compared with reference standards for adipose tissue, and regions corresponding to visceral, subcutaneous, and retroperitoneal fat depots are delineated by the operator. In order to simplify this task, automated methods have been proposed for identification of these compartments in the scan [27]. The accuracy of CT to predict visceral and total abdominal fat is very good. Using an 11-slice protocol, the correlation of CT abdominal fat measurements with cadaver analysis is >0.90 [30]. The reproducibility of the technique is also excellent, with <1 % variability in repeated measurements [31]. A single slice at the L4–L5 level also shows a high correlation with total visceral fat volume. However, Shen and colleagues [8] have shown that measuring visceral fat area 5 cm (women) or 10 cm (men) above the L4–L5 level provides better estimates of visceral fat volume. The authors suggest that this may be due to the fact that visceral fat consists of extra- and intraperitoneal fat, the latter being the largest contributor to intersubject variability.

CT and MRI in Clinical Trials

CT and MRI are broadly used in clinical trials in the delineation mode. There are less stringent requirements for quality control for delineation than for attenuation measures. In adult studies, CT is the preferred technique for reasons of speed, availability, and ease of protocol execution and analysis. CT is also calibrated to absolute standards of attenuation. MRI is the preferred method for research in children and adolescents where dose considerations are important [32]. Calibration phantoms are readily available for soft tissue and bone density. Image Analysis, Inc. (Image Analysis, Columbia, KY, USA) and Mindways (Mindways Software, Austin, TX) both provide commercial software to analyze CT images. Both also provide scan protocols and methods for exporting scans. MRI scans can only be analyzed for bone density and soft tissue areas using research software.

Radiation Protection Regulations for the Use of DXA and CT

All countries require that CT systems be operated by trained and certified personnel. Most countries require that the legal person responsible for the DXA facility unit applies to the radiation protection regulatory body for an authorization – either a

registration or a license. General requirements for protection and safety are given in the International Basic Standards for Protection against Ionizing Radiation and of the Safety of Radiation Sources [33], with more specific guidance in IAEA publications Radiological Protection for Medical Exposure to Ionizing Radiation [34] and Applying Radiation Safety Standards in Diagnostic Radiology and Interventional Procedures using X-rays [35]. Once satisfied, the regulatory body issues an authorization, typically with conditions or limitations that would need to be complied with.

Dose

Chapter 3 of this book goes into radiation dosage, but this section provides more granularity of radiation doses in body composition assessments. DXA systems generate ionizing radiation, and subjects being scanned and equipment operators consequently receive some (small) radiation dose as a result of any procedure. The absorbed dose to tissue is quantified as the amount of energy absorbed in a kilogram of tissue. The unit of measure is the gray (Gy) where 1 Gy is equivalent to 1 J/kg. Another useful quantity of dose is "effective dose" measured in sieverts (Sv). Effective dose takes account not only of the amount of energy absorbed but also of the type of radiation and the susceptibility of the tissue to radiation damage. Effective dose is used in assessing occupational and public exposure to radiation. It is also useful in characterizing the dose typically received by a patient from a given X-ray procedure. Patient effective doses in CT depend on the kVp and mAs used as well as the slice thickness and slice overlap. For CT protocols, dose should be calculated by a medical physicist. Effective dose for DXA depends on the type of unit (pencil beam, fan beam, cone beam), the protocol or mode used for the scan (scan area, tube current, scan speed), and the body region being scanned. However, DXA scan protocols are predefined and very low dose compared to CT or other imaging methods.

DXA scans of the whole body result in an effective dose of about 10 μSv for a fan-beam unit [36–41]. The patient dose may change by a factor of 1.5–3 [36], or more, between using the lowest and highest dose mode for the same examination. Pediatric patient effective doses, using an appropriate pediatric protocol, are similar to those for an adult [36, 37, 39]. However, adult protocols applied to children can lead to doses approaching 20 μSv [36]. CT doses reported in the literature for visceral fat measures are wide ranging. Table 12.3 is a sample of CT protocols used to measure visceral fat and range from 0.1 to 4.1 mSv. The range has to do with the volume of the abdomen scanned and the spatial contrast desired by the author. The minimum dose for optimum visceral measures is not known and should be the subject of standardization efforts.

To put these DXA and CT patient doses into perspective, it is helpful to consider exposure from other sources. Human beings are constantly exposed to ionizing radiation from natural sources including cosmic rays and naturally occurring radioactive materials in our foods, soil, water, and air. This is collectively referred to as

Table 12.3 Effective dose calculations for a sample of CT studies for visceral fat

Study	Population	Region	Technical factors	Estimated effective dose (mSv)	Reference
A	7 male; 32.7 +/− 9.3 years	Lower rib–iliac crest	120 kVp, 240 mAs, 12 mm slice width	0.4	Seidell et al. [42]
B	11 volunteers (21–49 years)	Umbilicus	130 kVp, 385 mAs, 10 mm slice width	0.8	Sobol et al. [43]
C	19 premenopausal women	L_2–L_4	120 kV, 450 mAs, 10 mm slice width	4.1 (0.7)	Greenfield et al. [44]
D	75 boys (6–14 years)	Umbilicus	120 kVp, 400 mAs, and 10 mm slice width	1.2 (6 years old)	Asayama et al. [45]
E	54 volunteers	L_2–L_3 interspace	120 kV, 150 mAs, 5 mm slice width	0.1	Potretzke et al. [46]

Table 12.4 Dose constraints for participants in research studies. These constrains are considered the upper limit of dose that should be asked of participants that receive no benefit from the research themselves

Participant category	Dose constraint	Equivalent number of WB DXA(<0.005 mSv)
Adult		
Total effective dose		
In any year	5 mSv	1,000
Over 5 years	10 mSv	2,000
Total effective dose in adult with life expectancy less than 5 years		
In any year	50 mSv	10,000
Equivalent dose to skin averaged over 1 cm^2		
In any year	200 mSv	n/a
Equivalent dose to any other organ or tissue		
In any year	100 mSv	n/a
Children and foetuses		
Total effective dose to age 18 years	5 mSv	1,000
Subject to:		
Effective dose from conception to birth	0.1 mSv	20
Effective dose in any year from birth to 18 years	0.5 mSv	100
Total equivalent dose to age 18 years to any organ or tissue	100 mSv	n/a

natural background radiation. The average annual natural background radiation dose to humans worldwide is about 2,400 μSv, but this can vary from 1,000 to 10,000 μSv with some populations receiving 20,000 μSv/year [47]. Furthermore, guidelines for dose constraints for participants in research are available [48]. See Table 12.4. These dose constraints in the form published by the government of Australia are for keeping the dose at minimal risk and reasonable to use in studies that are of limited benefit to the individual. Protocols that use doses above these levels should be held to higher scrutiny from Human Research Ethics Committees

by providing justification as to why higher doses are necessary and cannot be avoided before initiating the protocol.

Thus, in comparison, effective patient doses from DXA are small and are similar to that received on average from 1 or 2 day exposure to natural background radiation. From Table 12.4, DXA doses are very small in comparison to the allowed dose to research subjects. The dose of the CT visceral fat protocols is from 10 to 400 times higher than DXA. However, even the highest dose protocol is technically below the constraints for adults or children if no other research radiation procedures were performed. But care should be taken to minimize CT dose when possible.

Summary

In summary, body composition can be represented in terms of simple two-compartment models or more basic four- and five-compartment models. DXA, CT, and MRI methods all have unique capabilities for measuring body composition in clinical trials. The strengths of DXA include low radiation dose, high precision, low cost compared to CT and MRI, and wide availability. For these reasons, there are extensive DXA data and studies available for adults and children, for whole body and subregions, and for individuals up to 450 lb. CT and MRI are considered the gold standard for sub-compartments of fat, muscle, and organs that cannot be isolated in projectional DXA images including cross-sectional muscle areas, liver, and visceral fat. However, less standardization exists for CT and MRI than for DXA body composition protocols.

References

1. Wang ZM, Pierson Jr RN, Heymsfield SB. The five-level model: a new approach to organizing body-composition research. Am J Clin Nutr. 1992;56(1):19–28.
2. Flegal KM, Ogden CL, Yanovski JA, Freedman DS, Shepherd JA, Graubard BI, Borrud LG. High adiposity and high body mass index-for-age in US children and adolescents overall and by race-ethnic group. Am J Clin Nutr. 2010;91(4):1020–6.
3. Siri WE. Body composition from fluid spaces and density: analysis of methods. In: Brozek J, Henzchel A, editors. Techniques for measuring body composition. Washington, DC: National Academy of Sciences; 1961. p. 224–44.
4. Brozek J, Grande F, Anderson JT, Keys A. Densitometric analysis of body composition: revision of some quantitative assumptions. Ann N Y Acad Sci. 1963;110:113–40.
5. Van Der Ploeg GE, Withers RT, Laforgia J. Percent body fat via DEXA: comparison with a four-compartment model. J Appl Physiol. 2003;94(2):499–506.
6. Thomas LW. The chemical composition of adipose tissue of man and mice. Q J Exp Physiol Cogn Med Sci. 1962;47:179–88.
7. Abate N, Burns D, Peshoc RM, Garg A, Grundy SM. Estimation of adipose tissue mass by magnetic resonance imaging: validation against dissection in human cadavers. J Lipid Res. 1994;35(8):1490–6.

8. Shen W, Wang Z, Punyanita M, Lei J, Sinav A, Kral JG, Imielinska C, Ross R, Heymsfield SB. Adipose tissue quantification by imaging methods: a proposed classification. Obes Res. 2003;11(1):5–16.
9. Blunt B, Nevitt M, Fuerst T, Palermo L, Cauley J, Tylavsky F, Kelly T. Hologic whole body phantom: using in vitro data to correct in vivo whole body data. J Bone Miner Res. 2000;15 Suppl 1:S398.
10. Lovejoy JC, Smith SR, Rood JC. Comparison of regional fat distribution and health risk factors in middle-aged white and African American women: the healthy transitions study. Obes Res. 2001;9(1):10–6.
11. Smith SR, Lovejoy JC, Greenway F, Ryan D, de Jonge L, de la Bretonne J, Volafova J, Bray GA. Contributions of total body fat, abdominal subcutaneous adipose tissue compartments, and visceral adipose tissue to the metabolic complications of obesity. Metabolism. 2001;50(4):425–35.
12. Kelley D, Thaete F, Troost F, Huwe T, Goodpaster B. Subdivisions of subcutaneous abdominal adipose tissue and insulin resistance. Am J Physiol Endocrinol Metab. 2000;278(5):E941.
13. Norton K, Olds T. Anthropometrica. Sydney: UNSW Press; 1996.
14. Kelly TL, Slovik DM, Neer RM. Calibration & standardization of bone mineral densitometers. J Bone Miner Res. 1989;4(5):663–9.
15. Laskey M, Phil D. Dual-energy x-ray absorptiometry and body composition. Nutrition. 1996;12(1):45–51.
16. Kelly TL, Berger N, Richardson TL. DXA body composition: theory and practice. Appl Radiat Isot. 1998;49(5–6):511–3.
17. Pietrobelli A, Formica C, Wang Z, Heymsfield SB. Dual-energy x-ray absorptiometry body composition model: review of physical concepts. Am J Physiol. 1996;271(6 Pt 1):E941–51.
18. Blake GM, Fogelman I. Technical principles of dual energy x-ray absorptiometry. Semin Nucl Med. 1997;27(3):210–28.
19. Lehmann LA, Alvarez RE, Macovski A, Brody WR, et al. Generalized image combinations in dual KVP digital radiography. Med Phys. 1981;8(5):659–67.
20. Shepherd JA, Fan B, Lu Y, Wu XP, Wacker WK, Ergun DL, Levine MA. A multinational study to develop universal standardization of whole body bone density and composition using GE Healthcare Lunar and Hologic DXA systems. J Bone Miner Res. 2012;27(10):2208–16.
21. Lee K, Lee S, Kim YJ. Waist circumference, dual-energy X-ray absortiometrically measured abdominal adiposity, and computed tomographically derived intra-abdominal fat area on detecting metabolic risk factors in obese women. Nutrition. 2008;24(7–8):625–31.
22. Wiklund P, Toss F, Weinehall L, Hallmans G, Franks PW, Nordstrom A, Nordstrom P. Abdominal and gynoid fat mass are associated with cardiovascular risk factors in men and women. J Clin Endocrinol Metab. 2008;93(11):4360–6.
23. Micklesfield LK, Goedecke JH, Punyanitya M, Wilson KE, Kelly TL. Dual-energy x-ray performs as well as clinical computed tomography for the measurement of visceral fat. Obesity. 2012;20(5):1109–14.
24. Kaul S, Rothney MP, Peters DM, Wacker WK, Davis CE, Shapiro MD, Ergun DL. Dual-energy X-ray absorptiometry for quantification of visceral fat. Obesity. 2012;20(6):1313–8.
25. Kelly TL, Wilson KE, Heymsfield SB. Dual energy X-ray absorptiometry body composition reference values from NHANES. PLoS One. 2009;4(9):e7038.
26. Tataranni PA, Ravussin E. Use of dual-energy x-ray absorptiometry in obese individuals. Am J Clin Nutr. 1995;62(4):730–4.
27. Zhao B, Colville J, Kalaigian J, et al. Automated quantification of body fat distribution on volumetric computed tomography. J Comput Assist Tomogr. 2006;30:777–83.
28. Ellis KJ, Shypailo RJ. Bone mineral and body composition measurements: cross-calibration of pencil-beam and fan-beam dual-energy X-ray absorptiometers. J Bone Miner Res. 1998;13:1613–8.
29. Ravussin E, Smith SR. Increased fat intake, impaired fat oxidation, and failure of fat cell proliferation result in ectopic fat storage, insulin resistance, and type 2 diabetes mellitus. Ann N Y Acad Sci. 2002;967:363–78.

30. Freedland ES. Role of a critical visceral adipose tissue threshold (CVATT) in metabolic syndrome: implications for controlling dietary carbohydrates: a review. Nutr Metab (Lond). 2004;1(12):12.
31. Enzi G, Gasparo M, Biondetti PR, Fiore D, Semisa M, Zurlo F. Subcutaneous and visceral fat distribution according to sex, age, and overweight, evaluated by computed tomography. Am J Clin Nutr. 1986;44:739–46.
32. Owens S, Gutin B, Ferguson M, Allison J, Karp W, Le NA. Visceral adipose tissue and cardiovascular risk factors in obese children. J Pediatr. 1998;133(1):41–5.
33. International basic safety standards for protection against ionizing radiation and for the safety of radiation sources. Vienna: IAEA; 1996.
34. International Atomic Energy Agency. Radiological Protection for Medical Exposure to Ionizing Radiation, IAEA safety standards series. Vienna: IAEA; 2002.
35. Applying radiation safety standards in diagnostic radiology and interventional procedures using X rays, Safety report series, vol. 39. Vienna: IAEA; 2006.
36. Blake GM, Naeem M, Boutros M. Comparison of effective dose to children and adults from dual X-ray absorptiometry examinations. Bone. 2006;38(6):935–42.
37. Thomas SR, Kalkwarf HJ, Buckley DD, Heubi JE. Effective dose of dual-energy X-ray absorptiometry scans in children as a function of age. J Clin Densitom. 2005;8(4):415–22.
38. Steel SA, Baker AJ, Saunderson JR. An assessment of the radiation dose to patients and staff from a Lunar Expert-XL fan beam densitometer. Physiol Meas. 1998;19(1):17–26.
39. Njeh CF, Samat SB, Nightingale A, McNeil EA, Boivin CM. Radiation dose and in vitro precision in paediatric bone mineral density measurement using dual X-ray absorptiometry. Br J Radiol. 1997;70(835):719–27.
40. Bezakova E, Collins PJ, Beddoe AH. Absorbed dose measurements in dual energy X-ray absorptiometry (DXA). Br J Radiol. 1997;70:172–9.
41. Njeh CF, Apple K, Temperton DH, Boivin CM. Radiological assessment of a new bone densitometer – the Lunar EXPERT. Br J Radiol. 1996;69(820):335–40.
42. Seidell JC, Bakker CJ, van der Kooy K. Imaging techniques for measuring adipose-tissue distribution – a comparison between computed tomography and 1.5-T magnetic resonance. Am J Clin Nutr. 1990;51(6):953–7.
43. Sobol W, Rossner S, Hinson B, Hiltbrandt E, Karstaedt N, Santago P, Wolfman N, Hagaman A, Crouse 3rd JR. Evaluation of a new magnetic resonance imaging method for quantitating adipose tissue areas. Int J Obes. 1991;15(9):589–99.
44. Greenfield JR, Samaras K, Chisholm DJ, Campbell LV. Regional intra-subject variability in abdominal adiposity limits usefulness of computed tomography. Obes Res. 2002;10(4):260–5.
45. Asayama K, Dobashi K, Hayashibe H, Kodera K, Uchida N, Nakane T, Araki T, Nakazawa S. Threshold values of visceral fat measures and their anthropometric alternatives for metabolic derangement in Japanese obese boys. Int J Obes Relat Metab Disord. 2002;26(2):208–13.
46. Potretzke AM, Schmitz KH, Jensen MD. Preventing overestimation of pixels in computed tomography assessment of visceral fat. Obes Res. 2004;12(10):1698–701.
47. UNSCEAR focuses on Chernobyl accident in general assembly report. In: United Nations Scientific Committee on the Effects of Atomic Radiation (UNSCEAR). Vienna. 2000. Available from: www.unscear.org/unscear/publications/2000_1.html.
48. Loy J. Exposure of humans to ionizing radiation for research purposes, Radiation protection series. Publication no. 8. Barton: Australian Radiation Protection and Nuclear Safety Agency (ARPANSA); 2005.

Chapter 13
Optical Coherence Tomography as a Biomarker in Clinical Trials for Ophthalmology and Neurology

Robert C. Sergott

Abstract In the past, clinical trials in neuro-ophthalmology have focused upon functional endpoints including Snellen visual acuity and visual field testing, both of which may have significant test/retest variability because of their dependence upon subject alertness and cognitive function. The introduction of optical coherence tomography (OCT) into the practice of clinical ophthalmology has now extended into clinical trials in both ophthalmology and neurology. Recent advances in the resolution and reproducibility of spectral-domain OCT have now produced algorithms that measure the structural integrity of the optic nerve and retina. OCT is in vivo histopathology and the qualitative and quantitative metrics of this painless, noninvasive, non-contact technology can now be correlated with traditional functional measurements to provide a complete longitudinal profile of the afferent visual system for clinical trials.

Keywords Optical coherence tomography • Retina • Optic nerve • Multiple sclerosis • Neurodegeneration • Vitreo-macular traction

Introduction

Since its first commercial application in 1993, optical coherence tomography (OCT) has become a standard part of the practice of clinical ophthalmology. In addition, since the optic nerve and retina are integral components of the central nervous system, OCT naturally migrated into both the clinical and regulatory domains of primary neurological disease. OCT now permits safe, accurate, and precise evaluations of

R.C. Sergott, MD
Department of Neuro-Ophthalmology, Wills Eye Hospital,
840 Walnut Street, Philadelphia, PA 19107, USA
e-mail: rcs220@comcast.net

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_13,

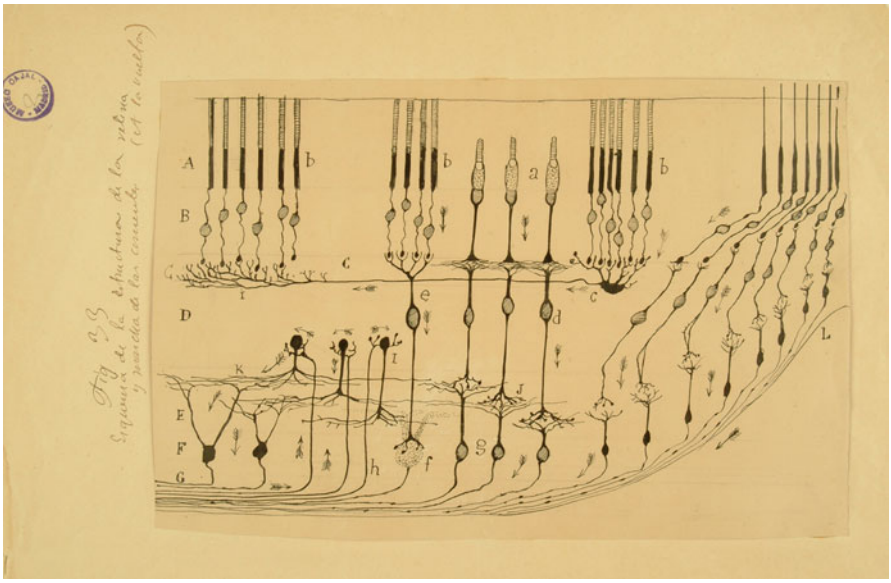


Fig. 13.1 One of the many illustrations of the human retina done by SR Cajal around 1880 that earned him the Nobel Prize in Medicine. While some details are incorrect, he did recognize the vertical orientation of the layers and their synaptic alignment (Courtesy of Instituto Cajal (CSIC) Madrid, Spain. Cajal Legacy)

the afferent visual system for both clinical care and clinical trials in ophthalmology and neurology. Because of the discovery that structural changes in the retinal nerve fiber layer and macula precede clinical changes in visual acuity and visual field in glaucoma and multiple sclerosis, OCT has rapidly evolved into a pivotal biomarker for many ophthalmic and neurologic diseases united by clinical and subclinical involvement of the optic nerve and retina. The surrogate biomarker may indeed be more reflective of the consequences of disease and its activity compared to traditional measurements of visual acuity and visual fields. Quite simply, it is both a qualitative and quantitative ophthalmoscope.

Paralleling the pathways of computerized tomographic (CT) scanning and magnetic resonance imaging (MRI), OCT first entered the clinical arena immediately adding new information for the clinical care of patients. Clinical ophthalmologists for the first time could visualize directly the layers of the retina designed for transforming light energy into remarkable achievement of sight. OCT represented the clinical reality of the beautiful and graceful histological drawings of Ramon y Cajal, for which he was awarded the Nobel Prize in Medicine, whose illustration of the retina is represented in Fig. 13.1.

As with CT and MRI, OCT unveiled a previously concealed, anatomical region protected from the in vivo observations. As the reproducibility and reliability of OCT improved, clinical trials were a natural added dimension to its portfolio (Fig. 13.2a–c). Biomedical engineering achievements have transformed OCT from

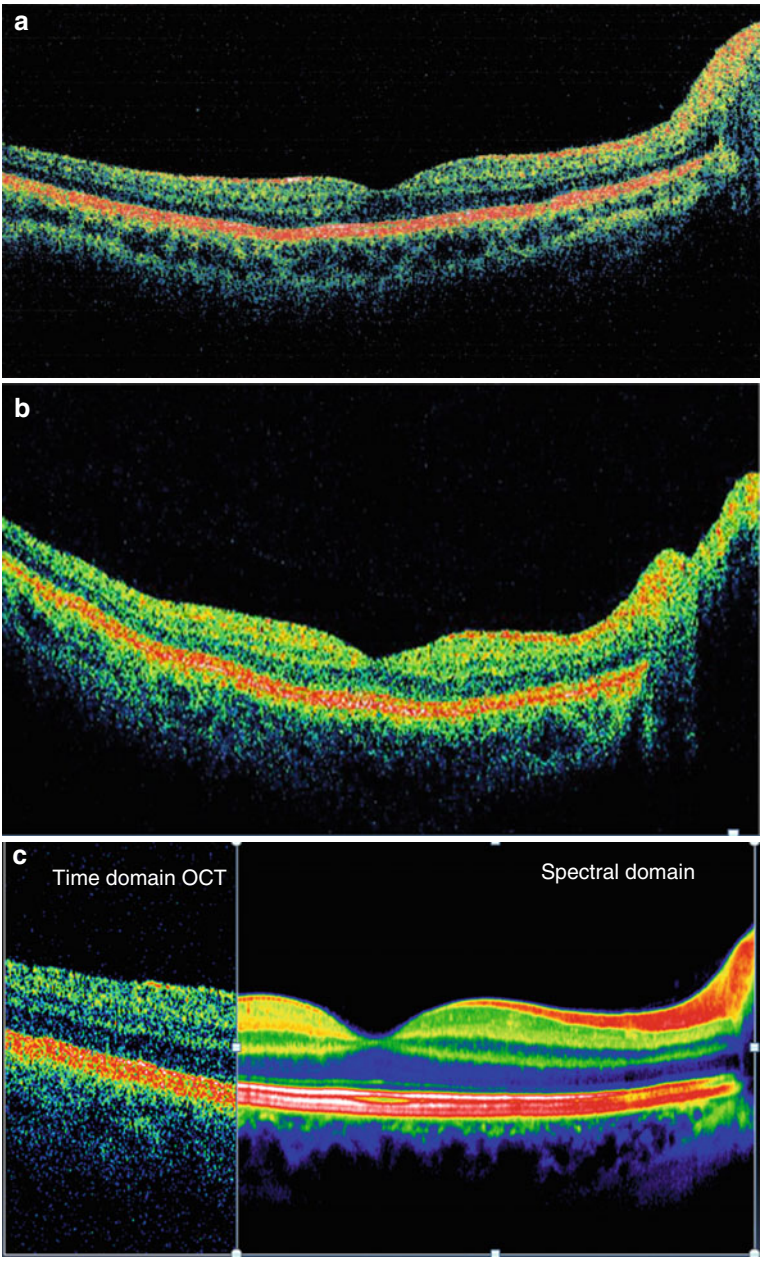


Fig. 13.2 (**a**) Time-domain OCT. (**b**) Spectral-domain scan. (**c**) Time domain versus spectral domain (**c** Courtesy of Heidelberg Engineering)

time-domain (TD) imaging, adept at capturing abnormalities of the vitreo-retinal interface, to spectral-domain (SD) technology that has defined in vivo imaging of the structure of the photoreceptors of the outer retina (Fig. 13.2a–c).

TD-OCT permitted only 2-dimensional imaging; however, SD-OCT, also referred to as Fourier-domain OCT, greatly improved data-acquisition speeds facilitating 3-dimensional imaging.

In addition, OCT is especially enticing for both clinical use and regulatory trials since the testing paradigm encompasses two perfect medical triads: one for the patient and one for clinical trial specialists and regulators. The patient triad includes its painless, non-contact, noninvasive nature. Light is the only imaging modality required and oral or intravenous contrast agents are not necessary. No complications of OCT have ever been reported. Clinical trial professionals profit in both the protocol and regulatory arenas by the remarkable correlation of qualitative findings with the in vivo histology of the retina and optic nerve as well as the highly refined reproducibility and reliability of the quantitative measurements.

OCT is not a replacement for fundus photography but rather incorporates an added dimension of visualizing living optic nerve and retina again and again without damage to the anatomy or physiology of these structures. Previously sequestered secrets and signatures of both ophthalmic and neurologic diseases are now lucid and transparent.

Therefore, in this volume dedicated to clinical trials, we applaud a new technology whose birth came through ophthalmology but whose offspring will foster major advances in every medical discipline in which we have access to in vivo tissue through endoscopes, catheters, or intraoperative probes. The initial lessons learned in ophthalmology have already found applications in neurology which will share the dais with ophthalmic diseases in this chapter. We, however, confidently predict that the next edition will embrace OCT as a clinical trial biomarker in gastroenterology, neurosurgery, cardiology, and pulmonary medicine. Through the collaborative efforts of basic researchers, innovative clinicians (not "providers"), and medical device makers, OCT has improved the care and quality of life of patients not only with potentially blinding ophthalmic diseases but also with potentially disabling and life-threatening conditions in the central nervous system.

Basic Technology

OCT utilizes the principles of interferometry and coherent light to produce images of both diagnostic and regulatory quality. However, the realm of OCT did not start with medical imaging elegance. The first commercial use of OCT was not in medicine but in the world of art history and old master painters to determine if multiple paintings were layered on a single canvas. Due to the expense of proper canvases, many Old World artists were inclined to paint over previous paintings, entombing the original work. On many occasions, the artwork concealed beneath the surface of the visible image was much more valuable and more important than the surface

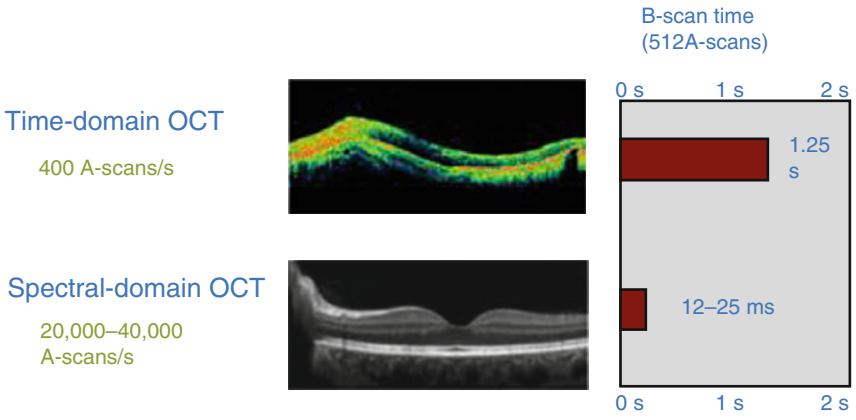


Fig. 13.3 Spectral-domain OCT greatly increased the number of A scans performed per second. That advance as well as improved lasers and minimizing movement artifacts increased both the imaging quality and coefficients of variability of OCT qualitative and quantitative analysis (Courtesy of Heidelberg Engineering)

facade. In a similar way, what lays beneath the surface of the retina and optic nerve are often more important than what is seen with the ophthalmoscope.

Before we explain how OCT has achieved its current sophistication, we will review the fundamental technology. When a wavelength of light encounters the boundary of two different media, the light may be reflected, refracted, and/or absorbed. The first step for OCT is the acquisition of the reflected light by a capture device.

Next, the algorithm of interferometry, an electromagnetic principle, is used. "Interference" can be either constructive or destructive for imaging purposes depending upon the relative phases of the waves. OCT employs the interaction of waves coherent with each other to form images by sectioning or "tomography," a physical application used in CT, MRI, and ultrasonography.

One of the limiting technological issues for the first OCT instruments was the number of A scans that could be performed in a finite time, thereby limiting image resolution (Fig. 13.3). Improvement in this engineering domain has increased the number of A scans and shortened the capture time for this biomechanical variable.

Another similarity with MRI, specifically orbital imaging, demanded correction for normal, continuous, physiological movements of the eyes in the resting stage, termed microsaccades. Until MRI software conquered this obstacle, orbital MRI scans were blurred and substandard due to movement artifacts. By using a variety of techniques to insure image stabilization and registration, SD-OCT devices have reduced this source of artifact (Fig. 13.4) and hence further improved image quality and decreased the coefficient of variability for measuring the dimensions of the retina.

Because OCT requires light to be precisely focused upon the retina, imaging is often limited by the dimensions of the eye. Somewhat smaller than normal eyes, hyperopia (far sightedness) greater than 6 diopters, and somewhat larger eyes, myopia (near sightedness) greater than 6 diopters, may induce errors in image resolution

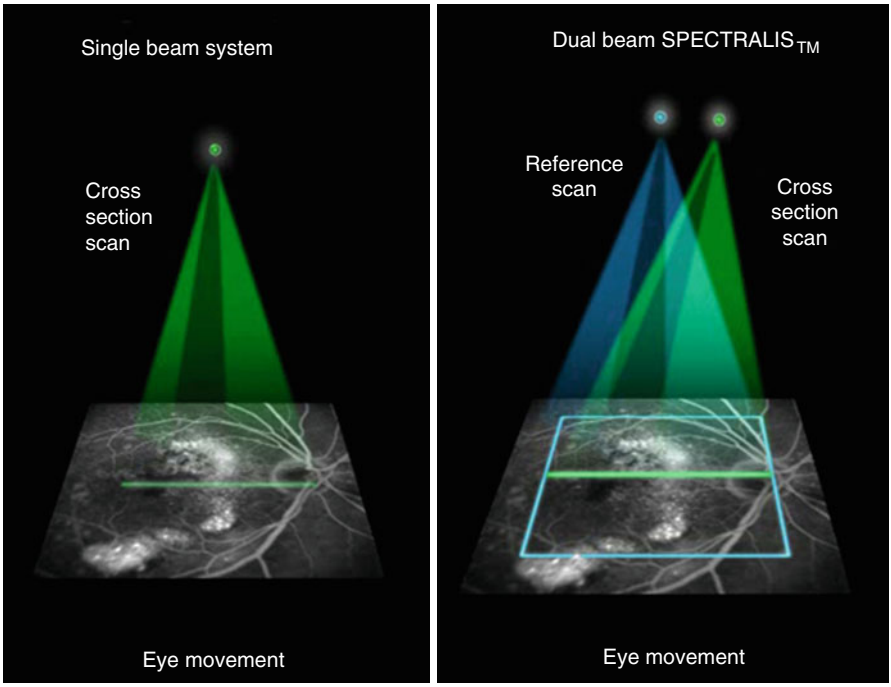


Fig. 13.4 With the single beam system used in TD-OCT and some SD-OCT devices (*left picture*), the eye location is not known unless a second reference scan is used (Courtesy of Heidelberg Engineering), as shown on the *right picture*

and the measurement algorithms of the peripapillary retinal nerve fiber layer thickness and the macular thickness.

What Is Normal?

When TD-OCT was introduced into the commercial market, the technology was patent protected. As SD-OCT emerged, patent protection was no longer possible because Fourier analysis was required to improve image quality. Fourier analysis is considered to be common scientific knowledge, not subject to patent enforcement. SD-OCT facilitated competition among device manufacturers in the commercial market, which has produced improved technology and helped to control price. The hardware and software platforms from various manufacturers differ fundamentally in methodology of image processing as well as the definition of the posterior border of the retina. Does the retina end at the inner border of the retinal pigment epithelium (RPE), the outer border of the RPE, or Bruch's membrane? The answer to that question depends upon which device a clinician or clinical trial selects (Tables 13.1 and 13.2). The most important conclusion from the different definitions of the

Table 13.1 Summary of specifications of the five commercial OCT instruments

OCT instrument	Axial resolution (μm)	A-scan speed (scans/s)	Macular thickness outer boundary[a]	Manufacturer	Software version, software protocol
Stratus	8–10	400–600	IS-OS junction	Carl Zeiss Meditec, Inc., Dublin, CA	v4.0; macular thickness map protocol
3D OCT-1000	5–6	25,000	Inner RPE	Topcon, Inc., Paramus, NJ	v2.12; 3D macular protocol
Optovue (RTVue-100)	5–6	26,000	Outer RPE	Optovue, Inc., Fremont, CA	v3.5; (E)MM5 and MM6
Cirrus	5	27,000	Outer RPE	Carl Zeiss Meditec, Inc., Dublin, CA	v3.0; macular cube 512×128
Spectralis	4–6	40,000	Bruch's membrane	Heidelberg Engineering, Inc., Heidelberg, Germany	v3.2; macular volume protocol

Used with permission of Elsevier from Kiernan DF, Mieler WF, Hariprasad SM. Spectral-domain optical coherence tomography: a comparison of modern high-resolution retinal imaging systems. Am J Ophthalmol. 2010;149:18–31

OCT optical coherence tomography, *IS* inner segment photoreceptor, *OS* outer segment photoreceptor, *RPE* retinal pigment epithelium

[a]Macular thickness inner segment boundary is inner limiting membrane across all instruments

Table 13.2 Average central macular thickness (microns) ± standard deviation across five OCT devices in normal and pathologic eyes

OCT device	Normal macula	Pathologic macula
Stratus	185.3 ± 16.7	256.2 ± 166.4
Cirrus	249.1 ± 26.5	295.2 ± 156.2
Optovue	248.3 ± 16.0	326.1 ± 150.6
Topcon	226.2 ± 18.7	278.1 ± 157.7
Spectralis	267.8 ± 17.1	336.2 ± 154.1

Used with permission of Elsevier from Kiernan DF, Mieler WF, Hariprasad SM. Spectral-domain optical coherence tomography: a comparison of modern high-resolution retinal imaging systems. Am J Ophthalmol. 2010;149:18–31

Stratus and Cirrus are manufactured by Carl Zeiss Meditec, Inc., Dublin, CA, Optovue is manufactured by Optovue, Inc., Fremont, CA, Topcon is manufactured by Topcon, Inc., Paramus, NJ, and Spectralis is manufactured by Heidelberg Engineering, Inc., Heidelberg, Germany

OCT optical coherence tomography

borders of the retina becomes the inability to transfer and merge qualitative data from one OCT platform to another for determining baseline measurements as well as changes over time during a clinical trial. Measuring percentage change over time may be a way to approach this conundrum but that is somewhat flawed since the

basic dimensions are different, and this approach has not been accepted by regulatory authorities. Tables 13.1 and 13.2 further illustrate this problem, which becomes very significant for international trials where use of a single device is rarely possible unless supplied by the sponsor. Ultimately, regulators may declare which platforms and measurements algorithms are acceptable for specific trials.

The "So What" Question

Does Structure Predict Functional Impairment for the Macula?

Just as questions have been raised by regulators concerning MRI changes in remitting relapses multiple sclerosis (RRMS), so OCT shares the same issue. To paraphrase the discomfort: "So we see new findings compared to baseline with both MRI and OCT. What does this mean to the patient's clinical status?" Briefly stated, "So what?"

In the RRMS clinical trial world, this issue forced MRI into the role of a secondary outcome measure following annualized relapse rate as the primary end point. The status of MRI has gradually escalated as clinicians reported that this technology was approximately seven times more sensitive than reports of patients and clinical examinations to detect disease activity. In addition, cerebral atrophy as detected by MRI appears to be the next MRI parameter on the horizon to gain regulatory blessing in RRMS. Having more brain instead of less brain does seem to be intuitively a more preferred neuroanatomic circumstance.

Regulators initially took a deservedly, subdued, unenthusiastic scrutiny of OCT in ophthalmology, due to the poor reproducibility of time-domain measurements and the imaging status of the outer retina which was based primarily on imagination. However, as we shall see in the next section, spectral domain has forced a reassessment of this position.

The Thrombogenics Story

The perfect answer to the "so-what" question was the situation with vitreo-macular adhesion and traction (VMA, VMT) and macular holes and the clinical trial program with ocriplasmin (Jetrea) from thrombogenics. This compound was approved by the FDA for treatment of symptomatic VMA and VMT and full thickness macular holes, October 18, 2012. A recent opinion from European regulatory authorities favors approval on that continent.

In VMA, VMT, and macular holes, the community of vitreo-retinal specialists had described the pathophysiology and natural history of these syndromes in precise detail from the original reports of the late Donald Gass through innumerable more recent publications.

While spontaneous resolution was possible in these conditions, the natural history was more frequently that of progressive traction with disruption of the

underlying photoreceptors. Vitrectomy surgery enjoyed a 90 % chance of successful repair but the postoperative care required a prolonged period of face down positioning and a 100 % chance of cataract development, necessitating a second surgical procedure. Surgery was often delayed until visual acuity was significantly reduced as clinicians and patients opted to avoid the risks of a delicate intraocular surgical procedure and the subsequent postoperative issues.

A group of very knowledgeable regulators realized that VMA and VMT were usually ophthalmic "time bombs" that given enough time could produce secondary retinal changes that could prove to be irreversible. Therefore, the two pivotal phase 3 trials for ocriplasmin were the first time that OCT was permitted to be a primary outcome biomarker. The regulators' judgment proved to be astute and both studies meet their primary end points, resulting in product approval and the promise that at least hundreds, and possibly thousands, of patients will avoid vitrectomy and cataract surgeries (Fig. 13.5a–c). Ongoing studies now with ocriplasmin include diabetic retinopathy and macular degeneration in which there may be an important VMA/VMT component in some patients.

We envision similar regulatory strategies and decisions in almost all retinal diseases where the thickness of the fovea and macular are associated with decreased central visual acuity. More precise overall thickness measurements coupled with measurement of the individual layers of the retina, so-called segmentation analysis, could prove to be a more accurate biomarker for predicting clinical improvement than high- or low-contrast visual acuity testing (Fig. 13.6). Also, the latest generation of SD-OCT has transformed this test in a completely objective compared to visual acuity, which can be subjective to a degree. Therefore, measurement of the thickness of these layers for diabetic retinopathy, macular degeneration, uveitis, and any entity involving the macula promises to create even more precise correlation with structure and function. Numerous trials have used OCT for these indications already; however, the spectral-domain technology adds more credibility to the results while simultaneously uncovering new findings.

The Neuro-ophthalmology and Neurology Chronicles

The legion of naysayers in neuro-ophthalmology and neurology far outnumbered the early advocates of OCT in these subspecialties after this technology reached the medical marketplace. The monotonous chorus of accusations of unnecessary testing, a test looking for an indication, and we are expert clinicians and we do not need this overpriced ophthalmoscope reverberated throughout the United States, Europe, and Asia.

However, the advances in the basic science of biomedical engineering and the fundamentals of neuropathology prevailed and confirmed the importance of OCT, especially when combined with sensitive electrophysiological assessments such as multifocal electroretinography and multifocal visual evoked potentials.

The engineers refined OCT to a point where the image quality and resolution defy any prior predictions and the future promises another paradigm shift. In contrast, the

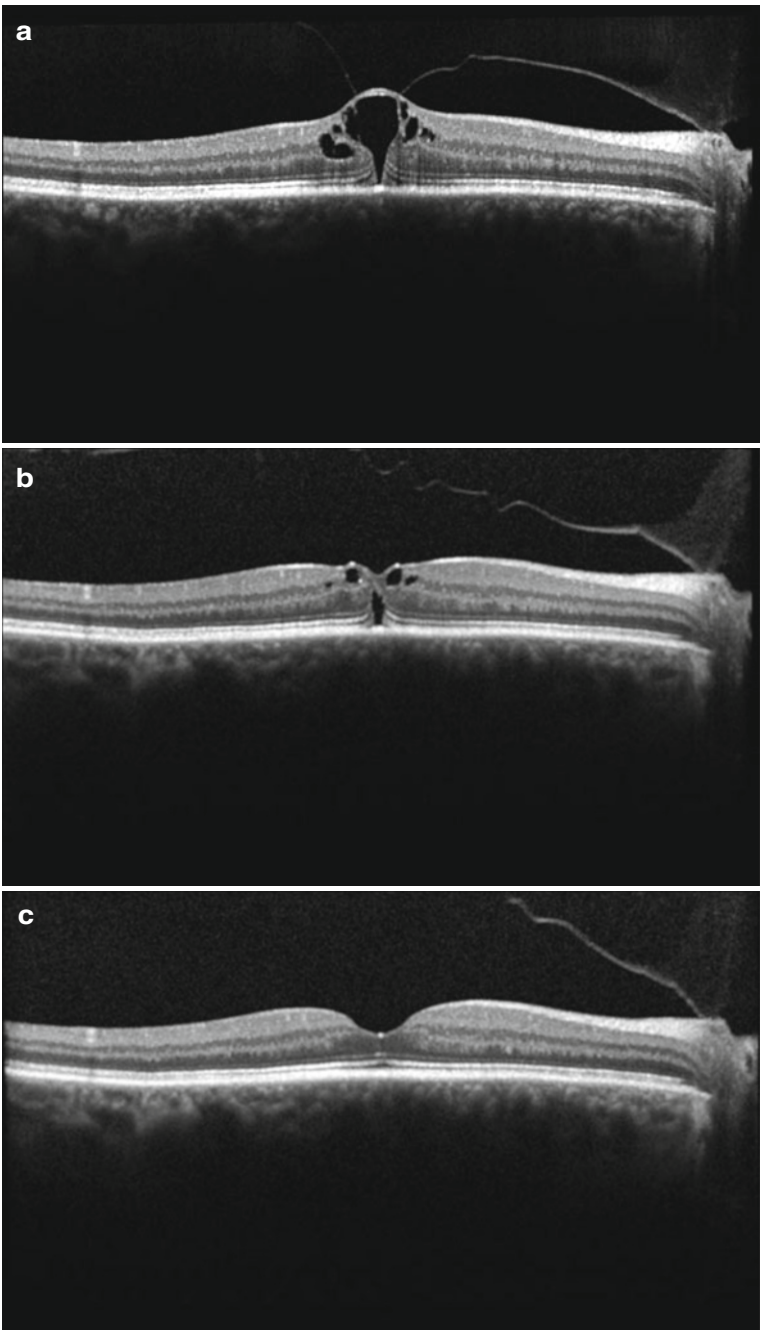


Fig. 13.5 (**a**) Preinjection longitudinal OCT of vitreo-macular adhesion and traction with underlying foveal cysts and underlying small area of RPE devoid of photoreceptors. (**b**) VMT has been released 28 days after injection of ocriplasmin (Jetrea). The fovea-macular contour is beginning to reform as the intraretinal cysts are decreasing in size. (**c**) 180 days following ocriplasmin injection; the foveal macular contour has completely reformed. The external limiting membrane, inner and outer segments, and outer nuclear layer of the photoreceptors have also returned to a normal appearance (All images courtesy of The Optic Nerve Research Center)

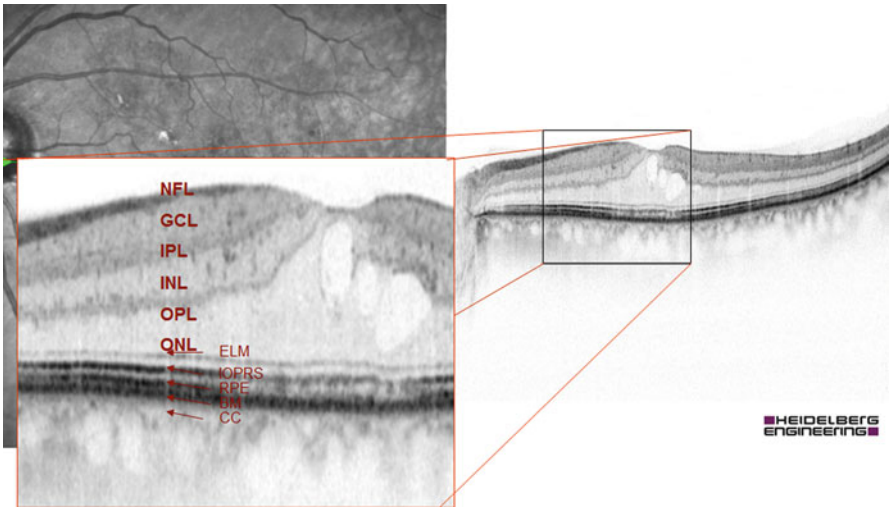


Fig. 13.6 High-resolution SD-OCT image of foveal and macular edema with segmentation of the layers of the retina as well as identification of fine structure at the junction of the photoreceptors and the RPE. *RNFL* retinal nerve fiber layer, *GCL* ganglion cell layer, *IPL* inner plexiform layer, *INL* inner nuclear layer, *OPL* outer plexiform layer, *ONL* outer nuclear layer, *ELM* external limiting membrane, *IOPRS* inner/outer segments of the photoreceptors, *RPE* retinal pigment epithelium, *CC* choriocapillaris (Courtesy of Heidelberg Engineering)

neuropathology side required rediscovery, reeducation, and new data to win the day. Multiple sclerosis (MS) and the classic neuro-ophthalmic disease of optic neuritis provided unequivocal proof that OCT permitted the evaluation of axonal loss in these conditions, a territory below the resolution of conventional MRI. First, however, the MS community had to be reeducated to the entire issue of axonal loss in this protean disease. Despite the recognition of axonal loss as part of the neuropathological spectrum by Charcot in his original description, this signature of the disease completely disappeared from a generation of clinical care and research as experimental allergic encephalomyelitis (EAE) ascended the position as an animal model of MS. While a near-perfect example of how the mammalian central nervous system responds to an immune challenge with myelin and its components, EAE falls short as an animal surrogate for MS, most specifically in its degree of axonal loss.

The almost simultaneous neurochemical studies of German investigators and elegant immunohistochemistry studies of Bruce Trapp coalesced to resurrect axonal loss as a key finding in MS and quite possibly the major determinant of physical disability. Unlike other neurodegenerative diseases, RRMS was often heralded by optic neuritis, a disorder of the afferent visual system lending itself perfectly to investigation with OCT and correlation with both clinical outcomes of visual function and MRI findings. Since that time, numerous well-done studies have agreed that following an event of optic neuritis that the average peripapillary retinal nerve fiber layer (RNFL) suffers a 20 % decline in thickness despite the paradoxical and deceptive recovery of almost complete visual function in approximately 85 % of patients. Therefore, the natural history of clinical improvement misled clinicians to conclude that events of optic neuritis, and by extrapolation attacks of MS elsewhere

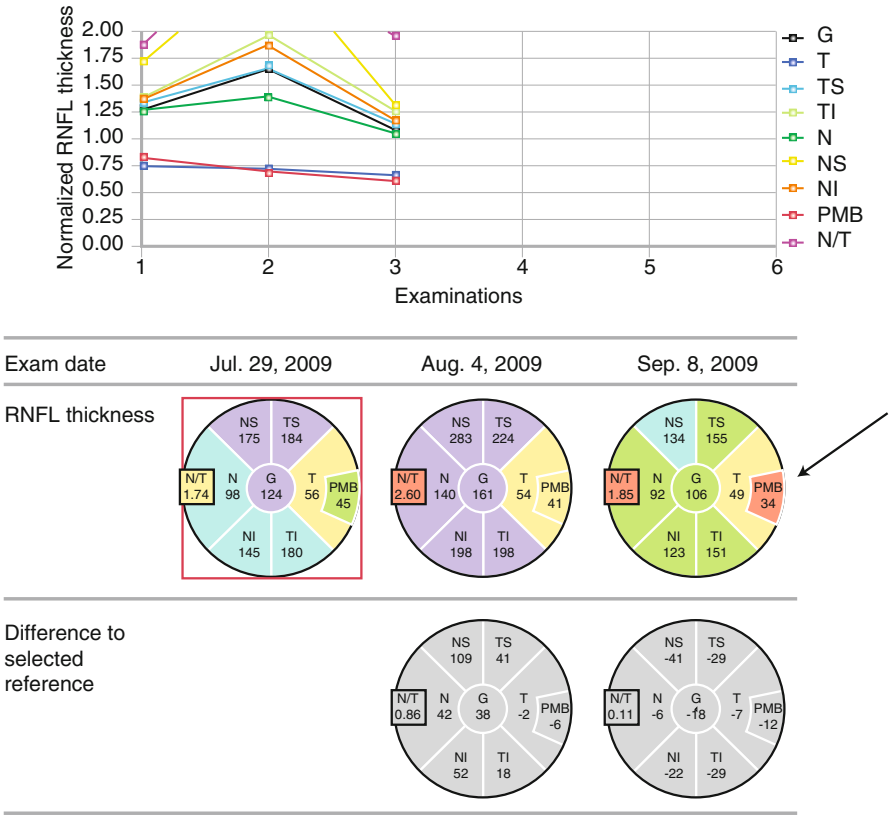


Fig. 13.7 The progression of retinal nerve fiber layer (unmyelinated central nervous system axons) from onset over a 5-week time frame. Notice how the papillomacular bundle arrow is preferentially affected in such a short period of time (*arrow*)

in the central nervous system, were benign. No conclusion could be further removed from the truth by realizing that the average aging change of the RNFL is 0.2 μm/year and that a loss of 20 % of the RNFL equals about 20 μm and is equivalent to 100 years of aging of the RNFL (Fig. 13.7).

More recently, attention has turned to the involvement of the retina in MS, a subject previously noteworthy for a paucity of clinical and pathological reports. Since the retina does represent a microcosm of both white and gray matter, clinical research with OCT and all forms of MS has yet another "hidden continent" to explore.

Neurodegeneration

All forms of MS may be categorized as neurodegenerative syndromes particularly focusing upon the somewhat divergent clinical and MRI findings of RRMS contrasted with secondary progressive MS and primary progressive MS. The question then must

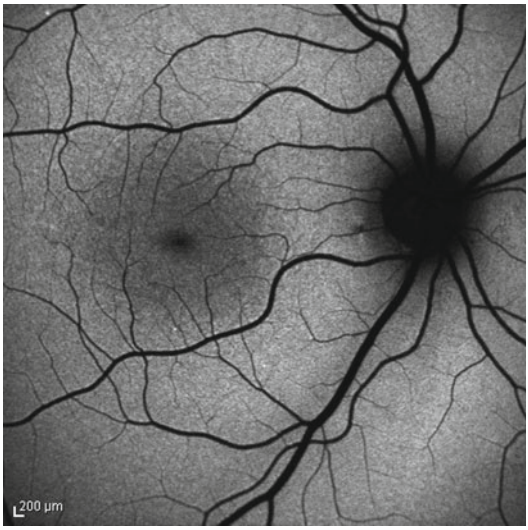


Fig. 13.8 Normal FAF image of the posterior pole. The hyperfluorescent signal emanates from lipofuscin. The causes of reduced and decreased signals are outlined in Table 13.3 (Courtesy of Heidelberg Engineering)

be asked whether or not OCT could become a validated biomarker for clinical trials in other neurodegenerative syndromes such as Parkinson's disease, Alzheimer's disease, traumatic brain injury (TB), frontotemporal dementia, and progressive supranuclear palsy. The answer from the current peer-reviewed literature is "Yes."

Given the luxurious amounts of dopamine present in the human retina, no sense of surprise or astonishment should occur that several studies have demonstrated thinning of the macula in these areas even though these patients rarely combine problems with their visual acuity or visual fields. The onset and tempo of these changes has yet to be described.

In a similar way, we have seen changes in Alzheimer's, TBI, and other dementias. No obvious pattern of abnormalities has yet evolved. Another imaging modality of the afferent visual system, fundus autofluorescence (FAF), which detects and measures the metabolism of lipofuscin in the retinal pigment epithelium, may prove to be quite useful (Fig. 13.8 and Table 13.3). Lipofuscin shares biochemical similarities with beta amyloid, a pathological finding in Alzheimer's. Focal zones of hypo-fluorescence have been seen in cases of Alzheimer's disease as well as profound retinal thinning by OCT (Sergott et al., manuscript in preparation 2014).

Because of its ability to define a topographic map of the retina compared to an isolated tomographic section, FAF has become an integral part of clinical trials for macular degeneration.

Summary

In the relatively brief span of 20 years in the history of medicine which usually measures time in terms of glacial equivalents, OCT has matured from a interesting curiosity able to detect only the most obvious abnormalities between the interface

Table 13.3 Causes of reduced and increased FAF signals

Causes for a reduced FAF signal	Causes for an increased FAF signal
Absence or reduction in RPE lipofuscin density	Excessive RPE lipofuscin accumulation
RPE loss or atrophy (e.g., geographic atrophy)	Lipofuscinopathies including Strargardt disease, Best disease, and adult vitelliform macular dystrophy
Absorption from extracellular material, cell, or intraretinal fluid (e.g., macular oedema)	Age-related macular degeneration (e.g., RPE in the junctional zone preceding enlargement of occurrence of geographic atrophy)
Fresh intraretinal and subrentinal haemorrhages	Subrentinal fluid leading to separation of the outer segments of the photoreceptors from the underlying RPE, which leads to improper outer segment turnover
Fibrosis, scar tissue, or borders of laser scars	Drusen in the sub-pigment epithelial space
Retinal vessels	Older intraretinal and subrential haemorrhages
Luteal pigment (lutein and zeaxanthin)	Lack of absorbing material
	Optic nerve head drusen

Reproduced with kind permission of Springer Science + Business Media from Holz FG, Schmitz-Valckenberg S, Spaide RF, Bird AC, editors. Atlas of fundus autofluorescence imaging. Berlin/ Heidelberg: Springer Science + Business Media; 2007

of the retina and the vitreous to both a qualitative and quantitative metric for clinical trials in both ophthalmology and neurology.

Naturally, the ophthalmology adoption of this technology for diabetic retinopathy, macular degeneration, vitreo-macular traction, and uveitis has been more rapid than for primary neuro-ophthalmic and neurologic diseases. MS has witnessed the most studies with OCT in neurology but considerable promise has developed for its use in all neurodegenerative syndromes.

Its use and validation for both clinical care and clinical trials testifies to a highly productive collaboration among basic researchers, clinicians, clinical trial experts, and regulatory authorities. While this chapter has ended, the saga of OCT is only beginning.

Acknowledgement Dr. Sergott is a paid consultant for Heidelberg Engineering and for Thrombogenics.

General Bibliography

Chen TC, Cense B, Miller JW, et al. Histologic correlation of in vivo optical coherence tomography images of the human retina. Am J Ophthalmol. 2006;141:1165–8.

Costello F, Coupland S, Hodge W, et al. Quantifying axonal loss after optic neuritis with optical coherence tomography. Ann Neurol. 2006;59:963–9.

Ferguson B, Matyszak MK, Esiri MM, Perry VH. Axonal damage in acute multiple sclerosis lesions. Brain. 1997;120(Pt 3):393–9.

Fisher JB, Jacobs DA, Markowitz CE, et al. Relation of visual function to retinal nerve fiber layer thickness in multiple sclerosis. Ophthalmology. 2006;113:324–32.

Frohman EM, Costello F, Stuve O, Calabresi P, Miller DH, Hickman SJ, Sergott RC, Conger A, Salter A, Krumwiede KH, Frohman TC, Balcer L, Zivadinov R. Modeling axonal degeneration within the anterior visual system. Implications for demonstrating neuroprotection in multiple sclerosis. Arch Neurol. 2008;65:26–35.

Kerrison JB, Flynn T, Green WR. Retinal pathologic changes in multiple sclerosis. Retina. 1994;14:445–51.

Kiernan DF, Mieler WF, Hariprasad SM. Spectral-domain optical coherence tomography: a comparison of modern high-resolution retinal imaging systems. Am J Ophthalmol. 2010;149:18–31.

Lalezary M, Medeiros FA, Weinreb RN, et al. Baseline optical coherence tomography predicts the development of glaucomatous change in glaucoma suspects. Am J Ophthalmol. 2006;142:576–82.

Naismith RT, Tutlam NT, Xu J, Klawiter EC, Shepherd K, Trinkaus SK, Song HK, Cross AH. Optical coherence tomography differs in neuromyelitis optica compared with multiple sclerosis. Neurology. 2009;72:1077–82.

Parisi V, Manni G, Spadaro M, et al. Correlation between morphological and functional retinal impairment in multiple sclerosis patients. Invest Ophthalmol Vis Sci. 1999;40:2520–7.

Pulicken M, Gordon-Lipkin E, Balcer LJ, et al. Optical coherence tomography and disease subtype in multiple sclerosis. Neurology. 2007;69:2085–92.

Sergott RC, Frohman E, Glanzman R, Al-Sabbagh A. The role of optical coherence tomography in multiple sclerosis: expert panel consensus. J Neurol Sci. 2007;263(1):3–14.

Spaide RF, Curio CA. Anatomical correlates to the bands seen in the outer retina by optical coherence tomography. Literature review and model. Retina. 2011;31:1609–19.

Spain RI, Maltenfort M, Sergott RC, Leist TP. Disease duration-dependent retinal nerve fiber layer loss in untreated multiple sclerosis: the natural history of retinal axonal degeneration in multiple sclerosis. J Rehabil Res Dev. 2009;46(5):633–42.

Trapp BD, Peterson J, Ransohoff RM, et al. Axonal transection in the lesions of multiple sclerosis. N Engl J Med. 1998;338:278–85.

Trip SA, Schlottmann PG, Jones SJ, et al. Retinal nerve fiber layer axonal loss and visual dysfunction in optic neuritis. Ann Neurol. 2005;58:383–91.

Trip SA, Schlottmann PG, Jones SJ, et al. Optic nerve atrophy and retinal nerve fibre layer thinning following optic neuritis: evidence that axonal loss is a substrate of MRI-detected atrophy. Neuroimage. 2006;31:286–93.

Chapter 14
Nuclear Medicine: An Overview of Imaging Techniques, Clinical Applications and Trials

Amit Mehndiratta, Prabu Anandaraj, Christian M. Zechmann, and Frederik L. Giesel

Abstract Nuclear medicine is a special division of nuclear physics that deals with the application of radioactivity in diagnostic and therapeutic medicine. This chapter will elaborate the basics of nuclear physics, concepts of nuclear imaging radioactive tracers used in nuclear imaging, and their mechanism. Radiation safety is a major concern while administering ionizing radiation for diagnostic or therapeutic purpose. An elaboration on the radiation safety measure for both patient and doctors followed by standardization of nuclear imaging in clinical practice is covered in this chapter. Following this multiple example will be discussed in cancer imaging of brain, lung, breast, GIT, ovary, prostate, etc. This chapter will conclude with future scope for research and outlook for clinical application in nuclear imaging.

Keywords Nuclear medicine • PET • PET/CT • Tumor imaging • Radiotracers

A. Mehndiratta, MBBS, DPhil (Oxon) (✉)
Department of Engineering Science, Institute of Biomedical Engineering,
Old Road Campus Research Building, University of Oxford,
Hedington, Oxfordshire OX3 7DQ, UK
e-mail: dramit.mehndiratta@gmail.com

P. Anandaraj, MBBS, MMST
Advanced Clinical Applications,
GE Healthcare, Chennai, India
e-mail: prarajya@gmail.com

C.M. Zechmann, MD
Department of Diagnostic Radiology and Nuclear Medicine,
Rinecker Proton Therapy Center, Schaeftlarnstr. 133, Munich 81381, Germany
e-mail: christian.zechmann@rptc-1.de

F.L. Giesel, MD, MBA
Department of Nuclear Medicine, University Hospital Heidelberg,
69120 Heidelberg, Germany
e-mail: frederik.giesel@med.uni-heidelberg.de

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_14, © Springer-Verlag London 2014

Introduction

Nuclear medicine is a fascinating application of nuclear physics. There are about 2,450 known isotopes of the 100 odd elements in the periodic table, out of which only about 300 are natural. The unstable isotopes of an atom attempts to reach the stability by a fission process and by emitting particles and/or energy in the form of radiation. This process is called radioactivity/nuclear decay processes. The unstable isotopes are those having too many protons/neutrons in their atomic nucleus which makes their bond energy unstable and tend to lose energy to the atomic electron or emitted as packet of radiation energy like gamma rays. Such isotope are called radionuclide or radioactive isotope or simply as radioisotope. In this chapter, we will describe the basics of nuclear physics and the various aspects of its application in diagnostic and therapeutic medicine. All aspects will be considered in this chapter, including research and clinical trials.

Basics of Nuclear Physics

The nuclear decay process emits energy in various forms which can be alpha, beta, or gamma radiations. When an unstable atom's nucleus emits two protons and two neutrons in a packet the process is called alpha decay. A proton can release a particle in a process called beta-plus decay, and a neutron can emit a particle in a process called beta-minus decay. Also some energy may emit from the nucleus of unstable atom which results from a process called gamma decay as well as an electron being attracted into the nucleus and being ejected again. Finally there is the rather catastrophic process where the nucleus breaks in smaller units called spontaneous fission.

The final expression is known as the Radioactive Decay Law. It describes the number of radioactive nuclei that will decrease in an exponential fashion with time with the rate of decrease being controlled by the decay constant. Half-life of a radionucleotide expresses the length of time it takes for the radioactivity of a radioisotope to decrease by a factor of two (Table 14.1). Some of the radionucleotides have a relatively short half-life. These tend to be the ones used for medical diagnostic purposes because they do not remain radioactive for very long following administration to a patient and hence result in a relatively low radiation dose. But isotopes with a relatively longer half-life have been used in the past for therapeutic applications in medicine.

Table 14.1 Half-life of radioisotopes

Radioisotope	Half-life (approx.)
^{81m}Kr	13 s
^{99m}Tc	6 h
^{131}I	8 days
^{51}Cr	1 month
^{137}Cs	30 years
^{241}Am	462 years
^{226}Ra	1,620 years
^{238}U	4.51×10^9 years

Nuclear Imaging Technique

The images are obtained by mapping the distribution of an administered radiopharmaceutical within the body. The radiation is emitted from within the patient and subsequently detected in the imaging device, unlike transmitted through the patient from an external X-ray source (CTs and radiographs). The specific organ function depicted is determined by the biological behavior of the radiopharmaceutical. Conventional imaging with the use of a gamma camera is referred to as planar imaging. More recently, the single photon emission computed tomography (SPECT) has been developed which produces axial slice imaging through the body. SPECT uses a gamma camera to record images at a series of angles around the patient, and the resultant data can be processed using filtered back projection and iterative reconstruction algorithms. SPECT gamma cameras can have one, two, or three camera heads. The more advanced imaging is positron emission tomography (PET) that is also an axial projection acquisition-based technique. PET exploits the positron annihilation process where two 0.51 MeV back-to-back gamma rays are produced. If these gamma rays are detected, their origin will lie on a line joining two of the detectors of the ring of detectors which encircles the patient.

It took more than 40 years for the PET to reach its current state as a clinically useful tool. It is primarily due to the challenge in engineering the electronic components of medical imaging instrument to be merged into the contemporary imaging modality. The major breakthrough in positron imaging heralded with the development of positron camera in 1960 which produce planar images [1]. Later the same year, true transaxial positron tomography utilizing a ring system of detectors was produced by Brookhaven National Laboratory group.

With the advent of advanced reconstruction techniques accompanying CT, the PET scanning took a giant leap ahead. The prototype of modern day positron emission computed tomography was first implemented by Phelps and colleagues in the mid-1970s [2].

More recently, the limited anatomical definition of radionuclide imaging has been addressed to some degree by the development of hybrid imaging techniques in which radionuclide imaging devices are combined with computed tomography in a single imaging system [3]. The resulting images display the functional data obtained from the radionuclide distribution (in color), overlaid on the anatomical information from CT (in gray scale) [3]. As the two image data sets are acquired almost simultaneously using the same imaging device, the two data sets can be co-registered very accurately. Not only do these hybrid systems allow abnormalities seen on radionuclide images to be assigned to precise anatomical structures, but they also enable the morphological appearances of disease processes depicted by CT to be assimilated into the interpretation of the findings on radionuclide images. For example, such combined interpretation can aid the distinction of malignant and inflammatory causes of uptake of the positron-emitting radiopharmaceutical ^{18}F-fluorodeoxyglucose (FDG) [4]. Further advantages of hybrid systems include the use of the CT data to correct radionuclide images for artifacts resulting from attenuation of photons travelling through the body and the ability to incorporate radionuclide image data into CT-based radiotherapy planning systems.

Table 14.2 Specific organ where radiopharmaceutical agent or radiotracer accumulates for a short period of time

Body organ	Radiotracer
Brain	^{99m}Tc-ceretec
Thyroid	$Na^{99m}TcO_4$
Lung (ventilation)	^{133}Xe gas
Lung (perfusion)	^{99m}Tc-MAA
Liver	^{99m}Tc-tin colloid
Spleen	^{99m}Tc-damaged red blood cells
Pancreas	^{75}Se-selenomethionine
Kidneys	^{99m}Tc-DMSA

Radiotracer in Human Studies

The radioactivity is generally administered to the patient in the form of a radiopharmaceutical agent or radiotracer. This follows some physiological pathway to accumulate for a short period of time in some specific organ of the body (Table 14.2). A good example is ^{99m}Tc-tin colloid which following intravenous injection accumulates mainly in the liver. The substance emits gamma rays, and we can produce an image of its distribution using a nuclear medicine imaging system. This image can tell us the physiological functional information of the liver or localize the diseased sections.

Nuclear medicine procedures are best served by tracers labeled with a radionuclide that has a physical half-life that is long enough to allow for imaging in a reasonable amount of time, but not so long as to continue to irradiate the patient much beyond that imaging. Thus, radiotracers cannot be stored but must be generated daily for immediate use. To provide for tracers labeled with short-lived radionuclides, generators containing the parent material are constructed to provide an extended source of the daughter; alternatively, the radionuclide may be generated in a medical cyclotron, used to label a tracer, and the tracer shipped to the nuclear medicine department for use. This latter method is generally employed for positron-emitting radionuclides, the exception being rubidium-82, a blood flow PET tracer produced in a generator. The most commonly used radionuclide in nuclear medicine procedures is Tc-99m, and the generation of this radionuclide is from molybdenum-99 (Mo-99)-Tc-99m generator. Discussion on moly-generator is beyond the scope of this book.

Other widely used radionuclide tracers which need to be discussed in detail are the PET radionuclides. The most commonly used radionuclides for PET include fluorine-18 (F-18), carbon-11 (C-11), nitrogen-13 (N-13), and oxygen-15 (O-15). In fact, the successful synthesis of F-18 and the application of 18F-FDG had provided another major drive for the advancement of PET [5, 6].

FDG is the most commonly used tracer for PET but is plagued with false-positivity. For instance, the overall accuracy of FDG-PET in detecting a solitary pulmonary nodule is in the order of 90 %, but false positivity due to granulomatous diseases like tuberculosis leads to incorrect diagnostic categorization due to similar uptake of FDG by affected cells. Tracers that use cellular mechanism for which tumor tissue will be different from that of normal tissue will reduce the incidence of

false positivity. One such mechanism that is prominent for tumor tissue is its high cellular proliferation. Tracer targeting this cellular feature will provide an appropriate diagnostic categorization of tumor tissue. F-18 fluoro-deoxy-L-thymidine (FLT) is a promising candidate which gets incorporated in the DNA in a fashion similar to that of natural thymidine in the cell nucleus. The cellular proliferation is earmarked by mitosis which leads to DNA synthesis and hence specific increased uptake of FLT which can be imaged using FLT-PET. On the downside, organs like the liver and bone marrow have high affinity for FLT even under normal conditions due to high cellular turnover. Hence, FLT is of less use in detecting primary or metastatic lesion in these sites.

Cellular Mechanism of Radiotracers

The success of nuclear medicine applications depends on the detection of signals emitted by an injected radionuclide that is concentrated in the pathological tissue under evaluation. For the radionuclide to concentrate in the specific target tissue, it should be tagged to a modified biomolecule for which the cells at the target location have high affinity for specific uptake. The substitution of radionuclide onto the biomolecules will not significantly alter the reaction time or mechanism of the molecule; hence it is taken up as if it is the natural substrate for the cell.

The common cellular mechanisms that are the targets for tracing the specific uptake of radionuclides are:

1. Glucose utilization of the cell targeting the glycolytic pathway: All cells utilize glucose as a substrate for energy. The glucose is taken up by the cell via GLUT membrane transporter. FDG is a glucose analogue that enters the cells via the same membrane transporters as glucose. Glucose as well as 18F-FDG is phosphorylated by the enzyme hexokinase. In contrast to glucose-6-phosphate, 18F-FDG-6-phosphate is not a substrate for further metabolism in the glycolytic pathway. Therefore, 18F-FDG-6-phosphate is trapped in the cells in proportion to their glycolytic activity [7].
2. Cellular proliferation mechanism: The proliferation rate of a normal cell is different from that of malignant cells. DNA synthesis is high in rapidly proliferating malignant cells. A carefully fluorinated analogue of a pyrimidine or a purine base will behave in the same manner as natural bases and gets incorporated into the DNA of rapidly proliferating tissues. F-18 fluoro-deoxy-L-thymidine (18F-FLT) is a fluorinated analogue of thymidine. The specificity of FLT is high for malignant tumors that have a high cellular proliferation rate relative to that of the nonmalignant tissue. They yield false-positive results for tumors in organs like liver and bone marrow which innately have high cellular turnover.
3. Protein synthesis machinery: Malignant tumors characteristically have high protein synthesis in comparison with benign tumors. Specific amino acids labeled with radionuclide will be taken up by malignant tumors for protein synthesis. A high concentration of radiolabeled amino acid analogues in a

tumor can be imaged using PET scanner. Examples include ^{11}C-methionine and F-18 fluoroethyltyrosine (^{18}F-FET).

4. Choline synthesis: Low-grade tumors have less affinity for ^{18}F-FDG and ^{18}F-FLT. Many low-grade tumors are characterized by high choline content. Presumably these cells are also associated with choline transport and involved in sterol metabolism. ^{18}F-fluorocholine, a radiolabeled choline analogue, has provided promising results in some ^{18}F-FDG false-negative cases [8].
5. Tumor vascularization: Proper vascularization is required for tumor growth and for metastasis. Inadequate vascularization especially in the core of the tumor results in hypoxia and necrosis. Tumor tissue hypoxia could be a good mechanism to target in several solid tumors. ^{18}F-fluoromisonidazole (FMISO) radionuclide tracer is a promising candidate specific for tissue hypoxia in vivo. FMISO in combination with ^{15}O-water perfusion imaging has been used to assess the presence and severity of intratumoral hypoxia, a major determinant of treatment resistance [9]. Another recent addition that targets tissue hypoxia is ^{18}F-fluoroazomycinarabinofuranoside (FAZA) [10].

Radiation Safety

Radiation exposure is a very critical issue which needs to be addressed in all radionuclide studies. If radiation exposure exceeds the permissible limit, it will have effects that may range from trivial to fatal with short-term and long-term sequela-like radiation sickness, vomiting, alopecia, radiation enteritis, GI bleeding, radiation burns, skin carcinoma. On an average, each individual is exposed to a natural background radiation of 3 mSv annually from naturally occurring radioactive materials and cosmic rays from outer space while the largest source of background radiation comes from radon gas. The maximum permissible dose (MPD) per year is for:

- Occupationally exposed individuals—50 mSv:
  - Optimal design goal for restricted area should not exceed 5 mSv/year.
- Individuals who are infrequently exposed or in contact with patients receiving radionuclide therapy—5 mSv.
- General Public—1 mSv.
- Individuals subjected to X-ray security screening—0.25 mSv.

370 MBq of 18F-FDG delivers a dose of 11 mSv to a patient. A patient who has undergone a therapeutic procedure with sealed or unsealed radionuclides may deliver a high radiation dose to people coming in contact with him/her. Hence, a guidance level of 1,000 MBq has been laid as a standard for discharge of patients who had recently undergone radionuclide therapy.

See Table 14.3 for the diagnostic CT effective radiation doses. See Table 14.4 for effective radiation doses for various radionuclide studies.

Table 14.3 Diagnostic CT effective radiation doses

Organ	Radiation dose (mSv)
Head[a,b]	2
Chest[a,c]	8
Abdomen[a,c]	10
Pelvis[a,c]	10

[a]IAEA. Radiation Protection for Patients. Available at: https://rpop.iaea.org/RPOP/RPoP/Content/InformationFor/Patients/patient-information-computed-tomography/
[b]Mettler et al. [107]
[c]Wall and Hart [108]

Table 14.4 Effective radiation dose for various radionuclide studies

Radionuclide study	Radiation dose (mSv)
Lung ventilation (Xe-133)[a]	0.4
Lung perfusion (Tc-99m)[a]	1.2
Kidney (Tc-99m)[a]	2.2–2.5
Thyroid (Tc-99m)[a]	2.6
Bone (Tc-99m)[b]	4.8
Cardiac gated study (Tc-99m)[a]	4.2
PET/CT whole body (F-18 FDG)[c]	7–8

[a]Part 6, Medical Exposure Protection of the Patient. IAEA Training Material on Radiation Protection in Nuclear Medicine. Available at: https://rpop.iaea.org/RPOP/RPoP/Content/Documents/TrainingNuclearMedicine/Lectures/RPNM_Part06_medical_exp_WEB.ppt
[b]IAEA. Radiation Protection for Patients. Available at: https://rpop.iaea.org/RPOP/RPoP/Content/InformationFor/Patients/patient-information-computed-tomography/
[c]IAEA Radiation Protection for Patients (RPOP), PET/CT Scanning. Available at: https://rpop.iaea.org/RPOP/RPoP/Content/InformationFor/HealthProfessionals/6_OtherClinicalSpecialities/PETCTscan.htm

Radiation Exposure to the Workers

The total average exposure for a radiology technician or technologist results in an annual dose equivalent of only 1–1.5 mSv, whereas for the nuclear medicine technologist, it is 2–2.5 mSv. The majority of whole-body radiation to the nuclear medicine worker comes from exposure to the dosed patient during imaging.

Radiation Exposure to the Patient

As most nuclear medicine procedures require an injection of radioactive material attached to a tracer, utmost care must be taken to ensure that the correct procedure has been selected to maximize diagnostic suitability, that the radioactivity of the

injected dose follows the principles of ALARA (as low as reasonably achievable) or ALARP (as low as reasonably practicable), and that the patient has been given adequate information prior to the procedure and such aftercare details as are appropriate. All nuclear medicine procedures require that precautions be taken when administering radioactive materials to females of child-bearing age.

Good Clinical Practice (GCP) and Good Manufacturing Practice (GMP)

Knowledge and compliance with the good clinical practice (GCP) is essential for everyone involved in clinical research trials for nuclear medicine. Clinical trials should be carried out within the framework of a good clinical practice environment in accordance with international guidelines and regulations as detailed in the Declaration of Helsinki [11, 12]. The International Commission of Radiological Protection and the World Health Organization (WHO) have publications that deal with this issue in clinical research. While GCP should form the backbone to successful nuclear medicine clinical studies, radiopharmaceuticals used in these trials need to be produced according to the good manufacturing practice (GMP) [13].

Radiopharmaceuticals for clinical research purposes must be manufactured in accordance with the basic principles of GMP. Due to their short half-lives, many radiopharmaceuticals are administered to patients shortly after their production, so some elements of the quality control may be retrospective. Therefore, strict adherence to GMP is essential. Special attention should be given to the production area environment and personnel, the two basic requirements of GMP production.

Quality Control

Since lots of confounding parameters exist in nuclear medicine, it is a good practice to stick to certain established standards laid by the regulatory bodies, advisory committee, professional organization, and manufacturer's guidelines to ensure quality control of the imaging equipment, radiopharmaceutical tracers, procedure, and imaging protocol. Quality control starts at the time of installation and continues throughout the life cycle of the equipment. It is an ongoing process involving measurements and analyses designed to ensure that the performance of a procedure or instrument is within a predefined acceptable range and to keep it operating at peak performance all the time [14]. Therefore, it is a critical component of routine nuclear medicine practice. A detailed explanation on the quality control program is beyond the scope of this book.

The common quality control measures followed routinely in nuclear medicine practice are:

- Uniformity calibration.
- Spatial linearity.
- Energy resolution and peaking.

- Pixel size calibration.
- Center of rotation correction.
- Tomography resolution.
- Normalization scan.

Uniformity Calibration

The uniformity in the performance of the crystal detector is important for the diagnostic quality of all images. Any nonuniformity in the performance can be demonstrated by an image called the flood image that is irradiated by a uniform distribution of radioactivity. From the high-count flood image of a particular radionuclide, the uniformity (or sensitivity) correction table is derived for that radionuclide which is essentially the pixel-by-pixel ratio of the calculated mean count per pixel to the actual count per pixel in the flood image. Causes of nonuniformity may be due to cracked crystals because of mechanical trauma or temperature excursion beyond threshold, improper tuning of photopeak of radionuclide with the photopeak energy window of camera, uncoupling of photomultiplier tubes from crystals, corrupted software correction tables, etc. Pixel-by-pixel multiplication of an uncorrected image by the ratio image thus calculated gives an image corrected for the nonuniformity.

Uniformity calibration could be evaluated either intrinsically by a point source of ^{99m}Tc placed at about 5 crystal dimensions in the z direction from the center of the uncollimated detector so that a uniform photon flux spreads over the detector (with a count rate of >25,000 cycles/second) or extrinsically by using a sheet source of ^{57}Co placed over the collimated detector (with a count rate of 10–15 million cps). This yields the integral and differential uniformity which expresses the deviation from flood image uniformity.

In current generation nuclear imaging systems, the integrated software solution provides uniformity correction tables that can be easily updated, processed, stored, and automatically applied.

Spatial Resolution

Spatial resolution is the power to discriminate two closely placed point sources distinctly. The overall resolution of a gamma camera is based on the sum of intrinsic spatial resolution of the detector system and the geometric resolution of the collimator. To evaluate the spatial resolution of a gamma camera, a wide variety of test patterns have been developed over the years. The most common test pattern is the four-quadrant bar phantom, which accounts for over 80 % of all resolution patterns used in nuclear medicine. Apart from this, several other test phantoms are available. All these provide a qualitative index of resolution. Correct use of these phantoms requires that all images be compared with a reference image. This reference image can be obtained during acceptance testing or when a quantitative measurement of intrinsic resolution is being performed on the system.

The modulation transfer function (MTF) is an index that exhibits the ability of a gamma camera to yield an image corresponding exactly with the physical distribution of the radionuclide.

Spatial Linearity

Spatial linearity may be affected with nonuniformity due to ill-defined factors like variations in crystal thickness. Spatial linearity correction is done by presenting the camera with an image consisting of a series of parallel straight lines aligned with either the *X* or *Y* axis of the camera. The deviation between the true position of each point on the line, as calculated from a best-fit straight line, and the image of the line, is recorded and stored as a correction factor to be applied to subsequently acquire clinical images.

The correction is generally performed on a weekly basis using either intrinsic or extrinsic techniques.

With this basic idea of nuclear physics, imaging, radiopharmaceuticals, and quality monitoring, we will now proceed to organ-based applications of the technique as a problem-solving tool in medicine.

Nuclear Imaging in Oncology

Positron emission tomography (PET) is an imaging technique based on nuclear physics principles that provides in vivo measurements in absolute units of a radioactive tracer. One of the attractive aspects of PET is that the radioactive tracer can be labeled with short-lived radioisotopes of the natural elements of the biochemical constituents of the body (^{18}F-fluoro-2-deoxy-glucose or [^{18}F]-FDG). This provides PET with a unique ability to detect and quantify physiologic and receptor processes in the body, particularly in cancer cells, which is not possible by any other imaging techniques today. Oncologic PET studies now represent almost 90 % of all clinical studies performed in clinical PET centers worldwide [15–19].

Although ^{18}F-FDG is the most commonly used positron-emitting tracer in PET imaging, there are measurement of tissue blood flow, oxygen metabolism, glucose metabolism, amino acid and protein synthesis and nucleic acid metabolism that have all been demonstrated in PET oncology clinical studies using other tracers [20–23]. Labeling of a large array of other compounds including hypoxic markers, amino acids, DNA proliferation markers, and chemotherapy drugs with ^{11}C and ^{18}F has been studied in various clinical trials (Table 14.5) [17, 22, 23].

The practical and ethical issues associated with PET examination poses difficulty in the conduct of randomized controlled trials; thus, the establishment of diagnostic accuracy and impact on patient management are mainly obtained from clinical practice [18, 24]. There is increasing evidence for the role of PET in staging, monitoring treatment response and biologic characterization of tumors [15–19, 21–23, 25–30].

Table 14.5 Positron-emitting radionuclides used in oncology clinical studies

Radionuclide	Half-life
^{15}O	122 s
^{13}N	9.97 min
^{11}C	20.4 min
^{18}F	109.8 min
^{124}I	4.17 days
^{86}Y	14.7 h
^{64}Cu	12.8 h

Brain Tumor

The evaluation of brain tumors with ^{18}F-FDG PET is a well-established oncologic application of PET. Tumor grade can be assessed accurately and noninvasively by ^{18}F-FDG PET, as the rate of glucose (tracer) utilization is directly proportional to the degree of malignancy [31]. This can be used in the planning of biopsies, and in monitoring high-grade recurrence, particularly in patients with low-grade glioma. Increased ^{18}F-FDG uptake is seen in high-grade glial tumors, as well as in primary cerebral lymphomas, pilocytic astrocytomas, and some unusual tumors (e.g., pleomorphic xanthoastrocytoma, low-grade gliomas). Primary brain tumors (e.g., meningiomas) do not usually show increased ^{18}F-FDG uptake except in more aggressive tumors and in postradiation meningiomas.

Secondary cerebral metastases occur in almost 20–40 % of systemic malignancies and may be the initial presentation of malignancy in 16–35 % of cases. ^{18}F-FDG PET has been extensively studied in these patients with a sensitivity ranging from 68 to 79 % [32]. The principal constraint in using FDG-PET for evaluation of secondary metastasis is the frequent hypometabolic nature of cerebral metastases, and in addition, metastatic lesions are often small (<1 cm in size), and because metastases most often occur at the interface between grey and white matter, identification of lesions can be challenging.

FLT could be used in a limited number of cases suspected of rapidly progressing proliferating brain tumor. C-11 methionine is another radionuclide that is being extensively evaluated for its application in the detection of brain tumor. Unlike FDG, the uptake of ^{11}C-methionine in the human brain is almost negligible. Since the proliferation rate of many brain tumors is relatively low, radiolabeled amino acid analogues are more sensitive than FLT [10]. Currently, the production yield of ^{11}C-labeled radioligand is far below its requirement, thus its practical application is faced with difficulty. F-18 fluoroethyltyrosine (F-18 FET) is a promising alternative fluorinated amino acid analogue whose preliminary studies in brain tumor evaluation appear promising [33].

Lung Carcinoma

There have been numerous studies examining the accuracy of ^{18}F-FDG PET in evaluating solitary pulmonary nodules [34–36]. Analyzing the published data has shown

a high sensitivity of 96 % and accuracy with 94 % in determining this malignancy [16, 23, 37]. Staging in non-small cell lung carcinoma (NSCLC) is very crucial in treatment planning. Studies that have evaluated the role of ^{18}F-FDG PET for staging have reported sensitivity of 82–100 % and specificity of 73–100 % [25, 30, 38–41]. In all series, ^{18}F-FDG PET has been shown to outperform CT in staging lymph node spread in up to 24 % of patients [39]. False-negative results often arise where small lesions are present (<0.6 cm), due to the resolution limitations of PET scanners, respiratory motion over the acquisition period or with certain types of lung cancer such as bronchoalveolar and carcinoid. Figure 14.1a, b shows PET/CT images with DOTATOC tracer from a patient with recurrent lymph node metastases with primary lung cancer. Figure 14.2a, b shows PET/CT image from a patient with pulmonary carcinoid tumor in left lower lobe with metastases to liver. DOTATOC (DOTA(0)-Phe(1)-Tyr(3)-octreotide) is a highly specific PET tracer to somatostatin receptor (SSR2) of neuroendocrine tumor [42, 43].

Colorectal Carcinomas

Primary colorectal cancers occasionally present as an incidental finding on ^{18}F-FDG PET, and ^{18}F-FDG uptake has been reported in adenomatous polyps, a precursor for colon cancer [44]. However, the presence of physiological gut uptake of FDG combined with false-positive uptake in inflammatory disease along with low sensitivity to lesions less than 1 cm precludes a significant role for FDG-PET in primary diagnosis or screening [45]. The role of PET in primary colon cancer remains limited and should be reserved for clinical situations where resection of metastatic disease requires accurate staging of distant spread. PET is an excellent tool in detecting secondary metastasis to liver or extrahepatic abdominal metastases. In published series, the accuracy of ^{18}F-FDG PET in identifying metastatic colorectal carcinoma in the liver has ranged from 90 to 98 % [23] and in extrahepatic diseases, it ranges from 92 to 93 %. In patients with elevated serum CEA markers, occult disease (often extrahepatic recurrence) has been identified accurately with ^{18}F-FDG PET [46].

In advanced rectal cancer, where ^{18}F-FDG PET has been shown to have a significant impact on management in up to one third of patients planned for preoperative adjuvant treatment (chemoradiation), indicating the potential role of ^{18}F-FDG PET in this clinical setting [47].

Lymphoma

The sensitivity and specificity of 18F-FDG PET in detecting the sites of lymphoma has been reported as 86 %–90 % and 93 %–96 % respectively and is considerably superior to CT scans. 18F-FDG PET aids in staging Hodgkin's and non-Hodgkin's lymphoma (NHL) with high acuity. [4, 16, 23]. ^{18}F-FDG PET has been also shown to change the

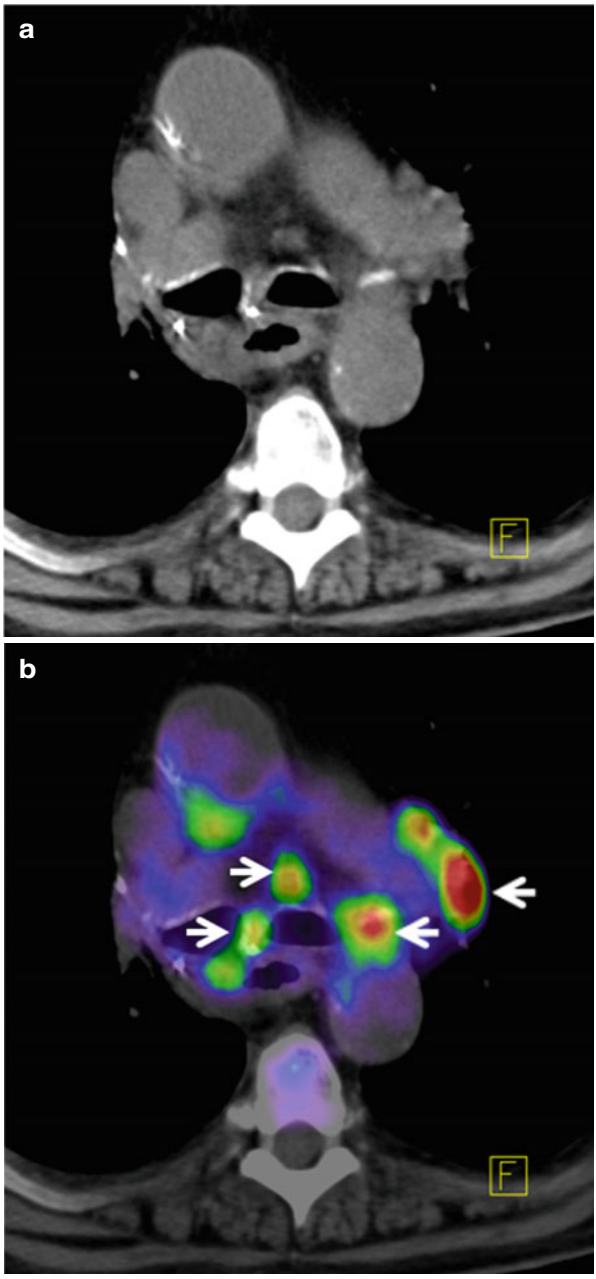


Fig. 14.1 (**a**) Axial slice from CT thorax level of the mediastinum presenting a patient with recurrent lymph node metastases from primary lung cancer. (**b**) Co-registered PET/CT images with high FDG tracer uptake indicating lymph node metastases shown with *arrows*

management in up to 40 % of patients undergoing staging at initial diagnosis [48–51]. In comparison to ^{67}Ga scans, ^{18}F-FDG PET has been shown to have a greater sensitivity for disease detection (particularly spleen) and in view of the potential advantage of a

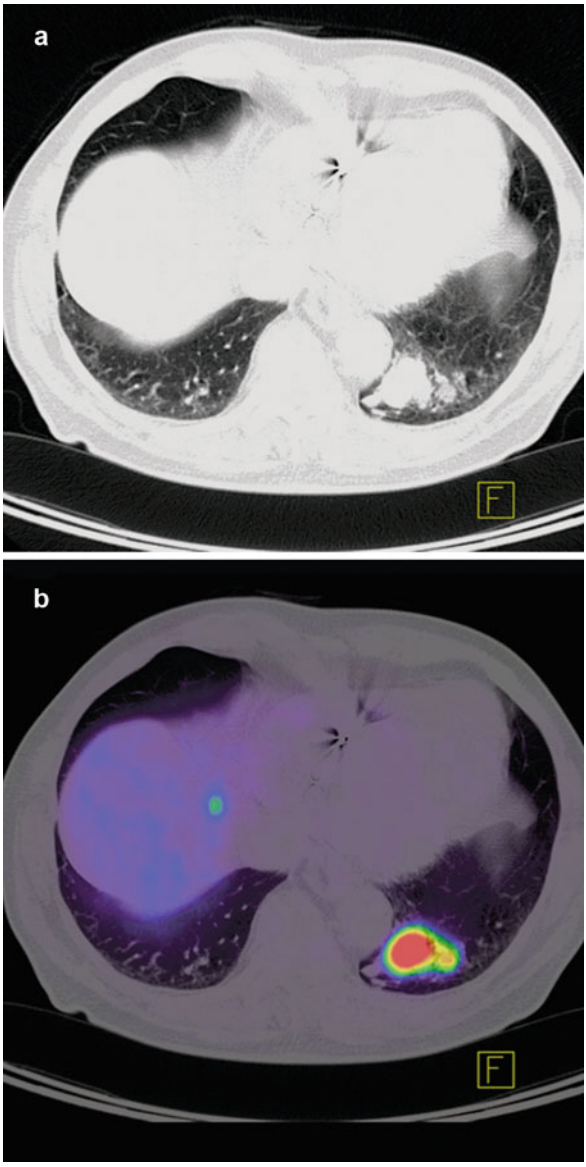


Fig. 14.2 CT (**a**) and PET-CT (**b**) axial slice of chest from a patient with pulmonary carcinoid tumor in the left lower lobe. The neuroendocrine differentiation of the tumor leads to an intense somatostatin receptor (SSR2) overexpression on which the somatostatin analogue 68Ga-DOTATOC tracer can bind. Physiologically SSR2 is expressed only in pituitary, thyroid, adrenals, and excretory organs; any other representation is considered pathological

same-day procedure has supplanted ^{67}Ga scans in many oncology centers [52]. ^{18}F-FDG PET is now routinely performed as part of the assessment of treatment response for NHL; it has also shown to be superior to conventional imaging and to be a strong prognostic indicator of response and progression-free survival [53, 54]. ^{18}F-FDG PET therefore has a major role in both the initial staging and restaging/therapy response assessment of patients with lymphoma. Figure 14.3 shows PET images before, during, and after chemotherapy from patient with Hodgkin's lymphoma.

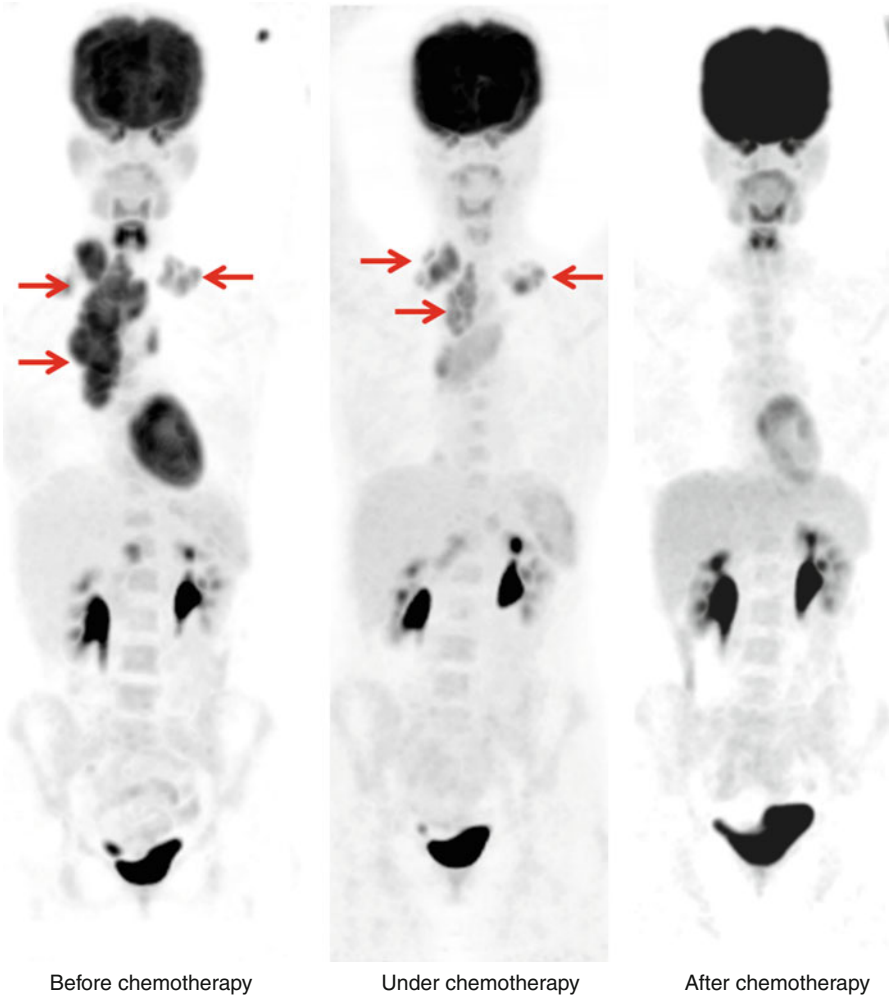


Fig. 14.3 Patient with Hodgkin's lymphoma before (*left panel*), under (*middle panel*), and after several cycles of chemotherapy (*right panel*) showing high FDG uptake of in the cervical and mediastinal region (*arrows*) which decline after several cycle of chemotherapy—metabolic component (FDG-PET) is the leading imaging modality in these tumor entities for tumor response assessment

Melanoma

Malignant melanoma can spread widely and unpredictably throughout the body, and median survival after the appearance of distant metastases is approximately 6 months [55] which makes it very critical to be detected at an early phase. The accuracy of ^{18}F-FDG PET in detecting metastatic melanoma has been reported to range

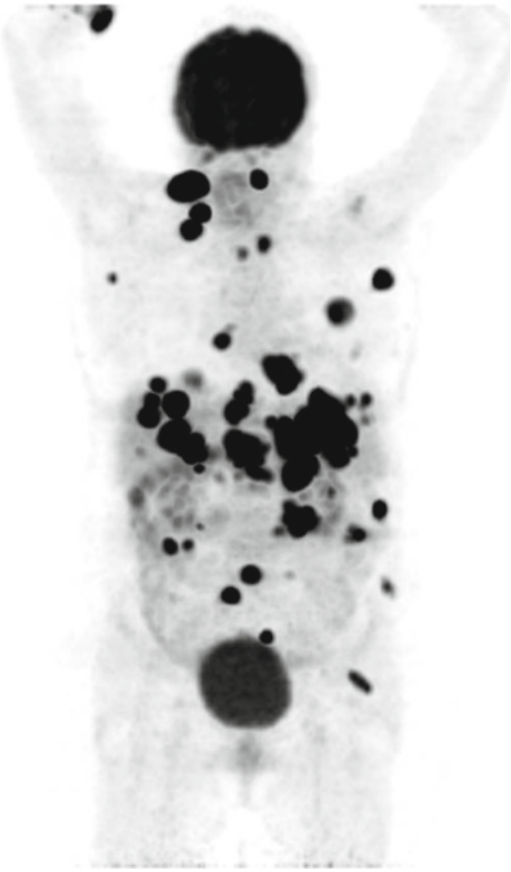


Fig. 14.4 Maximum intensity projection in coronal plane of body presenting FDG PET uptake in a patient with metastasized malignant melanoma

from 81 to 100 %, and in one series of 100 patients demonstrated a sensitivity of 93 % [27]. ^{18}F-FDG PET has been shown to be particularly sensitive in detecting subcutaneous and visceral metastases (Fig. 14.4). In published studies and meta-analyses of the literature, ^{18}F-FDG PET has been demonstrated to detect disease up to 6 months earlier than conventional techniques and alter management in 22–32 % of patients, principally by altering plans for surgical resection of metastatic disease [27, 56, 57]. The role of ^{18}F-FDG PET in melanoma is therefore principally in the evaluation of extent of metastatic disease, the accurate assessment of which can alter patient management particularly where surgery is planned.

Head and Neck Tumors

The presence of lymph node spread of head and neck tumors is associated with substantially worse prognosis which needs to be addressed in early phases. In patients with head and neck tumors studied prior to initial surgery, the sensitivity and

specificity of ^{18}F-FDG PET in detecting nodal metastases has been reported ranging from 71 to 91 % and 88 to 100 %, respectively [23, 58–60]. In patients studied after initial treatment of metastatic nodal disease with radiotherapy, ^{18}F-FDG PET is often accurate only after a 3-month period [35, 61–63]. In both patient groups, the accuracy of PET has the potential to direct surgeons to otherwise unexpected sites of metastatic disease, as well as in avoiding surgery at areas where the scan is negative. ^{18}F-FDG PET has also been shown to be a prognostic factor for radiotherapy response [64].

Breast Carcinoma

In primary breast tumors, ^{18}F-FDG PET has been shown to have a mean sensitivity and specificity for tumor detection of 88 and 79 % respectively in a recent meta-analysis [65]. Axillary nodal involvement is a critical issue in the management of patients with breast carcinoma and ^{18}F-FDG PET has been shown to have a sensitivity ranging from 57 to 100 % and specificity of 66–100 % across reported series [23] in detecting sentinel node involvement. One potential area where ^{18}F-FDG PET has shown great promise is in whole-body staging of metastatic breast cancer, where the accuracy of ^{18}F-FDG PET has been shown to be higher than conventional staging techniques [66, 67]. But certainly FDG-PET is not the best modality of choice for detecting primary carcinoma as most of primary breast tumors are FDG negative.

Gastroesophageal Carcinoma

^{18}F-FDG PET has been shown to significantly improve detection of hematogenous and distant lymphatic metastasis in carcinoma of the esophagus and gastroesophageal junction (GEJ) [34, 68–70]. There is no difference in accuracy in detecting squamous cell carcinoma or adenocarcinoma of the esophagus with ^{18}F-FDG PET. ^{18}F-FDG PET may not be as accurate as USG or CT in determining wall invasion or close lymph node spread of disease; however, the diagnostic specificity of lymph node involvement is greatly improved with PET [70]. ^{18}F-FDG PET is more accurate in detecting distant disease and is also highly accurate in the diagnosis of recurrent disease [68–71].

Ovarian Carcinoma

Ovarian carcinoma is one of the leading causes of death among gynecological malignancies [72]. The treatment of ovarian carcinoma primarily consists of surgical resection followed by chemotherapy or radiotherapy. Accurate staging is very much essential, particularly in the restaging of patients with elevated serum markers (CA-125). ^{18}F-FDG PET has been shown to have high accuracy in detecting in ovarian carcinoma lesions greater than 1 cm in size, but the detection of micrometastatic disease (one of the most

important issues in this disease) has been difficult [73–75]. The role of PET in ovarian carcinoma is very much restricted to post-therapy monitoring of lesion recurrence.

Prostate Cancer

Due to the slow growth and the accompanying low glucose metabolism, sensitivity of FDG-PET is low in diagnosing prostate cancer and metastases thereof (18–65 %; [76, 77]). Also in lymph node metastases smaller than 1 cm, sensitivity and specificity is low [78].

Above all, there is a great overlap of FDG uptake in tumors and benign prostate hyperplasia (BPH, [79]). Furthermore, the high residual activity in the bladder after renal excretion superimposes the uptake in the prostate.

Due to these limitations, other tracers were tested to improve sensitivity. These are ^{11}C- and ^{18}F-labeled cholines. Since cholines are metabolized to membrane phospholipids, the renal excretion is low, and the tracer is accumulated in dividing tumor cells. With choline, a higher sensitivity is possible in detection of metastasized prostate cancer [80]. In BPH the distribution is lower and more homogeneous compared to the tracer uptake in tumors.

^{18}F-fluoromethyl-dimethyl-2-hydroxyethyl-ammonium (FCH) is accumulated in primary tumor, as well as soft tissue and bone metastases [81]. It shows even in smaller lymph nodes a high sensitivity of 66 % and specificity of 96 % [82]. In one study it was shown that ^{11}C-acetate was superior to FDG-PET in detecting recurrent tumor (59 % vs. 17 %, [83]). However, there is an overlap in SUV values of tumor and BPH as well [84]. Figure 14.5a–c shows PET/CT image with FCH tracer from patient with recurrent prostate cancer. Prostrate-specific membrane antigen (PSMA) is a recently introduced small molecule which is specific to prostate cancer cells [85], however uptake by colon cancer and ENT tumors has also been reported. Figure 14.6a–c shows the PET image from one of the patients for whom total prostate resection was performed 3 years ago, now presenting with high PSA serum indicating recurrence of prostate cancer.

L-methyl-^{11}C-methionine is also superior to FDG-PET in lesion detection with a sensitivity of 72.1 % against 48 % for FDG-PET [86]. More than 95 % of metabolically active sites showed metabolism of 11C-methionine; whereas 18F-FDG showed metabolic activity in only 65 % of active sites [86]. However, a significant proportion of lesions (26 %) had no detestable metabolism with either of the two tracers [86].

Nuclear Imaging in Central Nervous System

SPECT and FDG-PET are used to provide complimentary information to the anatomical imaging offered by other conventional modalities (CT/MRI). New methods of imaging neurotransmitter receptors and transporters have been developed and offer expanded roles for brain SPECT. Nuclear medicine can make valuable contributions to the diagnosis and follow-up of patients with dementia, cerebrovascular disease, movement disorders, brain tumors, and other neurological diseases.

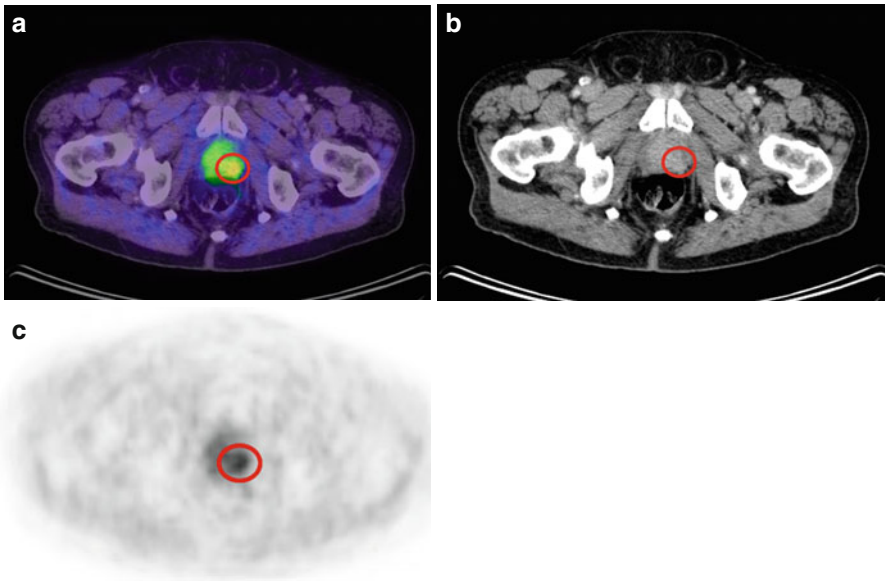


Fig. 14.5 FCh-PET/CT (**a**), CT (**b**), and FCh-PET (**c**) in transverse plane presenting high tracer uptake in a patient with recurrence of prostate cancer (red circle indicating the tumor in the left peripheral zone)

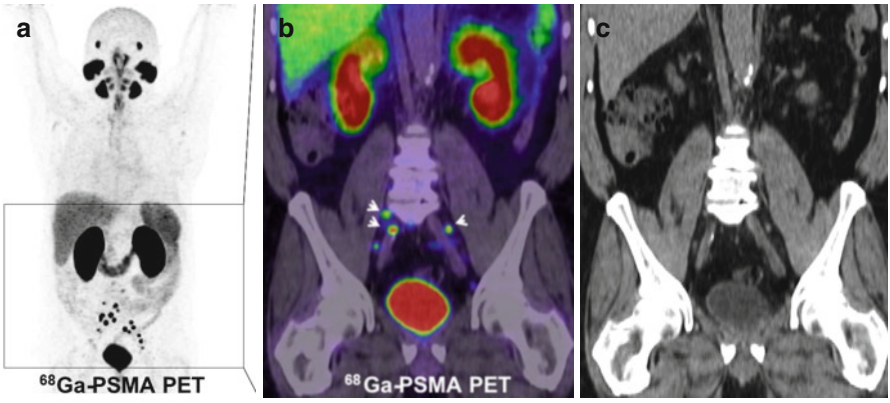


Fig. 14.6 A patient with an elevated PSA serum level as indicator of recurrence of prostate cancer after total prostate resection 3 years ago. Patient was imaged with a recent introduced new small molecule PET tracer called prostate-specific membrane antigen (*PSMA*). (**a**) Presenting whole-body distribution of ^{68}GaPSMA PET in MIP technique. (**b**) PSMA is specific and taken up by prostate cancer cells (presenting strong uptake in these lymph node (*arrow*) in the pelvis region) indicated recurrence of prostate cancer. The coronal CT (**c**) small lymph nodes close to the ilio-psoas muscle on both sides can be appreciated without suspicious enlargements (<1 cm)

The exclusive metabolic substrate of brain is glucose, but the brain has virtually no means of storing energy and is totally dependent on cerebral blood flow for the delivery of glucose and oxygen. Glucose is oxidized to yield ATP, the energy currency of the neurons. The metabolic demand owing to neuronal activity increases the blood flow delivering oxygen and glucose to the particular area of brain, and there exists a close coupling of neuronal activity and blood supply. This feature is exploited in brain activation studies with ^{15}O-H_2O PET, BOLD MRI, and occasionally activation studies with regional cerebral blood flow SPECT (rCBF SPECT).

Conversely, reduced function of parts of the brain will result in a reduction in regional cerebral blood flow, although such hypofunctioning areas may appear anatomically normal on structural imaging with CT or MRI. This ability of rCBF SPECT to reflect brain function is its key strength and it should be regarded as a complementary method of investigation to structural imaging methods.

Dementia

Alzheimer's disease (AD), predominantly a disease of elderly, presents clinically with memory and cognitive impairment. AD is primarily a clinical diagnosis. Pathologically it is characterized by neurofibrillary tangles comprising of abnormal hyperphosphorylated tau protein and extracellular beta amyloid. The classic metabolic abnormality associated with AD is bilateral temporoparietal hypometabolism (Fig. 14.7a, b). The FDG-PET showing hypometabolism of temporoparietal area confirms the diagnosis with 82 % diagnostic accuracy. If FDG-PET scans indicated a metabolic pattern other than bilateral temporoparietal hypometabolism, a cause of dementia other than AD should be suspected [87]. Newer radioligands like AZD2184 labeled with carbon-11 that binds to A beta deposits, associated with the pathogenesis of AD, provide improved contrast when compared with currently used PET radioligands for visualization of A beta deposits [88].

The rCBF SPECT study using the radiopharmaceuticals like hexamethyl propylene amine oxime (HMPAO) and ethylene cysteine dimer (ECD) may show hypometabolism of temporoparietal region in AD with predictive value ranging anywhere from 60 to 96 % [89].

It may not be superior to clinical diagnosis based on detailed neuropsychological testing using NINCDSADRDA (National Institute of Neurological and Communicable Disease and Stroke/Alzheimer's Disease and Related Disorders Association) criteria, but rCBF SPECT has higher specificity and can be used to distinguish dementia due to AD and other causes even without any established CSF biomarkers [90, 91].

Vascular dementia is the second most common dementia after AD caused by ischemic or hemorrhagic cerebrovascular disease (atherosclerosis of large cerebral vessels, hypertension, cerebral amyloid angiopathy, etc.) or by ischemic-hypoxic brain lesions of cardiovascular origin. rCBF SPECT is dependent on the perfusion of cells and is largely a gray matter imaging technique, but vascular dementia may be associated with white matter ischemia without any cortical infarct. Hence, SPECT is of limited use in white matter disease where the uptake of

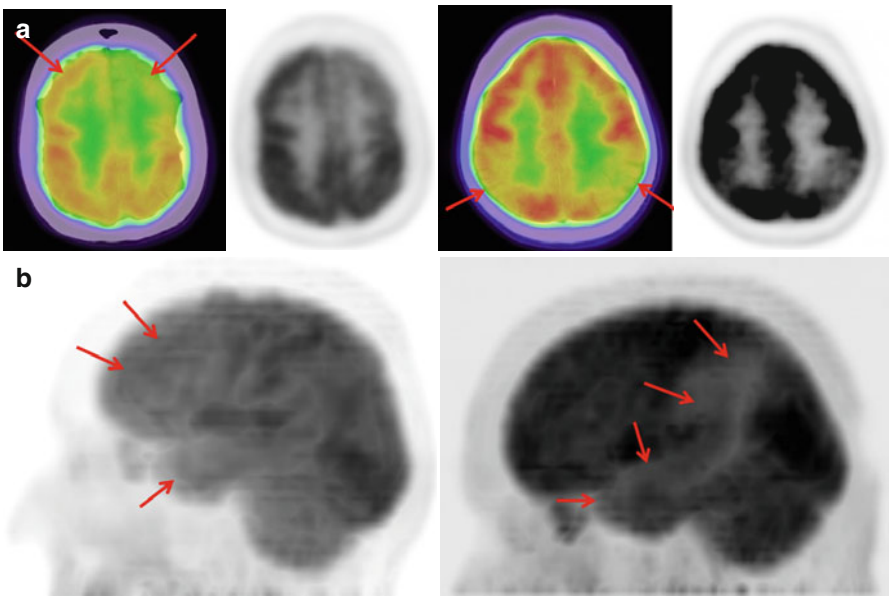


Fig. 14.7 Axial and MIP slice (**a**) and sagittal slice (**b**) from patient with dementia (Alzheimer's disease) showing hypometabolism (*arrows*) in the parietal and temporal region by sparing the motor cortex

radiopharmaceutical is minimal. HMPAO SPECT may demonstrate multiple patchy perfusion defects or reduced perfusion in one or more arterial territories. Amyloid imaging with ^{18}F-labeled radiotracers has recently been introduced for diagnosis of dementia, which is under further evaluations for clinical use. [J Nucl Med. 2011 Aug;52(8):1210-7. doi: 10.2967/jnumed.111.089730. Villemagne at al. Amyloid imaging with (18)F-florbetaben in Alzheimer disease and other dementias]

Epilepsy

In medically refractory epilepsy, surgery is contemplated to be of potential benefit to the patient. SPECT is used in cases where discrepancies arise between EEG finding and MRI.

During an episode of seizure, the metabolic activity of abnormally firing cells increases several fold which is associated with an increase in local blood flow. The radiopharmaceutical HMPAO can be injected during an ictal episode which reaches a steady state. The radiotracer remains bounded for about 6 h post-injection and can be imaged using rCBF SPECT demonstrating the epileptogenic focus [92]. Ictal SPECT shows good sensitivities in the correct lateralization of an electroencephalogram-defined epileptic focus in lesional and, to a lesser extent, non-lesional epilepsy.

Positron emission tomography (PET) using ^{18}F-FDG or ^{11}C-flumazenil will give a good detection rate of the epileptogenic zone in non-lesional cases and extratemporal epilepsy [93].

Nuclear Imaging in Cardiovascular System

Nuclear medicine imaging in the cardiovascular studies include gated/nongated myocardial perfusion imaging, myocardial viability studies, infarction imaging, ventricular function studies, and detection and quantitation of intracardiac shunts.

Exercise on a treadmill, or simulation of exercise by infusion of dipyridamole/adenosine/dobutamine, is used in conjunction with perfusion agents to increase radionuclide delivery to the normal myocardium. Stepwise increases in physical exercise are monitored by sequential electrocardiogram (ECG) and blood pressure and pulse measurements while the patient is queried for symptoms of angina. The radiopharmaceuticals widely used in myocardial perfusion scan are thallium-201 (Tl-201), Tc-99m sestamibi, and Tc-99m tetrofosmin. We will discuss the various aspects of myocardial perfusion imaging (MPI) with focus on Tl-201 imaging.

Radiopharmaceuticals

Thallium-201 is an analog of the potassium ion (K^+), which is delivered to capillary beds by regional blood flow and actively pumped into viable cells by the sodium/potassium (Na^+/K^+) adenosine triphosphatase pump. The effective half-life or 50 % washout of Tl-201 from the normal myocardium is about 4 h.

On the other hand, Tc-99m sestamibi is taken up by the perfused myocardium by passive diffusion and is bound in the myocyte, mostly within myocardial mitochondria. Tc-99m tetrofosmin is rapidly extracted from the blood by perfused myocardium in a fashion that resembles Tc-99m sestamibi. The two agents have proven to act clinically in a very similar manner, but availability and pricing make important considerations.

Clinical Application

Myocardial perfusion imaging demonstrates relative regional perfusion. Areas of myocardium with poor blood supply, usually because of atherosclerosis, fail to increase radiotracer uptake during the stress component. The most important feature of the myocardial perfusion test is comparison of the stress and rest images to detect areas of ischemia that are inadequately perfused at exercise yet still viable. These areas are redundantly called *reversibly ischemic*. A frequent location of ischemic tissue is immediately adjacent to an area of infarct. This is called *peri-infarct ischemia* and does not portend the same clinical significance as an ischemic or reversible zone. Abnormal anatomy in a coronary artery may not produce hemodynamically significant changes in blood flow to the myocardium, and not all ischemia is produced by large vessel atherosclerosis. Capillary disease in diabetics, left bundle branch block, vasospasm, vasculitis, or cardiomyopathy (dilated or hypertrophic) may produce ischemic myocardium even with normal arteries. Ischemia may

not be detected if there is inadequate exercise, inadequate pharmacologic challenge, or compensated triple-vessel disease. The myocardial perfusion scan also detects ischemia due to other causes (including left bundle branch block, coronary vasculitis, and small vessel disease) that cannot be seen on coronary arteriography and thereby reduces its apparent specificity.

Myocardial infarction produces layers of nonperfused scar tissue which are detected as areas of thin myocardium with decreased radiotracer uptake at both stress and rest imaging. The extent of an infarct, from subendocardial to transmural, is reflected by the size and degree of this perfusion defect. A single myocardial perfusion scan cannot determine the age of an infarct. Acute infarcts usually appear larger than old infarcts when imaged with Tl-201. Temporarily damaged cells around infarcted cells, referred to as stunned myocardium, will be hypokinetic/akinetic and will not hold on to the Tl-201 until recovered several weeks later.

PET is more expensive than standard myocardial perfusion imaging but offers the advantages of coincidence imaging, higher-energy photons, efficient attenuation correction, and different radiopharmaceuticals. PET agents can also be imaged on hybrid SPECT cameras or SPECT cameras with heavy collimators. PET scanning with coincidence detection allows high photon flux because collimators are not required. PET scans have higher-resolution images and fewer attenuation artifacts than standard MPI. Thus, PET scans may be the gold standard for MPI. PET perfusion is usually evaluated with rubidium-82 or ammonia-13 ($^{13}NH_3$), comparing the rest imaging with stress imaging, as in standard MPI. Other PET agents that are infrequently used for myocardial perfusion is ^{15}O water (^{15}O-H_2O). Viability of hibernating myocardium is evaluated with resting injection of fluorine-18 fluorodeoxyglucose (FDG). Myocardial oxygen consumption is often measured by ^{11}C-labeled acetate in conjunction with perfusion study [94–96]. Fatty acids undergo oxidation to yield energy for cardiac muscle, and its metabolism could be measured using ^{11}C-labeled palmitate radioligand.

Nuclear Imaging in Genitourinary System

The kidneys receive approximately 25 % of cardiac output and are one of the highly perfused organs. Though the widespread application and technical advancement of CT and MRI has reduced the use of nuclear medicine in assessing renal pathology, nuclear medicine is the choice for functional assessment and still is the gold standard for upper urinary tract obstruction and pyelonephritic scarring secondary to UTI.

The indications for diagnostic imaging usually depend upon the clinical presentation and the age of the patient. Frequently more than one imaging technique is required to fully evaluate the anatomy and physiology of the genitourinary tract.

The ^{99m}Tc-diethylenetriaminepenta-acetic acid (DTPA) and ^{51}Cr-ethylene diamine tetra-acetic acid (EDTA) are loosely bound to plasma protein and hence are freely filtered in the glomerulus. They are not reabsorbed from renal tubules back into the efferent vessels; the glomerular filtration represents its plasma clearance.

Hence, these can be used in the physiological study of glomerular filtration rate (GFR). Another radiopharmaceutical ^{99m}Tc 2, 3-dimercaptosuccinic acid (DMSA) is used for static and SPECT imaging of the kidneys.

Tubular transport tracers include technetium [^{99m}Tc], mercapto-acetyl-triglycine (MAG-3), and ^{99m}Tc dimercapto-succinic acid (DMSA); these agents identify renal cortical tissue and can localize ectopic renal tissue. The DMSA scan is the most accurate imaging modality for the diagnosis of acute pyelonephritis, the decreased accumulation of the tracer in the renal parenchyma is secondary to inflammatory edema, the resultant decreased blood flow and cellular enzymatic activity [97].

^{99m}Tc DTPA is used in diuretic renography. It is mainly used in the differential diagnosis of hydronephrosis/hydroureteronephrosis as the cause for obstructive nephropathy. To distinguish the conditions causing obstructive nephropathy is pertinent as medical management is usually indicated for a dilated/nonobstructed urinary tract, while surgery is recommended for a dilated/obstructed tract to improve the renal function [98].

Nuclear Imaging in Infection and Inflammation

Scintigraphic evaluation of infection and inflammation is a very broad topic in itself and beyond the scope of this chapter; we will briefly elaborate the numerous radiopharmaceuticals and imaging techniques like gallium-67 (Ga-67) citrate, radiolabeled leukocytes (WBC), and Tc-99m-fanolesomab imaging in evaluation of inflammation and infection.

Gallium-67

Gallium-67 has a half-life of 78.1 h, with principal photon energies of 93, 184, and 296 keV; it's been used for imaging over the last three decades, but a poor photon yield per disintegration makes it a suboptimal imaging agent [99].

About 90 % of circulating Ga-67 is in the plasma, nearly all transferrin bound. Increased vascularity and/or increased vascular membrane permeability result in increased delivery and accumulation of transferrin-bound Ga-67 at inflammatory foci. Ga-67 also binds to lactoferrin, which is present in high concentrations in inflammatory foci. Direct bacterial uptake may also account for some Ga-67 accumulation in infection. Siderophores, low-molecular-weight chelates produced by bacteria, have a high affinity for Ga-67. The siderophore-Ga-67 complex is presumably transported into the bacterium, where it eventually is phagocytosed by macrophages. Although some Ga-67 may be transported bound to leukocytes, it is important to note that, even in the absence of circulating leukocytes, Ga-67 accumulates in infection [99]. Imaging is usually performed 18–72 h after injection of Ga-67.

Various indications for Ga-67 imaging are as follows:

1. Ga-67 imaging is the radionuclide procedure of choice in the detection of infections unique to the immunocompromised patients.
2. Ga-67 is extremely sensitive to detection of pulmonary inflammations: sarcoidosis, interstitial pneumonitis, drug reactions, collagen vascular disease, and pneumoconioses.
3. Interstitial nephritis: Ga-67 can be helpful in differentiating interstitial nephritis from acute tubular necrosis in the acute setting.
4. Fever of undetermined origin (FUO) is an illness of at least 3 weeks duration with several episodes of fever exceeding 38.3 °C and no confirmed diagnosis. Ga-67 is typically reserved for those situations in which other imaging tests fail to localize the source of the fever. Since Ga-67 accumulates in foci of infection, inflammation, and tumor, it is often preferred over WBC imaging for this indication [99, 100].
5. Spinal osteomyelitis: Ga-67 imaging is frequently performed in conjunction with bone scintigraphy [101].

Radiolabeled Leukocytes

The only approved methods in radiolabeled leukocytes technique are the lipophilic compounds indium-111 (In-111) oxyquinoline and ^{99m}Tc-HMPAO (hexamethyl propyleneamine oxime). Advantages of the In-111 label are its stability and a virtually constant normal distribution of activity limited to the liver, spleen, and bone marrow. The 67-h physical half-life of In-111 permits delayed imaging, which is particularly valuable for musculoskeletal infection. Disadvantages of the In label include a low photon flux, less-than-ideal photon energies, and the fact that a 24-h interval between injection and imaging is generally required [100].

Advantages of Tc-99m-WBCs include a photon energy that is optimal for imaging using current instrumentation, a high photon flux, and the ability to detect abnormalities within a few hours after injection. The instability of the label and the 6-h half-life of Tc-99m are disadvantages when delayed 24-h imaging is needed. This occurs in those infections that tend to be indolent in nature and for which several hours may be necessary for accumulation of a sufficient quantity of labeled leukocytes to be successfully imaged [100].

Clinical Utility

Using Tc-99m-WBCs, diffuse pulmonary uptake on images obtained more than 4 h after injection of labeled cells is associated with opportunistic infection, radiation pneumonitis, pulmonary drug toxicity and ARDS. This pattern is almost never seen, however, in bacterial pneumonia [102]. Diffuse pulmonary uptake of WBCs is also seen in septic patients with normal chest radiographs and who have no clinical evidence of respiratory tract inflammation or infection.

^{111}In-WBCs do not accumulate in normal bowel. Such activity is always abnormal and is seen in antibiotic-associated or pseudomembranous colitis, infectious colitis, inflammatory bowel disease, ischemic colitis, and GI bleeding [100, 103]. Also radiolabeled WBCs do not accumulate in normally healing surgical wounds, so the presence of such activity indicates infection although there are certain exceptions.

Both Ga-67 and WBC imaging are confirmatory in detecting myocardial abscesses in patients with infective endocarditis [100]. WBC imaging is the radionuclide procedure of choice for diagnosing prosthetic vascular graft infection, with a sensitivity of more than 90 % [100].

WBC imaging is very sensitive for detecting inflammatory bowel disease and can be used as a screening test to determine which patients need to undergo more invasive investigation.

Tc-99m Fanolesomab

Tc-99m fanolesomab, a monoclonal murine M-class immunoglobulin, binds to CD15 receptors present on leukocytes. This agent presumably binds to circulating neutrophils that eventually migrate to the focus of infection, as well as to neutrophils or neutrophil debris containing CD15 receptors, already sequestered in the area of infection. At present, Tc-99m fanolesomab is approved for only diagnosis of equivocal appendicitis in patients older than 5 years.

Pediatric Nuclear Medicine

The nuclear scanning of children is not entirely equivalent in adults. The growth process and radiopharmaceutical distribution are different in children. Moreover, high-resolution images are necessary because an organ in a child may be just one fourth the size of an adult's yet have the same number of receptors attaching to the radiopharmaceutical. Children also suffer from different types of cancers with different disease patterns and are more likely to have more aggressive tumors. For this reason, most nuclear physicians have replaced localized bone scanning with whole-body images in children [104].

Nuclear medicine is only conservatively used in pediatric cases. It is used in detection of primary and secondary malignancies, pyrexia of unknown origin (PUO), cause of urinary retention, bone tumors, infection, trauma, GI bleeding, etc. ^{99m}Tc phosphate is the common tracer used for bone scans. FDG-PET is used in some common pediatric malignancies like brain neoplasms and lymphomas and in certain less common malignancies like neuroblastoma (even in cases that are not metaiodobenzylguanidine (MIBG) avid detected by SPECT), bone and soft tissue sarcoma, Wilm's tumor, and hepatoblastoma [105].

Future Trends and Outlook

Nuclear medicine is one of the fastest growing fields of medical science. The current progress in nuclear medicine is creating fresh directions in imaging of the body at the molecular level providing greater information about the cellular pathology, injurious signal, and its downstream pathway in the process of cell death including apoptosis and defects in metabolic pathways. Insight into the molecular nature of cancer and other disease processes provides a new dimension of target-oriented therapeutic options. Molecular imaging will leverage the future clinical trials involving molecular therapies. It may provide guidance for several treatment strategies that are targeted at the molecular level such as immune guided therapy, hormonal therapy, cancer chemotherapy, and gene therapy. In order to take nuclear medicine to that level, a synchronous advancement in synthetic chemistry, signal acquisition, and data processing technique and hardware are the prerequisites.

Currently, biopsy is considered the gold standard for establishing the diagnosis of several diseases, but biopsy findings are based on samples collected from small portion of a tumor at a single point in space and time. But the targets are known to change dynamically over time and space. Hence, biopsy-based clinical practice may not address the exact in vivo molecular abnormality that leads to the disease or tumor. Molecular imaging technology is seen as a promising tool in bridging this gap in knowledge.

Advancement in synthetic chemistry with newer radiopharmaceutical tracers having high specificity for the receptors or antigens in the diseased or malignant tissues is required for target identification for molecular therapies, tumor profiling, therapeutic selection, monitoring of early treatment response and monitoring of disease recurrence. In fact, the future trend of molecular medicine will be towards this direction extending far beyond what the currently available tracers are capable of. Most of the contemporary agents rely on processes that are fairly generic to malignant transformation and not necessarily specific to any particular cancer [10]. For instance, ^{18}F-FDG is taken up by almost all cells of the body that utilize glucose making it highly nonspecific but with high sensitivity. Newer agents like fluoroestradiol (FES) and fluorodihydroxytestosterone (FDHT) are specific ligands for estrogen receptor (ER) and androgen receptor (AR), respectively. FES is used to image the ER + breast carcinoma that influence the treatment modality as 30–77 % of ER + patients show a favorable response to hormonal therapy at far less morbidity than alternative chemotherapy [106]. An array of radiopharmaceutical tracers targeting various mechanisms specific for malignancy and diseases is being evaluated continuously, a discussion on which is beyond the scope of this book.

Molecular imaging provides an effective monitoring and early feedback of therapeutic effect on the tumor tissue. The early molecular response to therapy which could be identified by PET within weeks could influence the treatment plan, should the therapeutic agent used fails, which could be replaced with a more effective therapeutic agent. Conventional imaging may take months to provide feedback on therapeutic response based on morphologic information and may even lead to selection of cells resistant to chemotherapy. Several targeted therapy with specific ligands tagged with

radionuclide could be monitored for its therapeutic efficiency. Molecular imaging serves as a promising modality in monitoring the treatment efficacy and follow-up.

The hybrid PET/CT complimenting each other's deficiency offers excellent information about the function with good spatial and temporal resolution. But MRI provides greater soft tissue contrast and is the modality of choice for soft tissue imaging. Hybrid PET/MR was thus conceptualized but is currently not in routine clinical practice.

The combination of PET functional imaging and MRI structural imaging with excellent soft tissue contrast is the most yearned modality that is awaited to be added in the patient care diagnostic imaging chain. The PET/MR system might offer simultaneous information on anatomy, functionality, and biochemistry of the tissues and cells.

The scenario is optimistic and, PET/MR hybrid technology has the potential to become the imaging modality of choice for neurological studies, vast range of oncology studies, and may also have a great role in stem cell therapy. With this elaborate discussion on various aspects of nuclear medicine, we conclude that optimal diagnosis using radionuclides requires careful consideration of the patient, indications for the study, and the imaging modalities. Nuclear medicine imaging techniques and radionuclide imaging plays a pivotal role in the diagnosis of typical systemic diseases and will continue to do so in future, but still more research work is required to achieve its place in medical practice.

References

1. Brownell GL, Burnham CA, Chesler DA. Lateral and transverse section imaging with the MGH positron camera. Prog Nucl Med. 1978;4:158–64.
2. Phelps ME, Hoffman EJ, Mullani NA, Ter-Pogossian MM. Application of annihilation coincidence detection to transaxial reconstruction tomography. J Nucl Med. 1975;16:210–24.
3. Giesel FL, Mehndiratta A, Locklin J, McAuliffe MJ, White S, Choyke PL, Knopp MV, Wood BJ, Haberkorn U, von Tengg-Kobligk H. Image fusion using CT, MRI and PET for treatment planning, navigation and follow up in percutaneous RFA. Exp Oncol. 2009;31(2):106–14.
4. Gambhir SS, Czernin J, Schwimmer J, Silverman DH, Coleman RE, Phelps ME. A tabulated summary of the FDG PET literature. J Nucl Med. 2001;42:1S–93.
5. Gallagher BM, Ansari A, Atkins H, et al. Radiopharmaceuticals XXVII. 18F-labeled 2-deoxy-2-fluoro-d-glucose as a radiopharmaceutical for measuring regional myocardial glucose metabolism in vivo: tissue distribution and imaging studies in animals. J Nucl Med. 1977; 18:990–6.
6. Reivich M, Kuhl D, Wolf A, et al. The [18F] fluorodeoxyglucose method for the measurement of local cerebral glucose utilization in man. Circ Res. 1979;44:127–37.
7. Czernin J, Phelps ME. Positron emission tomography scanning: current and future applications. Ann Rev Med. 2002;53:89–112.
8. DeGrado TR, Baldwin SW, Wang S, et al. Synthesis and evaluation of (18)F-labeled choline analogs as oncologic PET tracers. J Nucl Med. 2001;42:1805–14.
9. Bruehlmeier M, Roelcke U, Schubiger PA, Ametamey SM. Assessment of hypoxia and perfusion in human brain tumors using PET with 18F-fluoromisonidazole and 15O-H2O. J Nucl Med. 2004;45:1851–9.

10. Hicks RJ. Beyond FDG: novel PET tracers for cancer imaging. Cancer Imaging. 2004; 4:22–4.
11. Otte A, Dierckx RA. Good clinical practice: a plea for nuclear medicine. Nucl Med Commun. 2005;26:561.
12. Otte A, Maier-Lenz H, Dierckx RA. Good clinical practice: historical background and key aspects. Nucl Med Commun. 2005;26:563–74.
13. De Vos FJ, De DM, Dierckx RA. The good laboratory practice and good clinical practice requirements for the production of radiopharmaceuticals in clinical research. Nucl Med Commun. 2005;26:575–9.
14. Zanzonico P. Routine quality control of clinical nuclear medicine instrumentation: a brief review. J Nucl Med. 2008;49:1114–31.
15. Strauss LG, Conti PS. The applications of PET in clinical oncology. J Nucl Med. 1991; 32:623–48.
16. Bar-Shalom R, Valdivia AY, Blaufox MD. PET imaging in oncology. Semin Nucl Med. 2000;30:150–85.
17. Tochon-Danguy HJ, Sachinidis JI, Egan GF, et al. Positron emission tomography: radioisotope and radiopharmaceutical production. Australas Phys Eng Sci Med. 1999;22:136–44.
18. Scott AM. Current status of positron emission tomography in oncology. Australas Radiol. 2002;46:154–62.
19. Hicks RJ, Binns DS, Fawcett ME, et al. Positron emission tomography (PET): experience with a large-field-of-view three-dimensional PET scanner. Med J Aust. 1999;171:529–32.
20. Som P, Atkins HL, Bandoypadhyay D, et al. A fluorinated glucose analog, 2-fluoro-2-deoxy-D-glucose (F-18): nontoxic tracer for rapid tumor detection. J Nucl Med. 1980;21:670–5.
21. Reutens DC, Bittar RG, Tochon-Danguy H, Scott AM. Clinical applications of [(15)O] H(2) O PET activation studies. Clin Positron Imaging. 1999;2:145–52.
22. Scott AM, Larson SM. Tumor imaging and therapy. Radiol Clin North Am. 1993;31: 859–79.
23. Conti PS, Lilien DL, Hawley K, Keppler J, Grafton ST, Bading JR. PET and [18F]-FDG in oncology: a clinical update. Nucl Med Biol. 1996;23:717–35.
24. Valk PE. Randomized controlled trials are not appropriate for imaging technology evaluation. J Nucl Med. 2000;41:1125–6.
25. Berlangieri SU, Scott AM, Knight SR, et al. F-18 fluorodeoxyglucose positron emission tomography in the non-invasive staging of non-small cell lung cancer. Eur J Cardiothorac Surg. 1999;16 Suppl 1:S25–30.
26. Findlay M, Young H, Cunningham D, et al. Noninvasive monitoring of tumor metabolism using fluorodeoxyglucose and positron emission tomography in colorectal cancer liver metastases: correlation with tumor response to fluorouracil. J Clin Oncol. 1996;14:700–8.
27. Damian DL, Fulham MJ, Thompson E, Thompson JF. Positron emission tomography in the detection and management of metastatic melanoma. Melanoma Res. 1996;6:325–9.
28. Kiffer JD, Berlangieri SU, Scott AM, et al. The contribution of 18F-fluoro-2-deoxy-glucose positron emission tomographic imaging to radiotherapy planning in lung cancer. Lung Cancer. 1998;19:167–77.
29. Shon IH, O'doherty MJ, Maisey MN. Positron emission tomography in lung cancer. Semin Nucl Med. 2002;32:240–71.
30. Berlangieri SU, Scott AM. Metabolic staging of lung cancer. N Engl J Med. 2000;343:290–2.
31. Di CG, Brooks RA. PET-FDG of untreated and treated cerebral gliomas. J Nucl Med. 1988; 29:421–3.
32. Jeong HJ, Chung JK, Kim YK, et al. Usefulness of whole-body (18) F-FDG PET in patients with suspected metastatic brain tumors. J Nucl Med. 2002;43:1432–7.
33. Weber WA, Wester HJ, Grosu AL, et al. O-(2-[18F] fluoroethyl)-L-tyrosine and L-[methyl-11C] methionine uptake in brain tumors: initial results of a comparative study. Eur J Nucl Med. 2000;27:542–9.
34. Gupta N, Bradfield H. Role of positron emission tomography scanning in evaluating gastrointestinal neoplasms. Semin Nucl Med. 1996;26:65–73.

35. Lowe VJ, Fletcher JW, Gobar L, et al. Prospective investigation of positron emission tomography in lung nodules. J Clin Oncol. 1998;16:1075–84.
36. Kratochwil C, Haberkorn U, Giesel FL. PET/CT for diagnostics and therapy stratification of lung cancer. Radiologe. 2010;50(8):684–91. doi:10.1007/s00117-009-1960-6. Review. German.
37. Gould MK, Maclean CC, Kuschner WG, Rydzak CE, Owens DK. Accuracy of positron emission tomography for diagnosis of pulmonary nodules and mass lesions: a meta-analysis. JAMA. 2001;285:914–24.
38. Vansteenkiste JF, Stroobants SG, De Leyn PR, et al. Mediastinal lymph node staging with FDG-PET scan in patients with potentially operable non-small cell lung cancer: a prospective analysis of 50 cases. Leuven Lung Cancer Group. Chest. 1997;112:1480–6.
39. Valk PE, Pounds TR, Hopkins DM, et al. Staging non-small cell lung cancer by whole-body positron emission tomographic imaging. Ann Thorac Surg. 1995;60:1573–81.
40. Bury T, Dowlati A, Paulus P, et al. Whole-body 18FDG positron emission tomography in the staging of non-small cell lung cancer. Eur Respir J. 1997;10:2529–34.
41. Gambhir SS, Hoh CK, Phelps ME, Madar I, Maddahi J. Decision tree sensitivity analysis for cost-effectiveness of FDG-PET in the staging and management of non-small-cell lung carcinoma. J Nucl Med. 1996;37:1428–36.
42. Giesel F, Stefanova M, Schwartz LH, Afshar-Oromieh A, Eisenhut M, Haberkorn U, Kratochwil C. Impact of peptide receptor radionuclide therapy on the 68Ga-DOTATOC-PET/CT uptake in normal tissue. Q J Nucl Med Mol Imaging. 2013;57(2):171–6.
43. Giesel FL, Kratochwil C, Mehndiratta A, Wulfert S, Moltz JH, Zechmann CM, Kauczor HU, Haberkorn U, Ley S. Comparison of neuroendocrine tumor detection and characterization using DOTATOC-PET in correlation with contrast enhanced CT and delayed contrast enhanced MRI. Eur J Radiol. 2012;81(10):2820–5.
44. Yasuda S, Fujii H, Nakahara T, et al. 18F-FDG PET detection of colonic adenomas. J Nucl Med. 2001;42:989–92.
45. Arulampalam TH, Costa DC, Bomanji JB, Ell PJ. The clinical application of positron emission tomography to colorectal cancer management. Q J Nucl Med. 2001;45:215–30.
46. Schiepers C, Penninckx F, De VN, et al. Contribution of PET in the diagnosis of recurrent colorectal cancer: comparison with conventional imaging. Eur J Surg Oncol. 1995;21:517–22.
47. Heriot AG, Hicks RJ, Drummond EG, et al. Does positron emission tomography change management in primary rectal cancer? A prospective assessment. Dis Colon Rectum. 2004;47:451–8.
48. Naumann R, Beuthien-Baumann B, Reiss A, et al. Substantial impact of FDG PET imaging on the therapy decision in patients with early-stage Hodgkin's lymphoma. Br J Cancer. 2004;90:620–5.
49. Delbeke D, Martin WH, Morgan DS, et al. 2-Deoxy-2-[F-18]fluoro-D-glucose imaging with positron emission tomography for initial staging of Hodgkin's disease and lymphoma. Mol Imaging Biol. 2002;4:105–14.
50. Jerusalem G, Beguin Y, Najjar F, et al. Positron emission tomography (PET) with 18F-fluorodeoxyglucose (18F-FDG) for the staging of low-grade non-Hodgkin's lymphoma (NHL). Ann Oncol. 2001;12:825–30.
51. Schoder H, Meta J, Yap C, et al. Effect of whole-body (18) F-FDG PET imaging on clinical staging and management of patients with malignant lymphoma. J Nucl Med. 2001;42:1139–43.
52. Foo SS, Mitchell PL, Berlangieri SU, Smith CL, Scott AM. Positron emission tomography scanning in the assessment of patients with lymphoma. Intern Med J. 2004;34:388–97.
53. Spaepen K, Stroobants S, Dupont P, et al. Prognostic value of positron emission tomography (PET) with fluorine-18 fluorodeoxyglucose ([18F] FDG) after first-line chemotherapy in non-Hodgkin's lymphoma: is [18F] FDG-PET a valid alternative to conventional diagnostic methods? J Clin Oncol. 2001;19:414–9.
54. Mikhaeel NG, Timothy AR, Hain SF, O'doherty MJ. 18-FDG-PET for the assessment of residual masses on CT following treatment of lymphomas. Ann Oncol. 2000;11 Suppl 1:147–50.
55. Balch CM, Soong SJ, Gershenwald JE, et al. Prognostic factors analysis of 17,600 melanoma patients: validation of the American Joint Committee on Cancer melanoma staging system. J Clin Oncol. 2001;19:3622–34.

56. Mijnhout GS, Hoekstra OS, van Tulder MW, Teule GJ, Deville WL. Systematic review of the diagnostic accuracy of (18) F-fluorodeoxyglucose positron emission tomography in melanoma patients. Cancer. 2001;91:1530–42.
57. Schwimmer J, Essner R, Patel A, et al. A review of the literature for whole-body FDG PET in the management of patients with melanoma. Q J Nucl Med. 2000;44:153–67.
58. Adams S, Baum RP, Stuckensen T, Bitter K, Hor G. Prospective comparison of 18F-FDG PET with conventional imaging modalities (CT, MRI, US) in lymph node staging of head and neck cancer. Eur J Nucl Med. 1998;25:1255–60.
59. Hannah A, Scott AM, Tochon-Danguy H, et al. Evaluation of 18 F-fluorodeoxyglucose positron emission tomography and computed tomography with histopathologic correlation in the initial staging of head and neck cancer. Ann Surg. 2002;236:208–17.
60. Paulus P, Sambon A, Vivegnis D, et al. 18FDG-PET for the assessment of primary head and neck tumors: clinical, computed tomography, and histopathological correlation in 38 patients. Laryngoscope. 1998;108:1578–83.
61. Keyes Jr JW, Watson Jr NE, Williams III DW, Greven KM, McGuirt WF. FDG PET in head and neck cancer. AJR Am J Roentgenol. 1997;169:1663–9.
62. Greven KM, Keyes Jr JW, Williams III DW, McGuirt WF, Joyce III WT. Occult primary tumors of the head and neck: lack of benefit from positron emission tomography imaging with 2-[F-18]fluoro-2-deoxy-D-glucose. Cancer. 1999;86:114–8.
63. Farber LA, Benard F, Machtay M, et al. Detection of recurrent head and neck squamous cell carcinomas after radiation therapy with 2-18F-fluoro-2-deoxy-D-glucose positron emission tomography. Laryngoscope. 1999;109:970–5.
64. Rege S, Safa AA, Chaiken L, Hoh C, Juillard G, Withers HR. Positron emission tomography: an independent indicator of radiocurability in head and neck carcinomas. Am J Clin Oncol. 2000;23:164–9.
65. Wahl RL, Siegel BA, Coleman RE, Gatsonis CG. Prospective multicenter study of axillary nodal staging by positron emission tomography in breast cancer: a report of the staging breast cancer with PET Study Group. J Clin Oncol. 2004;22:277–85.
66. Moon DH, Maddahi J, Silverman DH, Glaspy JA, Phelps ME, Hoh CK. Accuracy of whole-body fluorine-18-FDG PET for the detection of recurrent or metastatic breast carcinoma. J Nucl Med. 1998;39:431–5.
67. Bender H, Kirst J, Palmedo H, et al. Value of 18fluoro-deoxyglucose positron emission tomography in the staging of recurrent breast carcinoma. Anticancer Res. 1997;17: 1687–92.
68. Luketich JD, Friedman DM, Weigel TL, et al. Evaluation of distant metastases in esophageal cancer: 100 consecutive positron emission tomography scans. Ann Thorac Surg. 1999;68: 1133–6.
69. Meltzer CC, Luketich JD, Friedman D, et al. Whole-body FDG positron emission tomographic imaging for staging esophageal cancer comparison with computed tomography. Clin Nucl Med. 2000;25:882–7.
70. Lerut T, Flamen P. Role of FDG-PET scan in staging of cancer of the esophagus and gastroesophageal junction. Minerva Chir. 2002;57:837–45.
71. Fukunaga T, Okazumi S, Koide Y, Isono K, Imazeki K. Evaluation of esophageal cancers using fluorine-18-fluorodeoxyglucose PET. J Nucl Med. 1998;39:1002–7.
72. Einhorn N, Nilsson B, Sjovall K. Factors influencing survival in carcinoma of the ovary. Study from a well-defined Swedish population. Cancer. 1985;55:2019–25.
73. Karlan BY, Hawkins R, Hoh C, et al. Whole-body positron emission tomography with 2-[18F]-fluoro-2-deoxy-D-glucose can detect recurrent ovarian carcinoma. Gynecol Oncol. 1993;51:175–81.
74. Hubner KF, McDonald TW, Niethammer JG, Smith GT, Gould HR, Buonocore E. Assessment of primary and metastatic ovarian cancer by positron emission tomography (PET) using 2-[18F]deoxyglucose (2-[18F]FDG). Gynecol Oncol. 1993;51:197–204.
75. Stephens AW, Gonin R, Hutchins GD, Einhorn LH. Positron emission tomography evaluation of residual radiographic abnormalities in postchemotherapy germ cell tumor patients. J Clin Oncol. 1996;14:1637–41.

76. Yeh S, Imbriaco M, Larson S, Garza D, Zhang JJ, Kalaigian H, Finn RD, Reddy D, Horowitz SM, Goldsmith SJ, Scher HI. Detection of bony metastases of androgen independent prostate cancer by FDG–PET. Nucl Med Biol. 1996;23:693–7.
77. Shreve PD, Grossmann HB, Gross MD, Wahl RL. Metastatic prostate cancer: initial findings of PET with 2-deoxy-2-[F-18]-fluoro-D-glucose. Radiology. 1996;199:751–6.
78. Shvarts O, Han KR, Seltzer M, Pantuck AJ, Belldegrun AS. Positron emission tomography in urologic oncology. Cancer Control. 2002;9:335–42.
79. Effert PJ, Bares R, Handt S, Wolff JM, Bull D, Jakes G. Metabolic imaging of untreated prostate cancer by positron emission tomography with 18-fluorine-labeled deoxyglucose. J Urol. 1996;155:994–8.
80. Hara T. ^{18}F–fluorocholine: a new oncologic PET tracer. J Nucl Med. 2001;42:1815–6.
81. DeGrado TR, Coleman RE, Wang S, Baldwin SW, Orr MD, Robertson CN, Polascik TJ, Price DT. Synthesis and evaluation of ^{18}F–labeled choline as an oncologic tracer for positron emission tomography: initial findings in prostate cancer. Cancer Res. 2000;61:110–7.
82. Beheshti M, Imamovic L, Broinger G, Vali R, Waldenberger P, Stoiber F, Nader M, Gruy B, Janetschek G, Langsteger W. ^{18}F choline PET/CT in the preoperative staging of prostate cancer in patients with intermediate or high risk of extracapsular disease: a prospective study of 130 patients. Radiology. 2010;254:925–33.
83. Oyama N, Miller TR, Dehdashti F, Siegel BA, Fischer KC, Michalski JM, Kibel AS, Andriole GL, Picus J, Welch MJ. ^{11}C–acetate PET imaging of prostate cancer: detection of recurrent disease at PSA relapse. J Nucl Med. 2003;44:549–55.
84. Kato T, Tsukamoto E, Kuge Y, Takei T, Shiga T, Shinohara N, Katoh C, Nakada K, Tamaki N. Accumulation of [^{11}C]acetate in normal prostate and benign prostatic hyperplasia: comparison with prostate cancer. Eur J Nucl Med. 2002;29:1492–5.
85. Afshar-Oromieh A, Haberkorn U, Eder M, Eisenhut M, Zechmann CM. [68Ga]Gallium-labelled PSMA ligand as superior PET tracer for the diagnosis of prostate cancer: comparison with 18F-FECH. Eur J Nucl Med Mol Imaging. 2012;39(6):1085–6.
86. Nunez R, Macapinlac HA, Yeung HWD, Akhurst T, Cai S, Osman I, Gonen M, Riedel E, Scher HI, Larson SM. Combined ^{18}F–FDG and ^{11}C–methionine PET scans in patients with newly progressive metastatic prostate cancer. J Nucl Med. 2002;43:46–55.
87. Hoffman JM, Welsh-Bohmer KA, Hanson M, et al. FDG PET imaging in patients with pathologically verified dementia. J Nucl Med. 2000;41:1920–8.
88. Andersson JD, Varnas K, Cselenyi Z, et al. Radiosynthesis of the candidate beta-amyloid radioligand [(11)C]AZD2184: positron emission tomography examination and metabolite analysis in cynomolgus monkeys. Synapse. 2010;64:733–41.
89. O'Brien JT, Eagger S, Syed GM, Sahakian BJ, Levy R. A study of regional cerebral blood flow and cognitive performance in Alzheimer's disease. J Neurol Neurosurg Psychiatry. 1992; 55:1182–7.
90. Jagust W, Thisted R, Devous Sr MD, et al. SPECT perfusion imaging in the diagnosis of Alzheimer's disease: a clinical-pathologic study. Neurology. 2001;56:950–6.
91. Schmidt D, Zimmermann R, Lewczuk P, et al. Confirmation rate of blinded (99m)Tc-SPECT compared to neurochemical dementia biomarkers in CSF in patients with Alzheimer disease. J Neural Transm. 2010;117(9):1111–4.
92. Devous Sr MD, Thisted RA, Morgan GF, Leroy RF, Rowe CC. SPECT brain imaging in epilepsy: a meta-analysis. J Nucl Med. 1998;39:285–93.
93. Asenbaum S, Baumgartner C. Nuclear medicine in the preoperative evaluation of epilepsy. Nucl Med Commun. 2001;22:835–40.
94. Wadsak W, Mitterhauser M. Basics and principles of radiopharmaceuticals for PET/CT. Eur J Radiol. 2010;73:461–9.
95. Tu Z, Mach RH. C-11 radiochemistry in cancer imaging applications. Curr Top Med Chem. 2010;10:1060–95.
96. Schlyer DJ. PET tracers and radiochemistry. Ann Acad Med Singapore. 2004;33:146–54.
97. Rushton HG, Majd M. Dimercaptosuccinic acid renal scintigraphy for the evaluation of pyelonephritis and scarring: a review of experimental and clinical studies. J Urol. 1992;148:1726–32.

98. Fommei E, Volterrani D. Renal nuclear medicine. Semin Nucl Med. 1995;25:183–94.
99. Palestro CJ. The current role of gallium imaging in infection. Semin Nucl Med. 1994;24: 128–41.
100. Love C, Palestro CJ. Radionuclide imaging of infection. J Nucl Med Technol. 2004;32: 47–57.
101. Palestro CJ, Torres MA. Radionuclide imaging in orthopedic infections. Semin Nucl Med. 1997;27:334–45.
102. Love C, Opoku-Agyemang P, Tomas MB, Pugliese PV, Bhargava KK, Palestro CJ. Pulmonary activity on labeled leukocyte images: physiologic, pathologic, and imaging correlation. Radiographics. 2002;22:1385–93.
103. Palestro CJ, Torres MA. Radionuclide imaging of nonosseous infection. Q J Nucl Med. 1999; 43:46–60.
104. Kotz D. Pediatric nuclear medicine: special issues, unique clinical studies. J Nucl Med. 1998; 39:13N–4, 26N.
105. Jadvar H, Connolly LP, Fahey FH, Shulkin BL. PET and PET/CT in pediatric oncology. Semin Nucl Med. 2007;37:316–31.
106. Schellingerhout D, Gelovani J. Clinical trials in a molecular world. Neuroimaging Clin N Am. 2006;16:681–94, ix.
107. Mettler FA, Huda W, Yoshizumi TT, Mahesh M. Effective doses in radiology and diagnostic nuclear medicine: a catalog. Radiology. 2008;248(1):254–63.
108. Wall BF, Hart D. Revised radiation doses for typical x-ray examinations. Br J Radiol. 1997; 70(833):437–9.

Chapter 15
Contrast Agents in Radiology

Hendrik von Tengg-Kobligk, Amit Mehndiratta, and Frederik L. Giesel

Abstract To noninvasively diagnose disease and to describe response to therapy regarding morphology (e.g., size, structure) and (patho)physiology (e.g., blood perfusion) as well as cell function optimal image contrast is key. In our various body compartments in vivo contrast can be altered and improved by changing its intensity and distribution over time. This imaging fortune has been established by various contrast agents typically administered intravenously. Imaging methods in the field of ultrasound, magnetic resonance imaging, as well as computed tomography are continuously being improved by safe, valid, and efficient contrast agents. New targeted and specific agents are in the pipeline, but there are still a few more steps to go to reach market approval.

Keywords Different contrast media • Iodinated contrast media • MR contrast media • Ultrasound contrast media • Imaging • Imaging agents

---

H. von Tengg-Kobligk, MD, Dr. med., Priv-Doz (✉)
Institute for Diagnostic, Interventional, and Pediatric Radiology (DIPR),
Inselspital—University Hospital Bern, Freiburgstr. 10, Bern 3010, Switzerland

Department of Diagnostic and Interventional Radiology (DIR),
University Hospitals Heidelberg, Heidelberg 69120, Germany

Department of Radiology, German Cancer Research Center (DKFZ),
INF 280, Heidelberg 69120, Germany
e-mail: hendrik.vtk@web.de

A. Mehndiratta, MBBS, DPhil (Oxon)
Department of Engineering Science, Institute of Biomedical Engineering,
Old Road Campus Research Building, University of Oxford,
Hedington, Oxfordshire, OX3 7DQ, UK
e-mail: dramit.mehndiratta@gmail.com

F.L. Giesel, MD, MBA
Department of Nuclear Medicine, University Hospitals Heidelberg,
INF 400, Heidelberg 69120, Germany
e-mail: f.giesel@dkfz-heidelberg.de

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_15,

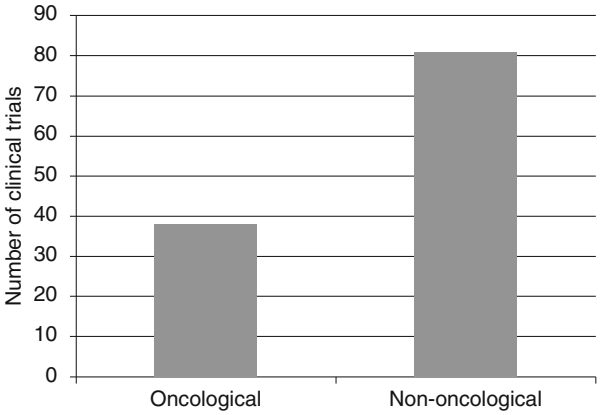


Fig. 15.1 Number of performed or ongoing oncological and non-oncological registered clinical trials from 01/2007 to 06/2010 using contrast media, (a) to investigate the CM itself or (b) to diagnose/monitor disease (Source: clinicaltrials.gov; search terms: contrast media)

Introduction

Contrast media development combined with improving equipment hardware of computed tomography (CT), magnetic resonance imaging (MRI), and ultrasound (US) systems helped to establish noninvasive imaging for therapy monitoring in the field of oncological and non-oncological clinical trials (Fig. 15.1). Currently, contrast media are not only being used to enhance static morphological structures but also to enable functional imaging with all these modalities, for instance, by using the methodologies such as dynamic contrast-enhanced (DCE) MRI [1–4] and contrast-enhanced ultrasound (CEUS). In addition to structure DCE, CEUS and perfusion CT patterns reflect the degree of vascularity, flow dynamics, and vascular perfusion. Contrast-enhanced techniques such as CT and MR angiography, perfusion imaging of macro- and microvasculature require efficient contrast media, and adequate knowledge about their individual characteristics is crucial [5].

At www.clinicaltrials.gov, analysis of registered trials shows the clinical importance of contrast-enhanced imaging with MRI being the most frequently used methodology in comparison to CT and US (Fig. 15.2). The number of registered clinical contrast media trials is highest in North America between 2007 and 2010 (Fig. 15.3).

This chapter briefly presents different classes of MR, CT, and US contrast agents available on the market, some with limited labeled indication and therefore still involved in clinical trials to achieve further labeled indications. Mechanisms of action and functionality of various contrast agents will also be reviewed in some detail. The described advantages and drawbacks of different contrast agents allow a better insight into the clinically approved indications and off-label usage. The chapter does not address safety issues of the various contrast media. Finally, we will outline some future scenarios of selected contrast agents in the field of radiology.

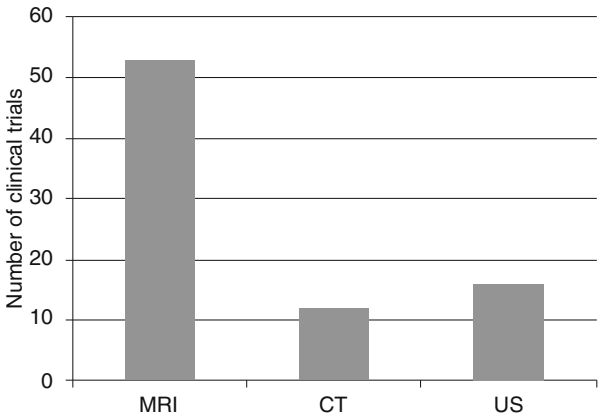


Fig. 15.2 Number of registered clinical trials using contrast media, which are, e.g., compared with existing CM or with native scans in order to investigate the CM itself (Source: clinicaltrials.gov; search terms: contrast media; from 01/2007 to 06/2010)

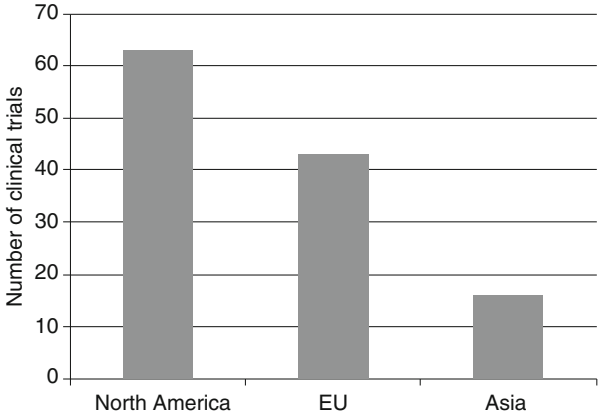


Fig. 15.3 Number of registered clinical trials from 01/2007 to 06/2010 using contrast media sort by region; source: clinicaltrials.gov; search terms: contrast media; (a) to investigate the CM itself or (b) to diagnose/monitor disease (Source: clinicaltrials.gov)

Groups of Contrast Agents

Contrast agents are used to enhance image contrast between pathological or anatomical structures of interest and their surrounding tissue or liquid. Intravenously (IV)-injected CM behave differently. First-generation MR contrast agents distribute within the intravascular space and due to its small molecule size diffuse into the interstitial space of extracellular fluid (ECF) and according to the law of equilibrium back into the intravascular compartment. These are often called "unspecific agents" and allow the evaluation of physiological parameters, such as the status or existence of the blood/brain barrier or the renal function. For this class of agents (ECF compounds), there is no intracellular uptake described. CM with intracellular uptake provide tissue-specific characteristics, e.g., they are taken up by lymphatic cells or hepatocytes, e.g., Gd-EOB-DTPA. Some MR contrast media have characteristics of two classes, e.g., the majority of the molecules take an extracellular path

being excreted via the kidney and the minor amount presents an intracellular pathway, e.g., liver-specific metallochelates are partially excreted via the biliary system [6–8].

Iodinated Contrast Media

Negative contrast agents are those which yield a lower X-ray attenuation than the body tissue and hence appear hypodense, e.g., gases such as air or CO_2. Positive contrast agents yield higher attenuation in comparison to adjacent structures. The positive contrast agents can be further divided into different groups. Iodine and barium are heavy elements and are comparatively less harmful to body giving excellent contrast. A positive IV or IA contrast agent must have a high concentration of iodine atoms with a comparable osmolality to blood. Since their introduction in the 1950s, organic radiographic iodinated contrast media (ICM) have been among the most commonly prescribed drugs in the history of modern medicine. The phenomenon of present-day radiologic imaging would be lacking without these agents.

Over the years quite a number of different iodinated contrast agents have been developed and are distributed by different manufactures. All agents consist of iodinated benzene ring derivatives. All ionic agents are typically formulated as sodium and/or meglumine salts and can be classified as high-osmolar contrast media (HOCMs, "ionics") and low-osmolar contrast media (LOCMs, "nonionics" and "ionics") (Appendix 15.1 at the end of this chapter).

ICM generally have a good safety record. In this context, nonionic ICM are generally safer than ionic CM, with less idiosyncratic (non-dose-dependent, e.g., allergy-like) adverse reactions [9] and are better tolerated if extravasation occurs; therefore, they are used almost exclusively today in clinical routine and trials.

Osmolality

Normal human reference range of osmolality in plasma is about 280–300 milliosmoles per kilogram. The ionic iodinated contrast agents contain three iodine atoms (in a monomer unit) and are having higher osmolality than blood. This means that each ionic monomer will dissociate into two particles due their ionicity, so the ratio is 2:3. These agents belong to the group of high-osmolality contrast medium, short HOCM. Then there are ionic dimers containing 6 iodine atoms (ratio 1:3) and having less osmolality than HOCM. Nonionic iodinated monomers (ratio 1:3) are also of less osmolar than ionic because of fewer particle number in water, i.e., nonionic agents do not require an accompanying cation and therefore have lower osmolality and belong to the group of low-osmolar CM, short LOCM [10]. The osmolality depends also on the iodine concentration, which typically ranges from 300 mg I/mL (=670 mOsm/kg H_2O) to 370 mg I/mL (=800 mOsm/kg H_2O)

(Appendix 15.1). In the end there are nonionic dimers containing 6 iodine atoms each molecule (ratio 1:6). An iso-osmolar contrast medium (IOCM) is considered to have least toxicity [9].

Iodine Content

For nonionic agents, the iodine (I) content is easy to determine because it is written on the label. For example, Omnipaque® 300 (GE Healthcare, Princeton, NJ, USA) contains 300 mg I/mL of solution. Vascular enhancement is proportional to the number of iodine molecules administered per unit of time [11]. Besides increasing the injection flow rate, the iodine concentration of the ICM can be adapted to the clinical need. If an iodine administration rate of, e.g., 1.5 g/s is desired, the injection rate needs to be 5 mL/s with a "standard" concentration (300 mg I/mL) compared with an injection rate of 4.3 mL/s using a higher iodinated compound (350 mg I/mL). The highest available concentration today is 400 mg I/mL reducing the flow rate to 3.8 mL/s in the given example. Due to their higher viscosity (Appendix 15.1), adequate cannula size and warming up the ICM to body temperature are highly recommended [12]. The maximum iodine concentration currently available in the USA is 370 mg I/mL.

There are studies showing advantages for higher concentrated ICM in vascular imaging [13]; other studies show that higher concentrated ICM may not be superior depending on the imaging protocol and the clinical indication and patient risk profile [14].

Meglumine Versus Sodium Salts

All HOCMs are ionic. They are organic acids consisting of an anion (radiodense iodinated benzoic acid derivative) and cation (sodium or meglumine). Sodium salts result in better renal opacification than meglumine salts. Therefore, e.g., sodium salts are used in diatrizoate meglumine (Urovist®, Bayer HealthCare, Montville, NJ, USA).

Nonionic LOCMs

Nonionic monomers (LOCMs) are currently the IV contrast media of choice. They have a lower incidence of adverse reactions (by a factor of 6 for all reactions and a factor of 9 for severe reactions) due to their lower osmolalities and potentially less chemotoxic than the ionic monomers [9]. Nonionic LOCMs are equally effective as contrast agents compared to HOCMs, but they are much higher in cost. The main

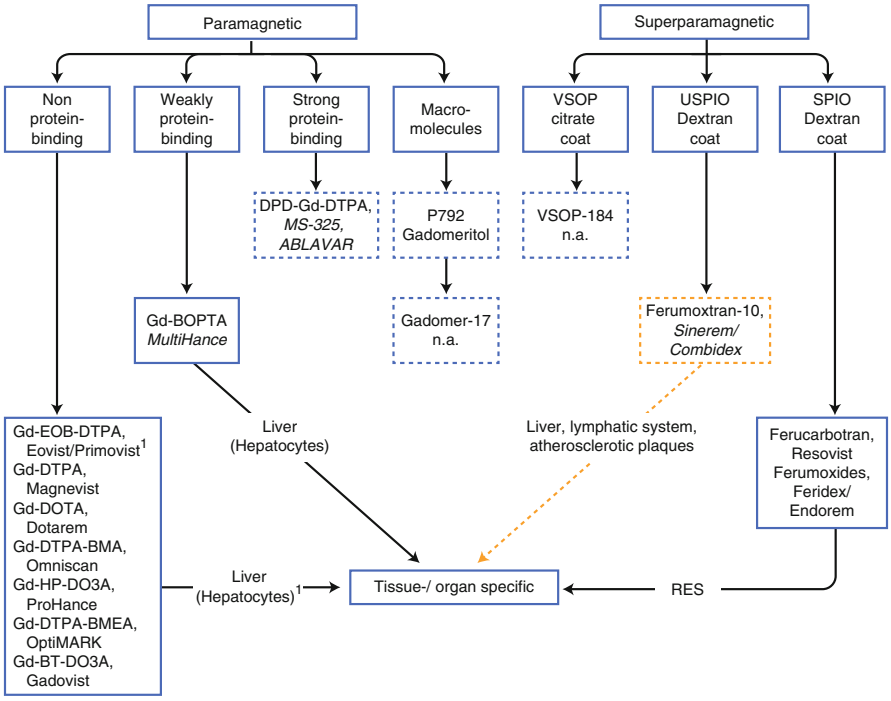


Fig. 15.4 Approved MR contrast media (CM) for intravenous administration. Development/marketing of Sinerem™/Combidex™ and Feridex® are discontinued; macromolecules and very small iron oxide particles are/were only in preclinical and clinical trials (phase II): dotted framed boxes. Most of small-molecular CM without any protein-binding capacity are nonspecific except Gd-EOB-DTPA. CM with prolonged intravascular half-life are framed in blue dotted-line boxes. SPIO/USPIO = small and ultrasmall particles of iron oxides. *RES* reticuloendothelial system

nonionic agents in the market are listed in Appendix 15.1. Common nonionic monomers are iohexol (Omnipaque®; GE Healthcare, Princeton, NJ, USA), iomeprol (Imeron®; Bracco, Princeton, NJ, USA), ioversol (Optiray®; Covidien, Mansfield, MA, USA), and iopromide (Ultravist®; Bayer HealthCare, Montville, NJ, USA) [15].

MR Contrast Media

There are four types of magnetic properties which MR contrast media fall into: paramagnetism, which is true for Gd and Mn agents; superparamagnetism produced by iron oxides; diamagnetism, which is generally not used for IV contrast agents; and ferromagnetism, also not used as IV contrast agents [16] (Fig. 15.4). Unlike CT

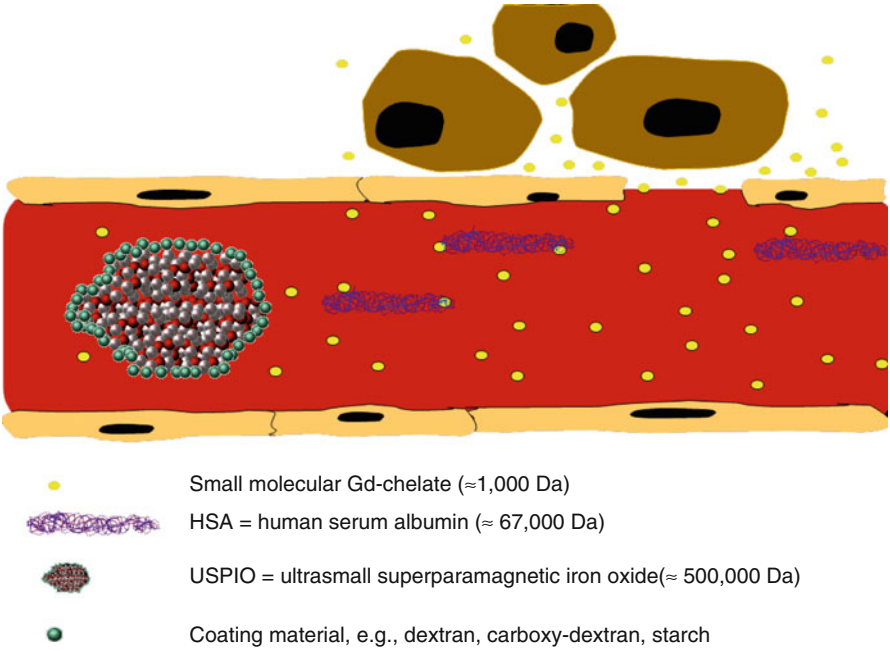


Fig. 15.5 Schematic healthy human capillary with pore size of 6–8 nm in diameter shows extravasation of extracellular fluid (ECF) agents like Gd-DTPA, whereas Gd-based molecules attached to HSA (e.g., Gd-BOPTA) remain intravascular. Also USPIOs (e.g., ferumoxtran-10) stay initially intravascular but leave the compartment by nonspecific vesicular transport or transendothelial channels

contrast media, the effect of MRI contrast agents is indirectly visualized via changes in nearby tissue proton behavior or magnetic susceptibility.

Contrast media reduce the T1, T2, and T2* relaxation times of the tissue surrounding the protons, thus producing a contrast enhancement compared to the surrounding tissue. MR contrast agents have different potencies in shortening relaxation times, which is expressed as relaxivity. According to the T1 and T2 relaxation times, there are two relaxivity constants, $r1$ and $r2$. The ratio between $r1$ and $r2$ determines if the contrast agent is more suitable for T1 or T2. A CM with a positive $r1/r2$ ratio is more suitable for T1-weithted imaging and with a positive $r2/r1$ ratio more effective for T2-weithted imaging [17]. In general, T1 agents lead to a signal enhancement and T2 to a signal loss.

Relaxivities of the classical extracellular fluid (ECF) agents (all ≤1,000 Da, e.g., pioneer Gd-DTPA) show rather little field dependency, whereas the slow tumbling compounds, e.g., MS-325 (non-covalent bound to plasma albumin) and gadomer or P972 (both macromolecules), as well as the iron oxide particles show strong field dependence [18] (Appendix 15.2 at the end of this chapter). Contrast media like Gd-BOPTA with an albumin binding of about 10 % present relaxivities intermediate between albumin-bound MS-325 and ECF agents. Electronic

and nuclear relaxation are described in more detail in various books and reviews [16, 18].

According to their magnetic behavior, approved MR contrast media can be divided into paramagnetic and superparamagnetic agents (Fig. 15.4). They are also classified according to their biodistribution: ECF, organ, or tissue specific as well as blood pool (intravascular). Targeting the liver cells, such as the hepatocytes and macrophages of the reticuloendothelial system (RES), is an effective approach for liver-specific agents (Fig. 15.5). Both the paramagnetic and the superparamagnetic agents target well-defined receptors or transporters that are uniquely expressed on the plasma membrane of specific liver cells [19], actually the first group of clinical available "molecular imaging agents."

Appendix 15.3 (at the end of this chapter) shows an overview of brand names with the physicochemical properties of the currently available MRCM with their application areas. For educational purposes non-approved contrast media gadomer-17, gadomeritol (both macromolecules), ferumoxtran-10, as well as VSOP are listed (Fig. 15.4). All marketed MR contrast media can be considered as safe drugs [20] following adequate patient selection with regard to NSF. The Gd-based agents are safe to use up to dose of 0.3 mmol/kg body weight.

Gadopentetate dimeglumine was the first MR contrast agent available but now several new agents with, e.g., higher relaxivity are available on the market with different imaging properties [21, 22]. The currently available agents are classified as shown in the next section of this chapter.

Type of Contrast Media

Paramagnetic Contrast Media

0.5 M contrast agents with paramagnetic properties were the first to be introduced for clinical use. The first MR contrast agent, gadopentetate dimeglumine (Gd-DTPA; Magnevist®, Bayer HealthCare, Montville, NJ, USA), entered clinical trials of MRI brain studies [23, 24] and was initially marketed in parts of Europe and Asia in 1998 and later on in the USA. Additional six Gd-based contrast agents (GBCA) were developed and are now routinely used in many countries. The ECF MRCM currently approved in Europe for the diagnosis of, e.g., CNS diseases and in regular use are Gd-DTPA, Gd-HP-DO3A (gadoteridol; ProHance®, Bracco, Princeton, NJ, USA), Gd-DTPA-BMA (gadodiamide; Omniscan®, GE Healthcare, Princeton, NJ, USA), Gd-DOTA (gadoterate meglumine; Dotarem®, Guerbet, Villepinte, France), Gd-BT-DO3A (gadobutrol; Gadovist®, Bayer HealthCare, Montville, NJ, USA), and Gd-DTPA-BMEA (gadoversetamide; OptiMARK®, Covidien, Mansfield, MA, USA). All of them are paramagnetic; i.e., they gain magnetic properties in a strong magnetic field reducing the T1 and T2 relaxation times of nearby water protons [17]. Using a T1-weighted MR sequence, these agents increase signal-to-noise ratio (SNR) of perfused tissue and improve contrast-to-noise ratio (CNR). They are

so-called positive enhancers. These Gd-chelates do not show any cellular uptake or any characteristics to bind serum proteins. Excretion is predominantly renal and 1 % or less via the hepatobiliary system [25]. Usually after intravenous bolus injection or infusion, GBCA do not cross the blood/brain barrier (BBB), but in the event of BBB disruption, increased extravasation of contrast medium into the CNS can occur quite rapidly [26].

The only paramagnetic non-Gd-based CM is Mn-DPDP (mangafodipir; Teslascan®, GE Healthcare, Princeton, NJ, USA) with a Mn^{2+} as a central moiety. It offers tissue-specific characteristics since peak liver enhancement occurs in ca. 15 min after injection and persists for several hours. Since manganese is a positive relaxation enhancer, a very small amount (5 μmol/kg) significantly enhances the contrast between healthy liver parenchyma and focal liver lesions. Due to safety reasons Mn-DPDP had to be slowely infused i.v., ie. high-contrasted first pass imaging was not labelled. Due to safety concerns Teslascan has been effectively withdrawn from the European and US market in 2012 (EMA/486286/2012).

Small-Molecular Contrast Media with Higher Molarity

The first Gd-chelates were all produced in the concentration of 0.5 M, which is considered as the standard. However, a compound with a higher concentration of 1.0 mol/L is now available in the market as gadobutrol (Gd-BT-DO3A; Gadovist®, Bayer HealthCare, Montville, NJ, USA) [27]. The size of this compound is comparable with that of conventional GBCA. Its in vitro relaxivity ($r1$) has been shown to be higher (approximately 20–25 % in plasma at 1.5 T) when compared with other non-protein-binding Gd-chelates [28]. The doubled concentration of this agent reduces the bolus volume by 50 %, which can be preferential for, e.g., neuroimaging depending upon the technique being used [29]. Imaging techniques such as T2* perfusion seem to benefit from this high-molar agent [30].

Small-Molecular Contrast Media with Weak Protein Affinity

Studies by Cavagna and colleagues [31] showed that Gd-BOPTA (gadobenate dimeglumine, MultiHance®, Bracco, Princeton, NJ, USA) leads to a stronger T1 shortening because of its weak affinity for human serum albumin (HSA) (Fig. 15.6) (Appendix 15.2). HSA (13 nm in diameter, 67,000 Da) serves as a macromolecule carrier for Gd-BOPTA demonstrating a stronger enhancing effect at equivalent dose and serum concentration to conventional extracellular contrast agents. This has been demonstrated in dedicated MR studies of the liver [32, 33], brain tumors [34], and the vascular system [35–37]. Compared to conventional ECF-CM, Gd-BOPTA allows a reduction in dose for vascular imaging that might be relevant for the discussion in the context of NSF [38]. Since it was proved in humans that approximately 0.6–6 % of the injected compound is excreted via hepatocytes into the hepatobiliary system [39], multiple studies on liver imaging using Gd-BOPTA

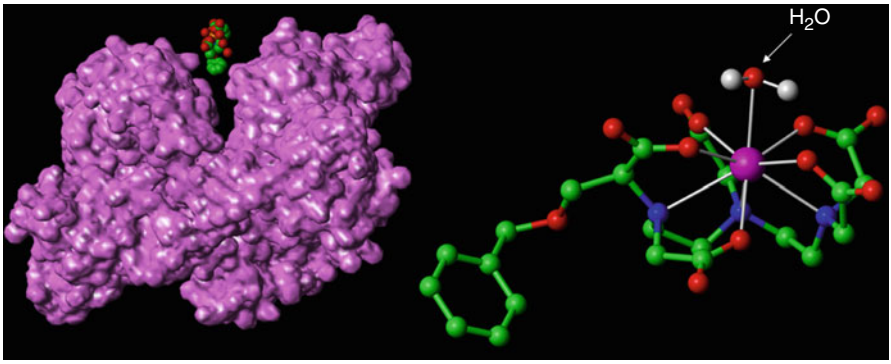


Fig. 15.6 Gadobenate dimeglumine [Gd-BOPTA]$^{2-}$ as a 3D model (*right*). The molecular structure is similar to that of Gd-DTPA with the exception of the extra benxiloxy-methyl group protruding from the molecule. This lipophilic structure leads to the reversible human serum albumin interaction, HSA (*left*), as well as its partial hepatobiliary excretion. The H_2O molecule (*white arrow*) binds temporarily to the contrast molecule. The serum albumin shows the surface structure which is accessible to water. Notice the different binding portions of HSA which are potentially available for certain contrast molecules. *Pink ball* = Gd^{3+}

have been published. Planchamp and colleagues showed that gadobenate dimeglumine accumulates in mildly cirrhotic rat liver, while, in severe disease, the expression of organic anion transporting peptides (OATPs) is so low that gadobenate dimeglumine behaves as an extracellular contrast [40]. Moreover, gadobenate dimeglumine remains trapped within hepatocytes in rats lacking the MRP2 transporter [41]. A dose of 50–100 μmol/kg of Gd-BOPTA produces significant liver enhancement in delayed imaging scans which peaks around 40–120 min after IV administration.

The protein-binding effect depends on the field strength [28] as well as on the amount of HSA, as proved in vitro by Giesel and colleagues [42]. The relaxivity of Gd-BOPTA measured in serum at 0.5 T is approximately twice that of currently available other gadolinium agents [43]. A docking study demonstrated potentially ten small hydrophobic pockets on the HSA surface where aromatic chains of gadobenate dimeglumine can attach for protein interaction and a strong non-covalent bonding can occur [42].

Recently, the small-molecular Gd-BOPTA has become available for contrast-enhanced CNS imaging. This agent was initially designed for advanced liver imaging, but later it has been proved to be useful for CNS imaging as well [34, 44].

Gadoxetic acid disodium (Gd-EOB-DTPA, Primovist®/Eovist®, Bayer HealthCare, Montville, NJ, USA) is now available in the European and US market and is proved to have a weak (non-covalent) HSA-binding capacity [7]. Primovist is marketed as 0.25 mol/L and labeled exclusively for liver imaging. Due to its lipophilic ethoxybenzyl (EOB) group added, about 40–50 % is taken up by hepatocytes and excreted via the hepatobiliary system [45], the rest via the renal system. In healthy livers, following extracellular distribution, gadoxetic acid enters into hepatocytes through the organic anion transporting peptides

(OATPs) located at the basolateral membrane and exits to bile through the canalicular transporter multidrug resistance associated protein 2 (MRP2). The two types of transporters are regulated differently [46]. Because of its hepatocellular accumulation and high relaxivity, a dose of only 10–25 mmol/kg was sufficient to selectively enhance the signal intensity of healthy liver parenchyma (starting at ~10 min postinjection) [19]. As described by Rohrer and colleagues, the effect of increased $r1$ due to HSA decreases with increasing field strength [28] (Appendix 15.2).

In summary, it needs to be emphasized that both agents, Gd-BOPTA and Gd-EOB-DTPA, show characteristics of an intracellular contrast agent due to uptake by hepatocytes. Therefore, they do not belong to the traditional group of "simple" ECF agents. Both agents belong to the group of liver-specific CM (LSCM) and are increasingly being used in clinical routine improving detection and characterization of liver masses during early dynamic perfusion phases and late hepatocyte phase in the cirrhotic and non-cirrhotic liver [47]. Gadoxetic acid-enhanced magnetic resonance (MR) imaging during the late hepatobiliary phase helps to characterize hepatocellular carcinomas (HCCs).

Small-Molecular Contrast Media with Stronger Protein Affinity

Recently MS-325 (gadophostriamine trisodium; ABLAVAR®, formerly Vasovist®, Lantheus Medical Imaging, N. Billerica, MA, USA) became available in selected markets globally. This compound is also a relatively small Gd-chelate, but has a much stronger affinity for HSA than Gd-BOPTA and Gd-EOB-DOTA [48]. The majority of the injected compound binds to HSA non-covalently [49]; depending on the administered dose and type of serum albumin (human or animal), the unbound fraction behaves as the a conventional ECF compound [25]. This "hermaphrodite" is defined as belonging to the group of intravascular contrast agents in spite of the fact that unbound fraction of this compound extravasates into the nonvascular space like any other small-molecular contrast agent. This strong binding effect can be explained by the so-called receptor-induced magnetization enhancement strategy [25]. The binding to HSA causes an increase in relaxivity of the molecule, an approximately ninefold increase at 20 MHz and four- to fivefold increase at 64 MHz (Appendix 15.2) [18, 28]. First published results in regard to this compound are for body contrast-enhanced MRA [50]. ABLAVAR® is currently only approved for peripheral vascular imaging (abdominal aorta and runoffs). Excretion of the compound is mainly via kidney and urinary system. Clinical studies need to prove how CNS imaging may benefit from its special pharmacokinetics and enhancement characteristics [51].

B-22956 is also an SMCM with a stronger affinity for HSA and hence exhibiting an increase T1 relaxivity (stronger binding with a bound fraction of HSA of approximately 94 %) as compared to MS-325 with an excretion pathway via the hepatobiliary (~45 %) and urinary system (~55 %) [52] but RND (phases II and III) has been discontinued after evaluation for coronary artery imaging.

Superparamagnetic Contrast Media

Historically, parental iron oxide particles have been used as contrast agents for a long time [53]. These colloid-based nanoparticles are produced with a core size of 50–180 nm as superparamagnetic or small particle of iron oxide (SPIO), with a core size of 10–50 nm as ultrasmall SPIO (USPIO) or even less than 10 nm as very small superparamagnetic iron oxide particle (VSOP) contrast agents. The core of the iron oxide colloids is composed of multiple 2–8 nm-sized iron oxide monocrystals and coated by, e.g., dextran or starch. Theses coated iron particles have a molecular weight of approx. 700,000 Da, similar to the endogenous ferritin with a diameter of approx. 11 nm (400,000–600,000 Da).

SPIOs present a high $r2$ during a short blood half-life time (within minutes). T2 relaxivities of >150 L/mmol/s are common with liver-specific SPIOs [19] (Appendix 15.2). Offering a stronger magnetic moment with a higher $r2$, they belong to the group of "negative enhancers" because in a stronger magnetic field, they predominantly cause strong susceptibility effects [54]. The apparent (hydrodynamic) size of the SPIO in blood is larger than the core because of hydration of the particle. The exact size of SPIO in blood is not known, probably in the order of 50–200 nm.

USPIOs present also a high $r1$ relaxivity offering a longer blood half-life (hours). Biodistribution and superparamagnetic effect varies strongly with size, electrical charge, and surface coating [55, 56].

In clinical routine there is not enough usage of this type of CM. Therefore, companies had to scale down or stop its production. However, in preclinical imaging of stem-cell tracking, many research groups are using (U)SPIO-based agents as described in more detail next.

SPIO

Depending on size and coating, SPIOs show specific uptake by the reticuloendothelial system (RES) of liver (~80 % of the injected dose) and spleen (6–10 % of the injected dose) within minutes after administration [57, 58]. Just 3 mg Fe/g liver will cause a signal decrease of 50 % in a standard T2-weighted sequence [19]. They have been available in the market as ferumoxides (Feridex®/Endorem®, Berlex, Guerbet), for example; however, they are only labeled for liver imaging [59] (Appendix 15.2). Ferumoxides were approved by the U.S. FDA in 1996 and is sold in the USA under the name Feridex®. It was originally developed as a liver contrast agent because it is taken up by Kupffer cells but did not live up to its promise and has been taken off the market because of lack of sales in November 2008. In Europe ferumoxides are still available under the name Endorem®. Ferucarbotran (Resovist®, Bayer HealthCare, Montville, NJ, USA) is the second available (Europe and Asia) SPIO, also coated with dextran with similar effects: using T2-weighted imaging techniques, a dose of ~8–15 μmol Fe/kg produced significant results in clinical studies.

The iron is handled by the usual metabolic pathways of the body after biodegradation of the compound. Once sequestered by phagocytic cells, the agent decreases liver and spleen signal intensity T2w (liver signal intensity is maximally decreased 0.5–6 h after administration and returns to normal within ~7 days). Malignant tumors are typically devoid of large number of phagocytic cells so that they appear as relatively hyperintense ("bright") lesions contrasted against hypointense ("black") background of normal liver parenchyma. Tumors with phagocytic elements (e.g., FNH) and hemangioma have been shown to also decrease in signal intensity [60].

USPIO

These agents have a blood half-life of over 20 h and thereafter accumulate in macrophages, e.g., lymph nodes, liver, and spleen (SPIOs, on the contrary, are too rapidly cleared by liver to be able to accumulate in lymph nodes). Because this group of agents has a significantly higher $r1$ compared to SPIO, it can also be used as T1-type blood-pool agents for imaging of tumor angiogenesis, atherosclerosis, and/or hepatic lesion characterization during the equilibrium phase [61, 62]. Perfused lesions (e.g., hemangioma) increase in signal intense on T12, whereas the same lesion decreases in signal intensity on T2w. Ferumoxtran-10 was in clinical trials (phase III) for lymph node imaging. The particles are phagocytosed by the macrophages of the lymphatic system [19].

The inhomogeneous distribution within nonmalignant lymph nodes increases their susceptibility effect (T2*) and leads to a reduction of the signal intensity. Unfortunately, it was not given market approval; therefore, no further clinical trials will be performed with support of the producer.

Ultrasmall and very small superparamagnetic particles of iron oxide have also been used for T1-weighted CE-MRA; however, for NC100150, clinical trials have been stopped [63, 64]. VSOP (very small superparamagnetic iron oxide particles) are a new class of contrast agents with smaller particle size than SPIO offering advantages for MR angiography. They have been tested in a phase I study [65]. SPIO particles are usually coated with an organic polymer such as dextran, carboxydextran, or polyethylene glycol, which limits the minimal overall particle size that can be obtained. VSOP-C184 consists of an aqueous solution of superparamagnetic iron oxide particles with a citrate coating with an overall particle size of only 4–8 nm.

In preclinical and clinical trials as well as in the clinical routine (as "off-label" use), USPIOs have been successfully also used for CNS imaging. In clinical studies, improved differentiation and characterization of the brain tumors using iron oxides have been shown as promising CNS applications [66].

Macromolecules

Macromolecular CM belong to the group of paramagnetic compounds. They are being tested worldwide (Weinmann, Corot, Choyke). Preclinical data show that

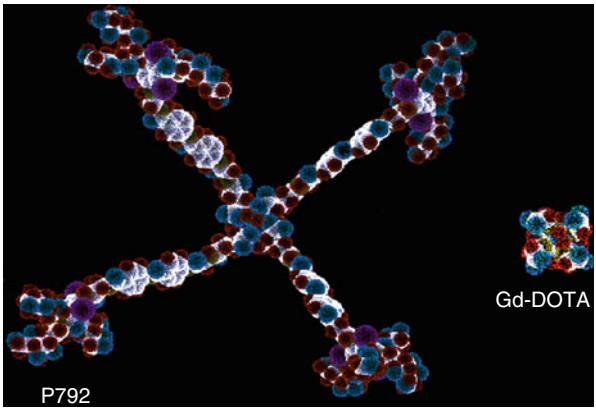


Fig. 15.7 Macromolecular compound P792 in comparison with Gd-DOTA. P792 itself has a macrocyclic Gd-DOTA molecule in its center. Both of them are paramagnetic contrast agents but with different diameters by molecular modelization, P792 = 29 nm and Gd-DOTA = 0.9 nm (Courtesy of Guerbet Research Group)

medium-sized polymeric gadolinium complexes demonstrate great promise for improving MR angiography and in quantificating capillary permeability and myocardial perfusion. One of the most advanced compounds is gadomer-17 which carries 24 gadolinium atoms with a numeric molecular weight of 17.5 kDa (Appendix 15.2 and Fig. 15.4). However, due to the large hydration radius of the molecule, the apparent molecular weight is 35 kDa. Gadomer-17 is small enough to guarantee glomerular filtration and large enough to ensure a slow diffusion through the bigger pores of the endothelial wall. The agent is eliminated by glomerular filtration within 24 h in an unmetabolized form [19].

P-792 (Vistarem®, Guerbet, Villepinte, France) is a Gd-based macromolecule, which is a hybrid between small agents and polymeric compounds (Fig. 15.7). The agent has a molecular weight of 5 kDa and exhibits some self-assembling characteristics, resulting in a larger molecular volume [67]. The concentration of the aqueous formulation is 35 mmol Gd/L, considerably lower than the other blood-pool agents, which have a 250 or 500 mmol/L strength and a very high relaxivity of 25 L/mmol/s at 1.5 T (Appendix 15.2). P-792 has a short half-life, but it distributes strictly in the vascular space during the first minutes after dosing [68]. The compound is cleared primarily by renal filtration and, to a minor extent, by hepatobiliary excretion.

Ventilation Agents

For ventilation perfluorinated agents, Gd-based aerosols as well as hyperpolarized gases (^{3}He and ^{129}He) can be used as well as oxygen [69].

Molecular or Cellular Imaging Agents

Magnetic resonance imaging (MRI) is a powerful imaging modality that can provide an assessment of function or molecular expression in tandem with anatomic

detail. Over the last 20–25 years, a number of gadolinium-based MR contrast agents have been developed to enhance signal by altering proton relaxation properties.

Gd-DTPA was the first approved MRI contrast agent and is the most widely used compound [70]. It is clinically safe and it seems logical to use it off label for clinical MRI cell tracking.

Unfortunately, there are some pitfalls with this famous compound; i.e., labeling cells with Gd-DTPA would result in long-term retention in the body with an uncertain clearing mechanism, i.e., it may not be renal. Although the pH of blood plasma is nearly neutral, sufficient stability of the chelate complex for its current clinical applications is challenged. The low pH in lysosomes and endosomes of cells may lead to rapid de-chelation once cells are labeled with a paramagnetic agent. There is potential toxicity concern about the existence of free Gd^{3+} ions leading to nephrogenic systemic fibrosis [71].

A further reason for not labeling cells with Gd-DTPA is that on the basis of the inherent physical principles of MR relaxation, intracellular Gd-DTPA has much reduced T1 relaxivity owing to differential water exchange and inner sphere relaxation [16]. Further challenges occur due to compartmentalization leading to local magnetic susceptibility effects that can produce ambiguous (e.g., positive vs. negative) contrast enhancement, particularly at higher field strengths.

SPIO and USPIO are a different class of MR contrast agent that leads to hypointense contrast enhancement of the cells of interest after proper labeling [72]. Ferumoxides are the only pharmaceutical-grade MR contrast agent that has been used for clinical cell tracking.

Many groups investigate which macromolecular structure might lead to an adequate MR-based labeling compound for clinical usage, which could be composed of albumin, polylysine, polysaccharides (dextran, inulin, starch), polyethylene glycol, copolymers of cystamine and cystine with GD-DTPA, and various dendritic structures based on polyamidoamine and polylysine (gadomers) [73, 74].

Gadofluorine M, Gf, is an amphiphilic Gd complex with a molecular weight of 1,528 g/mol and a concentration of 250 mmol Gd/L [75]. Gf binds to serum albumin at an affinity of kDa = 2 mmol/L. Moreover, Gf reveals a similar binding affinity to the extracellular matrix (ECM) components collagen, proteoglycan, and tenascin [68]. Unlike USPIOs, gadofluorines do not accumulate in the subendothelial monocyte-rich areas, but in deeper regions of the intima that are probably rich in foam cells and cellular debris. Porphyrins and other ring structures target atherosclerotic plaques too [19]. However, its localization of these compounds is not known.

About one-third of the dose is excreted by glomerular filtration, and two-thirds are excreted in the feces. Animal studies in VX2 tumor-bearing rabbits led by Misselwitz et al. showed that Gf accumulated in lymph nodes fascilitating the differentiation of metastastic tissue within the nodes. [75]. An in vitro and in vivo cell tracking study by Giesel and colleagues showed that Gf performs as a novel contrast agent with the capability of intracellular accumulation without an uptake mediator providing a T1-positive MRI signal at 1.5 T and may be suitable for cell tracking in animal models with intraparenchymal hemorrhages such as stroke or malignant tumors [76].

Gadofluorine uptake closely corresponded to inflammation and demyelination on tissue sections in a study performed by Bendszus and colleagues [68]. These unique features of gadofluorine M in visualizing inflammatory CNS lesions hold promise for future clinical development in multiple sclerosis.

Outlook

Four distinct future advantages of in vivo contrast agents can be defined: higher concentration, protein interaction, intracellular uptake in vivo, and intravascular characteristics although not exclusively found like ultrasound agents. Future MRI pulse sequence design will also need to take the potential benefits of the specific contrast agent into account for optimized results. The clinicians' dream of a perfect contrast media that accumulates highly and specifically in malignant tumors, allowing an accurate diagnosis at a stage when the disease is still treatable, is still possible. It is realistic to assume that new specific low-molecular peptides will be coupled to high-relaxivity moieties as a practical approach for moving molecular imaging into clinical MR reality [19].

The high number of ongoing preclinical and clinical trials will lead to the availability of further several different MR contrast agents in the future provided the excellent safety profile of current CM can be met and costs of extensive study designs can be compensated by successful business.

US Contrast Media

In the past, the potential of echo-enhanced US was underestimated. In addition, it took some years to overcome all development problems. All US contrast media have in common is that they are based on microbubbles restricted in their pharmacokinetics to the intravascular space (Appendix 15.4 at the end of this chapter).

Albunex® (Covidien, Mansfield, MA, USA) was the first agent that contained microbubbles coated with sonicated albumin with a mean size of 4 μm (95 % below 10 μm in diameter) with a short plasma half-life time of approx. 1 min [77]. Optison™ (GE Healthcare, Princeton, NJ, USA) (introduced by GE Healthcare in 1998) was then developed as a modification of Albunex, i.e., nearly the same microbubbles but instead of air the gas octafluoropropane was used [78]. The diameters are 3–4.5 μm (max. 32 μm; 95 % less than 10 μm). Echovist (Bayer HealthCare, Montville, NJ, USA) was developed based on galactose-coated microbubbles filled with air and, therefore, missing transpulmonary stability. Its successor Levovist (Bayer HealthCare, Montville, NJ, USA) (introduced by Schering in 1996) is coated in addition with palmitic acid as a surfactant offering better stability, echo increase is described up to 24 dB and duration of echo enhancement approx. 5 min [79]. SonoVue® (Bracco, Princeton, NJ, USA) (introduced by

Bracco in 2001) is a relatively new echo-contrast agent that is based on sulfur hexafluoride-filled microbubbles coated with phospholipids with a diameter of ca. 2.5 μm (90 % smaller than 8 μm) [80]. The stability lasts about 6 h after reconstitution, which allows a longer time window. SonoVue is currently being investigated in a number of clinical trials embracing the fields of cardiac as well as macro- and microvascular imaging in different organs and pathologies. It is already established in the diagnostic work-up of focal liver lesions. Perflutren protein type A microspheres (Definity®/Luminity® (Lantheus Medical Imaging, N. Billerica, MA, USA) introduced by Bristol-Myers Squibb in 2006) is based on lipid bilayers coating octafluoropropane with a size of 1.1–3.3 μm (98 % less than 10 μm) with a mean scan window is about 90s [81]. The shell of Imagent® (IMCOR Pharmaceutical, San Diego, CA, USA) (developed by Alliance Pharmaceutical Corporation and Schering AG; FDA approval 2002) is composed by surfactants coating perfluorohexane with a mean bubble diameter of 6 μm. Market approval was given for left ventricle echocardiography.

US Contrast Agents for Liver Imaging

Some US contrast agents have advantages in the liver-specific late phase where bubbles accumulate in normal liver parenchyma beginning approximately 2–5 min after injection when the vascular enhancement has faded, analogous to liver-specific contrast agents used in MRI. Microbubbles known to exhibit this behavior are Levovist and Sonazoid as well as, still under RND, Sonavist (GE Healthcare, Princeton, NJ, USA) and BR14 (Bracco, Princeton, NJ, USA) [82]. The mechanism of hepatic accumulation is not completely understood. Several experimental agents such as Sonazoid or BR14 combine the advantage of good enhancement at low MI with strong liver-specific properties. Uptake by the Kupffer cells of the liver has been demonstrated for Sonavist and Sonazoid. Sonazoid™ (Daiichi Sankyo, Parsippany, NJ, USA) (approved in Japan since 2006) is a lipid-stabilized suspension of perfluorocarbon microbubbles with a median diameter of 2.4–2.5 μm. The product is characterized by a prolonged ability to maintain its contrasting attributes with indication to improve differential diagnosis of hepatic lesions. In clinical trials, perflubutane-enhanced ultrasound was shown to have the same diagnostic ability as computed tomography for tumor detection and characterization of hepatic nodules.

Contrast Mechanism

Echo enhancement is managed by several factors. First of all, these gas bubbles, which are less than 10 μm in diameter, are able to enhance the echo strength (up to 25 dB; greater than 100-fold increase) by their compressibility under acoustic waves

leading to impedance between the gas bubbles and the surrounding tissue or fluid. Another factor is that these bubbles are able to oscillate, which leads to a better sound energy absorption and reradiation [83].

But the most important factor for echo enhancement is the so-called nonlinear oscillation of the microbubbles. In higher acoustic powers, the microbubbles will enlarge more than they are compressed thus leading to reradiation, not only of the frequency with which the bubbles were excited but also containing harmonics [83]. SonoVue provides strong and persistent signal enhancement due to its strong harmonic resonance at low ($\leq$0.2) and very low ($\leq$0.1) mechanical index (MI), where minimal or no bubble destruction occurs. Real-time low-MI imaging is a preferred method in many instances, although SonoVue has weaker liver-specific properties than Levovist. Levovist, however, requires high-MI imaging (MI $\geq$0.7). The disadvantage of high-MI imaging is the highly transient nature of the signals, which persist only for a few frames.

An increase of the acoustic power leads to a destruction of the bubbles. This effect can be used for diagnostic purposes but also for targeted drug therapy.

Outlook

Microbubbles used as blood-pool tracer have also great potential in targeted molecular imaging [84]. This is possible by attaching molecules to the bubbles that are targeted to specific receptor sites. Currently such microbubbles are investigated in the field of angiogenesis, atherosclerosis, and inflammation.

BR14™ is a new experimental ultrasound contrast agent, consisting of bubbles containing a high molecular weight filling gas enclosed by a flexible phospholipid monolayer shell a few nanometers thick. This agent shows significant nonlinear scattering and agent modification even at low insonation pressures, the detection pulses do not destroy the contrast bubbles. The results obtained with harmonic power Doppler (HPD) before the release burst show that the BR14 bubbles are efficient scatterers that can be modified and, thus, detected by low-power insonation. First application targets are myocardial and microcirculatory kinetics [85]. BR55™ (Bracco, Princeton, NJ, USA) is a lipopeptide-based VEGFR2-targeted ultrasound contrast agent for molecular imaging of angiogenesis [86].

Imagify™ (Acusphere, Lexington, MA, USA) is an ultrasound imaging agent comprised of perflubutane gas in a synthetic biodegradable microsphere whose primary components are already used in other approved pharmacologic products such as surgical sutures. Imagify is administered intravenously and is intended to allow visualization of perfusion deficits in the myocardium in addition to movement of the muscular walls of the heart

BiSphere™ (Point Biomedical, USA) is a technology for drug delivery applications by ultrasound. The flexibility in size control in the BiSphere™ technology has enabled the construction of submicron ultrasound contrast agents suitable for lymphatic imaging.

Approval Status of Contrast Agents and Their Off-Label Use

If the question "what is off-label drug use?" arises, the answer may be that "off-label" use is defined as the use of an approved drug in an unapproved manner or more specifically one indication not addressed in the package insert (drug labeling) [87]. However, unapproved use does not imply improper use, disapproved for use, or any contraindication to use but merely indicates a lack of formal approval. Contrast agents are considered drugs and as such are regulated, e.g., in the USA by FDA. Once a drug is available in market for use, clinicians may lawfully deviate from the circumstance for use or dose as mentioned in the package insert as long as it is not contraindicated [88].

Generally spoken approval of drugs for its particular clinical usage is a long-term process and expensive so that approval status is usually outdated. Thus it is not astonishing that many contrast agents are not used in their approved indication or dose as contained in the package insert but off label. For instance, administration of an MR contrast agent to perform MRA is considered "off-label" use in the USA. In this case the radiologist is ultimately responsible for selecting the contrast agent and dose for a particular indication.

Another fact is that approval of drug for clinics differs from region to region. In Europe many CM are approved, whereas in the USA they are not permitted, important reasons being the safety, reimbursements, and legal issues [89].

The most common MR contrast agent Magnevist® may illustrate an example. Off-label use of contrast media in MR is a common practice and in many instances is considered the standard of practice. Intra-articular injection, which is a common accepted practice in MR, is off-label use. Although there is little concern regarding clinical safety with this route of administration, the gadolinium chelates have only been evaluated in rigorous clinical trials for intravenous injection. Clinical uses of intra-articular injection include MR exams of the shoulder, wrist, hip, knee, and ankle [89]. Intrathecal usage is described by Siebner and colleagues [90]. Intrathecal administration is considered experimental in nature and by definition research. Intrathecal use should, however, be considered in a completely different category as possibility of major adverse reactions is high neurotoxic and neuropathologic effects of gadolinium chelates following intrathecal injection have been observed in preclinical studies [91].

The responsibilities of a physician increases when an off-label drug is prescribed. Appropriateness, safety and efficacy, scientific data, and relevant published literature are factors prescribers have to consider before use. A manufacturer is unlikely to be liable for injury resulting from the decision of the physician to use the product off label. A physician will be liable for negligence if he fails to take reasonable or proper care and a patient is injured as a result [89].

The patient has to be adequately informed about the potential benefit, risk, and an absence of authorized product information regarding the planned prescription. There is the possibility that off-label drug might not be reimbursed. In general, only medicine administered according to their label is reimbursed based upon current

knowledge of an expert group. Positive advice will not be given if an alternative medicine is available. If the manufacturer accepts liability for a new indication, the national health-care system may accept the indication and subsequently has to reimburse for the administration [92]. It is generally understood and very well accepted that medicines may be used off label in some clinical situations in the best interest of the patient when no alternative medicine is available.

Appendix 15.1: Overview of Available Nonionic, Low-osmolar, and Iso-osmolar Contrast Media and Their Physicochemical Properties

Generic name (mg I/mL)	Trade name(s)	Manufacturer	Labeled indications[a]	Type, ionicity	Osmolality (mOsm/kg H_2O)[b, c]	Viscosity at 37 °C (mPa×s)	Molecular weight (Da)
Iohexol (300/350)	Omnipaque Omnitrast Omnigraf	GE Healthcare Bayer HealthCare Juste	Myelography, cisternography, ventriculography (intrathecally), head, body, GI tract	LOCM nonionic (monomer)	672/844	6.3/10.4	821
Iopentol (300)	Imagopaque	GE Healthcare	Whole body, angiography, GI tract	LOCM nonionic (monomer)	640/818	6.5/12.0	835
Iopamidol (300/370)	Isovue Iopamiron Iopamiron Niopam Solutrast	Bracco	Head, whole body	LOCM nonionic (monomer)	616/796	4.7/9.4	777
Iopromide (300/370)	Ultravist	Bayer HealthCare	Whole body CT and angiography	LOCM nonionic (monomer)	607/774	4.9/10.0	791
Ioversol (300/350)	Optiray	Covidien	Head, body	LOCM nonionic (monomer)	651/792	5.5/9.0	807
Iobitridol (300/350)	Xenetix	Guerbet	Head, whole body	LOCM nonionic (monomer)	695/915	6.0/10.0	835
Ioxilan (300/350)	Oxilan	Guerbet	Head, body	LOCM nonionic (monomer)	585/695	5.1/8.1	791
Iomeprol (300/400)	Imeron	Bracco	Brain, body, cavernosography, myelography, urography	LOCM nonionic (monomer)	521/726	4.5/12.6	777
Iodixanol (320)	Visipaque	GE Healthcare	Head, body	IOCM nonionic (dimer)	290	11.8	1,550

IOCM iso-osmolar contrast media, *LOCM* low-osmolar contrast media

[a]Labeled indications may vary in different in countries/continents. Some CM are not even available in certain countries

[b]Osmolality values differ with measuring technique

[c]Values can be measured in a pure solution of the active substance in water or in the commercial product

Appendix 15.2: Physicochemical Properties of Tissue-specific and Nonspecific MR Contrast Agents

Lab name	Ionicity	Chemical structure	log *K*	log *K′*	Osmolality osmol/kg H_2O (37 °C)	Viscosity (mPa s) (37 °C)	WS *r*1	HP *r*1	HP *r*2	HP 3.0T[13y] *r*1/*r*2
[Gd-DTPA][2–]	2–	Linear	22.1[4]	17.7[4]	1.96[2]	2.9[2]	3.8[3q]	4.9[3z]	6.3[3z]	3.7/5.2
Gd-BT-DO3A	Neutral	Macrocyclic	21.8[1]	17.1[1]	1.60[2]	5.0[2]	3.6[2r]	5.6[8z]	6.5[8z]	n.a.
Gd-HP-DO3A	Neutral	Macrocyclic	23.8[4]	16.9[2]	0.63[4]	1.4[4]	3.7[7q]	4.6[7c]	5.3[7c]	3.7/5.7
[Gd-BOPTA][2–]	2–	Linear	22.6[3]	18.4[2]	1.97[2]	5.3[2]	4.4[2]	9.7[2z]	12.5[2z]	5.2/11.0
[Gd-DOTA][1–]	1–	Macrocyclic	25.4[5]	18.2[5]	1.35[10]	2.0[10]	3.5[7q]	4.3[7c]	5.0[7c]	3.5/4.9
Gd-DTPA-BMA	Neutral	Linear	16.9[4]	14.9[4]	0.65[4]	1.4[4]	3.9[7r]	4.6[7c]	5.1[7c]	4.0/5.6
Gd-DTPA-BMEA	Neutral	Linear	16.8[6]	15.0	1.10[6]	2.0[6]	4.6	3.8[c]	n.a.	n.a.
Gd-EOB-DTPA	2–	Linear	23.5[9]	18.7[11]	0.88[11]	1.2[11]	5.5[11m]	8.6	n.a.	6.2/4.0
[Mn-DPDP][2–]	2–	Linear	n.a.	n.a.	0.30[2]	0.8[2]	n.a.	2.3[12]	4.0[12]	n.a.
[Gadophostriamine trisodium][3–]	3–	Linear	22.4[11]	18.9[11]	0.83[11]	1.8[11]	6.6[11]	47.2[13]	57.6[13]	9.9/73
SHU555A	Neutral	Coated with carboxydextran	–	–	0.32[11]	1.03[11]	23.6[13]	15[13]	101[13]	3.3/160
AMI-25	Negat. charged	Coated with dextran	–	–	0.34[10]	1.3[10] (20 °C)	20.2[13]	24[14]	107[14]	2.7/45
AMI-227	Negat. charged	Coated with dextran	–	–	0.36[10]	1.3[10] (25 °C)	28[10]	23.0[14]	51.0[14]	n.a.
[P792][1–]	1–	Macrocyclic	–	–	0.30[10]	7.2[10]	42[10]	48[10]	n.a.	n.a.
Gadomer-17	Neutral	Polyamide	–	–	0.38[11]	7.0[11]	17.3[11]	18.7[11]	23[11]	13/25

Relaxivities *r*1 and *r*2 (mM^{-1}/s) were measured under the following conditions; note indicated exceptions: *WS* water solution, 0.47 T (20 MHz); pH 7.3; 0.15 M NaCl solution: [q]39 °C; [r]40 °C; [t]distilled water; *HP* human plasma, 0.47 T (20 MHz); [y]plasma, 37 °C, [z]heparinized, 39 °C; [e]40 °C; [f]1.0 T; [g]1.5 T, [k]in PBS (37 °C), [c]in human serum (Seronorm®) at 39 °C; [i]39 °C, bovine plasma. [j]Rheolab DSR 4000 (Paar Physica, LTD)

log_{term}: thermodynamic stability for each ion and each chelate; log_{cond}: conditional stability constant for each ion and each chelate: pH 7.4, 25.0±0.1 °C, I=0.1 M NaClO. (Per: Cacheris WP et al. The relationship between thermodynamics and the toxicity of gadolinium complexes. Magn Reson Imag 1990;8:467–481)

Literature

1. Staks T, Schuhmann-Giampieri G, Frenzel T, Weinmann HJ, Lange L, Platzek J. Pharmacokinetics, dose proportionality, and tolerability of gadobutrol after single intravenous injection in healthy volunteers. Invest Radiol. 1994;29(7):709–15.
2. Package Insert, Bayer Healthcare, Germany.
3. Kirchin MA, Pirovano GP, Spinazzi A. Gadobenate dimeglumine (Gd-BOPTA). An overview. Invest Radiol. 1998;33(11):798–809.
4. Tweedle MF. The ProHance story: the making of a novel MRI contrast agent. Eur Radiol. 1997;7 Suppl 5:225–30.
5. Corot C, Idee JM, Hentsch AM, Santus R, Mallet C, Goulas V, Bonnemain B, Meyer D. Structure-activity relationship of macrocyclic and linear gadolinium chelates: investigation of transmetallation effect on the zinc-dependent metallopeptidase angiotensin-converting enzyme. J Magn Reson Imaging. 1998;8(3):695–702.
6. Rubin DL, Desser TS, Semelka R, Brown J, Nghiem HV, Stevens WR, Bluemke D, Nelson R, Fultz P, Reimer P, Ho V, Kristy RM, Pierro JA. A multicenter, randomized, double-blind study to evaluate the safety, tolerability, and efficacy of OptiMARK (gadoversetamide injection) compared with Magnevist (gadopentetate dimeglumine) in patients with liver pathology: results of a Phase III clinical trial. J Magn Reson Imaging. 1999;9(2):240–50.
7. de Haën C, Cabrini M, Akhnana L, Ratti D, Calabi L, Gozzini L. Gadobenate dimeglumine 0.5 M Solution for injection (MultiHance(R)): pharmaceutical formulation and physicochemical Properties of a new magnetic resonance imaging contrast medium. J Comput Assist Tomogr. 1999;23(1):S161–S168.
8. Benner T, Reimer P, Erb G, Schuierer G, Heiland S, Fischer C, Geens V, Sartor K, Forsting M. Cerebral MR perfusion imaging: first clinical application of a 1 M gadolinium chelate (Gadovist 1.0) in a double-blinded randomized dose-finding study. J Magn Reson Imaging. 2000;12(3):371–80.
9. Speck U. Contrast media: overview, use, and pharmaceutical aspects. 4th ed. Berlin: Springer Science + Business Media; 1994.
10. Guebert, Corot C. Personal Communication. 2000.
11. Bayer Healthcare. Personal Communication. 2001.
12. http://www.mr-tip.com/.
13. Modo MMJ, Bulte JWM. Molecular and cellular MR imaging. Boca Raton: CRC Press; 2007.
14. Rubin, GD, Rofsky NM. CT and MR angiography. Philadelphia: Lippincott Williams & Wilkins; 2009.

Appendix 15.3: Overview of Approved and Selected Experimental MR Contrast Media with Approval Status, Distribution, and Physicochemical Properties

Agent name	Lab name	Trade name	Company	Body part (Country)	MW (g/mol)	C (molar)	T ½ in blood (min)	Distribution
Gadopentetate dimeglumine	Gd-DTPA	Magnevist	Bayer HealthCare	CNS, WB (EU, USA, Japan)	938	0.5	~90	Extracellular
Gadobutrol	Gd-BT-DO3A	Gadovist	Bayer HealthCare	CNS (EU, USA)	605	1.0	~90	Extracellular
Gadoteridol	Gd-HP-DO3A	ProHance	Bracco	CNS, WB (EU, USA, Japan)	559	0.5	~90	Extracellular
Gadobenate dimeglumine	Gd-BOPTA	MultiHance	Bracco	CNS, liver (EU, USA)	1058	0.5	~90	Liver/extracellular
Gadoterate meglumine	Gd-DOTA	Dotarem	Guerbet	CNS, WB (EU)	558	0.5	~90	Extracellular
Gadodiamide	Gd-DTPA-BMA	Omniscan	GE Healthcare	CNS, WB (EU, USA, Japan)	574	0.5	~90	Extracellular
Gadoversetamide	Gd-DTPA-BMEA	OptiMARK	Covidien	CNS, liver (EU, USA)	662	0.5	~90	Extracellular
Gadoxetic acid disodium	Gd-EOB-DTPA	Primovist/ Eovist	Bayer HealthCare	Liver (EU, USA)	726	0.25	~90	Liver/extracellular
Gadofosveset	Diphenylcyclo-hexyl phospho-diester-Gd-DTPA	ABLAVAR	Lantheus Medical Imaging, Inc.	USA, EU (abd, limbs)	957	0.25	224 ± 30	Intravascular/ extracellular
SHU555 A	Ferucarbotran	Resovist/ Cliavist	Bayer HealthCare	Discontinued	~700,000	0.5 M Fe/l	144–216/234–348	Liver/RES

AMI-25	Ferumoxides	Feridex IV/ Endorem	Berlex/Guerbet	Liver (Japan, EU, USA)	n.a.	0.2 M Fe/l	144	Liver/RES
AMI-227	Ferumoxtran-10	Combidex/ Sinerem	Cytogen/Guerbet	Discontinued	n.a.	0.36 M Fe/l	>1,000	LN, RES
Dendrimer	Gadomer-17	n.a.	Bayer HealthCare	Phase II	17,453	0.05	120–180	Intravascular
P792	Gadomeritol	Vistarem	Guerbet	Phase II	6,473	0.035	~120	Intravascular

CNS central nervous system, *EU* European Union, *USA* United States of America, *LN* lymph nodes, *RES* reticuloendothelial system

Appendix 15.4: Overview of Available US Contrast Media with Current Approval and Distribution Status

Trade name(s)	Manufacturer	Labeled indications[a]	Shell composition	Gas	Availability
Albunex	Covidien	Cardiac	Albumin	Air	Discontinued since availability of Optison
Optison	GE Healthcare	Cardiac	Albumin	Perfluoropropane	US, EU
Echovist	Bayer HealthCare	Hysterosalpinx, cardiac	D-Galactose	Air	Discontinued since availability of Levovist
Levovist	Bayer HealthCare	Cardiovascular, focal liver lesions	Galactose/palmitic acid	Air	EU, Canada, Japan, China[a]
Imagent	Targeson, Inc	Echocardiography	Surfactant	Perfluorohexane	China
Definity/Luminity	Lantheus Medical Imaging	Cardiac; liver, kidney	Liposome	Octafluoropropane	EMEA=EU Definity=[b]
SonoVue	Bracco	CEUS	Surfactant/Phospholipid	Sulfur hexafluoride	EU
Sonazoid	Daiichi Sankyo	Liver, spleen, myocardial perfusion	Lipid stabilized	Perfluorocarbon	Japan
Br14	Bracco	Microcirculation, cardiac	Phospholipid	Perfluorocarbon	Phase III
Imagify	Acusphere	Cardiac	Poly-L-lactide co-glycide	Perflubutane	Phase III
Sonavist/Sonovist	GE Healthcare	Uptake by Kupffer cells in RES	Polycyanoacrylate	Air	RND
Echogen	Sonus	Heart	Surfactant	Dodecafluoropentane	European approval then withdrawn from market in 2000

[a]Labeled indications may vary in different in countries/continents. Some CM are not even available in certain countries

[b]India, North America, Australia and New Zealand, parts of the Pacific Rim, and several countries in the Middle East. *RND* research and development

References

1. Barrett T, Brechbiel M, Bernardo M, Choyke PL. MRI of tumor angiogenesis. J Magn Reson Imaging. 2007;26:235–49.
2. Knopp MV, von Tengg-Kobligk H, Choyke PL. Functional magnetic resonance imaging in oncology for diagnosis and therapy monitoring. Mol Cancer Ther. 2003;2:419–26.
3. Ah-See ML, Makris A, Taylor NJ, Harrison M, Richman PI, Burcombe RJ, Stirling JJ, d'Arcy JA, Collins DJ, Pittam MR, Ravichandran D, Padhani AR. Early changes in functional dynamic magnetic resonance imaging predict for pathologic response to neoadjuvant chemotherapy in primary breast cancer. Clin Cancer Res. 2008;14:6580–9.
4. Giesel FL, Choyke PL, Mehndiratta A, Zechmann CM, von Tengg-Kobligk H, Kayser K, Bischoff H, Hintze C, Delorme S, Weber MA, Essig M, Kauczor HU, Knopp MV. Pharmacokinetic analysis of malignant pleural mesothelioma-initial results of tumor microcirculation and its correlation to microvessel density (CD-34). Acad Radiol. 2008;15:563–70.
5. Heiland S, Erb G, Ziegler S, Krix M. Where contrast agent concentration really matters – a comparison of CT and MRI. Invest Radiol. 2010;45:529–37.
6. Knopp MV, Giesel FL, von Tengg-Kobligk H, Radeleff J. 3D MR colonography after exclusive intravenous administration of a hepatobiliary contrast agent. Eur Radiol. 2001;11:170.
7. Zech CJ, Herrmann KA, Reiser MF, Schoenberg SO. MR imaging in patients with suspected liver metastases: value of liver-specific contrast agent Gd-EOB-DTPA. Magn Reson Med Sci. 2007;6:43–52.
8. Davies BE, Kirchin MA, Bensel K, Lorusso V, Davies A, Parker JR, LaFrance ND. Pharmacokinetics and safety of gadobenate dimeglumine (multihance) in subjects with impaired liver function. Invest Radiol. 2002;37:299–308.
9. Katayama H. Survey of safety of clinical contrast media. Invest Radiol. 1990;25 Suppl 1:S7–10.
10. Pugh ND. Haemodynamic and rheological effects of contrast media: the role of viscosity and osmolality. Eur Radiol. 1996;6 Suppl 2:S13–5.
11. Rubin G, Rofsky N. CT and MR angiography – comprehensive vascular assessment. Philadelphia/Baltimore/New York/London/Buenos Aires/Hong Kong/Sydney/Tokyo: Wolters Kluwer/Lippincott Williams & Wilkins; 2009.
12. Knopp M, Kauczor HU, Knopp MA, Manella P, Delorme S, Wenz F. Effects of viscosity, cannula size and temperature in mechanical contrast media administration in CT and magnetic resonance tomography. Rofo. 1995;163:259–64.
13. Giesel FL, Essig M, Zabel-Du-Bois A, Bock M, von Tengg-Kobligk H, fshar-Omarei A, Debus J, Kauczor HU, Krix M. High-contrast computed tomographic angiography better detects residual intracranial arteriovenous malformations in long-term follow-up after radiotherapy than 1.5-Tesla time-of-flight magnetic resonance angiography. Acta Radiol. 2010;51:64–70.
14. Behrendt FF, Pietsch H, Jost G, Sieber MA, Keil S, Plumhans C, Seidensticker P, Gunther RW, Mahnken AH. Intra-individual comparison of different contrast media concentrations (300 mg, 370 mg and 400 mg iodine) in MDCT. Eur Radiol. 2010;20:1644–50.
15. Rutten A, Prokop M. Contrast agents in X-ray computed tomography and its applications in oncology. Anticancer Agents Med Chem. 2007;7:307–16.
16. Lauffer RB. MRI contrast agents: basic principles. In: Edelman R, Hesselink JR, Zlatkin MB, editors. Clinical magnetic resonance imaging. Philadelphia: WB Saunders Company; 1996. p. 177–91.
17. Koenig SH. From the relaxivity of Gd(DTPA)2- to everything else. Magn Reson Med. 1991;22:183–90.
18. Caravan P, Cloutier NJ, Greenfield MT, McDermid SA, Dunham SU, Bulte JW, Amedio Jr JC, Looby RJ, Supkowski RM, Horrocks Jr WD, McMurry TJ, Lauffer RB. The interaction of MS-325 with human serum albumin and its effect on proton relaxation rates. J Am Chem Soc. 2002;124:3152–62.

19. Weinmann HJ, Ebert W, Misselwitz B, Schmitt-Willich H. Tissue-specific MR contrast agents. Eur J Radiol. 2003;46:33–44.
20. Runge VM. Safety of magnetic resonance contrast media. Top Magn Reson Imaging. 2001;12:309–14.
21. Knopp MV, von Tengg-Kobligk H, Floemer F, Schoenberg SO. Contrast agents for MRA: future directions. J Magn Reson Imaging. 1999;10:314–6.
22. Giesel FL, Mehndiratta A, Essig M. High-relaxivity contrast-enhanced magnetic resonance neuroimaging: a review. Eur Radiol. 2010;20(10):2461–74.
23. Claussen C, Laniado M, Schorner W, Niendorf HP, Weinmann HJ, Fiegler W, Felix R. Gadolinium-DTPA in MR imaging of glioblastomas and intracranial metastases. AJNR Am J Neuroradiol. 1985;6:669–74.
24. Runge VM, Schoerner W, Niendorf HP, Laniado M, Koehler D, Claussen C, Felix R, James Jr AE. Initial clinical evaluation of gadolinium DTPA for contrast-enhanced magnetic resonance imaging. Magn Reson Imaging. 1985;3:27–35.
25. Essig M, Weber MA, von Tengg-Kobligk H, Knopp MV, Yuh WT, Giesel FL. Contrast-enhanced magnetic resonance imaging of central nervous system tumors: agents, mechanisms, and applications. Top Magn Reson Imaging. 2006;17:89–106.
26. Provenzale JM, Mukundan S, Dewhirst M. The role of blood–brain barrier permeability in brain tumor imaging and therapeutics. AJR Am J Roentgenol. 2005;185:763–7.
27. Huppertz A, Rohrer M. Gadobutrol, a highly concentrated MR-imaging contrast agent: its physicochemical characteristics and the basis for its use in contrast-enhanced MR angiography and perfusion imaging. Eur Radiol. 2004;14 Suppl 5:M12–8.
28. Rohrer M, Bauer H, Mintorovitch J, Requardt M, Weinmann HJ. Comparison of magnetic properties of MRI contrast media solutions at different magnetic field strengths. Invest Radiol. 2005;40:715–24.
29. Tombach B, Bohndorf K, Brodtrager W, Claussen CD, Duber C, Galanski M, Grabbe E, Gortenuti G, Kuhn M, Gross-Fengels W, Hammerstingl R, Happel B, Heinz-Peer G, Jung G, Kittner T, Lagalla R, Lengsfeld P, Loose R, Oyen RH, Pavlica P, Pering C, Pozzi-Mucelli R, Persigehl T, Reimer P, Renken NS, Richter GM, Rummeny EJ, Schafer F, Szczerbo-Trojanowska M, Urbanik A, Vogl TJ, Hajek P. Comparison of 1.0 M gadobutrol and 0.5 M gadopentetate dimeglumine-enhanced MRI in 471 patients with known or suspected renal lesions: results of a multicenter, single-blind, interindividual, randomized clinical phase III trial. Eur Radiol. 2008;18:2610–9.
30. Essig M, Lodemann KP, Le-Huu M, Bruning R, Kirchin M, Reith W. Intraindividual comparison of gadobenate dimeglumine and gadobutrol for cerebral magnetic resonance perfusion imaging at 1.5 T. Invest Radiol. 2006;41:256–63.
31. Cavagna FM, Dapra M, Maggioni F, de Haen C, Felder E. Gd-BOPTA/Dimeg: experimental disease imaging. Magn Reson Med. 1991;22:329–33.
32. Grazioli L, Morana G, Kirchin MA, Schneider G. Accurate differentiation of focal nodular hyperplasia from hepatic adenoma at gadobenate dimeglumine-enhanced MR imaging: prospective study. Radiology. 2005;236:166–77.
33. Helmberger T, Semelka RC. New contrast agents for imaging the liver. Magn Reson Imaging Clin N Am. 2001;9:745–66, vi.
34. Knopp MV, Runge VM, Essig M, Hartman M, Jansen O, Kirchin MA, Moeller A, Seeberg AH, Lodemann KP. Primary and secondary brain tumors at MR imaging: bicentric intraindividual crossover comparison of gadobenate dimeglumine and gadopentetate dimeglumine. Radiology. 2004;230:55–64.
35. Knopp MV, Schoenberg SO, Rehm C, Floemer F, von Tengg-Kobligk H, Bock M, Hentrich HR. Assessment of gadobenate dimeglumine for magnetic resonance angiography: phase I studies. Invest Radiol. 2002;37:706–15.
36. Herborn CU, Lauenstein TC, Ruehm SG, Bosk S, Debatin JF, Goyen M. Intraindividual comparison of gadopentetate dimeglumine, gadobenate dimeglumine, and gadobutrol for pelvic 3D magnetic resonance angiography. Invest Radiol. 2003;38:27–33.

37. von Tengg-Kobligk H, Floemer F, Knopp MV. Multiphasic MR angiography as an intra-individual comparison between the contrast agents Gd-DTPA, Gd-BOPTA, and Gd-BT-DO3A. Radiologe. 2003;43:171–8.
38. Giesel FL, Runge V, Kirchin M, Mehndiratta A, Gerigk L, Corell B, von Gall C, Kauczor HU, Essig M. Three-dimensional multiphase time-resolved low-dose contrast-enhanced magnetic resonance angiography using TWIST on a 32-channel coil at 3 T: a quantitative and qualitative comparison of a conventional gadolinium chelate with a high-relaxivity agent. J Comput Assist Tomogr. 2010;34:678–83.
39. Kirchin MA, Pirovano GP, Spinazzi A. Gadobenate dimeglumine (Gd-BOPTA). An overview. Invest Radiol. 1998;33:798–809.
40. Planchamp C, Montet X, Frossard JL, Quadri R, Stieger B, Meier PJ, Ivancevic MK, Vallee JP, Terrier F, Pastor CM. Magnetic resonance imaging with hepatospecific contrast agents in cirrhotic rat livers. Invest Radiol. 2005;40:187–94.
41. Planchamp C, Hadengue A, Stieger B, Bourquin J, Vonlaufen A, Frossard JL, Quadri R, Becker CD, Pastor CM. Function of both sinusoidal and canalicular transporters controls the concentration of organic anions within hepatocytes. Mol Pharmacol. 2007;71:1089–97.
42. Giesel FL, von Tengg-Kobligk H, Wilkinson ID, Siegler P, von der Lieth CW, Frank M, Lodemann KP, Essig M. Influence of human serum albumin on longitudinal and transverse relaxation rates (r1 and r2) of magnetic resonance contrast agents. Invest Radiol. 2006;41:222–8.
43. de Haën C, Cabrini M, Akhnana L, Ratti D, Calabi L, Gozzini L. Gadobenate dimeglumine 0.5 M solution for injection (MultiHance) pharmaceutical formulation and physicochemical properties of a new magnetic resonance imaging contrast medium. J Comput Assist Tomogr. 1999;23 Suppl 1:S161–8.
44. Cotton F, Hermier M. The advantage of high relaxivity contrast agents in brain perfusion. Eur Radiol. 2006;16 Suppl 7:M16–26.
45. Saito K, Kotake F, Ito N, Ozuki T, Mikami R, Abe K, Shimazaki Y. Gd-EOB-DTPA enhanced MRI for hepatocellular carcinoma: quantitative evaluation of tumor enhancement in hepatobiliary phase. Magn Reson Med Sci. 2005;4:1–9.
46. Pastor CM. Gadoxetic acid-enhanced hepatobiliary phase MR imaging: cellular insight. Radiology. 2010;257:589.
47. Morana G, Salviato E, Guarise A. Contrast agents for hepatic MRI. Cancer Imaging. 2007;7(Spec No A):S24–7.
48. Lauffer RB. Targeted relaxation enhancement agents for MRI. Magn Reson Med. 1991;22:339–42.
49. Caravan P. Protein-targeted gadolinium-based magnetic resonance imaging (MRI) contrast agents: design and mechanism of action. Acc Chem Res. 2009;42:851–62.
50. Goyen M, Edelman M, Perreault P, O'Riordan E, Bertoni H, Taylor J, Siragusa D, Sharafuddin M, Mohler III ER, Breger R, Yucel EK, Shamsi K, Weisskoff RM. MR angiography of aortoiliac occlusive disease: a phase III study of the safety and effectiveness of the blood-pool contrast agent MS-325. Radiology. 2005;236:825–33.
51. Essig M, Rohrer M, Giesel F, Tuttenberg J, Weber MA, Michaely H, Gerigk L, Voth M. Human brain tumor imaging with a protein-binding MR contrast agent: initial experience. Eur Radiol. 2010;20:218–26.
52. Lorusso V, Pascolo L, Fernetti C, Visigalli M, Anelli P, Tiribelli C. In vitro and in vivo hepatic transport of the magnetic resonance imaging contrast agent B22956/1: role of MRP proteins. Biochem Biophys Res Commun. 2002;293:100–5.
53. Bonnemain B. Contrast products in magnetic resonance imaging. Ann Pharm Fr. 1994;52:229–39.
54. Chachuat A, Bonnemain B. European clinical experience with Endorem. A new contrast agent for liver MRI in 1000 patients. Radiologe. 1995;35:274–6.
55. Koenig SH, Kellar KE. Blood-pool contrast agents for MRI: a critical evaluation. Acad Radiol. 1998;5 Suppl 1:200–5.

56. Koenig SH, Kellar KE. Theory of 1/T1 and 1/T2 NMRD profiles of solutions of magnetic nanoparticles. Magn Reson Med. 1995;34:227–33.
57. Hemmingsson A, Carlsten J, Ericsson A, Klaveness J, Sperber GO, Thuomas KA. Relaxation enhancement of the dog liver and spleen by biodegradable superparamagnetic particles in proton magnetic resonance imaging. Acta Radiol. 1987;28:703–5.
58. Stark DD, Weissleder R, Elizondo G, Hahn PF, Saini S, Todd LE, Wittenberg J, Ferrucci JT. Superparamagnetic iron oxide: clinical application as a contrast agent for MR imaging of the liver. Radiology. 1988;168:297–301.
59. Reimer P, Tombach B, Daldrup H, Hesse T, Sander G, Balzer T, Shamsi K, Berns T, Rummeny EJ, Peters PE. New MR contrast media in liver diagnosis. Initial clinical results with hepatobiliary Eovist (gadolinium-EOB-DTPA) and RES-specific Resovist (SH U 555 A). Radiologe. 1996;36:124–33.
60. Mller M, Reimer P, Wiedermann D, Allkemper T, Marx C, Tombach B, Rummeny EJ, Shamsi K, Balzer T, Peters PE. T1-weighted dynamic MRI with new superparamagnetic iron oxide particles (Resovist): results of a phantom study as well as 25 patients. Rofo. 1998; 168:228–36.
61. Sigal R, Vogl T, Casselman J, Moulin G, Veillon F, Hermans R, Dubrulle F, Viala J, Bosq J, Mack M, Depondt M, Mattelaer C, Petit P, Champsaur P, Riehm S, Dadashitazehozi Y, De Jaegere T, Marchal G, Chevalier D, Lemaitre L, Kubiak C, Helmberger R, Halimi P. Lymph node metastases from head and neck squamous cell carcinoma: MR imaging with ultrasmall superparamagnetic iron oxide particles (Sinerem MR) – results of a phase-III multicenter clinical trial. Eur Radiol. 2002;12:1104–13.
62. Ruehm SG, Corot C, Vogt P, Kolb S, Debatin JF. Magnetic resonance imaging of atherosclerotic plaque with ultrasmall superparamagnetic particles of iron oxide in hyperlipidemic rabbits. Circulation. 2001;103:415–22.
63. Bremerich J, Bilecen D, Reimer P. MR angiography with blood pool contrast agents. Eur Radiol. 2007;17:3017–24.
64. Kellar KE, Fujii DK, Gunther W, Bjornerud A, Spiller M, Koenig SH. NC100150 Injection, a preparation of optimized iron oxide nanoparticles for positive-contrast MR angiography. J Magn Reson Imaging. 2000;11:488–94.
65. Taupitz M, Wagner S, Schnorr J, Kravec I, Pilgrimm H, Bergmann-Fritsch H, Hamm B. Phase I clinical evaluation of citrate-coated monocrystalline very small superparamagnetic iron oxide particles as a new contrast medium for magnetic resonance imaging. Invest Radiol. 2004;39:394–405.
66. Manninger SP, Muldoon LL, Nesbit G, Murillo T, Jacobs PM, Neuwelt EA. An exploratory study of ferumoxtran-10 nanoparticles as a blood–brain barrier imaging agent targeting phagocytic cells in CNS inflammatory lesions. AJNR Am J Neuroradiol. 2005; 26:2290–300.
67. Port M, et al. P792: a rapid clearance blood pool agent for magnetic resonance imaging: preliminary results. MAGMA Magn Reson Mater Phys Biol Med. 2001;12(2–3):121.
68. Bendszus M, Ladewig G, Jestaedt L, Misselwitz B, Solymosi L, Toyka K, Stoll G. Gadofluorine M enhancement allows more sensitive detection of inflammatory CNS lesions than T2-w imaging: a quantitative MRI study. Brain. 2008;131:2341–52.
69. Kauczor HU. Helium-3 imaging of pulmonary ventilation. Br J Radiol. 1998;71:701–3.
70. Knopp MV, Balzer T, Esser M, Kashanian FK, Paul P, Niendorf HP. Assessment of utilization and pharmacovigilance based on spontaneous adverse event reporting of gadopentetate dimeglumine as a magnetic resonance contrast agent after 45 million administrations and 15 years of clinical use. Invest Radiol. 2006;41:491–9.
71. Edward M, Quinn JA, Burden AD, Newton BB, Jardine AG. Effect of different classes of gadolinium-based contrast agents on control and nephrogenic systemic fibrosis-derived fibroblast proliferation. Radiology. 2010;256(3):735–43.
72. Kiessling F, Huppert J, Zhang C, Jayapaul J, Zwick S, Woenne EC, Mueller MM, Zentgraf H, Eisenhut M, Addadi Y, Neeman M, Semmler W. RGD-labeled USPIO inhibits adhesion and endocytotic activity of alpha v beta3-integrin-expressing glioma cells and only accumulates in the vascular tumor compartment. Radiology. 2009;253:462–9.

73. Kiessling F, Morgenstern B, Zhang C. Contrast agents and applications to assess tumor angiogenesis in vivo by magnetic resonance imaging. Curr Med Chem. 2007;14:77–91.
74. Bumb A, Brechbiel MW, Choyke P. Macromolecular and dendrimer-based magnetic resonance contrast agents. Acta Radiol. 2010;51(7):751–67.
75. Misselwitz B, Platzek J, Weinmann HJ. Early MR lymphography with gadofluorine M in rabbits. Radiology. 2004;231:682–8.
76. Giesel FL, Stroick M, Griebe M, Troster H, von der Lieth CW, Requardt M, Rius M, Essig M, Kauczor HU, Hennerici MG, Fatar M. Gadofluorine m uptake in stem cells as a new magnetic resonance imaging tracking method: an in vitro and in vivo study. Invest Radiol. 2006;41:868–73.
77. Barnhart J, Levene H, Villapando E, Maniquis J, Fernandez J, Rice S, Jablonski E, Gjoen T, Tolleshaug H. Characteristics of Albunex: air-filled albumin microspheres for echocardiography contrast enhancement. Invest Radiol. 1990;25 Suppl 1:S162–4.
78. Shaw LJ, Gillam L, Feinstein S, Dent J, Plotnick G. Use of an intravenous contrast agent (Optison) to enhance echocardiography: efficacy and cost implications. Optison Multicenter Study Group. Am J Manag Care. 1998;4(Spec No):SP169–76.
79. Bhutani MS, Hoffman BJ, van Velse A, Hawes RH. Contrast-enhanced endoscopic ultrasonography with galactose microparticles: SHU508 A (Levovist). Endoscopy. 1997;29:635–9.
80. Schneider M. SonoVue, a new ultrasound contrast agent. Eur Radiol. 1999;9 Suppl 3: S347–8.
81. Kitzman DW, Goldman ME, Gillam LD, Cohen JL, Aurigemma GP, Gottdiener JS. Efficacy and safety of the novel ultrasound contrast agent perflutren (definity) in patients with suboptimal baseline left ventricular echocardiographic images. Am J Cardiol. 2000;86:669–74.
82. Quaia E. Microbubble ultrasound contrast agents: an update. Eur Radiol. 2007;17: 1995–2008.
83. Furlow B. Contrast-enhanced ultrasound. Radiol Technol. 2009;80:547S–61.
84. Kiessling F, Huppert J, Palmowski M. Functional and molecular ultrasound imaging: concepts and contrast agents. Curr Med Chem. 2009;16:627–42.
85. Fisher NG, Christiansen JP, Leong-Poi H, Jayaweera AR, Lindner JR, Kaul S. Myocardial and microcirculatory kinetics of BR14, a novel third-generation intravenous ultrasound contrast agent. J Am Coll Cardiol. 2002;39:530–7.
86. Pochon S, Tardy I, Bussat P, Bettinger T, Brochot J, von Wronski M, Passantino L, Schneider M. BR55: a lipopeptide-based VEGFR2-targeted ultrasound contrast agent for molecular imaging of angiogenesis. Invest Radiol. 2010;45:89–95.
87. Runge VM, Knopp MV. Off-label use and reimbursement of contrast media in MR. J Magn Reson Imaging. 1999;10:489–95.
88. Torres A. The use of food and drug administration – approved medications for unlabeled (off-label) uses. The legal and ethical implications. Arch Dermatol. 1994;130:32–6.
89. Reimer P, Vosshenrich R. Off-label use of contrast agents. Eur Radiol. 2008;18:1096–101.
90. Siebner HR, von Grafin EH, Conrad B. Magnetic resonance ventriculography with gadolinium DTPA: report of two cases. Neuroradiology. 1997;39:418–22.
91. Ray DE, Cavanagh JB, Nolan CC, Williams SC. Neurotoxic effects of gadopentetate dimeglumine: behavioral disturbance and morphology after intracerebroventricular injection in rats. AJNR Am J Neuroradiol. 1996;17:365–73.
92. Raine J. Off-label use of medicines – legal aspects. In: Thomsen HS, editor. Contrast media safety issues and ESUR guidelines. Berlin/Heidelberg/New York: Springer Science + Business Media; 2006. p. 5–8.

Part III
The Future of Imaging in Clinical Trials

Chapter 16
Pharmaceutical Industry Perspective Regarding Imaging Techniques

Joel Krasnow

Abstract Product development within the pharmaceutical and medical device industries is shaped by market and regulatory forces. Patients and payers are only willing to pay for products that represent true innovation and value. As good branded products lose exclusivity, the bar with respect to product performance is raised for new products. The healthcare industry has successfully demonstrated value for new products by identifying populations with significant unmet need through companion diagnostics or orphan diseases. The regulatory hurdle for product approval has steadily increased with the key metric being the product benefit compared to its risk. Clinical trial size and complexity has increased markedly in order to meet the regulatory hurdles for risk assessment. Imaging can document the pathophysiology of many disorders. Through various methods, imaging can be employed as a biomarker in early development, as a tool for identifying a population for a specific product, or as an endpoint in registration trials. This chapter reviews the changing landscape of product development and the role for effective implementation of imaging technology within clinical development.

Keywords Benefit-risk • Molecular imaging • Biomarker • Regulatory • Unmet need

Introduction

In order for a pharmaceutical company to succeed, it needs to generate products that satisfy currently unmet medical needs. In this chapter, we will outline the evolving market-driven standards within the pharmaceutical industry. This will demonstrate that the trend towards increasingly personalized healthcare is driven both by

J. Krasnow, MD, MBA
Department of Strategy and Regulatory Affairs, TherapeuticsMD,
Boca Raton, FL, USA
e-mail: jkras@optonline.net

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3_16, © Springer-Verlag London 2014

scientific advances as well as market forces. We will then discuss how imaging will play an important role in the progression of personalized medicine.

Evolving Regulatory Standards

Over the past several decades there has been an increasing focus on product safety such that many in the industry joke that aspirin would not be approved as a new chemical entity if its new drug application was submitted today. The key metric for product approval has evolved from efficacy centric to safety centric to the current benefit-risk ratio paradigm. Interpretation of the BRR can vary considerably between health authorities. The increased focus on patient risk coupled with technological advances has resulted in an increased percentage of recent drug approvals coming from monoclonal antibodies, largely due to their high specificity and consequent relative paucity of off-target effects.

Health authorities such as the FDA and EMEA play key roles in the drug approval process. The overall assessment of a new therapeutic entity is evaluated according to the benefits that the product provides and the risks that are associated with its use. Pharmaceutical products with favorable benefit-risk ratios (BRR) will be approved [1]. When products are already on the market, a potential new market entrant will be evaluated in reference to these existing products from both a regulatory and from a payer perspective. Therefore, the currently available medicines in effect dictate the attributes of new pharmaceutical entities under development. Table 16.1 divides pharmaceutical markets into four levels based on existing therapeutic options and unmet medical need. It becomes apparent that in more mature (advanced) markets, niche populations may offer the greatest opportunity. Identification of these populations through various disease parameters can be achieved through clinically available or evolving imaging techniques that can be adapted for clinical trial purposes. We will first discuss the different types of pharmaceutical markets then move on to the evolving area of personalized medicine including the role of imaging.

Evolution of Pharmaceutical Markets

Each therapeutic area and medical condition has evolved in a different manner depending upon scientific advances and the commercial market environment. Table 16.1 presents strategic considerations typical for that of a mid- to large-size pharmaceutical company. Four phases of market evolution are characterized. In the first phase, no medical treatments are available for the disease state. Presently, the number of disease states for which this situation exists is limited primarily to diseases with low prevalence (orphan diseases) or disease states where the pathophysiology is not sufficiently understood to have yielded successful treatments. Historically, inherited diseases caused by deficiencies in a single enzyme or substrate have been

Table 16.1 Effect of pharmacological advances within a therapeutic area

Therapeutic options	Unmet medical need	Strategic options	Patient response
Phase 1 market			
No available treatment	Very high (results in high morbidity and mortality)	Develop a product with some therapeutic efficacy	All patients with the disease will be candidates for using the therapy as there are no alternatives
No available treatment	Moderate to high (moderate morbidity with limited mortality)	Develop a product with some therapeutic efficacy Given the moderate morbidity, a reasonable safety profile will be required for significant uptake	Most patients with high morbidity will use the therapy. Patients with less sequelae may embark on a trial of use depending on the safety profile
Phase 2 market			
Treatments with limited effectiveness are available (number of products is limited)	Moderate to high (existing treatments provide satisfactory efficacy to a limited number of patients)	Due to limited available options, products with comparable benefit-risk ratio (BRR) to marketed compounds will gain some market share which will be a function of its advantages relative to existing alternatives	Will initiate with preferred compound and some will switch in search of best BRR
Treatments with limited effectiveness are available (number of products is significant or some products are generic)	Moderate (existing treatments provide reasonable BRR to a majority of patients)	Develop a product which has a superior BRR relative to existing options	Depending on the delta in BRR between the new product and the relative cost and access, the product will be used as either 1st line or following the first treatment failure
Phase 3 market			
Generic (low-cost) treatments with good effectiveness are available	Modest (existing treatments provide very high BRR to a majority of patients)	Don't invest Develop for salvage therapy (negative selection) Demonstrate markedly higher BRR Identify a subpopulation that will have a greater BRR than existing treatments (positive selection)	Those patients who can be identified to have a higher BRR will access this therapy early in the treatment process

(continued)

Table 16.1 (continued)

Therapeutic options	Unmet medical need	Strategic options	Patient response
Phase 4 market			
Generic (low-cost) treatments with good effectiveness are available, and personalized medicine therapies are utilized	Low overall Moderate to high only for identified subpopulations	Identify a subpopulation that will benefit from a novel therapeutic approach	Those patients who can be identified to have a higher BRR will access this therapy early in the treatment process
Most patients are satisfied with generic treatments	Still present but hurdle for demonstrating improved BRR is extremely high	Very limited	Limited new therapies (gout)

successfully targeted. Fabry disease is an X-linked genetic disorder of glycosphingolipid metabolism resulting in a deficiency of the lysosomal enzyme α-galactosidase A. Treatment with an exogenous source of α-galactosidase (Fabrazyme®, Genzyme, Cambridge, MA, USA) improves outcomes [2]. A similar strategy of replacing the missing substrate has been successfully employed by treatment with an analog of ß-glucocerebrosidase to replace its deficiency in Gaucher's disease [3], an exogenous source of GAA (acid α-glucosidase) in Pompe disease [4] where the enzyme is deficient, for mucopolysaccharidosis I and VI and phenylketonuria. As effective treatments get developed for diseases with well-known disease mechanisms, fewer such opportunities remain. Currently, most of the diseases for which effective treatments are not available have less well-understood pathophysiology.

If a pharmaceutical company develops a strategy to target a particular disease, they may elect to conduct research into the pathophysiology of the disease in order to develop a competitive advantage with respect to early identification of potential therapeutic targets. When a disease-focused approach is adopted, the development of imaging endpoints to assess disease progression becomes cost-effective. Once developed, this competitive advantage may enable the organization to become a leader in the field.

Developing the first effective treatment for a disease state is usually extremely rewarding both scientifically and financially. Prior to 1996, no effective treatment for Alzheimer's disease was available. When Aricept® (Eisai Inc., Woodcliff Lake, NJ, USA), the first product demonstrating efficacy in well-controlled trials, emerged, it was welcomed by patients and families. A similar situation existed for acquired immune deficiency syndrome treatments one decade earlier.

When an unmet medical need exists for a disease or condition with low morbidity that is not usually fatal, treatments need to clear a higher hurdle with respect to safety. This may also include a longer duration of follow-up to lower the likelihood of rare side effects emerging following market introduction. For these reasons, an increasing number of products that are solutions to unmet medical needs are being approved first outside of the USA. In 2010, Asclera® (Merz Aesthetics, San Mateo, CA, USA)

(polidocanol) which is a sclerosing agent indicated to treat uncomplicated spider veins (varicose veins ≤1 mm in diameter) and uncomplicated reticular veins (varicose veins 1–3 mm in diameter) in the lower extremity was approved by the FDA. This followed extensive experience in Europe that demonstrated its favorable safety profile [5].

In phase 2 pharmaceutical markets, treatments with limited effectiveness exist for the disease state; therefore, significant opportunity exists for new therapeutic entrants with improved efficacy or safety profiles. Pulmonary arterial hypertension (PAH) is associated with high morbidity and mortality. Prior to the introduction of effective pharmacotherapy, PAH would progress to death 2–3 years from the time of initial diagnosis. Currently available treatments include diuretics, vasodilators, anticoagulants and antiproliferative agents. Many of these agents have significant side effects and require close management. Due to the limited effectiveness of any one agent, most patients are treated with a combination of therapeutic agents. Healthcare providers and patients are eager for additional agents that will demonstrate efficacy. Most available agents have demonstrated an increase in the 6-min walking test but have not been demonstrated to alter the course of the disease. In summary due to the ongoing high medical need in PAH, there is significant opportunity for a new agent that can demonstrate a favorable BRR.

A category of highly successful pharmaceutical agents that entered into a phase 2 market is the biologics for rheumatoid arthritis. Prior to 1999, treatments for rheumatoid arthritis included analgesics such as NSAIDs and corticosteroids, and disease-modifying antirheumatic drugs such as methotrexate, hydroxychloroquine, and sulfasalazine. These treatments provided a reasonable solution for patients with mild disease, but significant unmet medical need remained for patients with moderate to severe disease. In 1999, infliximab (Remicade®, Janssen Biotech, Horsham, PA, USA) was approved based on placebo-controlled results from 428 patients which demonstrated an ACR 20 of approximately 50 % in moderate to severe RA patients when used in conjunction with methotrexate, compared to approximately 20 % when methotrexate was used alone following 30 weeks of treatment [6]. By demonstrating a clinically meaningful improvement in BRR, infliximab was used either as first-line treatment for severe patients or following failure of traditional oral DMARD therapy primarily as a consequence of the difference in cost. Subsequently other TNF antagonists were developed which provided advantages in administration (subcutaneous), less immunogenicity, and improved efficacy.

Some therapeutic areas are in or approaching a phase 3 market. At this juncture, generic options are available that demonstrate a reasonably high BRR. The cost of bringing a new product to market will still be at or in excess of 500 million US dollars. One example is the area of gastroesophageal reflux disease (GERD) where most patients are satisfied with existing products. Currently proton pump inhibitors with a greater duration of effectiveness and similar adverse events are being developed. Although these agents will have an improved BRR relative to generic alternatives, payers are likely to reserve these newer options for patients who do not respond adequately to generics. Thus, due to the cost of developing these agents, even a successful approval may not recoup the substantial development cost. Since it is very difficult to develop a new agent that has a major increase in BRR which

payers will authorize for first-line therapy, many pharmaceutical companies have exited GERD as a therapeutic area.

Many of the cardiovascular areas such as hypertension and hypercholesterolemia currently fall into phase 3 with the availability of several highly effective generic products. For hypercholesterolemia, large costly clinical outcome studies will be required to demonstrate an increased benefit relative to patients treated with statins alone as the current standard of care. Cholesterol esterase transfer protein is being targeted in the hope of demonstrating a markedly high BRR relative to statins.

The number of clinical development trial in many phase 3 markets such as those listed previously has declined relative to the activity in past decades. This is a consequence of the fact that the probability of success associated with these programs relative to the magnitude of the investment is lower than that for investments in other therapeutic areas.

In some phase 3 market areas of oncology such as breast cancer (HER2) or colon cancer (k-ras), specific tumor antigens (HER2) are being utilized to identify specific populations that will benefit from an increased BRR due to highly targeted therapy [7, 8]. This trend is accelerating as advances in the understanding of tumor biology emerge.

In phase 4 markets, low unmet need is present for the overall disease population because generic treatments with very good BRR are available. This was the reason why following the approval of allopurinol, it took more than 40 years for the next products (pegloticase, febuxostat) to treat gout to be commercially available. For phase 4 markets, remaining unmet medical need is in the very high-risk population or those with advanced disease. One common strategy is to identify subpopulations with more severe disease through biomarkers and to develop specific products for this group.

As therapeutic areas shift from phase 1 to more advanced stages of healthcare management, the challenges of satisfying remaining unmet medical needs become progressively more challenging. This is one of the reasons why the number of new molecular entities approved by health authorities has declined over the past decade. The reason why most pharmaceutical companies are developing biomarker strategies to identify the patients who will benefit from their therapeutic approach is that it is required to be competitive in the marketplace. It is also beneficial to patients. Imaging techniques can represent viable biomarker strategies across many therapeutic areas. A brief outline of biomarkers is useful to put the opportunity for imaging in context.

Biomarkers

Biomarkers are measureable biologic parameters that are markers of a medical condition or disease state. Biomarkers can be classified according to the parameter being measured. These include:

- Structural biomarkers
- Whole-cell biomarkers
- Physiological biomarkers
- Biochemical and proteomic biomarkers
- Genomic biomarkers

Imaging is frequently used to assess structural biomarkers. These include the measurement of atherosclerotic plaque within a vessel wall, tumor volume, bone mineral density, and other measurements of anatomically detectable structure. Whole-cell biomarkers are where whole cells such as cervical cells are assessed as in a "pap smear." Other examples include biopsies for the detection of tumors or other pathology. Physiological biomarkers include blood pressure, heart rate, nerve conduction, and other parameters representative of an organ system in an intact individual. Biochemical biomarkers represent substrates or products of chemical reactions in humans, while proteomic biomarkers represent the effect of proteins. These can be structural differences, different quantities of protein, or functional differences which may include enzymatic activity or protein interactions. Genomic markers include DNA at the structural level and RNA at the transcriptional level.

While historically imaging has been used primarily as a structural biomarker, the area of physiological imaging is currently the most rapidly expanding area both from a research and a clinical use perspective. Functional MRI and FDG-PET are early examples of physiological or functional imaging technologies.

From a clinical care perspective, biomarker use can be classified into four broad categories. These are (a) screening, (b) diagnosis, (c) monitoring disease, and (d) prognosis. Imaging is used across all four categories. Screening is the process whereby one distinguishes those who do not have any evidence of a disease versus those who may have a disease. Examples include pap smears, mammograms, and skin testing (PPD) for tuberculosis. For many diseases, biomarkers are either supportive or required for establishing the diagnosis. Examples include bone mineral density for osteoporosis, troponin for myocardial infarction, DNA for sickle cell disease. Virtually all diseases are monitored through biomarkers. These include echocardiography for cardiac disease, pulmonary function tests or imaging for respiratory disease, and serum ALT/AST and bilirubin for hepatic disease. These biomarkers can also be used for (1) monitoring of disease progression and (2) determining the effects of therapeutic interventions which include both positive effects as well as potential toxicity. The final category of prognosis aims to predict future events. This includes the outcome of the disease without intervention. An expanded definition includes predicting treatment efficacy or toxicity, possibility of developing a disease, or specific sequelae of a disease. Within clinical trials, biomarkers are commonly used for establishing the diagnosis and monitoring disease. As a consequence of the trial, prognostic factors may be identified. In a phase 3 or phase 4 market, the objective is to be able to identify a population with the diagnosis that will have a good prognosis when treated with the new pharmaceutical candidate.

Imaging Biomarkers

Many imaging techniques can be used to generate outputs that function as biomarkers. Within imaging, there are two levels that must be defined, the imaging technology and the imaging technique as listed in Table 16.2.

Table 16.2 Selected imaging technologies and techniques

Imaging technology platform	Related technologies	Imaging technique utilizing associated technology
X-ray	Dual-photon absorptiometry	Intravenous pyelogram
	Dual-energy X-ray absorptiometry	Cardiac angiography
	Single-photon Absorptiometry	
	Single X-ray absorptiometry	
	X-ray fluorescence	
Computed tomography	Dual-energy CT	Optical coherence tomography
	Electron beam CT	
	Multi-detector CT	
	Quantitative CT	
	Spiral CT	
Magnetic resonance	Electron paramagnetic resonance imaging	Double inversion-recovery MR
	Magnetic resonance-gated CSF dynamics	Functional MRI
	Proton-electron double-resonance imaging	Magnetic field correlation
		Diffusion tensor imaging
		Diffusion-weighted imaging
		Magnetic field correlation
		Susceptibility-weighted imaging
Ultrasound	Echo planar imaging	Intravascular ultrasound
		Perfusion-weighted imaging
		Tissue Doppler imaging
Radionucleotide	Positron emission tomography	FDG-PET
	Single-photon emission CT	

Evolving technologies can be organized into the categories of structural, dynamic, and functional imaging. Structural imaging refers to the measurement of anatomical features such as tumor size, volume, and density. It is performed on static structures at a particular point in time. Dynamic imaging looks at non-static effects over a specific time period. Examples include cardiac output, biophysical profile performed by prenatal ultrasound, and receptor number and location prior to and following pharmacological intervention. Functional imaging enables visualization at the molecular or cellular level. Neuronal activity associated with specific tasks is providing powerful insights in neurobiology. One can view blood flow and tumor dynamics which will be used more often as personalized oncology care evolves.

Historically, medical treatment has been based on disease that was both visible and symptomatic. While we will still persist in treating disease at this level, advanced disease is often less responsive to intervention. In addition, it is difficult to know if treatments are "working" other than following patients for clinical progression. Imaging offers two important biomarker advantages.

Advantages of Imaging Biomarkers

The advantages of imaging biomarkers include:

- Ability to detect change in disease (e.g., disease progression or regression) much earlier than standard clinical endpoints
- Provide timely, functional information at the molecular, cellular, or tissue level regarding the impact of pharmacological intervention in a disease process

The ability to accurately assess responsiveness to therapy at the cellular level within hours to days following pharmacological treatment will shift the treatment paradigm in many areas of oncology [9, 10]. Functional imaging will evolve across most therapeutic areas. It has advanced more rapidly in neurology driven by clinical unmet needs. In summary, structural, dynamic, and functional imaging should be considered as part of a comprehensive disease area strategy. External imaging consultants are readily available to assist in strategic planning.

Biomarkers and Qualification

Imaging biomarkers can be used in discovery and clinical development if they are deemed to add value independent of whether they are recognized by health authorities as validated biomarkers. However, in order to be able to use a particular biomarker to identify patients who would benefit from a specific treatment in clinical practice, the biomarker must be certified as appropriately validated. The FDA and EMEA have issued guidelines for qualification of pharmacogenomic biomarkers. The qualification of a particular biomarker is for the context in which it will be used. When a biomarker assessment can reliably and reproducibly reflect a biological process, response, or event, then it can be validated for that biological process. Once validated, the biomarker can support product development. The context may be narrow such as for a single drug for a single disease or may be broader encompassing several drug classes. Some common principles apply. The biomarker must be demonstrated to be on the causal pathway of the disease process. It cannot simply correlate statistically to the disease outcome. The biomarker must account for the full clinical effect. If a biomarker can only explain 50 % of the effect of a pharmaceutical intervention, while it may be used in early development, it cannot be used as a surrogate for the clinical endpoint. Biomarkers should be assessed in more than a single study. Prospective biomarkers may be identified during retrospective analysis of clinical trial data. Subsequent prospective clinical trials can then be used to validate the biomarker. K-ras is an example of a biomarker for the identification of patients who will respond to EGFR inhibitors for colon cancer. Table 16.3 places biomarkers into the categories of exploratory, probable, and validated. In order for a biomarker to attain validation status, it must be made available for peer scientific

Table 16.3 Sequence of characteristics in biomarker development

Exploratory	Probable valid	Valid
Based on ongoing research Test system is being developed	Measured in an analytical test system with well-established performance characteristics[a]	Measured in an analytical test system with well-established performance characteristics
The significance of the test has not yet been established	There is a scientific body of evidence that appears to elucidate the physiologic, toxicologic, pharmacologic, or clinical significance of the data[a]	Well-established and widespread agreement in the scientific community regarding the interpretation and significance of the data
	Not (yet) broadly used in decision making	Used in decision making by regulators

[a]The scientific information may have been generated within a single organization and has not been made available for public review. Independent verification of the results by second parties may not yet have been performed

review. It is likely that health authorities will require demonstration of its performance characteristic in well-established systems by more than a single scientific group.

Similar guidance can be applied to the qualification of imaging biomarkers. Imaging biomarkers are used as the gold standard for approving drugs across many of the musculoskeletal diseases. These include the assessment of vertebral fractures for osteoporosis and the assessment of joint space narrowing and erosions for rheumatoid arthritis.

Advances in physics, information technology, and molecular medicine have come together to create the discipline of molecular imaging. Molecular imaging utilizes imaging technologies to assess biological activity in the intact organism. This can be performed at the cellular or tissue level. The advantage is that the changes in cell biology occur well before anatomical changes can be detected by MRI or CT. This diagnostic methodology enables early detection of changes occurring in the tissue of interest that can determine whether a therapeutic intervention is having its intended effect. The FDA has developed industry guidance for the development of diagnostic radiopharmaceuticals. Most pharmaceutical companies want to have qualified imaging agents so that they can use them to improve their success in drug development. Diagnostic imaging companies are moving quickly to bring many new agents to commercialization. This area will see exponential growth over the next decade.

As these molecular imaging agents (Table 16.4) get incorporated into clinical practice because of their increased sensitivity, they will also need to be incorporated into the drug development process. It is worthwhile to briefly review the molecular heterogeneity observed in many cancer types. This will serve to highlight the direction of current research and future clinical care in oncology.

The majority of targeted oncology drugs approved in the past 5 years have in their product labeling the requirement to have demonstrated the target in excised tissue specimens. While for those who respond to these therapies it is wonderful, it

Table 16.4 Representative molecular imaging agents

Imaging agent	Mechanism of action	Indications	Regulatory status
Fluorodeoxyglucose (^{18}F), (abbreviated ^{18}F-FDG)	FDG, as a glucose analog, is taken up by high-glucose-using cells such as brain, kidney, and cancer cells, where phosphorylation prevents the glucose from being released again from the cell, once it has been absorbed. The 2′ hydroxyl group (—OH) in normal glucose is needed for further glycolysis (metabolism of glucose by splitting it), but FDG is missing this 2′ hydroxyl; thus, FDG cannot be further metabolized in cells. The ^{18}F-FDG-6-phosphate formed when ^{18}F-FDG enters the cell thus cannot move out of the cell before radioactive decay. As a result, the distribution of ^{18}F-FDG is a good reflection of the distribution of glucose uptake and phosphorylation by cells in the body	1. Identification of regions of abnormal glucose metabolism associated with foci of epileptic seizures 2. Assessment of abnormal glucose metabolism to assist in the evaluation of malignancy in patients with known or suspected abnormalities found by other testing modalities or in patients with an existing diagnosis of cancer 3. In patients with coronary artery disease and left ventricular dysfunction, when used together with myocardial perfusion imaging, for the identification of left ventricular myocardium with residual glucose metabolism and reversible loss of systolic function [11]	Approved
[18F]-3′-fluoro-3′-deoxy-L-thymidine	The pyrimidine analog is phosphorylated by the enzyme thymidine kinase 1, which leads to intracellular trapping. Thymidine kinase 1 concentration increases nearly tenfold during DNA synthesis; therefore, ^{18}F-FLT uptake may accurately reflect cellular proliferation	Cell proliferation index	In multicenter trials. Consistency and standardization challenges are being pursued
[^{18}F]-fluorocholine	Choline is taken up into cells by specific transport mechanisms, phosphorylated by choline kinase, metabolized to phosphatidylcholine, and eventually incorporated into the cell membrane. Increased choline uptake has been shown in tumor cells and activated macrophages through which it is taken up into the athlerosclerotic plaque	Brain and prostate cancer	Investigational

is apparent that from an overall population perspective, these represent incomplete solutions. When the 30 % of breast cancer patients who are HER2 positive are treated with Herceptin® (Genentech, San Francisco, CA, USA), approximately 30 % will have a significant positive response [12]. While this is a major advance, it must also be acknowledged that this represents about 9 % of the total population of breast cancer patients. This is a lower rate than one would expect given this targeted approach. In metastatic colon cancer, Avastin® (Genentech, San Francisco, CA, USA) and Erbitux® (ImClone, Somerville, NJ, USA) have reported 10–15 % improvement in response rates relative to standard of care [13, 14].

Antibody-drug conjugate (ADC) technologies are markedly improving response rates over the first generation of tumor antigen-directed monoclonal antibodies. The antibody-drug conjugate attaches a cytotoxic agent to the monoclonal antibody. This attachment is with a highly stable linker such that the cytotoxic agent does not separate from the monoclonal antibody prior to target engagement thereby preventing the cytotoxic agent from interacting with off-target cells. Upon binding of the antibody to its tumor target cell, it will be internalized leading to release of the cytotoxic agent which results in a much higher rate of cell death and hence improved tumor response [15].

While some cancers represent a single mutation followed by clonal expansion, many tumors are heterogeneous in that different mutations are active in different parts of the cancer. This has been demonstrated for many tumor types including breast [16] and prostate cancer [17]. Epigenetic changes such as DNA methylation, chromatic modifications, and genomic imprinting also play important roles in susceptibility to therapeutic agents [18]. Given that not all tumor metastases are biologically equivalent and that even within an anatomical mass of tumor cells, biological heterogeneity is common, molecular imaging provides a powerful tool to begin to define the tumor characteristics associated with response to a particular therapeutic intervention. There is a movement towards understanding the proliferative capacity of the tumor cells prior to, during, and following treatment. Radiotracers are being developed to track this biological process. These techniques can be initiated in preclinical animal models and extended into clinical trials [19].

As we evolve towards personalized medicine, it becomes apparent that molecular imaging will play an increasing role in characterizing the specific biology of the disease process. This will not be restricted to oncology but will occur in all therapeutic areas where there is significant unmet medical need.

Opportunity for Effective Implementation of Imaging Technology

Within the pharmaceutical development process, imaging technologies represent a powerful tool to enable improved decision making from target selection through personalized healthcare in the clinic. Within a drug discovery program, imaging can be used to establish the validity and clinical relevance of novel biological targets.

For novel biological targets, testing of the hypothesis that target modulation has the desired impact on the disease process in an intact biological system such as a small animal is very important. Technological advances have resulted in improved image resolution to the extent that anatomical imaging in small animals such as rodents is both practical and highly informative. With most modalities, image resolution and sensitivity is comparable to that achieved in humans. Functional imaging, while more specialized, is also both applicable and available in mice and other animals.

Target validation, lead optimization, and characterization of biological activity are key steps in the discovery process that can be enhanced through functional and anatomical imaging. A key decision for pharmaceutical and biotechnology companies involves deciding which capabilities to develop internally and which to outsource. For animal research, optical imaging provides a logical starting point to develop internal resources. Additional capabilities will be dependent upon the areas of research within the organization. For organizations engaged in the identification and validation of novel targets, developing in-house expertise should improve decision making within discovery biology. The biologists involved in discovery research should develop an understanding of the capabilities available to them either internally or externally and should incorporate them into their project plans.

When imaging technologies provide useful information in the preclinical setting, they should be considered for incorporation in early clinical development to determine if they can be translated to humans. The advantages of:

- Greater sensitivity relative to clinical signs and symptoms
  - Smaller sample size and shorter trial duration required for clinical development
- The potential for objective quantification of the detected changes
- Obtaining physiological information

are so important in a clinical development program that imaging should be incorporated into routine training within biopharmaceutical organizations in a manner similar to good clinical practice, ICH guidelines, etc. Several anatomical and functional imaging techniques represent validated biomarkers that already improve drug development. The proliferation of newer functional imaging markers in conjunction with improvements in sensitivity and resolution of current modalities will provide even more tools to those in drug development. As the practice of medicine evolves to more personalized care, diseases such as breast and colon cancer are being further segregated according to the biologic properties defined by expression of specific biomarkers. This process is beginning to occur in several other therapeutic areas as well.

The time from initiation of drug development to commercial approval routinely exceeds 10 years. Therefore, depending on the particular phase of pharmaceutical development in which one is engaged, it is necessary to consider the state of clinical medicine, 5, 10, and sometimes 20 years into the future. Imaging technologies represent powerful tools for both research purposes and for clinical trials. They are being increasingly used throughout the pharmaceutical development process because of their potential value in decision making. Their role both in clinical

practice and in pharmaceutical development will increase significantly in the next decade. Acquiring and further developing imaging expertise within biotechnology and pharmaceutical organizations is and will become an even greater factor in bringing products to market in a cost-effective manner.

Summary

Competitive forces will continue to shape the healthcare markets. Pharmaceutical treatments are generally cost-effective solutions especially following patent expiration. Biophysical improvements leading to enhanced resolution is occurring with imaging technologies. These advances coupled with the development of agents that track biological processes are leading to important new applications in biological imaging. Although the principles of drug development have remained consistent over time, we will need to continually improve through the incorporation of technological advances. Imaging technologies have the potential to enhance efficiency throughout the development process. Their use should be considered as part of strategic planning. Implementation will need to be individualized based on the organizations objectives, expertise, and resources.

References

1. Sarac SB, Rasmussen CH, Rasmussen MA, Hallgreen CE, Soeborg T, Colding-Jorgensen M, Christensen PK, Thirstrup S, Mosekilde EA. Comprehensive approach to benefit-risk assessment in drug development. Basic Clin Pharmacol Toxicol. 2012;111(1):65–72.
2. Ramaswami U. Update on role of agalsidase alfa in management of Fabry disease. Drug Des Devel Ther. 2011;5:155–73.
3. Elstein D, Cohn GM, Wang N, Djordjevic M, Brutaru C, Zimran A. Early achievement and maintenance of the therapeutic goals using velaglucerase alfa in type 1 Gaucher disease. Blood Cells Mol Dis. 2011;46(1):110–23.
4. van Capelle CI, van der Beek NA, Hagemans ML, Arts WF, Hop WC, Lee P, Jaeken J, Frohn-Mulder IM, Merkus PJ, Corzo D, Puga AC, Reuser AJ, van der Ploeg AT. Effect of enzyme therapy in juvenile patients with Pompe disease: a three-year open-label study. Neuromuscul Disord. 2010;20(12):775–82.
5. Rabe E, Schliephake D, Otto J, Breu FX, Pannier F. Sclerotherapy of telangiectases and reticular veins: a double-blind, randomized, comparative clinical trial of polidocanol, sodium tetradecyl sulphate and isotonic saline (EASI study). Phlebology. 2010;25(3):124–31.
6. Lipsky PE, van der Heijde DM, St Clair EW, Furst DE, Breedveld FC, Kalden JR, Smolen JS, Weisman M, Emery P, Feldmann M, Harriman GR, Maini RN. Infliximab and methotrexate in the treatment of rheumatoid arthritis. Anti-tumor necrosis factor trial in rheumatoid arthritis with concomitant therapy study group. N Engl J Med. 2000;343(22):1594–602.
7. Ross JS, Slodkowska EA, Symmans WF, Pusztai L, Ravdin PM, Hortobagyi GN. The HER-2 receptor and breast cancer: ten years of targeted anti-HER-2 therapy and personalized medicine. Oncologist. 2009;14(4):320–68.
8. Ballestrero A, Garuti A, Cirmena G, Rocco I, Palermo C, Nencioni A, Scabini S, Zoppoli G, Parodi S, Patrone F. Patient-tailored treatments with anti-EGFR monoclonal antibodies in advanced colorectal cancer: KRAS and beyond. Curr Cancer Drug Targets. 2012;12(4):316–28.

9. Debergh I, Vanhove C, Ceelen W. Innovation in cancer imaging. Eur Surg Res. 2012;48(3):121–30.
10. Nishino M, Jagannathan JP, Krajewski KM, O'Regan K, Hatabu H, Shapiro G, Ramaiya NH. Personalized tumor response assessment in the era of molecular medicine: cancer-specific and therapy-specific response criteria to complement pitfalls of RECIST. AJR Am J Roentgenol. 2012;19(4):737–45.
11. SteriPET. Summary of product characteristics. www.mhra.gov.uk/home/groups/l-unit1/documents/websiteresources/con025955.pdf.
12. Herceptin® Prescribing information. www.gene.com/gene/products/information/pdf/herceptin-prescribing.pdf.
13. Avastin® Prescribing information. www.gene.com/gene/products/information/pdf/avastin-prescribing.pdf.
14. Erbitux® Prescribing information. http://packageinserts.bms.com/pi/pi_erbitux.pdf.
15. LoRusso PM, Weiss D, Guardino E, Girish S, Sliwkowski MX. Trastuzumab emtansine: a unique antibody-drug conjugate in development for human epidermal growth factor receptor 2-positive cancer. Clin Cancer Res. 2011;17(20):6437–47.
16. Torres L, Ribeiro FR, Pandis N, Andersen JA, Heim S, Teixeira MR. Intratumor genomic heterogeneity in breast cancer with clonal divergence between primary carcinomas and lymph node metastases. Breast Cancer Res Treat. 2007;102(2):143–55.
17. Shah RB, Mehra R, Chinnaiyan AM, Shen R, Ghosh D, Zhou M, MacVicar GR, Varambally S, Harwood J, Bismar TA, Kim R, Rubin MA, Pienta KJ. Androgen-independent prostate cancer is a heterogeneous group of diseases: lessons from a rapid autopsy program. Cancer Res. 2004;64:9209–16.
18. Iacobuzio-Donahue CA. Epigenetic changes in cancer. Ann Rev Pathol Mech Dis. 2009;4:229–49.
19. Kapty J, Murray D, Mercer J. Radiotracers for noninvasive molecular imaging of tumor cell death. Cancer Biother Radiopharm. 2010;25(6):615–28.

Appendix

Appendix 1: FDA Draft Guidance Document – Guidance for Industry Standards for Clinical Trial Imaging Endpoints

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3, © Springer-Verlag London 2014

Guidance for Industry
Standards for Clinical Trial Imaging Endpoints

DRAFT GUIDANCE

This guidance document is being distributed for comment purposes only.

Comments and suggestions regarding this draft document should be submitted within 60 days of publication in the *Federal Register* of the notice announcing the availability of the draft guidance. Submit electronic comments to http://www.regulations.gov. Submit written comments to the Division of Dockets Management (HFA-305), Food and Drug Administration, 5630 Fishers Lane, rm. 1061, Rockville, MD 20852. All comments should be identified with the docket number listed in the notice of availability that publishes in the *Federal Register*.

For questions regarding this draft document contact (CDER) Dr. Rafel Rieves at 301-796-2050 or (CBER) Office of Communication, Outreach, and Development at 301-827-1800 or 800-835-4709.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

August 2011
Clinical/Medical

Guidance for Industry

Standards for Clinical Trial Imaging Endpoints

I:\9676dft.doc
08/08/11

Guidance for Industry

Standards for Clinical Trial Imaging Endpoints

Additional copies are available from:

Office of Communications, Division of Drug Information
Center for Drug Evaluation and Research
Food and Drug Administration
10903 New Hampshire Ave., Bldg. 51, rm. 2201
Silver Spring, MD 20993-0002
Tel: 301-796-3400; Fax: 301-847-8714; E-mail: druginfo@fda.hhs.gov
http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm

or

Office of Communication, Outreach, and Development, HFM-40
Center for Biologics Evaluation and Research
Food and Drug Administration
1401 Rockville Pike, Suite 200N, Rockville, MD 20852-1448
Tel: 800-835-4709 or 301-827-1800; E-mail: ocod@fda.hhs.gov
http://www.fda.gov/BiologicsBloodVaccines/GuidanceComplianceRegulatoryInformation/default.htm

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Drug Evaluation and Research (CDER)
Center for Biologics Evaluation and Research (CBER)

August 2011
Clinical/Medical

TABLE OF CONTENTS

Contains Nonbinding Recommendations

Draft — Not for Implementation

Guidance for Industry[1]
Standards for Clinical Trial Imaging Endpoints

This draft guidance, when finalized, will represent the Food and Drug Administration's (FDA's) current thinking on this topic. It does not create or confer any rights for or on any person and does not operate to bind FDA or the public. You can use an alternative approach if the approach satisfies the requirements of the applicable statutes and regulations. If you want to discuss an alternative approach, contact the FDA staff responsible for implementing this guidance. If you cannot identify the appropriate FDA staff, call the appropriate number listed on the title page of this guidance.

I. INTRODUCTION

The purpose of this guidance is to assist sponsors in the use of endpoints that depend on the results of imaging tests in clinical trials of therapeutic drugs and biological products.[2] This guidance focuses on the imaging standards that we regard as important when imaging is used to assess a primary endpoint, or an endpoint component, in a clinical trial intended to confirm a drug's efficacy.[3] These standards can be used by sponsors to ensure that the imaging data are obtained in a manner that complies with a trial's protocol, that the quality of imaging data is maintained within and among clinical sites, and that there is a verifiable record of the imaging process. By considering the topics highlighted within this guidance, sponsors can obtain clinical trial imaging data in a manner that minimizes variability and enhances data quality and the ability to detect drug treatment effects.

This guidance describes the procedures recommended for collecting and interpreting medical images in efficacy trials. The guidance does not address whether or not specific measurements are clinically meaningful and are acceptable for drug approval.

[1] This guidance has been prepared by the Division of Medical Imaging Products in the Center for Drug Evaluation and Research (CDER) in cooperation with the Center for Biologics Evaluation and Research (CBER) at the Food and Drug Administration.

[2] For the purposes of this guidance, all references to *drugs* include both human drugs and therapeutic biological products unless otherwise specified.

[3] A confirmatory trial is an adequately controlled trial in which the hypotheses are stated in advance and evaluated. Most phase 3 trials are confirmatory trials that use designs intended to confirm a drug's efficacy. Additional characteristics of a confirmatory trial are described within the guidance for industry *E9 Statistical Principles for Clinical Trials*. We update guidances periodically. To make sure you have the most recent version of a guidance, check the FDA Drugs guidance Web page at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

Even though many of the concepts within this guidance also can be applied to clinical trials of diagnostic products and devices, those clinical trials often involve more technical considerations. We encourage sponsors to consult guidances directed toward those types of products. For considerations involving development of imaging drugs, see the guidance for industry *Developing Medical Imaging Drug and Biological Products (Parts 1, 2, and 3)*.

As part of the reauthorization of the Prescription Drug User Fee Act (PDUFA 4), we committed to certain performance goals (see letters from the Secretary of Health and Human Services to the Chairman of the Committee on Health, Education, Labor, and Pensions of the Senate and the Chairman of the Committee on Energy and Commerce of the House of Representatives, as set forth in the Congressional Record).[4] This draft guidance addresses one of these goals with the creation of a guidance document that addresses the "imaging standards for use as an endpoint in clinical trials." Although this guidance addresses imaging standards, it does not address the use of any specific imaging endpoints nor does it address a process of qualification of imaging biomarkers for use in clinical drug development. For issues that may be relevant to such a process, see the draft guidance for industry *Qualification Process for Drug Development Tools*.[5]

FDA's guidance documents, including this guidance, do not establish legally enforceable responsibilities. Instead, guidances describe the Agency's current thinking on a topic and should be viewed only as recommendations, unless specific regulatory or statutory requirements are cited. The use of the word *should* in Agency guidances means that something is suggested or recommended, but not required.

II. BACKGROUND

Imaging has long been used in therapeutic drug development, particularly in the early phases of drug development (e.g., phase 1 and phase 2 trials). More recently, imaging studies have been proposed for use in phase 3 trials, often as a component of the primary or secondary endpoints.

Imaging most commonly provides an assessment of human anatomy and/or physiology in the form of a pictorial assessment. If the clinical implications are not understood, simply generating an image may not confer benefit to a patient, and an outcome dependent on the interpretation of an imaging test may not be accepted by the Food and Drug Administration (FDA) as an appropriate endpoint for showing efficacy in a clinical trial. We addressed the evidentiary standards for imaging products in Parts 2 and 3 of the guidance for industry *Developing Medical Imaging Drug and Biological Products (Parts 1, 2, and 3)*. As stated in that guidance, acceptable indications for medical imaging agents include the following categories: structure delineation, disease or pathology detection or assessment; functional, physiological, or biochemical assessment; and diagnostic or therapeutic patient management.

[4] See Section A: PDUFA Reauthorization Performance Goals and Procedures Fiscal Years 2008 Through 2012 (http://www.fda.gov/ForIndustry/UserFees/PrescriptionDrugUserFee/ucm119243.htm).

[5] When final, this guidance will represent the FDA's current thinking on this topic. For the most recent version of a guidance, check the FDA Drugs guidance Web page at http://www.fda.gov/Drugs/GuidanceComplianceRegulatoryInformation/Guidances/default.htm.

In this guidance, we address the imaging standards for obtaining and interpreting medical images used to measure efficacy endpoints in confirmatory clinical trials. To illustrate the procedures applicable to imaging in a confirmatory clinical trial, we can divide imaging acquisition and interpretation standards into either a *medical practice* standard or a *clinical trial* standard, as follows:

- **Medical practice imaging standard.** For a medical practice imaging standard, the imaging acquisition and interpretation methods used in a clinical trial do not exceed those used in medical practice. For example, the imaging data incorporated into the clinical trial's final database may rely solely upon an investigator's response to a question about the report of a cardiac ejection fraction. This ejection fraction could be determined by any available medical practice method, depending upon the protocol's expectations (e.g., routine echocardiography or radionuclide imaging). Similarly, an adverse event report that cites a computed tomography finding of an intracranial hemorrhage is generally recognized as based upon the imaging standards typically used in the practice of medicine (i.e., a medical practice standard). A medical practice standard may prove useful for eligibility determination as well as safety monitoring and exploratory endpoint assessments in a confirmatory clinical trial. Sponsors are required to provide justification for the use of a medical practice standard when the imaging data form a component of a confirmatory trial's primary endpoint. The objective is to provide adequate assurance that the imaging methods for the assessment of the endpoint are well-defined and reliable.[6]

- **Clinical trial imaging standard.** With a clinical trial standard for image acquisition and interpretation, sponsors should address the features highlighted within the subsequent sections of this guidance. These features, including various aspects of data standardization, exceed those typically used in medical practice. A trial standard for image acquisition and interpretation is particularly important when an imaging outcome defines a primary endpoint in a phase 3 trial or when important quantitative outcomes are obtained from images. A clinical trial standard enhances the ability to detect a drug effect because of a reduction in the variability of the imaging data, and it also enhances the ability to verify data integrity.

In the following sections, we outline the topics sponsors should address when imaging is used within a clinical trial's primary endpoint to assess a drug's therapeutic efficacy. We emphasize here the nature of the processes that should be standardized. We further recommend that sponsors, in their materials being submitted for discussions with review divisions, describe specific technical aspects in great detail.

[6] See 21 CFR 314.126(b)(6).

III. INITIAL CONSIDERATIONS

Logistical and technical factors may limit the ability to use imaging in a confirmatory clinical trial, regardless of whether the trial relies upon a medical practice standard or a clinical trial standard for imaging acquisition and interpretation. The use of imaging within clinical trials may be limited by the availability of imaging technology. Some clinical sites may lack the resources to support a trial's imaging expectations. Similarly, the frequency of imaging and the distance to a qualified imaging facility may preclude or limit a patient's participation in a clinical trial. These factors may discourage the use of imaging in a clinical trial or limit the role of imaging within the trial. Nevertheless, imaging data may provide particularly persuasive evidence of a drug's bioactivity and also demonstrate a mechanism to help monitor drug effects in clinical practice. The following questions illustrate some of the factors a sponsor may wish to consider before proposing the use of imaging in a confirmatory clinical trial.

A. Why Use Imaging in a Confirmatory Trial?

Imaging may assist in the assessment of efficacy and safety as well as patient eligibility. The value of an imaging-based efficacy endpoint is dependent upon the investigational drug's proposed benefit, the nature of the underlying clinical condition, and the precedents for use of imaging in the specific drug development therapeutic area, as well as unique trial design features. Sponsors should consult with individual review divisions when considering the use of imaging to measure an important endpoint in a confirmatory clinical trial.

We anticipate that a medical practice standard for image acquisition and interpretation will prove sufficient for many clinical trial eligibility and safety assessments. However, in some situations, even if the use of imaging does not involve assessment of efficacy, the use of a clinical trial standard should be considered. For example, a clinical trial standard for image acquisition and interpretation would probably apply to the eligibility criteria for a clinical trial of a drug to be used solely among patients with certain quantitative nuclear imaging features of metastatic disease. In this case, detailed imaging methods may be needed to ensure that all patients meet the quantitative imaging expectations for enrollment. Indeed, clinical use of the drug might ultimately require the use of the specialized imaging technology.

B. Are Imaging Standards Important?

The use of imaging within a clinical trial necessitates some form of standardization. For many trials, a medical practice imaging standard alone is sufficient such that no imaging methods (beyond those typically used in medical practice) need to be described in the clinical protocol or supportive trial documents. The importance of the imaging-based eligibility criteria or outcome is the key consideration in determining the extent of imaging standardization needed for a clinical trial.

C. Is Centralized Image Interpretation Important?

The need for a centralized (*core*) image interpretation process is contingent upon the role of imaging within the trial. In situations where image interpretation results in measurements

representing important components of trial eligibility determination or safety or efficacy endpoints, and these measurements are vulnerable to considerable variability among clinical sites, a centralized image interpretation process is needed. A centralized image interpretation process also is critical to controlling bias in open label trials. In general, compared to a site-based image interpretation, the centralized process can better provide verifiable and uniform reader training as well as ongoing management of reader performance, ensuring that the process is accurate and that bias and variability are minimized.

There are, however, situations where a site-based image interpretation might provide sufficient assessment of the images, even when these data define the trial's primary endpoint. For example, a site-based image interpretation may be reasonable in a randomized, double-blinded clinical trial of an investigational therapeutic drug where the imaging technology is widely available, the image is easily assessed by a clinical radiologist, and the investigational drug has shown little or no evidence of unblinding effects.[7] In this situation, the use of randomization and blinding controls bias in image interpretation.

D. Should Image Interpretation Be Blinded to Clinical Data?

The extent of blinding of readers depends upon the role of imaging in the clinical trial. Blinding is of little importance for images used to determine clinical trial eligibility in a controlled trial, because randomization follows this determination. However, in single-arm trials even image-based eligibility should be blinded to clinical data because unanticipated factors may inadvertently bias image interpretations and select patients who are not representative of the desired patient population.

In some situations, image interpretations should be performed with no knowledge of clinical data, including date of the image acquisition or knowledge of prior imaging observations. In other situations, a primary endpoint may require integration of clinical data into an image interpretation (Sargent, Rubinstein, et al. 2009). This determination requires a solid knowledge of the underlying clinical condition and the precedent for the use of imaging within a primary endpoint, as well as multiple logistical considerations, but it is critical that the image interpretation can be blinded to knowledge of treatment.

E. How Often Should Imaging Evaluations Be Performed?

The timing of imaging evaluations depends upon the role and nature of the primary endpoint, the feasibility of the imaging schedule, and overall trial design features, including the potential for unscheduled (*off-protocol*) imaging and the potential effect of missing data upon the primary endpoint. For a primary endpoint that uses a time-to-event analytical approach, imaging evaluations should be performed at baseline and at sufficient frequency to provide a reasonably precise measure of the time to the expected clinical event. For example, imaging evaluations performed as infrequently as every 6 months may prove sufficient to assess progression-free survival among patients with a cancer known to have a slow progression and prolonged survival. However, in certain situations, relatively long intervals between scheduled imaging evaluations

[7] See the guidance for industry *Clinical Trial Endpoints for the Approval of Cancer Drugs and Biologics*.

might predispose the trial to bias if unscheduled imaging evaluations occur earlier in one of the treatment arms, potentially resulting in earlier disease detection (Amit, Bushnell, et al. 2010).

F. How Quickly Should Images Be Interpreted?

In clinical practice, images are typically interpreted on-site within minutes to several hours following acquisition for the purpose of clinical management. In a clinical trial, this *turnaround time* by a central image interpretation facility may prove impractical or inappropriate for the design.

The rapidity of image interpretation in a trial varies with the role of imaging in the trial. For example, when specialized, quantitative imaging is important for eligibility determination, a rapid turnaround time in image interpretation from a centralized image interpretation facility would be important for ensuring adequate enrollment. Inability to complete this turnaround on time may make the trial unfeasible.

Less urgency may accompany the turnaround time for image interpretation of efficacy endpoints, although these images too may need prompt evaluation by a centralized facility in certain trial designs. For example, the determination of cancer *progression* by a centralized image interpretation facility may be required to verify the appropriateness of initiation of a new therapy or cross-over administration of the investigational anti-neoplastic drug. Similarly, interim trial efficacy analyses that rely upon centralized image interpretations may necessitate a rapid turnaround in image interpretation.

G. What Procedures Should Be Standardized if Imaging Is an Important Aspect of a Clinical Trial Endpoint?

The procedures that should be standardized are determined by the role of imaging in the clinical trial. Therefore, no single set of detailed standards is readily applicable to clinical trials that differ in design and imaging objectives.

IV. BEFORE IMAGING: DEVELOPING A CHARTER

Sponsors should generally develop a document that provides a comprehensive, detailed description of the clinical trial imaging methodology if a trial standard for image acquisition and interpretation applies to the imaging data. We suggest sponsors refer to this document as an *imaging charter* and develop the document with the close vetting typically applied to the main components of a clinical protocol. Indeed, sponsors should generally regard the imaging charter as an integral component of the protocol, much as a statistical analysis plan is often developed as a component of a clinical protocol. The imaging charter can be attached to a clinical protocol as an appendix or developed as a section within the clinical protocol. For FDA review, we encourage submission of the charter simultaneously with a complete clinical protocol, including the final statistical analysis plan and any important supportive documents.

When imaging forms an important part of a clinical trial, we also encourage sponsors to discuss the imaging charter expectations at an end-of-phase 2 meeting. At this meeting, the sponsor can request advice on the development of an imaging charter and its role in a special protocol assessment submission.

Listed below are the suggested headings and subheadings for the elements within an imaging charter. Some of these elements may not apply to a particular clinical trial, while others may need considerable expansion to sufficiently describe the imaging methods. We encourage sponsors to list each of the elements within the imaging charter and elaborate upon the methods that address the element or briefly describe why the element does not apply to the trial. Compliance with the imaging charter can form an important aspect of the trial conduct verification process as well as the data quality assessment following completion of the trial.

Imaging technology rapidly evolves, can be highly technical, and varies markedly from measurement to measurement. For example, the technical specifications for obtaining reproducible echocardiographic measures of cardiac function profoundly differ from the methods essential to intercenter standardization of F18 fludeoxyglucose standard uptake value measures (Shankar, Hoffman, et al. 2006; Douglas, DeCara, et al. 2009). Imaging professional societies have developed or are developing publications that detail modality-specific standards and we encourage sponsors to become familiar with these documents when developing an imaging charter (Frank 2008; Boellaard, Oyen, et al. 2008). The complexity of technical standardization may preclude or markedly limit the use of imaging in a multicenter clinical trial even though the imaging methods have well-recognized value in clinical medicine (Keen, Mease, et al. 2010).

A. An Executive Summary of the Trial Design and the Role of Imaging in the Trial

Sponsors should summarize the role of imaging within the clinical trial and provide a detailed description of the imaging database variables (*deliverables*) to be incorporated into the analysis of the primary endpoint. Sponsors should describe how important trial design features may affect the proposed imaging database variables (e.g., procedures to minimize missing data, and to handle missing data in the analysis plan and plans for the use of off-protocol images).

Sponsors should provide an overview of the major aspects of the image acquisition, interpretation, and reader-defined deliverables to the sponsor. Presentation of a flow chart that identifies the specific steps in the process can be especially useful in summarizing the flow of the imaging information.

B. Image Acquisition Standards

Development of image acquisition standards involves having a broad knowledge of imaging modalities, including anticipation of imaging equipment upgrades or malfunction during the conduct of the clinical trial. In some situations, exploratory clinical trials may be needed to identify the most important imaging technical details, including those vulnerable to technical failure and charter noncompliance. For example, an explicit description of the imaging acquisition time may be critical when rapid dynamic cardiac arteriography is used to assess

cardiac function; in this situation, the X-ray energy (kVp) must be standardized and appropriate for imaging iodinated contrast agent within the heart. Similarly, optimization of X-ray energy is essential for breast imaging because a high kVp will obscure the signal intensity differences between adipose, glandular, or cancerous tissue and variations in kVp among clinical sites may increase variability in the imaging endpoint. The feasibility of maintaining technical consistency within and among clinical sites is particularly important when choosing and optimizing the imaging modality.

1. Equipment Standardization and Operation

The charter should identify the following.

a. Vendor-specific equipment/platforms (e.g., injectors, scanners, software)

The charter should identify the use of any investigational equipment. We recommend the use of only FDA-approved or cleared and marketed imaging equipment. The use of investigational equipment, including software, within a confirmatory clinical trial may necessitate special review and qualification considerations and, in some situations, may necessitate a process for obtaining FDA clearance or marketing approval of the equipment coincident with (or before) marketing approval of the investigational drug.

Sponsors should specify the important imaging equipment for the trial, including the imaging drug (contrast) injectors, scanners, and software. The importance of the equipment specifications varies with the role of imaging in the trial and may importantly limit the number of qualifying clinical sites. For example, imaging scanners may differ in technical details that can influence image quality, such as image reconstruction software programs and techniques for respiratory and cardiac gating, patient positioning, scan times, probe positioning, and technician-dependent procedures. We anticipate the need, in some situations, for detailed specification not only of the acceptable vendor-specific scanners but also of the model as well as any requisite upgrades to the equipment. We encourage the use of a tabular listing of the acceptable imaging equipment, including the key characteristics of the acquisition, processing, and display components of each scanner. Another approach could identify the physical benchmarks and testing parameters that must be met by the imaging equipment in accordance with a prespecified protocol for the acquired images to be used in the trial.

Most three-dimensional imaging currently requires the raw data to be processed using proprietary software algorithms. Unknown, unplanned, or inadvertent software upgrades may affect how images are generated. Changes in an image may be caused by these unknown software changes, and be incorrectly attributed to actual clinical changes. The charter should specify important software and also identify any situations when alternatives are acceptable.

Occasionally, requisite imaging equipment may become unavailable at a qualified site because of equipment malfunction or unavailability of technical support. In these situations, a clinical site can choose to substitute one imaging modality for another (such as magnetic resonance for computed tomography). The charter should identify the situations when these changes are acceptable. We anticipate that, in many situations, modalities will not prove interchangeable

(such as arteriography for ultrasound) when the endpoint assessment involves a quantitative imaging measurement. Indeed, ad hoc, unplanned interchange of modalities (including substitution of film for digitized imaging data) may compromise the objectives of a trial.

b. Equipment technical settings to be used at each site

The charter should state the technical settings for image acquisition for each type of important imaging equipment and identify any acceptable deviations from these settings. We encourage sponsors to identify these settings based upon the findings from exploratory clinical trials or other trials that attempted to standardize the technology among multiple clinical sites. Details critical to quantitative imaging, such as tomographic slice thickness, pulse sequence, and contrast agent injection time (especially for dynamic imaging), may importantly require specification.

c. The role of site imaging technicians in equipment operation, including identification of faulty or unacceptable images and the need to repeat imaging

The charter should describe the role of the imaging technician in the image acquisition process, including the minimum qualifications and the role of the technician, if any, in the initial assessment of image quality. Situations should be identified when repeat imaging is critical because of technical failure. In some situations, such as ultrasound imaging, detailed procedures should describe the technician's role in manipulation of the imaging probe and opportunity to deviate from these minimum expectations. Depending upon the imaging modality and the technical demands, the charter may need to describe a technician training process that will help ensure consistency in image acquisition.

d. Phantoms to be used for site qualification and image quality monitoring

We regard the use of phantoms (i.e., prespecified objects for scanning) as a critical part of site qualification and image quality monitoring during the conduct of a clinical trial. Phantoms can simulate a variety of conditions and have been developed for a range of imaging modalities (e.g., magnetic resonance, nuclear medicine, radiography). The choice of the specific phantom type depends upon the imaging objectives as well as the specific imaging modality. In general, we do not regard equipment specifications and image acquisition details as a sufficient substitute for the use of a phantom. Standardization of image acquisition using imaging and dosimetry phantoms will likely enhance the consistent performance of the imaging equipment during the course of the trial.

e. Patient preparation, positioning, and comfort measures

Many imaging modalities require specific patient preparation (e.g., fasting or special dietary limitations), positioning (e.g., supine, right lateral decubitus), preparation (e.g., removal of jewelry and eyeglasses), and comfort measures (e.g., ear plugs or sedation). These common aspects of imaging may vary markedly among clinical sites. Allowing significant site-to-site variations in patient preparation can result in unacceptable levels of image data variability. Patient preparation might also be based on patient-specific factors, such as age, weight, and

physical condition; the importance of standardization of these aspects may widely vary. For example, a trial conducted among pediatric patients may necessitate some form of sedation and description of the acceptable sedatives (including doses, route of administration, and potential for repeat dosing) may prove essential to quality imaging as well as the avoidance of missing images.

f. The date and time for imaging and alternatives

The charter should identify the planned dates and, if necessary, the times for imaging. In some situations, patients may need to be imaged at a specific time of day or night or following the development of certain clinical features (such as pain in a joint). The charter should describe these expectations and also identify the date and time windows that represent acceptable alternatives to the planned imaging evaluations.

g. Handling of off-protocol images

Patients in a clinical trial lasting many months are likely to undergo imaging examinations in addition to the ones intended to assess the response to therapy or to detect disease progression. The charter should specify the handling of these off-protocol images. In some situations, the off-protocol images are essential for inclusion within a trial's imaging database (e.g., liver computed tomographic images obtained in response to patient signs or symptoms that develop during a trial of an antineoplastic drug), whereas other situations may justify exclusion of these images from a trial's imaging database (e.g., hand radiographs obtained following a motor vehicle accident for a patient enrolled in a trial that assesses ultrasound peripheral artery intimal thickness).

h. Imaging risks

Imaging may involve many important risks to patients, such as exposure to radiation and contrast agents. The charter should describe these risks and specifically identify the radiation dose to be administered during imaging as well as the risks associated with administration of imaging drugs. Additional risks may relate to noise exposure, thermal energy, or magnetic fields. The charter should briefly describe the extent to which these risks are to be described in the trial consent process.

Occasionally, imaging detects incidental findings that are important for further clinical evaluation. Some of these findings may represent false signals of disease and expose patients to invasive evaluations that would have otherwise been avoided. Some of these findings may also provide the first important signals of a clinically important condition. The charter should identify the process for handling these situations, including the areas to be highlighted within the trial consent process. In general, we anticipate that all incidental imaging findings that may have clinical consequences will be reported to the patient and the patient's physician.

i. Site qualification process

The charter should describe the process used to qualify clinical sites for trial participation, specifically describing the tests to be performed to verify equipment performance, technical

support, and capability for compliance with the charter expectations. We anticipate that phantom imaging, on-site inspection, and training will provide sufficient site qualification for many trials. In some situations, the site qualification process may need to build upon these expectations by imaging patients as part of a qualifying clinical trial. These types of site qualification can be particularly important for highly technical imaging modalities or international trials that include countries where the imaging technology might be uncommon in clinical practice.

j. Acquisition quality control monitoring process

The charter should describe the plan for periodic, on-site quality control monitoring of imaging acquisition, storage, and transfer, including the plan for repetitive phantom imaging and the correction of deviations from the quality expectations. The importance and nature of this type of monitoring varies, depending upon the nature of the imaging technology, but, at a minimum, it will probably involve some form of episodic imaging quality reporting from clinical sites. In general, we anticipate the need for periodic on-site inspection by trial monitors to assess the imaging technical compliance of each clinical site or a subset of all the sites. Situations should be identified in which sites will be requalified or terminated because of failure to comply with image quality expectations. Any requalification procedures should be described.

k. Data storage, transfer, and site display

The charter should describe the expectations for imaging data storage, transfer to any separate facility (e.g., core laboratory or the sponsor) from the imaging site, and the plans for image display and interpretation at the clinical sites. In general, the charter should:

- Specify the storage of imaging data at the clinical site
- Describe any and all plans for transfer and storage of imaging data outside the clinical site
- Describe any image alteration procedures to be performed at the site (such as removal of all patient-identifying information)
- Specify the time period for storage of images at clinical sites and the format for data storage

2. Imaging Drug Standardization

Drugs are commonly used as a component of imaging and often require administration procedures intimately related to the scanning of a patient. Most notable are:

- Preparative drugs
- Contrast agents
- Radionuclide agents

The charter should identify the important aspects of drug selection, dosage, and administration for each of these agents, as exemplified below. When describing the drug doses, the charter should state that the drugs should be administered in accordance with approved labeling or state justification for alternative dose regimens.

a. Preparative drugs

The charter should identify acceptable and/or requisite pre-imaging drugs, including sedatives, stimulants, intravenous fluids, or contrast agents. In some situations, the drugs may need to be identified by brand name and, in most situations, by dosages and routes of administration. These specifications can be particularly important for trials that enroll pediatric patients and for the imaging of patients following administration of drugs that may alter images (such as drugs essential for cardiac stress testing). For international trials, the charter may need to identify nation-specific drug options.

b. Contrast agents

Many modality-specific contrast agents are not interchangeable and differ importantly in doses, techniques for administration, and risks. The charter should identify acceptable and/or requisite contrast agents, including specific brand names if essential. The charter should also identify the doses, routes of administration, rates of administration, and any special administration procedures (such as automatic injectors or administration times that may trigger scanning).

Some contrast agents can be safely administered only to patients with acceptable renal function or other characteristics. The charter should identify any laboratory tests and outcomes critical for supporting the administration of contrast agents.

c. Radionuclide agents

In addition to specification of the dose and route of administration, the charter may need to briefly identify the major drug quality features for any clinical trial radionuclide agents manufactured at the site. Unlike preparative drugs and contrast agents, some radionuclides (e.g., positron emission tomography (PET) agents) are commonly produced at clinical sites and the composition as well as the quality of these drugs may importantly vary from site to site. Standardization of these drug attributes may be important in achieving the trial's imaging objectives. The charter should identify any site-specific production considerations for site qualification.

C. Clinical Trial Standards for Image Interpretation

Image interpretation generally is carried out by trained readers, such as radiology and/or nuclear medicine specialists, who review and interpret, or *read*, images obtained in the course of a clinical trial. For the purposes of this guidance, terms such as *image interpretation*, *image review*, or *image read* are used interchangeably, and image readers are sometimes referred to as image reviewers.

The following elements pertain predominantly to the use of a core (centralized) facility for image interpretation in a clinical trial. Whether images are interpreted (or read) solely at the clinical site or at both the clinical site and a core facility, we regard these elements as important aspects to address within the charter.

1. Image Transfer, Receipt Documentation, and Initial Quality Assessment

The charter should identify the process for transfer of imaging data from each clinical site to the core image interpretation facility, including the plan for:

- Verification of the image technical adequacy as defined in the protocol
- Transfer of images and supportive information to the core facility
- The core facility process for querying sites for missing images, data, or imaging technical problems
- Obtaining repeat images of patients
- The logging of images received at the core facility, including the patient-specific tracking system
- The format for image data transfer (e.g., Digital Imaging and Communications in Medicine compact disc sent by courier)
- Digitization of received images or data
- Any technical evaluation (or pre-interpretation) or alteration of images, including de-identification of patient information, biasing marks, or other undesired image signals
- Monitoring compliance with the transfer, receipt, and initial image assessment process
- Correction of deficiencies and failures in the transfer, receipt, or initial image assessment process

The process should be highlighted for removal of all patient-identifying information from images relayed over electronic communication (e.g., Internet or laptop computers) or other pathways that are vulnerable to a security breach.

2. Image Display and Interpretation

The paradigm shift from film-based to filmless imaging has redefined clinicians' processes of image display, and interpretation of images within a clinical trial may critically depend on the quality of the displayed image. Image display in many digital systems is a flexible and dynamic process whereby radiologists directly interact with the soft-copy image, which is displayed on a computer workstation. The hardware component of a display system is usually composed of a

display device and a display driver or graphics card. The specifications given for a system are valid only for that particular combination of devices. Another important aspect of the display system is the hardware and software components used for maintaining the display presentation mapping between image values and luminance levels under a desired calibration model. Information regarding the calibration hardware, software, and procedures, including frequency and nature of the performed tests, should be identified.

a. Selection of images for interpretation, display sequence, and randomization

The charter should identify the nature and extent of images to be interpreted (e.g., all scheduled images as well as off-protocol images) as well as any important sequence aspects (e.g., baseline images followed by subsequent time point images). The appropriateness of excluding images from the interpretation (read) process should be emphasized and justified. The charter should prespecify the following:

- Criteria for classifying an image as a *technical failure* or other classification that leads to the exclusion of an image from the interpretation process

- The qualification of individual(s) who are to make the determination of whether an image is included or excluded in the reading queue

- If individual(s) other than the actual image interpreters have the responsibility of excluding certain images from the interpretation process, whether the image interpreters can also determine that an image is uninterpretable and the criteria used to make this decision

- Criteria for excluding images from the analysis and how missing imaging data will be accounted for (imputation scheme)

If images (or image sets for a patient at any specific time point) are to be randomized for display to readers, the charter should describe the randomization process. For example, one trial's image interpretation process may involve the time-sequential presentation of a patient's complete image set (from baseline through the follow-up evaluations) while another may involve the randomization of a patient's single time point image sets among those for many other patients' image sets. The randomization process is a key component of the overall image assessment plan.

b. Number of readers and their background qualifications

Sponsors should identify the number of image readers and their requisite background qualifications. In development of the plan, sponsors should consider:

- The extent of technical knowledge essential to image interpretation

- The avoidance of any other reader involvement in the clinical trial

- The avoidance of reader financial conflicts of interest with the sponsor[8]

- The need for confidentiality of image reads and/or the reading process

- The potential for reader fatigue and the need for substitute readers

- The time commitment of readers and reader availability

- Any unique considerations for identification of an adjudicating reader

- Any need for clinical readers (i.e., image interpretation by clinicians aware of non-imaging clinical trial information)

- The compensation plan for readers and avoidance of a compensation plan that may compromise or bias the quality of the read

The plan for documenting reader qualification should be described, including attestation of the extent of any conflicts of interest. The guidance for industry *Financial Disclosure by Clinical Investigators* describes the types of financial disclosure information, as well as the format, for submission within a marketing application.

c. Reader training and qualification

The reader training process should be described, emphasizing the use of any specific training materials (e.g., a training manual or training images), image display training sessions, any image read testing process, and the training documentation process. The origin (e.g., other clinical trials) of training images should be described, especially any images of patients anticipated for enrollment into the confirmatory clinical trial. In addition, the charter should prespecify whether any performance criteria will be used to qualify readers after training.

Sponsors should consider the importance of the following items in the development of the reader training process:

- **An overview of the major goals of the image interpretation.** In general, reader training should emphasize only the image-specific aspects of the image interpretation process unless the process also requires the integration of clinical information into the image interpretation process. The process should also minimize the potential for introduction of bias into image interpretation through knowledge of any potential image signatures that may break the desired blind to treatment assignment (e.g., if a PET ligand uptake is more common among the elderly, the co-registration of PET-computed tomography may bias the PET assessment because of recognition of aging-related cerebral atrophy on the tomogram).

[8] Under the applicable regulations (21 CFR parts 54, 312, 314, 320, 330, 601, 807, 812, 814, and 860), a sponsor is required to submit to the FDA a list of clinical investigators who conducted covered clinical trials and certify and/or disclose certain financial arrangements. Additional information is available in the guidance for industry *Financial Disclosure by Clinical Investigators*.

- **An overview of the major expectations for image manipulation, lesion measurement, and other image evaluations.** Readers may need special training in computer-assisted interpretation, measurement, or other analysis tools, and the process for performing and recording measurements, especially if this process involves unique software data lock features and password-protected features. The reading process may require knowledge of unique assessment tools, such as Response Evaluation Criteria in Solid Tumors (RECIST) outcome expectations (Eisenhauer, Therasse, et al. 2009). The charter should describe these expectations in detail and address situations when images may not be conducive to the requisite lesion measurement or other tool expectations.

- **Identification of any unique read definitions and/or criteria, including the use of image case report forms.** Some clinical trials may require predefined criteria for reads (e.g., identification of the specific basis for an unreadable image) and these criteria may differ from commonly used clinical criteria. Training and verification of training (with mock image reads) may be important in documenting reader proficiency.

- **Description of any reader retraining procedures.** Some image interpretation processes may include the use of test images intermixed among the clinical trial images such that readers are intermittently tested as to the proficiency and/or consistency in their reads. Failure to sustain proficiency may necessitate replacement of a reader with another trained and qualified reader. The charter should describe the reader testing and retraining or replacement procedure.

d. Timing of image reads and the read process

The charter should describe the timing of image reads with respect to the clinical trial conduct. In some situations, prompt interpretation of images is important for determining trial eligibility. In other situations, images are interpreted only following completion of all patient evaluations. Perhaps most commonly, readers can interpret images in batches periodically during the trial. If readers interpret images in batches, the size of the batches should be specified and the batch size justified. The allowable time interval between the batch sessions should also be predefined.

The charter should provide a detailed description of the image review process. Among many other items, the following should be identified:

- The review setting (e.g., a room with a controlled lighting system that allows for minimizing ambient illumination to a certain level, with eight computer display panels of a certain size and available only to the reader)

- Whether readers interpret images independent of any other individuals; if not, the individuals who may be present during the read should be specified and their role in image interpretation described; any consensus read process should be detailed

- A description of any image adjudication process

- Detailed use of any clinical information in the read

- A definition of the read outcome information to be described on case report forms and any special procedures in this process (e.g., an initial read followed by a redisplay of images to form a global reassessment)

- The assessment tools and qualitative and/or quantitative measurements to be performed during the image read (e.g., modified RECIST criteria assessment of each image set)

- A description of any computer software or other electronic processes involved in image interpretation, such as an automatic calculation of progression

- Any lesion tracking system (e.g., certain requisite target lesions), particularly any nuances related to the appearance of new lesions for tracking or inability to identify any previously tracked lesions (e.g., imaging problems or lesion resolution)

- Options and/or requirements for image manipulation, including application of calibers, zoom, pan, alteration of window/level, and application of spatial features and adjustment of contrast or image inversions

- A description of any process for re-read of images

- The reader's role in citation of missing images or technical deficiencies within the images

- A description of the plan to ensure that all original read outcome information is locked and available for subsequent verification and comparison to any re-read outcomes

When developing the image display process, sponsors should consider, as appropriate for the chosen modality, the key performance characteristics of medical displays such as luminance range, viewing angle, contrast ratio, reflection coefficients, grayscale, spatial and temporal (for image stacks) and color resolution, and spatial and temporal noise. The charter may need to specify these details as well as other modality-specific items, such as the process for displaying dynamic images in relation to static images and any software manipulations of images, for instance, for the minimization of degradations that may occur along the imaging process or transfer chain.

Computer-assisted image interpretation may form an important component of the read process and, in some situations, may generate all the information subsequently transferred to the imaging analytical database. The extent of computer assistance may vary widely but should be described explicitly within the charter, including a plan for quality-control checks upon any critical software functions. For example, the image interpretation may be driven primarily by a reader who then uses a computer-generated analysis tool to complement the reader's assessment. Such reliance on computer assistance can be algorithmic, with prespecified parameters for the use of a tool, or can be elective. In either case, such use should be defined within the charter in a manner that results in a sufficient audit trail. To evaluate for systematic errors, we suggest that a subset of computer-generated analyses be verified by blinded external readers.

Sponsors should use an FDA-approved computer-assisted interpretation tool or a tool justified for use with a given imaging modality (for additional advice on investigational devices, see section IV.B.1.a., Vendor-specific equipment/platforms (e.g., injectors, scanners, software)). If there is a specific tool that is required for image interpretation for assessment of response to therapy or other patient monitoring, the use of this tool might need to be included in the eventual labeling for the investigational drug. The same computer-assisted interpretation tool should be available to all readers at a centralized read.

e. Imaging case report forms

We anticipate the need for specific imaging case report forms for many clinical trials, particularly trials that involve quantitative imaging within endpoint construction. The charter should briefly describe the content of the case report form and emphasize the specific data content or notations that will subsequently be transferred to the sponsor to form the imaging database for the trial's endpoint analyses. We encourage the attachment of a case report form example to the charter. On this case report form, sponsors should denote the specific items to be transferred to the sponsor to form the imaging analytical database. In some situations, the case report form may consist of a tabular display of numbers (such as lesion measurements) or categories (such as predefined categories of bone erosion). An example of the tabular display within the charter may help lessen the potential for errors during the imaging flow process.

f. Imaging data lock process

At a predetermined point during the image review process, the image interpretation data (case report form information and any other important reader notations, including notations on images) generated by the readers should be *locked*. Locking data means that no further modification of image assessment is allowed. Predetermination of the data locking process and timing should be closely linked with the image read process. Data can be automatically locked by the imaging display equipment or triggered in response to reader notations. In some situations, the reading process may necessitate a re-read of previously interpreted images, including access to locked data. In all situations, the charter should describe the locking and any potential re-reads.

In general, we encourage the use of a sequential, locked approach to the read process whereby readers interpret the assigned image (or image set) and lock their read (e.g., lesion measurements, response category, lesion severity) such that the read outcome is documented and not altered.

g. Quality control of the image display and interpretation process

The charter should describe the process for monitoring compliance with the image display and interpretation process. This monitoring should include technical assessment of equipment, such as display systems and data locking software, as well as the reader interpretation process.

Digital test patterns for quality control purposes can be used on a daily basis to ensure consistent performance and to detect changes in the hardware or software that can degrade image quality.

In some instances, automatic luminance corrections might compensate for the reduction in luminance that is expected over time. Some of these quality control approaches offer the convenience of centralized reporting that facilitates the comparison of different display systems used in a given trial. In some circumstances, these automatic adjustment features may actually complicate measurements if they are unaccounted for. In either case, knowledge of such automatic compensation should be known and accounted for.

We recommend evaluating reader performance with defined and prespecified metrics. Evaluation should be ongoing during the interpretation process as well as retrospective.

Intra-reader variability as a measure of reader performance should, in many situations, be assessed by periodic blinded testing of the reader with a preselected set of images randomly interspersed with the clinical trial images. It is important that images from the trial being assessed are not used for reader testing. A drift in reader performance is not infrequently observed in clinical trials, therefore necessitating a periodic reader re-training and re-qualification. All details of reader testing, retraining, re-qualification, and possible replacement should be prespecified within the charter.

Image interpretation is inherently subjective. Therefore, inter-reader variability and the resulting need for adjudication are expected. The degree of variability among central readers leading to a certain adjudication rate observed in a given trial depends on multiple factors. Similarly, the same images might be interpreted differently by central as opposed to local readers at a clinical site. We recommend the use of quantitative measurement of reader variability as a valuable index of reader performance.

D. Charter Modifications Before Imaging

The charter should briefly describe the process for modifying the charter in response to potential deficiencies within the imaging process or need to improve the process. The plan for submitting charter modifications to the FDA and other regulatory authorities should be described. In general, we anticipate charter revisions to be uncommon, particularly if imaging has been used in exploratory clinical trials and the imaging processes follow precedents.

E. Imaging Data Transfer Process to the Sponsor

Image interpretation should result in the completion of a case report form and/or tabular display of numbers, measures, or categories of responses. The charter should describe the process for transfer of this information to the sponsor and the time point(s) for transmission of this information. The charter should describe how the sponsor will use the transferred information to establish the variables used in the analysis of the primary endpoint.

F. Archiving of Images and Image Interpretations

Images should be archived as a usual component of patient care as well as for use as the source documentation in clinical trials. Electronic source data should meet the same elements of data quality that are expected of paper records and should comply with all applicable statutory and

regulatory requirements. The FDA's acceptance of data from clinical trials for decision-making purposes relies upon verification of the quality and integrity of data, generally based upon the findings from audits and inspections.[9] In addition to images themselves, the image interpretations (case report forms or assessment tabulations) represent source data and should be retained for potential inspection and auditing. All source records, whether electronic or paper, must be retained (by the site investigator for site-specific information and by the sponsor for all trial information) for a period of no less than 2 years following approval of a marketing application or termination of drug development, as described in 21 CFR 312.57(c) and 21 CFR 312.62(c).

The charter should describe the process for archiving imaging information by the site investigator as well as the sponsor. In some situations, the sponsor may choose to archive the imaging at a core contractual facility or institution. Regardless of the physical storage route, the archiving process should address the following items:

- Limiting access to ensure images and data are retained in their original form
- Back-up storage
- Archiving in a manner conducive to a clear audit trail, including date and time stamps

Additional information regarding systems and personnel controls for computerized source data are described in the guidance for industry *Computerized Systems Used in Clinical Investigations*.

V. DURING IMAGING: MONITORING PLANS AND CHARTER MODIFICATIONS

A. Monitoring Plans

The charter should outline the complete plan for monitoring the imaging process. The extent of monitoring is anticipated to vary widely, dependent upon the use of imaging within a trial. In some situations, monitoring will be minimal, while in other trials, intense monitoring (to include requalification of equipment with phantoms and periodic retesting of readers) will be critical. Sponsors should comply with the monitoring plan described within a charter and verification of this compliance may prove an important component of the assessment of imaging data integrity.

B. Charter Modifications

During the clinical trial, circumstances may necessitate modification of the imaging procedures. For example, unanticipated technical features may obscure a portion of an image or preclude one of the expected quantitative assessments. In these situations, we anticipate the need to revise the charter to correct the problem and to maintain a record of the modification. The revision should identify any potential effect of the modification upon the trial's important endpoint analyses. In some situations, modification of the charter may affect the definition of the primary endpoint (e.g., alteration of the method for lesion measurement may call into question the clinical

[9] See the guidance for industry *Computerized Systems Used in Clinical Investigations* (http://www.fda.gov/regulatoryInformation/Guidances/ucm122046.htm).

meaningfulness of any size changes) and require reconsideration of the role of imaging in the trial as well as premature termination of the trial. To avoid these difficulties, we encourage sponsors to thoroughly consider the role of imaging (including the technical aspects) in a clinical trial, especially if the imaging is highly technical and/or relies upon quantitative assessments that require vigilant patient and site cooperation with the imaging process. The use of imaging in early phases of drug development may help lessen the challenges associated with wider use of the technology within confirmatory trials.

VI. AFTER IMAGING: DATA TRANSFER, ARCHIVING, ANALYSIS, AND INTERPRETATION OF IMAGING INFORMATION

A. Data Transfer

It is important for sponsors to document fidelity to the charter-specified process of imaging information transfer from a site to a core facility and from the core facility to the sponsor throughout a clinical trial. Many clinical trials are likely to require transfer of imaging data to the sponsor only following completion of all image assessments and interpretations and some may require image data modification, tabulation, or even reinterpretation of images before this transfer. For example, the sponsor may supply certain prespecified clinical information for readers to consider as they reinterpret images. In these unique situations, audit trails can be especially critical and will likely form an integral component of data quality assessment.

B. Archiving

Sponsors and investigators should comply with the charter-specified plan for imaging source data archiving. Deviations from this plan and/or loss of imaging information may compromise the ability of the FDA to verify data quality and/or necessitate reassessment of images. We do not accept images as a component of new drug applications or biologics license applications. However, we may require sponsors to display images during inspections of the core image laboratory, or in presentations to FDA review staff or for use on laptop computer screens by individual reviewers (21 CFR 312.58(a)).

C. Analysis and Interpretation of Image Information

We anticipate that most analyses of imaging information will be performed by the sponsor in accordance with the clinical protocol specifications. In some situations, clinical sites or a core facility may analyze certain aspects of imaging (such as the determination of reader interpretation consistency) as a quality control measure. Sponsors should specify these site and core facility roles in the charter. Clinical trial imaging data should not be analyzed in an ad hoc, unplanned manner.

Imaging processes that had taken place during the conduct of the trial, such as image acquisition, image interpretation, data transfer and other processes described in this guidance, should all be thoroughly presented in the final study report submitted for review to the FDA.

REFERENCES

Amit, O, W Bushnell, L Dodd, N Roach, D Sargent, 2010, Blinded Independent Central Review of the Progression-Free Survival Endpoint, The Oncologist, 15:492-95.

Boellaard, R, W Oyen, C Hoekstra, O Hoekstra, E Visser, A Willemsen, B Arends, F Verzijlbergen, J Zijlstra, A Paans, E Comans, J Pruim, 2008, The Netherlands Protocol for Standardization and Quantification of FDG Whole Body PET Studies in Multi-Centre Trials, Eur J Nucl Med Mol Imaging, 35:2320-2333.

Douglas, P, J DeCara, R Devereux, S Duckworth, J Gardin, W Jaber, A Morehead, M Picard, S Solomon, K Wei, N Weissman, American Society of Echocardiography Standards, American College of Cardiology Foundation, 2009, Echocardiographic Imaging in Clinical Trials: American Society of Echocardiography Standards for Echocardiography Core Laboratories: Endorsed by the American College of Cardiology Foundation, J Am Soc Echocardiograph, 22(7):755-65.

Eisenhauer, E, P Therasse, J Bogaerts, et al., 2009, New Response Evaluation Criteria in Solid Tumors: Revised RECIST Guideline (version 1.1), Eur J Cancer, 45:228-247.

Frank, R, 2008, Quantitative Imaging Biomarkers Allicance FDG-PET/CT Working Group Report, Mol Imaging Biol, 10(6):305.

Keen, H, P Mease, C Bingham, J Giles, G Kaeley, P Conaghan, 2010, Systematic Review of MRI, Ultrasound, and Scintigraphy as Outcomes Measures for Structural Pathology in Interventional Therapeutic Studies of Knee Arthritis: Focus on Responsiveness, J Rheumatol, (epub), 38(1):1-13.

Sargent, DJ, L Rubinstein, L Schwartz, JE Dancey, C Gatsonis, LE Dodd, LK Shankar, 2009, Validation of Novel Imaging Methodologies for Use as Cancer Clinical Trial Endpoints, Eur J Cancer, 45:290-99.

Shankar, L, J Hoffman, S Bacharach, M Graham, J Karp, A Lammertsma, S Larson, D Mankoff, B Siegel, A Van den Abbeele, J Yap, D Sullivan, 2006, Consensus Recommendations for the Use of F18-FDG PET as an Indicator of Therapeutic Response in Patients in National Cancer Institute Trials, J Nucl Med., 47(6):1059-66.

Appendix 2: Imaging Core Lab Lexicon

This lexicon was developed in conjunction with the Drug Information Association (DIA), the Pharmaceutical Research and Manufacturers Association (PhRMA), and the Food and Drug Administration (FDA). This public document was still in draft state as version 6.0 dated 26 December 2007. However, a significant amount of time and work had gone into this document, and it is therefore presented as a useful reference document with thanks to all the unacknowledged contributors.

Term/Definition

Adjudication or Adjudication Review The process of decision making that involves an independent party with the authority to determine a binding resolution.

Adjudication Rate The number of cases that is adjudicated divided by the total number of cases evaluated.

Anonymization The process of de-identification and further removal or ambiguation of information to reduce the probability of re-identification of the image despite access to other information sources.

Baseline Followed by Randomized Temporal Image Presentation The sequence of image presentation such that the baseline (earliest) time point is shown to the reviewer for the purpose of identifying regions of interest, such as selecting neoplastic masses as target lesions. Subsequent time points are presented in a random order with respect to the date.

Blinded Review The analysis of images to determine results of the testing in which the reviewer is unaware of any subject or site information.

Blinding A procedure in which one or more parties to the trial are kept unaware of the treatment assignments and other information that might introduce bias.

Burned-In Information Information that is part of the actual pixel data as opposed to present in the image header.

Comment Generally referring to a text field that can capture additional reviewer insight into the review process or reviewer thought processes. Comments are generally required when the reviewer indicates an image is Not Evaluable or Uninterpretable or their opinion differs from the derived response.

Computer-Generated Quantitative Image Analysis An analysis performed automatically by a computer with little or no human interaction using signal processing algorithms to quantify an image analysis result. This type of analysis should be deterministic (always produce identical output from the same input) or have low variability.

Confirmation Review Generally referring to a central review that occurs based on an "on-site" event. Confirmation reviews are associated with eligibility criteria, disease progression, or other events that may benefit from a third party confirmation.

Data Lock The point and method when the image analysis result(s) are considered final and are protected. This must be predefined in the analysis. Locking must not be construed to mean an assessment cannot be overturned as indicated by emerging data as long as (1) the process is predefined in the Independent Review Charter (2) the process is driven by data that, by design, emerges after the initial assessment and (3) there are adequate audit trails that can substantiate the changes.

De-identification The process of removing real patient identifiers or the removal of all subject demographics from imaging data for anonymization.

De-personalization The process of completely removing any subject-related information from an image, including clinical trial identifiers.

Derived Response An outcome measure algorithmically derived based on information from the image analysis result.

End of Review Data Lock In this scenario, the data are locked when the reviews of all the time points for the subject have been completed.

Evaluable or Interpretable Generally referring to image quality as assessed by the blinded reviewer. Based on presence or absence of necessary imaging and the associate image quality. The response generated when an image and/or time point can be assessed. Grounds for the assessment are commonly captured.

Human Interfaced Image Analysis Image analysis that is driven primarily by a human reviewer who may use computer-generated analysis tools to quantify an image analysis result.

Hybrid Randomized Image Presentation In this paradigm, the first stage of the assessment is fully randomized or the post-baseline scans are randomized. Once the results have been locked for each time point, the images are re-presented in known chronological order for reconsideration. Changes in any of the randomized assessments are tracked and highlighted in the final assessment. In within-patient-control trials (e.g., comparative imaging), images obtained before and after the test agent should be presented in randomized unpaired fashion. The minimum size of the randomization block necessary to minimize recall should be considered.

Image Analysis Procedures and processes that culminate in the generation of results, such as brain volume, cardiac output, or tumor response criteria. Reviews can be performed for eligibility, safety, or efficacy. The review paradigm may be context specific and dependent on the specific aims of a trial, the imaging technologies in play, and the stage of drug development, among other parameters.

Image Analysis Results Variables derived from the image review or quantitative or qualitative variables resulting from the image review. Such variables may be used to assess eligibility for study and treatment response, or information that results from or is produced by the image analysis or review processes (such as lesion selection and their associated spatial measurements), or from algorithmically derived assessments specified in the protocol. In this context, the term also refers to "marks" placed on images, such as regions of interest boundaries, annotations such as "Target Lesion 4," etc.

Image Header The part of the file or dataset containing the image other than the pixel data itself.

Image Review Plan or Radiology Technical Manual A document that summarizes the plan for the acquisition of imaging data.

Imaging Case Report Forms IRC-specific forms designed to capture elements of image acquisition and/or image interpretation and/or derived responses for enrollment and/or eligibility review and/or confirmation review and/or efficacy review.

Imaging Endpoint Endpoint based on objective image features chosen to evaluate the activity of a study treatment (e.g., retardation of joint destruction in patients with rheumatoid arthritis).

Imaging Examination A single set of intimately related images acquired contemporaneously with a single technology, such as a CT scan of the chest, a whole body bone scintigram, or an echocardiogram, as described in the Independent Review Charter.

Imaging Examination Level Data Lock In this scenario, the data are "locked" in "final form" after each imaging examination is assessed. The purpose is to assess the differential contribution of each imaging examination to the overall assessment.

Imaging Phantoms Devices used for periodic testing and standardization of image acquisition. This testing must be site specific and equipment specific and conducted prior to the beginning of a trial (baseline), periodically during the trial, and at the end of the trial.

Imaging Surrogate Endpoint Imaging endpoint that is correlated with a clinical outcome but is not sufficient to show clinical benefit.

Independent Review A review that is completely unaware of findings of other reviews (including findings from other blinded reviews or on-site reviews) and is not otherwise influenced by the findings of other reviews.

Independent Review Charter (IRC) An ensemble of text describing processes and procedures that govern the use of images in a clinical trial.

Individually Identifiable Information Data that alone or in combination may be used to identify an individual.

Interobserver Variability or Inter-Reviewer Variability The variability in the interpretation of a set of images by different reviewers.

Intra-Observer Variability or Intra-Reviewer Variability The variability in the interpretation of a set of images by the same reviewer after an adequate period of time inserted to reduce recall bias.

Investigator Assessment Image interpretation at the clinical site for patient care purposes.

"N" Time Point Data Lock In this scenario, a variable number of time points "N" can be combined and shown together at a particular stage of the review process. For example, the baseline/screening and the first subsequent time point after baseline/screening may be reviewed together to establish the baseline extent of disease.

Not Evaluable or Uninterpretable Generally referring to image quality. Based on presence or absence of necessary imaging and the associate image quality. Not Evaluable or Uninterpretable is the response generated when an image and/or time point cannot be interpreted. May be assessed in real time by a blinded third party quality assessor independently of image reviewer. Provision for reimaging (where

feasible) should be prespecified. Listing of criteria is provided and responses are captured in the imaging CRF.

Off-Protocol Imaging Imaging that may have been performed during a trial and should not be reviewed by the IRC or imaging which is performed during a trial but not required by the protocol.

On-Protocol Imaging Imaging that is performed during a trial as required by and defined in the protocol or imaging that is performed during a trial as required by and defined in the protocol that should be reviewed by the IRC.

Order of Image Presentation The sequence that images are presented to reviewers for formal review and generation of the image analysis results.

Personal Information Data related to person identification – see EU guidance (e.g., age).

Primary Review The blinded review of imaging data in which one or more independent reviewers review images to generate the image analysis result associated with the efficacy endpoint.

Pseudonymization The process of de-identification and replacement of identifiers with a pseudonym that is unique to the individual and known within the context of a trial but not linked to the individual in the external world.

Randomized Independent Temporal Image Presentation The sequence of image presentation that each time point is presented alone, in a random order with respect to the date of acquisition, and reviewed independently without access to other time points.

Scheduled Imaging Imaging that is performed during a trial at one or more of the time points (or window assigned to a time point) designated for imaging assessment in the protocol. Applies to either on-protocol or off-protocol imaging or imaging examinations that are routine assessments.

Secondary Review A blinded review of imaging data in which one or more independent reviewers review images to generate outcome data that is not part of the efficacy endpoints. An example would be a review that is part of intra-reviewer variability analysis.

Sensitive Personal Information Data related to personal preferences and disposition. – see EU guidance (e.g., Ethnicity).

Sequential Chronologic The sequence of image presentation in the order in Image Presentation which they were actually acquired. In this format, the reviewer should not know the total number of time points to be assessed unless that information has been prespecified in the Independent Review Charter. (For example, prespecification is usual and customary in imaging studies of neurodegenerative disorders, arthritis, osteoporosis, and congestive heart failure, among others.)

Sequential Unblinding Consecutive interpretation of images with and without clinical information (e.g., demography, clinical assessments).

Simultaneous Chronological The sequence of image presentation that all images Image Presentation associated with a subject are shown to the reviewer at the same time without blinding the date or sequence or total number.

Simultaneous Randomized Temporal Image Presentation The sequence of image presentation that all images associated with a subject are shown to the reviewer at the same time in a random order with respect to the date but without blinding to total number.

Statistical Analysis Plan for Medical Imaging Analysis plan focused on primary efficacy analysis and including statement of null hypothesis, study power, statistical test, efficacy population, handling of missing or uninterpretable images, and sensitivity analyses.

Time Point A discrete period during the course of a clinical trial when groups of imaging examinations are scheduled as defined in the study protocol.

Time Point Data Lock In this scenario, the data are locked after all prespecified information associated with each time point is assessed. In some paradigms, the time points are known to be presented in chronological order in others, the time points may be randomized during the early stages of the image analysis process (vida infra).

Truth Standard Provides an independent way of evaluating the same variable being assessed by the investigational medical imaging agent. Believed to give the true state of a patient or true value of a measurement. Used to demonstrate that the results obtained with the medical imaging agent are valid and reliable and to define summary test statistics (sensitivity, specificity, positive and negative predictive value).

Unique Identifiers (UIDs) Globally unique identifier used to identify images, sets of images, or components within an image.

Unscheduled Imaging Examination that is performed during a trial at a time/date outside the window assigned to a time point designated for imaging assessment in the protocol. It may be ad hoc imaging performed to evaluate an unscheduled clinical circumstance. It may be an on-protocol or off-protocol examination.

Appendix 3: Quantitative Imaging Biomarkers Alliance (QIBA)

The information provided here is taken directly from the QIBA WIKI: http://qibawiki.rsna.org/index.php?title=Main_Page. This and additional information may also be found at http://rsna.org/QIBA_Protocols_and_Profiles.aspx.

QIBA Mission

Improve the value and practicality of quantitative biomarkers by reducing variability across devices, patients, and time.

QIBA is an initiative to advance quantitative imaging and the use of imaging biomarkers in clinical trials and clinical practice by engaging researchers, healthcare professionals, and industry. This involves:

Collaborating to identify needs, barriers, and solutions to develop and test consistent, reliable, valid, and achievable quantitative imaging results across imaging platforms, clinical sites, and time

Accelerating the development and adoption of hardware and software standards needed to achieve accurate and reproducible quantitative results from imaging methods

QIBA develops profile specifications and coordinates the necessary research and qualification groundwork.

QIBA Modality Committees

QIBA currently has six active technical committees:

- Perfusion, Diffusion, and Flow-MRI tech
- fMRI tech
- FDG-PET tech
- CT Volumetry
- COPD-Asthma tech
- Ultrasound SWS tech

Index

C.G. Miller et al. (eds.), *Medical Imaging in Clinical Trials*,
DOI 10.1007/978-1-84882-710-3, © Springer-Verlag London 2014