1

Concepts: Measurement in MR[1]

Paul Tofts

Department of Medical Physics, NMR Research Unit, Institute of Neurology, University College London, Queen Square, London WCIN 3BG, UK

1.1 INTRODUCTION

1.1.1 Measurement Science and MRI Come Together

Measurement science has been around a long time; MRI[2] has been around for about 20 years. This book is about the blending of the two paradigms.

We have come to expect to be able to measure certain quantities with great accuracy, precision and convenience. Instruments for mass, length and time are all conveniently available, and we expect the results to be reproducible when measured

[1] Reviewed• by ... Sections of this chapter were reviewed by ... Jeffrey T. Yap, Department of Radiology, University of Pittsburgh, Pittsburgh, PA, USA.

[2] Magnetic resonance imaging (MRI) is a term invented by US radiologists to describe nuclear magnetic resonance (NMR) imaging. The phenomenon of NMR is described in Chapter 6. The 'nuclear' part was removed from the name NMR to prevent the public being alarmed. Spectroscopy (Chapter 9) was originally concerned with identifying chemical compounds, and there was no spatial information contained in the data. It developed separately from imaging, on different machines, and is often referred to as MRS. Modern MRS is carried out largely on MRI machines, and uses the imaging gradients to localize the spectra to particular parts of the body. For these reasons, MRI is now considered to include spectroscopy. MR is a more correct term, and refers to MRI and MRS together.

Quantitative MRI of the Brain. Edited by Paul Tofts
 ISBN: 0-470-84721-2

again, and also to be comparable with measurements made by others in other locations. In the human body we expect to measure some parameters (height, weight, blood pressure) ourselves, recognizing that some of these parameters may have genuine biological variation with time. More invasive measurements (e.g. blood alcohol level or blood sugar level) are also expected to have a well-defined normal range, and to be reproducible. In physics, chemistry, electrical engineering and manufacturing there is a strong tradition of measurement, international agreements on standards and training courses for laboratory practitioners. International standards of mass, length and time have been in existence for many years. Secondary standards have been produced which can be traced back to the primary standards. National and international bodies provide coordination.

As individual scientists we may have a passionate desire to use our talents for the benefit of mankind, preferring to devote our energy to finding better ways of helping our fellow humans to be healthy than to improving weapons for their destruction. In this context, developing measurement techniques in MRI constitutes a perfect application of traditional scientific skills to a modern problem.

MRI is now widespread, and accepted as the imaging method of choice for the brain (and for many body studies). It is generally used in a qualitative way, with the images being reported non-numerically from light boxes[3] by radiologists. Many MRI machines now have independent workstations connected to the scanner and the database of MR images which enable and encourage simple quantitative analysis of the images in their numerical (i.e. digital) form. However the data collection procedure often prevents proper quantification being carried out; machine parameters such as transmitter gain flip angle value (and its spatial variations), receiver gain and image scaling may all be acceptable for qualitative analysis, but cause irreversible confusion in images to be quantified. Researchers may be unaware of good practice in quantification, and collect or analyse data in an unsuitable way, even though the MRI machine is capable of more.

The process of quantifying, or measuring, parameters in the brain necessarily takes more time and effort than a straightforward qualitative study. More MRI scanner time is needed, and considerable physics development effort and computing resources may be needed to set up the procedure. In addition, analysis can be very time-consuming, and support of the procedure is required to measure and maintain its reliability over time. Procedures have to be found[4] which are insensitive to operator procedure (whether in the data collection or image analysis) and to scanner imperfections (such as radiofrequency nonuniformity from a particular head coil), which provide good coverage of the brain in a reasonable time, and which are stable over study times which may extend to decades.

The benefits of quantification are that fundamental research into biological changes in disease, and their response to potential treatments, can proceed in a more satisfactory way. Problems of bias, reproducibility and interpretation are substantially reduced. MRI can move from a process of picture-taking, where reports are made on the basis of unusually bright, dark, small or large objects, to a process of measurement, in the tradition of scientific instrumentation, where a whole range of quantities can be tested to see whether they lie in a normal range, and whether they have changed from the time of a previous examination.

In this book, the intention is to merge these two traditions, or paradigms, of measurement and of MRI to form the field of quantitative MRI, or qMRI. The MRI measurement process is analysed, often in great detail. Limits to accuracy and precision are identified as far as possible, with the intention of identifying methods that are reliable and yet practical in a clinical MRI scanning environment.

[3] A light box illuminates from behind a film (size approximately 14×17 inches, or 35×43 cm), which may contain traditional X-ray images or several MR images. As MRI produces progressively more slices per study (a three-dimensional image dataset may contain 128 slices), the desire to print these all onto a few films has resulted in progressively smaller images, in which the relevant detail cannot be seen without using a magnifier.

[4] The author's website (www.ion.ucl.ac.uk/nmrphysics/index.html) contains more links and references to qMR.

The biological meaning of the many MR parameters that are available is explored, and many clinical examples are given where MR parameters are altered in disease. Often these changes have been observed qualitatively, and they serve to encourage us to improve the measurement techniques, in order that more subtle effects of disease can be seen earlier than is currently possible, and in tissue that is currently thought to be normal as judged by conventional MRI. The ideal is to obtain push-button (turnkey) techniques for each of the many MR parameters in this book, such that an MRI radiographer (technologist) can measure each of these parameters reliably and reproducibly with a minimum of human training or intervention, in the same way that we can currently step onto a weighing machine and obtain a digital readout of our mass. In the case of qMR the output would be considerably richer, perhaps showing images of abnormal areas (computed from large databases of normal image datasets), changes from a previous MRI exam, possible interpretations (diagnoses), and an indication of certainty for each piece of information. The advances in the pre-scan and the spectroscopy MR procedures, which used to be very time-consuming and operator-dependent and are now available as fully automated options, show how this might be possible.

Thus MRI may be undergoing a *paradigm shift*[5] in how it is viewed and used. In the past it was used for forming qualitative images (the 'happy-snappy MRI camera', taking pictures); in the future it may be increasingly used as a scientific instrument to make measurements of clinically relevant quantities. The dichotomy can be seen in this book. Clinical descriptions will often speak of signal hyperintensity in a sequence with a particular weighting, whilst elsewhere (idealized) physical measurement methods are described, with talk of localized concentration values, normal ranges, age and gender effects, and reproducibility. As measurement becomes more precise, and analysis enables clinically relevant information to be extracted from myriad information, in will become possible in principle to make measurements on an individual patient to characterize the state of their tissue, guiding the choice of treatment and measuring its effect. The issues involved in bringing qMR into the radiological clinic are well summarized in an Editorial in the *American Journal of Neuroradiology* (McGowan, 2001).

As part of this ongoing paradigm shift, our view of what MRI can tell us is changing. When it started, information was largely anatomical (*anatomical MRI*), in the sense that relatively large structures would be observed. Changes in their geometric characteristics (usually size), compared with normal subjects or a scan carried out in previous weeks or months, would be noted. Quantitative examples would be volume and atrophy. *Functional MRI* (fMRI) claimed the complementary ground, studying short-term changes in tissue arising from carrying out particular (neural) functions. *Micro-structural MRI* occupies a third role, as shown in this book. Many MR parameters [such as diffusion, magnetization transfer ratio (MTR), spectroscopy] show structural changes in tissue arising from damage caused by disease. To observe these changes directly would require imaging resolution of the order of 1–100 μm,[6] since they generally involve a variety of biological changes at the cellular level. These can be observed by pathologists in post-mortem tissue, using optical or electron microscopy and special staining techniques (histopathology). This resolution is well below the spatial resolution of MRI (which is about 1 mm). However changes at the microscopic level (e.g. in cellular structure) give changes in the MR parameters (e.g. in water diffusion); these can be observed at coarser spatial resolution (of about 1 mm). Thus structural changes of sizes well below those that would be called anatomical can be detected. In addition, the concentrations of chemical compounds (metabolites) in cells, and their changes, can be measured with spectroscopy. The physiological permeability of the endothelial membrane around blood

[5] Thomas S. Kuhn, in *The Structure of Scientific Revolutions* (3rd edn, University of Chicago Press, Chicago, IL, 1996), first introduced the idea of paradigm shifts. An example would be the move from a classical physics to a quantum physics view of the world. A paradigm is a pattern or model, a way of viewing the world or part of it, a point of view, a mindset.

[6] 1 micron (μm) is 10^{-3} mm or 10^{-6} m.

vessels can be measured using dynamic imaging of gadolinium (Gd)-contrast agent. These microstructural changes are generally more quantitative than fMRI in terms of their reproducibility and how well we can relate them to underlying physiological changes.

These changes may occur both in a 'lesion', which is tissue seen at post-mortem and in conventional MRI to be visibly different from the surrounding tissue, and in the 'normal-appearing' tissue, which appears normal at post-mortem[7] and in conventional MRI. Lesions are usually described as 'focal', meaning that the change is localized to a relatively small area (a few mm or cm) with a distinct boundary; thus its different brightness in an image distinguishes it from the surrounding tissue (considered normal). In contrast, a diffuse change may extend over more area, has no distinct boundary, and is harder to detect by simple visual observation of the image. Diffuse changes are often well characterized by quantification, since this is the absolute value of quantities within the area that is measured, without reference to surrounding tissue, or the need for a distinct boundary.

1.1.2 Limits to Progress

It may appear that qMR research proceeds under its own impetus. However the current state and rate of progress in developing reliable qMR methodology are determined by several factors: MRI manufacturers, research institutions, pharmaceutical companies, computer technology and publicly funded research councils.

MRI machine manufacturers (vendors) will take on some of the measurement procedures over time, incorporating them into their research and development programmes, and then offering them as turnkey (push-button) products. The speed of this process is driven by demand from clinical purchasers, by whether competing manufacturers offer such facilities, and by whether public medical funding bodies such as the US Food and Drugs Administration (FDA) is likely to approve reimbursement of the cost of such procedures from medical insurance policies. The existence of a large and growing installed base of high-quality, reliable and ever improving MRI machines, primarily designed for routine clinical use, largely in environments where they can be run as parts of profitable businesses, has enabled and encouraged the development on these machines of qMR techniques, which are still of interest to only a minority of users.

Research institutions have particular structural strengths and weaknesses. qMR needs input from chemists, computer scientists, neurologists, physicists, radiologists and statisticians. There may be good career support for those applying methods to study clinical problems, but none for those basic scientists inventing and developing the methods. There may be a clash of paradigms or traditions between those who have been educated in a hierarchical environment where asking questions is considered to be irrelevant or subversive, and those who consider asking questions to be an absolute basic necessity of undertaking modern high-quality scientific research. The availability of talented researchers in turn depends on how much value is placed on science in society, schools and universities, and whether appropriate postgraduate training opportunities exist. The International Society for Magnetic Resonance in Medicine (ISMRM) is a powerful force bringing together researchers from different institutions who are working on similar methodologies, through both its journals and its scientific meetings.

The demand from pharmaceutical companies and neurologists for qMR measurements to be used in drug trials is large and likely to increase (Miller, 2002; Filippi and Grossman, 2002; Filippi *et al*., 2002; McFarland *et al*., 2002). The traditional double-blind placebo-controlled phase III trial involves many patients (typically 100–1000), who are studied for several years in order to obtain enough statistical power to determine whether a drug is effective. The large sample size is needed to deal with the variability of disease in the absence of treatment, and the imperfect treatment effect (which may vary according to patient subgroup). Such trials typically cost several US$100 million. qMR can potentially shorten the procedure

[7] However, pathologists do report cellular abnormalities in normal-appearing white matter in multiple sclerosis.

by identifying treatment failures early on in the testing process, on a smaller sample. If there is no observed biological effect from the treatment, it may be considered unlikely that the drug is working (this will depend on the particular way the drug has been postulated to act). For example, if a potential treatment for multiple sclerosis (MS) showed no effect on all the MR measures that are known to be abnormal in MS, it would probably be dropped in favour of other drugs. With new biotechnology and gene-based treatments being developed, the number of candidate drugs for evaluation will increase by a large factor, and traditional trials will become too expensive and slow to evaluate all of them. Thus direct *in vivo* qMR observation of treatment effect will become increasingly valued.

The rapid increase in power and availability of computing technology has also been key in enabling data acquisition and image analysis techniques to be realized. Numerically designed magnets, coils and radiofrequency pulses, digital receivers and rapid image registration and analysis have all changed the way that MRI is carried out.

The resources available from pharmaceutical companies to drive the process of developing and supporting reliable qMR measures may exceed those available from traditional publicly funded research sources. Traditional research council sources have been willing to support the application of qMR methods to study particular diseases, but often unwilling to support the development of new quantitative methods, sometimes claiming that MRI manufacturers should be doing this.

1.1.3 Using this Book

This book can be used in many ways. Those interested in each MR parameter can read each chapter in turn. Physicists will be more interested in the details on how to implement measurement techniques for that parameter, and what can go wrong in a practical situation. Attempts at multicentre studies pinpoint the most important issues that can prevent good quantification. Clinicians interested in a particular disease can look this up in the index, and find sections, often substantial, in most chapters that define how MR parameters have been used to characterize that particular disease. The lay reader, with a basic knowledge of science and mathematics, may wish to skim through the book, perhaps concentrating on the boxes, figure, tables and their accompanying captions. These are intended to tell their own story of qMR, without the detail of the full text. Each section on physical principles, or at least the first part of each one, is intended to be accessible to a non-MRI specialist.

In this chapter the issues in measurement that occur repeatedly throughout the book as each MR parameter is considered are examined in more detail. These are grouped into the processes of data collection, data analysis and quality assurance, all of which crucially affect how well MR quantities can be measured. Units are usually given in SI (System International), and conventions used in this book for physical units and symbols (e.g. TR, TE, T_1, T_2) are those recommended in the style guide for the journal *Magnetic Resonance in Medicine*, published for the ISMRM.[8] Most of the focus is on techniques which can be implemented on standard clinical MRI scanners; techniques (e.g. ^{31}P spectroscopy or ^{23}Na imaging) which use nonstandard hardware have not generally been considered.

The reader should ideally have basic knowledge of how MRI works, and some knowledge of the brain, including the major diseases (cancer, epilepsy, stroke, MS, dementia). Basic books[9] that can be recommended are given in Table 1.1. One MRI book and one or two brain books will provide the background that is needed. Several of the brain books are paperback course texts that have gone into several editions and provide very good value.

1.2 HISTORY OF MEASUREMENT

Early quantitative techniques focussed around the desire to measure distance, mass, monetary value

[8] See www.ismrm.org/journals.htm

[9] Websites can also be very useful; the ISMRM site is a good place to start: www.ismrm.org/

Table 1.1. Recommended books for background information on brain MRI

Title	Authors	Date published	Number of pages	Description
Magnetic Resonance Imaging Physical Principles and Sequence Design	E.M. Haacke, R.W. Brown, M.R. Thompson, R. Venkatesan	1999	914	Thorough exposition of MRI principles
Principles of Nuclear Magnetic Resonance Microscopy	Paul T. Callaghan	1993	510	Excellent description of principles of NMR and imaging. Paperback
MRI from Picture to Proton	D.W. McRobbie, E.A. Moore, M.J. Graves, M.R. Prince	2002	350	Written by physicists; not seen
Magnetic Resonance Imaging	David D. Stark, William G. Bradley	1998	2800 (3 volumes)	A marathon tour of MRI, much of it clinical
Functional MRI	Peter Jezzard (Editor), Paul M. Matthews (Editor), Stephen M. Smith (Editor)	2001	404	Contains good descriptions of fundamentals
Barr's The Human Nervous System	John A. Kiernan	1998	518	Includes complete description of the brain 7th edn, paperback
Concise Text of Neuroscience	Robert Kingsley	1999	679	Approachable paperback, 2nd edn
The Nervous System	Peter Nathan	1997	342	Neurology in a paperback, 4th edn
Clinical Neuroanatomy and Related Neuroscience	M.J.T. Fitzgerald, J. Folan-Curran	2001	336	Paperback, 4th edn
Biomedical Imaging, Visualization, and Analysis	R.A. Robb	2000	356	General book on principles
Medical Imaging Physics	William R. Hendee and E. Russell Ritenour.	2002	512	Comprehensive hardback, 4th edn
Physics of Diagnostic Imaging	D.J. Dowsett, P.A. Kenny, R.E. Johnston	1997	609	Covers all imaging methods
Physics for Medical Imaging	R.F. Farr, P.J. Roberts, J. Weir (Editor)	1996	288	Gives basics, written by physicists for radiologists
MRI: the Basics	Ray Hashemi and William G. Bradley	1997	360	Good introduction for nonphysicists
Questions and Answers in MRI	Allen D. Elster, Jonathan Burdette	2001	352	Simple approach by two MDs
MRI for Technologists	Peggy Woodward	2000	432	For radiographers (technologists)
MRI Principles	Donald G. Mitchell	1999	288	Basics described by a radiologist
MRI: Basic Principles and Applications	Mark A. Brown and Richard C. Semelka	1999	222	Not seen (new edition 2003)
Handbook of Medical Imaging	Isaac Bankman	2000	1000	Excellent overview of many image analysis techniques

and time. Developed in about 3000 BC in ancient Egypt, the cubit was a ubiquitous standard of linear measurement, equal to 524 mm. It was based on the length of the arm from the elbow to the extended fingertips and was standardized by a royal master cubit of black granite, against which all cubit sticks used in Egypt were to be measured at regular intervals.[10] The precision of the thousands of cubit sticks used in building the great Pyramid of Giza is thought to have been very high, given that the sides of the pyramid are identical to within 0.05 %.

Early astronomers developed remarkably precise measurement methods (as demonstrated at Stonehenge); their ability to guide navigation and predict eclipses brought them fame. In the sixteenth century, precise calculations of planetary orbits by Copernicus, Kepler and Galileo challenged the intellectual dominance of the Catholic Church, bringing an end to the idea that all heavenly bodies rotate around the earth.

In 1581 the word 'quantitative' was first used, meaning 'involving the measurement of quantity or amount'. Quantity means 'size, magnitude or dimension', from Middle English. In 1847 'quantitative analysis' was first used, meaning 'chemical analysis designed to determine the amounts or proportions of the components of a substance'. In 1878 'quantify' was used to mean 'to determine the quantity of, to measure', and hence 'quantification' is 'the operation of quantifying'. In 1927 'quantitate' was first used to mean 'to measure or estimate the quantity of, especially to measure or determine precisely'. However Webster's dictionary calls this term a 'back-formation',[11] which is probably as derogatory as a dictionary compiler can be, and this term is not used in this book, nor is it in the *Oxford English Dictionary*.

Francis Bacon (1561–1626) had a great influence on the generations of British scientists who followed him.[12] He stressed collecting as much data as possible, then setting out to explain the observations, instead of dreaming up an idea and then looking for facts to support it. Science must be built on the foundation provided by the facts. What would he say about the modern 'hypothesis-driven' research? In 1662, the Royal Society of London for the Promotion of Natural Knowledge received its charter from King James II, as one of the first, and best known, scientific societies.

In the eighteenth century the problem of navigation around the globe was severe. Although latitude (distance from the equator) could be measured accurately, using the elevation of the sun above the horizon at noon (the time of maximum altitude), longitude[13] (the easterly or westerly distance around the globe, now measured from Greenwich, London, UK) could not be. Samuel Pepys, commenting on the pathetic state of navigation, wrote of 'the confusion all these people are in, how to make good their reckonings, even each man's with itself', recognizing the distinction between intra- and inter-observer variation. Newton wrote of the sources of error involved in trying to measure time at sea 'One [method for determining longitude] is by a Watch to keep time exactly. But, by reason of motion of the Ship, the Variation of Heat and Cold, Wet and Dry, and the Difference of Gravity in different Latitudes, such a watch hath not yet been made'.

As a result many lives were lost at sea, through shipwreck and failure of supplies, and navigation was such a sensitive issue that sailors were forbidden to carry out their own calculations, for fear that they would show up errors in those of their superior officers. The growth of vastly profitable world trade was held back. In this context, the Longitude Act of 1714 was passed in the British Parliament, offering a reward of £10 000[14] to anyone who could devise a method of measuring longitude accurately.

[10] Much of the historical material in this chapter comes from the *Encyclopaedia Britannica*.

[11] A back-formation is a word formed by subtraction of a real or supposed affix from an already existing longer word. Thus from *quantitation* was created *quantitate*.

[12] From *Science: a History 1543–2001*, by John Gribbin (Penguin, London, 2002).

[13] See *The Illustrated Longitude*, by Dava Sobel and William J.H. Andrewes (Fourth Estate, London, 1999).

[14] The sum was graded according to the accuracy that could be achieved.

The challenge of solving the 'longitude problem', as it came to be known, was taken up by an English clockmaker, John Harrison, who lived near the port of Hull, and had heard the stories of souls going to their deaths, and the reward offered. The Longitude Board paid expenses and could give incentive awards to help impoverished inventors bring promising ideas to fruition. He built four clocks altogether. The first kept good time on land (better than one second per month) and in small trips out to sea. He succeeded in getting a full trial at sea with the navy on a voyage to Lisbon in 1736; his clock showed unexpected error at sea, being susceptible to an artefact caused by accelerations in the motion at sea. His own perfectionism and obstinacy all round delayed matters, and the next trial, taking his fourth clock to the West Indies, did not take place for another 25 years. The watch performed well but insufficient data points were produced to satisfy the Longitude Board. The Longitude Board was dominated by eminent astronomers and others from the Naval establishment, and repeatedly refused to give Harrison his payment, requiring that the chronometer should first be taken from prototype into mass production. The Board realized that replicate voyages and clocks were needed to establish the reproducibility, without which the accuracy could not be guaranteed. A single measurement could not establish the maximum error. Harrison's son William took up his case, and the Royal Society offered Harrison a Fellowship. It was only intervention by King George III and the passing of a second act by Parliament that gave Harrison his recognition, at the age of 80, 46 years after he had built his first sea clock.

This story, of finding a scientific solution to a human problem, has all the elements of the struggles that modern scientists may have to go through to develop a technique that they believe will save lives, and many parallels can be seen. Harrison's clocks are preserved in the old Royal Observatory at Greenwich.

The Lunar Society of Birmingham (England) was a group of forward-thinking scientists who met between 1766 and 1791. They met on the day of the full moon (so that travel would be easier), and flourished independently of the Royal Society (in London). Birmingham was the location of much inventive scientific activity stimulated by the industrial revolution. Both of Charles Darwin's grandfathers (Josiah Wedgewood, the pottery manufacturer, and Erasmus Darwin, the naturalist) were members, as were Matthew Boulton (the manufacturer), Joseph Priestly (who discovered oxygen) and James Watt (who invented the steam engine). The Industrial Revolution in Britain and the rest of Europe gave commercial impetus to the invention of a variety of measuring instruments to be used in the manufacturing process. Lord Kelvin, delivering a lecture on electrical units of measurement in 1883, expressed the desire of his time to quantify:

> *When you can measure what you are speaking about, and express it in numbers, you know something about it; but when you cannot measure it, when you cannot express it in numbers, your knowledge is of a meagre and unsatisfactory kind: it may be the beginning of knowledge, but you have scarcely, in your thoughts, advanced to the stage of science, whatever the matter may be.*

However, he might have added a caveat about the danger of numbers giving a pseudo-scientific respectability to some studies.

In the newly formed United States of America, it was found to be impossible to reform the archaic system of weights and measures inherited from the British, in spite of the Napoleonic metric system that had recently been adopted in France. The Office of Weights and Standards became the National Bureau of Standards, then the National Institute of Standards and Technology (NIST). In 1960 the 11th General Conference of Weights and Measures, meeting in Paris, established the International System of Units, based on the metre, kilogram, second, ampere, degree Kelvin and candela. These units are often called the *SI units*, after the French expression *Système Internationale*, and are preferred in the scientific community.[15] The kilogram is represented by a cylinder of platinum–iridium alloy

[15] The engineering community in the USA still uses units based on the British Imperial system (although these are not

kept at the International Bureau of Weights ands Measures in France,[16] with a duplicate in the USA; the other units are defined with respect to natural standards (e.g. the metre is defined by the wavelength of a particular visible atomic spectral line). National centres such as the US NIST and the UK National Physical Laboratory (NPL) are now centres of expertise in measurement science.

In parallel with the development of physical instruments was the invention of mathematical techniques. Ancient Babylonians, Egyptians, Greeks, Indians (Harappans) and Chinese all had mathematics, originally used for computing areas and volumes of regular objects, and also used for handling monetary currency. In the sixth century BC, Pythagoras established the link between the musical note of a string and its length. This bridge between the world of physical experience and that of numerical relationships has been called the birth of mathematic physics, where numbers explain the origin of physical forms and qualities. Newton's differential calculus and Fourier's transform are essential tools used by our current MRI scanners. Early digital computers, most famously used to decipher the Enigma code used by submarines during the Second World War, developed to the stage we take for granted today.

In medicine the concepts of the new scientific methods, including quantification, were applied. William Harvey (1578–1657) was a physician and scientist who studied the blood circulation extensively, and was the first to measure the cardiac volume and estimate the total blood volume in the human body. In 1833 William Beaumont, a US army surgeon, published a series of studies[17] on a soldier who had been wounded in the stomach and then developed a flap that could be opened. Beaumont could watch food in the stomach, and extract gastric juice. Nowadays we have more convenient ways of making *in vivo* studies.

In the late 1970s scientists started connecting medical imaging hardware to computers that look extremely basic by modern standards, motivated by the desire to manipulate and interrogate the images. Sophisticated medical imaging instruments were produced, in nuclear medicine, ultrasound, X-ray computed tomography, and NMR.

In the early 1980s the annual meetings on Information Processing in Medical Imaging (IPMI) started taking place. In 1989 it was argued that (Tofts *et al*., 1991a,b):

> *The history of image processing in nuclear medicine shows that collection of good quality image data is at least as important as access to image processing techniques. Even now one could argue that real improvements in the usefulness of image data come from instrumental improvements rather than from more sophisticated ways of image processing. However in the case of large datasets that are already of good quality, the problem is then one of data presentation and reduction, rather than correcting images to compensate for errors in data collection. With this philosophy we have initially concentrated on collecting good quality data, that are sensitive to the clinical question being studied. For example T_2 weighted images of the brain can show Multiple Sclerosis (MS) lesions, and one could develop sophisticated algorithms for measuring lesion volume to assess disease and therapies; however the images show oedema and scar tissue, which are secondary to the disease process. Primary visualisation of the disease is shown by the newer technique of GD-DTPA scanning, and therefore we have developed this data collection technique in preference. A second example is the use of expensive classification techniques on image data clearly showing gross nonuniformity which can be removed relatively simply. Having taken care of the instrumental aspects and obtained good quality data, the processing requirements may become less expensive, and mostly consist of PACS,[18] 3D display, calculation of functional*

used in the UK any more). Incompatibility between Imperial and Metric units was blamed for a space vehicle failure in recent years.

[16] The BIPM, Bureau International des Poids et Mesures (www.bipm.fr/enus/).

[17] From 'The Man with a Lid on his Stomach', in the *Faber Book of Science*, edited by John Carey (Faber & Faber, London).

[18] PACS is picture archival and computing system, and refers to computer-based systems to store, display and interrogate large quantities of medical images. By 'functional images' was meant parametric maps of any kind (e.g. permeability).

images, and segmentation algorithms. Where sophisticated forms of information processing are required, to make full use of them they must be integrated into a programme that includes aspects of data collection such as sequence design, quality control of instrumental parameters, validation of the quantitative results, and good experimental design. In summary, we believe that data must be appropriate, and of good quality, before undertaking any processing.

Thus the notion that good quantification required attention to both *data collection* and *image analysis* techniques was born at the start of the last decade, and this complementarity can be seen in the structure of this book. Experience has shown that advances are often made by groups who have access to both data collection (so that the acquisition technique can be optimized for the job in hand) and to advanced analysis techniques (to obtain the most from the data). Computing groups working isolated from the clinical questions and acquisition hardware may produce solutions to nonexistent problems, or use data that are degraded by poor acquisition technique.

Quantification was recognized by some radiologists as having a potential role in studying disease (Tofts and du Boulay, 1990):

Serial measurements in patients and correlation with similar studies in animal models, biopsy results and autopsy material taken together have provided new knowledge about cerebral oedema, water compartmentation, alcoholism and the natural history of multiple sclerosis. There are prospects of using measurement to monitor treatment in other diseases with diffuse brain abnormalities invisible on the usual images.

When making quantitative measurements, the physicist can adopt the paradigm of the scientific instrument designer, who is presented with a sample (the patient) about which he or she wishes to make the most careful, detailed measurements possible, in a nondestructive way, using the infinitely adjustable instrument (the imager). The biological question to be answered and thus the bio-physical feature to be measured need very careful choice after discussion by all concerned.

Quantitative magnetic resonance was the subject of a small meeting organized by the UK Institute of Physics and Engineering in Medicine in 1997 at Dundee, Scotland, and it is here that the expression 'qMR' was first used. qMR is now coming to denote that part of MR concerned with quantitative measurements, in the same way that fMRI (functional MRI), MRA (MR angiography), MRS (spectroscopy) and qMT (quantitative magnetization transfer) denote subspecialties of MR.

1.3 GENERAL CONCEPTS OF MEASUREMENT IN MEDICAL IMAGING

Physical quantities can be *intensive* or *extensive*, and when we are considering various properties and manipulations to quantities, it can be helpful to be aware of these differences. Intensive quantities[19] can describe a piece of tissue of any size, and they do not alter as the tissue is subdivided (assuming it is uniform). Examples are density, temperature, colour, concentration, magnetization, membrane permeability, capillary blood volume and perfusion per unit volume of tissue, texture and the MR parameters proton density, T_1, T_2, the diffusion coefficient of a liquid and magnetization transfer. Extensive quantities refer to a piece of tissue as a whole, and subdivision reduces (or at least changes) the value of the quantity. Examples are mass, volume, shape and total blood supply to an organ.

Some intensive quantities, such as metabolite concentration, local blood flow or local permeability, can be expressed either per unit mass of tissue or per unit volume of tissue. Traditionally, physiologists have used the former system, since the mass of a piece of excised tissue is more easily determined that its volume. In qMR, where the volume of each voxel is well defined, the latter system is more natural. Conversion from per mass to per volume can be achieved by multiplying by the *density of brain*[20] ($1.04\ \mathrm{g\,ml^{-1}}$ or $1040\ \mathrm{kg\,m^{-3}}$ for both white and grey matter (see Chapter 4, Table 4.2).

[19] Intensive: of or relating to a physical property, measurement etc. that is independent of mass; extensive: a property that is dependent on mass.

[20] For example the normal concentration of water in white matter is about 0.690 g water per g tissue (0.690 kg water per

1.3.1 Images, Partial Volume and Maps

Images and maps are terms used to mean different things. An *image* is produced by the MRI scanner, and has an intensity[21] that depends on a variety of parameters, including some that describe the tissue (e.g. PD, T_1, T_2 and combinations of these), and some that are characteristic of the scanner (e.g. the scanner transmit flip angle and receiver gain). The image consists of a two-dimensional matrix of numbers stored in a computer. Each location in the matrix is called a *pixel* (picture element), which is typically square and 1–2 mm wide. The image data come from a slice of brain tissue which has been interrogated, or imaged. This slice has a specified thickness (usually 1–10 mm), and each pixel in the image in fact derives from a cuboidal box-shaped piece of tissue, called a *voxel*. The first and second dimensions of the voxel are those of the image pixel, and the third dimension is the *slice thickness*.

The interplay between pixels and voxels is subtle. At times when we are thinking of images, pixels are more natural, and in fact the term originates from the science of interpreting images of two-dimensional surfaces (e.g. in robot vision or remote sensing of the Earth by satellite). Yet when we are thinking of the cuboids of three-dimensional tissue from which the pixel intensities originate, voxels are more natural, and serve to remind us to think about the tissue, not the image. Slices of voxels are inside the object, whilst surfaces of pixels are outside the object. Some imaging procedures will use very small pixels ('in-plane' resolution), yet set a large slice thickness (in order to retain signal-to-noise ratio). An extreme example would be a voxel of size $0.7 \times 0.7 \times 10$ mm, which appears to have the ability to resolve small structures, yet any structures that do not lie close to the perpendicular to the slice plane would be blurred by the large slice thickness. In this case the voxel would be shaped like a matchstick (i.e. have a large *aspect ratio*); a more appropriate voxel size would be $1.5 \times 1.5 \times 2.2$ mm, which has the same volume (and hence signal-to-noise ratio, for a given imaging time), but is more likely to resolve small structures. Three-dimensional imaging sequences can give us voxels which are *isotropic* (i.e. have the same dimensions in all three directions).

Structures in the brain have very fine detail; thus a voxel with dimensions 1–10 mm cannot capture thus fine detail, and very often there are two (or more) types of tissue inside the voxel. The resulting NMR signal from this voxel is simply a combination, or weighted average, of what each individual tissue would give if it filled the whole voxel. Thus if we are trying to measure the T_1 of brain tissue, near to CSF (cerebrospinal fluid) in the ventricles the value measured will be somewhere between that of pure white matter and pure CSF, depending on the relative proportions of white matter and CSF in the voxel. This is called the *partial volume effect*, and is a major source of error when making measurements in brain tissue at locations near to boundaries with other tissue types. The value measured in the tissue is altered by its proximity to another tissue, and the determination of boundaries and of volumes is brought into error. Partial volume errors can be reduced by using smaller voxels, although the price paid is that of a worsening of the signal-to-noise ratio. An inversion pulse before data collection can remove signal from a tissue with a particular T_1 value (as in the FLAIR and STIR sequences, which null the signal from CSF and fat, respectively). Alternatively, image analysis techniques that take account of the effect of partial volume on signal intensity can be used (e.g. Tofts, 1998).

A parametric *map* can be calculated from two (or more) images of the same piece of tissue. A simple example would be to collect two images with differing amounts of T_2-weighting. The ratio of these two images then only depends on the tissue parameter T_2, and is independent of scanner parameters (such as transmitter or receiver settings). By calculating this ratio for each pixel, a third matrix, or map, can be formed, which has

kg tissue), equivalent to 0.718 g water per ml tissue (718 kg water per m^3 tissue) (see Chapter 4, Table 4.2).

[21] Often called the 'signal intensity', since it is proportional to the signal voltage induced in the RF coil by precessing magnetization in that piece of tissue seen in that voxel of the image.

the appearance of an image (brain structures can be identified), but is conceptually different from an image, in that individual pixel values now have a numerical meaning (such as value of T_2, in milliseconds, at each location in the brain), rather than representing signal intensity on an arbitrary scale.

1.3.2 Study Design

Many studies set out to compare groups of subjects. Typically, a new MR parameter will be evaluated in a particular disease by measuring it in a group of patients and in a group of controls. Other differences between the groups should be removed as much as possible, hence the need for age and gender matching. The scanning should be carried out at the same time, using *interleaved controls*, rather than leaving the controls until the end of patient scanning (when a step change in the measurement procedure could produce an artificial group difference). Some patients may be on treatment which alters the MR results. Matching can be improved by *dynamic matching*, carried out as part of subject recruitment as the study proceeds. Thus if normal controls are in short supply, but patients plentiful, then each time a control is recruited, a matched patient is selected from the available patients. In placebo-controlled trials, allocation of a patient to the placebo or treated group can be decided at the time of recruitment, to keep the groups matched at all times. Double-blinding[22] is a powerful way of reducing bias in treatment trials. The person giving the treatment, the person making the measurement[23] and the patient are all blinded as to whether they are receiving a genuine treatment or a placebo.

Inexperienced researchers should beware of 'stamp collecting' when 'interesting patients' are studied, almost at random, with no hypothesis or controls. To design high-quality investigations that will be accepted for publication by the best international journals, the investigator should be aware of what work has already been published or presented at international scientific conferences. A literature search[24] should be carried out. Studies should not be replicated unless there is a case for confirming the results with a different group of patients. Methodological pitfalls, as illustrated by existing published work, should be identified before the study begins. Some errors (for example estimating reproducibility without rescanning, or scanning controls after an upgrade, not interleaved with the patients) will irreversibly destroy the value of the data.

Selection of MR parameters requires thought. To acquire all the parameters discussed in this book would require more time than can be fitted into one examination (although acquisition times will probably come down as the techniques are refined). Parameters should be selected according to the biological changes that are expected in the particular disease being studied. Combined parameter studies can be powerful (see Section 2.2.2.5 in Chapter 2). Mixed-parameters acquisition can address specific questions (for example diffusion weighted spectroscopy, or MT-prepared multiecho measurements [see Chapter 8, Figure 8.6(b)].

[22] The double-blind design is not suitable for treatments where the practitioner plays an essential part in the treatment. This is particularly relevant in 'alternative' therapies (e.g. acupuncture, homeopathy, osteopathy, psychotherapy and reiki). Although a placebo cannot be given, different treatments can be compared. More research on methodology may be needed to find suitable study designs for such treatments.

[23] Ideally this includes both the radiographer making the scan, and the observer analysing the MR data.

REFERENCES

Filippi, M. and Grossman, R. I. 2002, MRI techniques to monitor MS evolution: the present and the future, *Neurology*, **58**, 1147–1153.

Filippi, M., Rocca, M. A. and Rovaris, M. 2002, Clinical trials and clinical practice in multiple sclerosis: conventional and emerging magnetic resonance imaging technologies, *Curr. Neurol. Neurosci. Rep.*, **2**, 267–276.

McFarland, H. F., Barkhof, F., Antel, J. and Miller, D. H. 2002, The role of MRI as a surrogate outcome measure in multiple sclerosis, *Mult. Scler.*, **8**, 40–51.

McGowan, J. C. 2001, On the use of quantitative MR imaging, *Am. J. Neuroradiol.*, **22**, 1451–1452.

[24] For example using PubMed, from the USA National Library of Medicine, available free of charge on-line at www.ncbi.nlm.nih.gov/entrez/query.fcgi

Miller, D. H. 2002, MRI monitoring of MS in clinical trials, *Clin. Neurol. Neurosurg*., **104**, 236–243.

Tofts, P. S. 1998, Object strength correlates closely with OCCA MS brain phantom lesion volume, and can accurately predict lesion volume, *Proc. Intl. Soc. Mag. Reson. Med.*, **6**, 2074.

Tofts, P. S. and du Boulay, E. P. 1990, Towards quantitative measurements of relaxation times and other parameters in the brain, *Neuroradiology*, **32**, 407–415.

Tofts, P. S., Wicks, D. A. and Barker, G. J. 1991a, The MRI measurement of NMR and physiological parameters in tissue to study disease process, *Prog. Clin. Biol. Res*., **363**, 313–325.

Tofts, P. S., Wicks, D. A. G., and Barker, G. J. 1991b, The MRI measurement of NMR and physiological parameters in tissue to study disease process, in *Information Processing in Medical Imaging*, Ortendahl, D. A. and Llacer, J. (eds). Wiley-Liss, New York pp. 313–325.

Queries in Chapter 1:

Q1. Reviewer details needed in footnote 1