Dosimetry using MRI: can it really be that difficult?

Simon J Doran

CRUK Cancer Imaging Centre, Institute of Cancer Research, London, UK

E-mail: Simon.Doran@icr.ac.uk

Abstract. Magnetic resonance imaging (MRI) has been used in gel dosimetry from the very first studies back in the 1980s. Almost all the imaging problems that we still encounter were known about at that time and many were described in the proceedings of the very first International Workshop on Radiation Therapy Gel Dosimetry (DOSGEL ‘99). The quality of MRI scanners has improved enormously in the intervening two decades, so why are there still issues? This review will draw on the wealth of previously presented material from the literature to attempt to answer this question. The reference list provides a “starter pack” that should be viewed as a “jumping off point” for further investigations by the reader, rather than an exhaustive summary of what is a large domain of work.

1. Introduction

Whilst there were a number of earlier publications that described the modification of nuclear magnetic resonance (NMR) relaxation times by absorbed radiation dose, most commentators would likely agree that the seminal work in the field of gel dosimetry is the 1984 paper by Gore and Kang [1]. It lays out a methodology for measuring a 3-D dose distribution in a tissue-equivalent phantom, by using magnetic resonance imaging (MRI). Although taken for granted nowadays, one of the novel features about this work was that it proposed to use imaging to make quantitative measurements of a physical property. This was one of a number of early publications that sought to move MRI on from a modality used purely for visualising the presence of pathology, to one with which accurate science could be performed.

The measurement of NMR relaxation times themselves, however, predated Gore and Kang by more than three decades. By 1948, just a few years after the discovery of the NMR phenomenon [2, 3], Bloembergen, Purcell and Pound had already produced an extensive description of NMR relaxation and had designed apparatus to measure it [4]. By 1950, ideas about radiofrequency (RF) pulses were becoming familiar [5], and in 1954, Carr and Purcell introduced their eponymous pulse sequence for measuring $T_2$ [6]. Arguably, by the late 1950’s, most of NMR had already been invented and everything that followed was simply the out-working of these formulae. Imaging introduced an entirely new concept in the 1970’s [7-9], but the factors contributing to measurement errors in spatially-resolved relaxation time mapping were easily incorporated into the framework already developed. Certainly, by the time of the 1st International Workshop on Radiation Therapy Gel Dosimetry (DOSGEL ‘99), the majority of the effects described later in this article had been thoroughly analysed in a variety of other contexts and are neatly summarised in Table 1 of De Deene’s review article [10] from the workshop.

Much has changed since 1999: MRI scanners have improved and continue to do so; image reconstruction techniques have become more sophisticated; and successive waves of disruptive technology (parallel imaging [11, 12] simultaneous multislice imaging [13], compressed sensing [14] and MRI fingerprinting [15]) have changed the discipline to the extent that many of the pulse sequences of today would be almost unrecognisable to the early pioneers. So, with rapid, high-quality imaging at...
our fingertips, why is MRI-based 3-D dosimetry not already a solved problem? In the remainder of this article, I discuss the various sources of error and try to highlight, with particular reference to previous articles in the DOSGEL and IC3DDose conference series, the reasons why 20 years of effort by the gel dosimetry community have yet to produce a turn-key solution.

2. A question of philosophy
Let me first pose a philosophical question: what we do with MRI-derived dose maps when we have generated them?

A fundamental problem of calibration, as has been pointed out by De Deene [16], is that we have no access to a gold standard dose distribution against which to validate our gel dosimetry. All we can do is to compare the output of one technique against another, without any guarantee that either is “the truth”. For simple irradiations, we expect different techniques (e.g., gel, film, diode array, treatment planning system (TPS) and Monte Carlo) to agree closely. And so they do, as it turns out, even for some relatively complex situations [17]. However, our problem comes precisely with the extrapolation to the cases in which we are most interested, that is those where we are unsure which method to trust. It should be obvious that it is logically incoherent to argue: “My gel dosimetry system is proven to work because it matches the TPS for Irradiation A. I am therefore now going to use my gel to investigate whether the TPS produces the correct result for Irradiation B.” Which modality is validating which?

Two aspects of this dilemma been addressed by Schreiner in recent IC3DDose articles: firstly, what are the scenarios in which gel dosimetry may be used most profitably [18] and, secondly, how best should we compare two measurements of a 3-D dose distribution [19]. A related question with which we, as a community, are perhaps still struggling, is “What do we do when our two methods give us different results?” In commissioning, how long should we spend agonising about small differences between outputs from different techniques? In clinical practice, what action levels should we set based on MR gel dosimetry and what action should we take when they are breached?

3. Chemical dosimetry
A second overarching issue is the nature of chemical dosimetry. A very large section of the gel dosimetry literature has been concerned with the ways in which the NMR properties of different types of gel are affected by radiation dose and the ways in which gel properties interact with the MRI measurement. Thus, there are studies that relate errors in MR-derived dose maps to oxygen sensitivity [20], ion diffusion [21, 22], dose rate [23], dose fractionation [24], post-irradiation evolution of the chemical system [23, 25, 26], temperature during storage and irradiation [27] and temperature during scanning [28], container size and calibration methodology [29, 30]. Ref. [31] is a topical review of the area of polymer gel dosimetry that provides an excellent starting point for further study.

The source of dose-related contrast in polymer gel [31] imaging is primarily the change in transverse relaxation resulting from radiation-induced polymerisation, whilst in Fricke gels [32-35], absorbed radiation causes the oxidation of Fe$^{2+}$ ions to Fe$^{3+}$, which changes both longitudinal and transverse relaxation rates. Thus, both quantitative $R_1$ and $R_2$-mapping sequences are of interest in gel dosimetry. All of the effects described above change the actual $R_1(=1/T_1)$ or $R_2(=1/T_2)$ value of the dosimeter, which is what the quantitative MRI sequences measure. Other effects, such as water diffusion [6], have the potential to affect the measurement process itself and need to be accounted for.

These aspects will not be discussed further here, except to make one general point. A very significant disadvantage of gel dosimetry is that, although multi-use dosimeters have been proposed for optical CT readout [36] and at least one company has marketed Fricke gels that can be reused (TruView™, Modus Medical Devices Inc.), it is still generally true that each gel dosimeter is an individual system. By its very nature, the dosimeter cannot be characterised by repeated test irradiations, as can an ion chamber or a diode array. Absolute measurements are problematic [37] and most results are reported as “relative dosimetry”. Measurements are sometimes found not to be repeatable to within the tolerances required for precision radiotherapy, and inter- and intra-batch differences are common. The most careful study of such effects thus far reported is that of Vandecasteele and De Deene [38].
Despite the speed of modern pulse sequences and steadily improving access to MR scanners through the advent of MR-based radiotherapy treatment planning, the “fabricate-irradiate-image-analyse” cycle is much more arduous on an ongoing basis than the use of a calibrated diode- or ion chamber array phantom with stable properties. At each stage of the process, errors can be introduced and these are summarised in Fig. 18 of [31].

4. Sources of error related to the MRI measurement
The high quality of images created by modern MR scanners has already been alluded to above, and this, paradoxically, leads to some concerns about their use in radiotherapy dosimetry that were less apparent in the early days of gel dosimetry. Standard clinical software can now produce, relatively quickly, quantitative maps that “look good”: that is, they are free of obvious artefact and have good signal-to-noise ratio. It may thus be tempting to assume that they must also now be more accurate measurements than hitherto. Whilst this may be true in some cases, rigorous evaluation using well established and recently calibrated relaxation test objects is still vital. Throughout, it should be borne in mind that quantitative relaxation time measurement is a minority use case for scanner manufacturers: systems are optimised for routine clinical imaging whose requirements are different.

Useful summaries of the type of image artefact that can occur in MRI in general may be found in [39-42]. The remaining sections of this article provide an overview of the specific issues that arise in quantitative MRI measurements. Many of these have been discussed by previous DOSGEL and IC3DDDose reviews [10, 16, 43, 44], but potential new sources of error are associated with recent developments in imaging and these are discussed briefly at the end of the article.

4.1. RF flip angle
A number of methods are available for $R_1$-mapping. These include, among others, inversion-recovery [42] and Look-Locker [45] techniques (both often employed in association with an echo-planar imaging readout for speed) and rapid 3-D gradient echo sequences (run several times with different flip angles [46, 47]). The method historically used to measure $R_2$ has generally been a multi-echo sequence, which is the imaging equivalent of the Carr-Purcell sequence [6]. Sequences based on the steady-state free precession paradigm, including DESPOT [48] and inversion-recovery trueFISP [49] allow both $R_1$ and $R_2$ to be calculated. However, in all cases, it is of great importance that the RF flip angle is well calibrated. As an example, the exponential signal decay in a Carr-Purcell sequence, relies on the accuracy of the repeated 180° pulse. Deviations in flip angle lead to incorrect echo amplitudes which do not match the mathematical model used to fit the data. Hence, systematically incorrect $R_2$ values may be generated [50], and this may occur without there being obvious image artefacts.

Why have the problems associated with RF pulses not been eradicated along with the general improvement in MR technology? The answer is that the laws of physics make it very difficult to create RF probes that give a constant flip angle over the entire volume of the sample. Similarly, slice-selective RF pulses do not impart exactly a 180° flip to magnetisation within the slice and 0° to magnetisation outside; slices do not have an ideal “square-cornered” profile, but instead a transition region at the edge, in which the flip angle varies between the two extremes [51]. At high field (e.g., 3T clinical scanners), a further complication arises in the form of incomplete RF penetration into the sample, caused by so-called standing wave effects [52, 53]. Such variations of RF flip angle across the imaging volume means that a homogeneous sample can give rise to $R_1$ and $R_2$ maps that vary systematically. Furthermore, sequences that are nominally of the same type may be implemented differently on two different MR platforms, potentially leading to differing results for the same physical sample.

4.2. Ghosting artefacts and eddy currents
Ghosting artefacts are regions of MRI signal that appear systematically “in the wrong place” in images. Typically, coherent ghosts are entire copies of the correct image that appear with a lower intensity and are shifted in the phase-encoding direction. Ghosting artefacts can have many sources, but often occur for techniques whose pulse sequences include gradient reversals and zig-zag trajectories in k-space.
Implementations of the echo-planar sequence commonly suffer from the so-called “Nyquist N/2 ghost” [54]. Simple Fourier transform theory shows that if the forward and backward echoes are not perfectly aligned and there is an alternating pattern, then the reconstructed image consists of the superposition of two separate copies of the desired image, separated by half the field-of-view.

Eddy currents are one cause of such echo misalignment in k-space. Eddy currents are a consequence of Lenz’s Law: when an MRI gradient is switched rapidly, by changing the current through the gradient coil, unwanted electric currents are induced in conducting surfaces in the magnet assembly in such a way as to set up magnetic fields that oppose the original change [55, 56]. This is an area that has received much attention from equipment manufacturers and gradient performance has improved markedly since the early years of gel dosimetry. Nevertheless, ghost artefacts still have the potential to reduce the accuracy of quantitative $R_1$ and $R_2$ sequences by mixing signals from two separate regions of the irradiated sample, which might have different values of relaxation time (arising from different doses). The effect may be particularly pernicious for a sample that fills the field-of-view with a smoothly varying dose distribution (i.e., no obvious aliased boundaries between regions), and for ghosts of relatively low intensity that influence results subtly.

4.3. Spatial distortion

The basis of MR imaging is the observation that application of a magnetic field gradient across the sample maps the frequency of NMR spin precession onto a spatial location. Data are acquired in the temporal domain and converted (traditionally by Fourier transformation, but nowadays by a number of more complex post-processing signal models) into the spatial frequency domain. To make the jump to real space, we need to know the mapping between space and frequency. If this mapping function is incorrect, then the images created will exhibit spatial distortions. It is normally assumed that the mapping is linear, but this is not always the case. A brief summary follows, but for more detail, see [57, 58].

Non-linearities in magnetic field gradient are an inevitable result of the need to build gradient coils of a limited size to fit within the bore of the main magnet. The gradient system of a modern magnet is a sophisticated compromise between a number of competing requirements: linearity is traded off against eddy current performance, maximum gradient strength, self-inductance, gradient slew rate, performance of the gradient amplifiers, and bore length [59]. In practice, both the manufacturer and those performing local quality assurance go to considerable lengths to try and map and gradient-induced distortions and correct them.

Inhomogeneities in the static magnetic field ($B_0$) lead to local frequency perturbations and, hence, localised shifts in the apparent positions of features in images. These effects may be caused by poor shimming, but are often the result of the composition of the sample, which may contain regions of different magnetic susceptibility. One example of great relevance to 3-D dosimetry is the study of dose effects at air-tissue interfaces and in low density tissues such as the lungs [60]. We typically wish to measure accurate values of dose in tumours near such an interface, yet it is precisely in this region that the spatial accuracy and quantification of the MR measurement are called into question. A related effect is the so-called chemical shift artefact, whereby a change in chemical composition is associated with Larmor precession at a different frequency, and hence an apparent spatial position that is shifted from the true position. The effects in this paragraph can often be corrected using the “forward and reverse gradient” technique introduced by Chang and Fitzpatrick [61].

4.4. Experimental design: SNR, resolution and sampling

Compared with other imaging modalities (e.g., x-ray computed tomography), the range of options available in the design of an MRI experiment is very large. As detailed in the next section, a number of new methods are being investigated. However, for quantitative $R_2$ mapping, the majority of genuinely quantitative work requiring high accuracy is probably still performed using pulse sequences consisting of “trains” of spin echoes. In [16], De Deene summarises strategies for optimising this type of experiment from a signal-to-noise ratio (SNR) perspective, and a more detailed analysis is presented in
As in other applications of MRI, there is a trade-off between acquisition time, SNR and spatial resolution. If the image resolution is too poor, then partial volume effects may lead to inappropriate dose values being measured. On the other hand, if fine detail is required, then either the pulse sequence takes longer to run, or the images are more noisy, resulting in lower dosimetric precision. It is still a challenge to obtain good $R_2$ data for a large number of slices or an isotropic 3-D volume and this may be limiting for some polymer gel applications. 3-D quantitative $R_1$ measurements, which can be used for the readout of Fricke gel dosimeters, are somewhat easier to perform, and have been widely employed in dynamic contrast-enhanced MRI [63]).

4.5. Potential issues related to advances in MR image acquisition and reconstruction

The acquisition time of an MRI sequence is largely governed by the number of times one has to repeat the basic unit of (RF excitation – spin evolution – readout – recovery). Arguably, the important paradigm shift in MRI during the early years of the 21st century has been the realisation that one can acquire images of similar quality with fewer such loops.

Parallel reception (e.g., SMASH [11], SENSE [12]) makes use of the properties of multi-coil receiver arrays. In Section 2.1, I described the problem of spatial non-uniformity of transmission by an RF probe. The flip-side of this is that, in receive mode, the probe sensitivity varies spatially, too. Since each coil of a receiver array is located in a slightly different position in relation to the sample, the different spatial sensitivity patterns encode the NMR signals differently. Incorporated correctly into the reconstruction algorithm this spatial information can replace the data from a fraction of the phase-encoding steps, without the need to develop any new imaging sequences. Significant acceleration factors can be obtained from large arrays of coils. However, possibly the biggest drawback of the technique from a quantitative imaging point of view is the issue of spatially inhomogeneous noise amplification (so-called g-noise) [12].

Although the ideas underlying simultaneous multislice (SMS) imaging have been in development since the 1980s [64-67] it is only recently that technology has been available for the widespread deployment of the method. Instead of increasing speed by reducing the number of phase-encoding steps, one effectively acquires two or more images at once, by exciting several slices with the same, specially designed, RF pulse. Image reconstruction is achieved via one of a number of different techniques (some related to SENSE), as described in the review by Barth et al. [13]. This family of techniques potentially allows a signal-to-noise gain over parallel imaging, but its use in quantitative imaging may be limited by issues such as “leakage” of signal between slices.

Compressed sensing [14] is a class of method that seeks to reduce MR acquisition times by exploiting the sparsity that is “implicit” in MR images. By this, what is meant is that for some mathematical transforms, the transformed MR signal will be close to zero for certain parts of the domain, benefiting the sparsity that is “implicit” in MR images. By this, what is meant is that for some mathematical transforms, the transformed MR signal will be close to zero for certain parts of the domain, “do not need to be acquired”. This topic is very strongly related to our ability to compress such images and, for exactly this reason, the types of issues that may arise are similar to the image degradations seen when performing JPEG (or similar) compressions on image data. Whilst these methods may be widely acceptable for clinical imaging, care is needed when applying them to quantitative measurements where a high degree of accuracy is required. Furthermore, Lustig et al. [14] note that “the most distinct artifacts in [compressed sensing] are not the usual loss of resolution or increase in aliasing interference, but rather loss of low-contrast features in the image” and this may be a problem for dose-mapping.

Magnetic resonance fingerprinting (MRF) was first introduced in 2013 [15] and uses an entirely new image reconstruction approach. Its key assumption is that, by simulating the Bloch equations, unique patterns of evolution of the MR signal may be designed for each combination of proton density, $T_1$ and $T_2$. These patterns are stored in a data dictionary and the process of reconstruction involves searching this dictionary for the closest match to the observed signals. As this is a relatively new technique, relatively few multi-centre trials have yet been undertaken to standardise acquisitions or determine the accuracy of the parametric maps generated.
5. Conclusion

There is no doubt that the capabilities of modern MRI scanners are significantly greater than they were 20 years ago, and this includes the ability to make better imaging measurements of relaxation times. As a community, we are also much more aware of the pitfalls in using MRI for 3-D dosimetry and how to avoid them. Nevertheless, the demands of radiotherapy are exacting: our quantitative measurements are not likely to be useful if we cannot guarantee an accuracy of 5% or better and even greater precision.

Obtaining high-quality and consistent results is possible, but it is not a turn-key operation and is not likely to be so in the near future. It requires not just experience but also a high level of understanding of the physics and chemistry underlying the measurements. For the moment, expertise in all of these aspects together remains localised within a small number of research groups and even fewer commercial companies. To the knowledge of this author, no large-scale, multicentre trials have been funded in order to standardise gel measurements. Results are thus potentially dependent on individual scanners within an institution, and these tend to be replaced on timescales of 5 – 10 years, after which protocols need to be re-established.

To date, measurement standardisation in MRI has failed to catch the imagination of funders. However, there are indications that this state of affairs may be beginning to change, at least as far as mainstream radiology is concerned. Two organisations, the Quantitative Imaging Biomarkers Alliance (QIBA), established by the Radiological Society of North America (RSNA), and the Quantitative Imaging Network (QIN), which has grown out of the National Cancer Institute (NCI) cancer imaging programme, are making slow but steady progress.

Thus, the answer to the question posed in the title - “Can it really can be that difficult?” - is “yes.” But we are getting there!

6. References

[1] Gore J et al 1984 Phys. Med. Biol. 29 1189
[2] Bloch F 1946 Phys. Rev. 70 460
[3] Purcell E M et al 1946 Phys. Rev. 69 37
[4] Bloembergen N et al 1948 Phys. Rev. 73 679
[5] Hahn E L 1950 Phys. Rev. 80 580
[6] Carr H Y et al 1954 Phys. Rev. 94 630
[7] Lauterbur P C 1973 Nature 242 190-191
[8] Mansfield P et al 1973 Journal of Physics C: Solid State Physics 6 L422
[9] Garwood A N et al 1974 J. Phys. C: Solid State Phys. 7 457-462
[10] De Deene Y et al 1999 1st International Workshop on Radiation Therapy Gel Dosimetry 75-89
[11] Sodickson D K et al 1997 Magn. Reson. Med. 38 591-603
[12] Pruessmann K P et al 1999 Knowledge Creation Diffusion Utilization 962 952-962
[13] Barth M et al 2016 Magn. Reson. Med. 75 63-81
[14] Lustig M et al 2007 Magn. Reson. Med. 58 1182-1195
[15] Ma D et al 2013 Nature 495 187-192
[16] De Deene Y 2006 J. Phys.: Conf. Ser. 56 72-85
[17] Ceberg S et al 2010 Phys. Med. Biol. 55 4885
[18] Schreiner L 2008 J. Phys.: Conf. Ser. 164 012001
[19] Schreiner L et al 2013 J. Phys.: Conf. Ser. 444 012016
[20] Hepworth S et al 1999 Phys. Med. Biol. 44 1875
[21] Baldock C et al 2001 Aust. Phys. Eng. Sci. Med. 24 19-30
[22] Schulz R et al 1999 Phys. Med. Biol. 35 1611
[23] De Deene Y et al 2006 Phys. Med. Biol. 51 653
[24] Karlsson A et al 2007 Phys. Med. Biol. 52 4697
[25] De Deene Y et al 2000 Phys. Med. Biol. 45 859
[26] McJury M et al 1999 Phys. Med. Biol. 44 1863
[27] Vandecastelee J et al 2012 Phys. Med. Biol. 58 43
[28] Vandecastelee J et al 2012 Phys. Med. Biol. 58 63
[29] Dumas E-M et al 2006 J. Phys.: Conf. Ser. 56 239-41
[30] Oldham M et al 1998 Phys. Med. Biol. 43 2709
[31] Baldock C et al 2010 Phys. Med. Biol. 55 R1-63
[32] Audet C 1999 1st International Workshop on Radiation Therapy Gel Dosimetry (DOSGEL) 31-45
[33] Båck S Å J et al 1999 1st International Workshop on Radiation Therapy Gel Dosimetry (DOSGEL) 47-61
[34] Schreiner L J 2001 2nd International Conference on Radiotherapy Gel Dosimetry (DOSGEL)
[35] Schreiner L J 2004 3rd International Conference on Radiotherapy Gel Dosimetry (DOSGEL) 10-22
[36] Juang T et al 2015 J. Phys.: Conf. Ser. 573 012039
[37] Cosgrove V et al 2000 Phys. Med. Biol. 45 1195
[38] Vandecasteele J et al 2012 Phys. Med. Biol. 58 19
[39] Gaillard F et al https://radiopaedia.org/articles/mri-artifacts access date 6 October 2016
[40] Paudyal S http://www.slideshare.net/SudilPaudyal/mri-artifacts access date 6 October 2016
[41] Morelli J N et al 2011 Radiographics 31 849-866
[42] Doran S J et al 2012 Ch. 7 in Webb's Physics of Medical Imaging 487-621
[43] De Deene Y 2009 J. Phys.: Conf. Ser. 012033
[44] De Deene Y 2010 J. Phys.: Conf. Ser. 012015
[45] Look D C et al 1970 Rev. Sci. Instrum. 41 250-251
[46] Wang H Z et al 1987 Magn. Reson. Med. 5 399-416
[47] Wang J et al 2006 J. Magn. Reson. 182 283-292
[48] Deoni S C et al 2005 Magn. Reson. Med. 53 237-241
[49] Schmitt P et al 2004 Magn. Reson. Med. 51 661-667
[50] Majumdar S et al 1986 Magn. Reson. Med. 3 397-417
[51] Elster A http://mri-q.com/dielectric-effect.html access date 7 October 2016
[52] Webb A G et al 2010 Int. J. Imag. Syst. Tech. 20 2
[53] Buonocore M H et al 1997 Magn. Reson. Med. 38 89-100
[54] Elster A http://mriquestions.com/eddy-current-problems.html access date 7 October 2016
[55] Le Bihan D et al 2006 J. Magn. Reson. Imag. 24 478-488
[56] Sumanaweera T S et al 1994 Neurosurg. 35 696-704
[57] Doran S J et al 2005 Phys. Med. Biol. 50 1343-1361
[58] Turner R 1993 Magn. Reson. Imag. 11 903-920
[59] De Deene Y et al 2006 Med. Phys. 33 2586-2597
[60] Chang H et al 1992 IEEE Trans. Med. Im. 11 319-329
[61] Leach M et al 2012 Eur. Radiol. 22 1451-1464
[62] Souza S et al 1988 J. Comp. Asst. Tomog. 12 1026-1030
[63] Larkman D J et al 2001 J. Magn. Reson. Imag. 13 313-317
[64] Maudsley A 1980 J. Magn. Reson. 41 112-126
[65] Müller S 1988 Magn. Reson. Med. 6 364-371