Eur Radiol
DOI 10.1007/s00330-012-2500-8

MAGNETIC RESONANCE

# A systematic review of the utility of 1.5 versus 3 Tesla magnetic resonance brain imaging in clinical practice and research

**Joanna M. Wardlaw · Will Brindle · Ana M. Casado · Kirsten Shuler · Moira Henderson · Brenda Thomas · Jennifer Macfarlane · Susana Muñoz Maniega · Katherine Lymer · Zoe Morris · Cyril Pernet · William Nailon · Trevor Ahearn · Abdul Nashirudeen Mumuni · Carlos Mugruza · John McLean · Goultchira Chakirova · Yuehui (Terry) Tao · Johanna Simpson · Andrew C. Stanfield · Harriet Johnston · Jehill Parikh · Natalie A. Royle · Janet De Wilde · Mark E. Bastin · Nick Weir · Andrew Farrall · Maria C. Valdes Hernandez · The SINAPSE Collaborative Group**

Received: 19 January 2012 / Revised: 5 April 2012 / Accepted: 9 April 2012


**Abstract**
*Objective* MRI at 3 T is said to be more accurate than 1.5 T MR, but costs and other practical differences mean that it is unclear which to use.

*Methods* We systematically reviewed studies comparing diagnostic accuracy at 3 T with 1.5 T. We searched MEDLINE, EMBASE and other sources from 1 January 2000 to 22 October 2010 for studies comparing diagnostic accuracy

**Electronic supplementary material** The online version of this article (doi:10.1007/s00330-012-2500-8) contains supplementary material, which is available to authorized users.

J. M. Wardlaw · K. Shuler · M. Henderson · J. Macfarlane · S. Muñoz Maniega · K. Lymer · C. Pernet · T. Ahearn · A. N. Mumuni · C. Mugruza · J. McLean · G. Chakirova · Y. Tao · J. Simpson · H. Johnston · J. Parikh · N. A. Royle · J. De Wilde · M. E. Bastin · A. Farrall · M. C. Valdes Hernandez
Scottish Imaging Network, A Platform for Scientific Excellence (SINAPSE) Collaboration, Edinburgh, UK

J. M. Wardlaw · W. Brindle · A. M. Casado · K. Shuler · M. Henderson · B. Thomas · S. Muñoz Maniega · K. Lymer · Z. Morris · C. Pernet · Y. Tao · J. Parikh · N. A. Royle · J. De Wilde · M. E. Bastin · A. Farrall · M. C. Valdes Hernandez
Division of Clinical Neurosciences, University of Edinburgh, Edinburgh, UK

J. Macfarlane
University of Dundee,
Dundee, UK

W. Nailon
Edinburgh Cancer Centre, University of Edinburgh,
Edinburgh, UK

T. Ahearn
Institute of Medical Sciences, University of Aberdeen,
Aberdeen, UK

A. N. Mumuni
Department of Clinical Physics,
University of Glasgow,
Glasgow, UK

C. Mugruza
School of Psychology, University of Dundee,
Dundee, UK

J. McLean
Institute of Neurological Sciences, University of Glasgow,
Glasgow, UK

G. Chakirova · A. C. Stanfield
Division of Psychiatry, University of Edinburgh,
Edinburgh, UK

J. Simpson
Department of Psychology, University of Stirling,
Stirling, UK

H. Johnston
Department of Psychology, University of St Andrews,
St Andrews, UK

N. Weir
Department of Medical Physics, Royal Infirmary of Edinburgh,
Edinburgh, UK

J. M. Wardlaw (✉)
SINAPSE Collaboration, Brain Research Imaging Centre,
Division of Clinical Neurosciences, Western General Hospital,
Crewe Road,
Edinburgh EH4 2XU, UK
e-mail: joanna.wardlaw@ed.ac.uk

*Present Address:*
J. De Wilde
The Higher Education Academy,
York, UK

Published online: 09 June 2012

at 1.5 and 3 T in human neuroimaging. We extracted data on methodology, quality criteria, technical factors, subjects, signal-to-noise, diagnostic accuracy and errors according to QUADAS and STARD criteria.

*Results* Amongst 150 studies (4,500 subjects), most were tiny, compared old 1.5 T with new 3 T technology, and only 22 (15 %) described diagnostic accuracy. The 3 T images were often described as "crisper", but we found little evidence of improved diagnosis. Improvements were limited to research applications [functional MRI (fMRI), spectroscopy, automated lesion detection]. Theoretical doubling of the signal-to-noise ratio was not confirmed, mostly being 25 %. Artefacts were worse and acquisitions took slightly longer at 3 T.

*Conclusion* Objective evidence to guide MRI purchasing decisions and routine diagnostic use is lacking. Rigorous evaluation accuracy and practicalities of diagnostic imaging technologies should be the routine, as for pharmacological interventions, to improve effectiveness of healthcare.

*Key Points*

- *Higher field strength MRI may improve image quality and diagnostic accuracy.*
- *There are few direct comparisons of 1.5 and 3 T MRI.*
- *Theoretical doubling of the signal-to-noise ratio in practice was only 25 %.*
- *Objective evidence of improved routine clinical diagnosis is lacking.*
- *Other aspects of technology improved images more than field strength.*

**Keywords** Magnetic resonance imaging · Sensitivity and specificity · Brain · Neuroimaging · Systematic review

## Introduction

Magnetic resonance imaging (MRI) has revolutionised modern medicine and research, particularly in neurosciences [1]. Since its introduction to human use in the mid 1980s, the operating field strengths have progressively increased, from very low [e.g. 0.15 Tesla (T)] to 1.5 T or 3 T for clinical use and up to 7 T for research.

Imaging at higher field strengths should increase the signal detected from tissue, with relatively less increase in background noise (higher signal-to-noise ratio; SNR) compared with lower field strengths, thereby producing better images with higher spatial resolution faster, increasing diagnostic accuracy [2]. For imaging methods where inherent signal strength is low, e.g. spectroscopy or functional MRI (fMRI), the higher field strengths could confer considerable advantage. However, the potential price is more artefacts [3], greater magnetic field heterogeneity, increased vigilance required for patient screening, increased biological effects [4, 5] and increased cost of equipment purchase, installation, maintenance and operation.

How should the prospective user, or purchaser, decide what field strength to use? Throughout medicine, treatment decisions are based on objective evidence from randomised clinical trials, without which no drug can obtain a licence. But diagnostic imaging is poorly served by randomised trials, or even comparative studies, for several reasons: perceived lack of need, high expense and constantly changing technologies. However, the costs of imaging are huge. In 2008, the diagnostic imaging machine market of the USA totalled over US$5 billion [6] with 27 MRIs per million population (only Japan had more at 40 per million population) and rising [7]. Images have disproportionately persuasive powers [8, 9] which can influence attitudes, perceptions and decisions regarding the value of imaging well beyond the validity, relevance or accuracy of the actual image [10]. Consequently there is little objective evidence to guide decisions regarding which imaging method to use, which may distort diagnoses and influence, possibly adversely, treatment decisions.

To address a fundamental question in clinical practice and research—which MR field strength to use and when—we systematically reviewed the literature on the use of 1.5 or 3 T MRI in clinical practice and research. We restricted this to neuroimaging, as the area with the longest use in clinical practice and research.

## Materials and methods

The work was undertaken by 28 individuals from the SINAPSE Collaboration (Scottish Imaging Network, A Platform for Scientific Excellence, http://www.SINAPSE.ac.uk) and Cochrane Stroke Group, from six different centres and seven main disciplines (neuroradiology 4, medical physics 12, psychiatry 2, image analysis 2, neuroscience 1, psychology 3 and general medical/data management/bibliometrics 4).

We used QUADAS [11] and STARD [12] criteria for assessing the studies and guidance on diagnostic test systematic reviews from the Cochrane Database of Systematic Reviews (http://srdta.cochrane.org/handbook-dta-reviews).

### Search strategy

We searched Medline and EMBASE from 1 January 2000 to 22 October 2010 using a search strategy (details in Electronic Supplementary Material) developed and refined with the aid of the Cochrane Stroke Group literature search coordinator (B. Thomas). We sought studies that imaged humans at 1.5 and 3 T. We excluded studies published only in abstract; those that only included animals, phantom objects or simulated data; review articles; and most studies not published in English

(owing to limited resources). Two researchers ran the search strategies and screened studies for further evaluation. Disagreements were resolved by discussion with a third reviewer. We obtained all potentially relevant papers and distributed them to 28 researchers for evaluation and data extraction. Six researchers (K.L., K.S., M.H., W.B., A.C., J.M.W.), including three neuroradiologists, reviewed the completed forms centrally for consistency of data extraction and clinical accuracy, and arbitrated on any differences in interpretation and abstraction of data. In the case of multiple publications on the same subjects, we took care only to count subjects once.

### Data extraction

We designed a dedicated data extraction form to record the study population, design characteristics (prospective, retrospective, etc.), body area examined, technical aspects of the equipment and sequences (including duration of imaging), analysis of the image data (blinding, number of observers, etc.), any information on sensitivity or specificity, the study's conclusions, the subjects' ability to tolerate the imaging and any adverse events or artefacts.

### Data synthesis

We entered the data into a dedicated, purpose-designed spreadsheet. We analysed information on study quality characteristics (sample size, blinding, prospective/retrospective, etc.), participants, whether the same subjects were definitely examined at both field strengths, evidence of bias in attributing a reference standard diagnosis, technical factors including the specifications of the equipment, whether the sequences had been optimised for each field strength, data on (or to calculate) sensitivity and specificity, artefacts, SNR and the authors' general impression of the diagnostic value. Our original aim was to perform a meta-analysis, but insufficient studies provided data on sensitivity or specificity. We therefore focussed on descriptive analyses. Where there were two or more relevant studies on a particular condition (e.g. stroke or tumours) or sequence (e.g. spectroscopy or angiography), we summarised all available data on diagnostic accuracy or SNR, as appropriate.

## Results

We identified 6,894 papers published between 1 January 2000 and 22 October 2010 (Fig. 1). After exclusion of 882 duplicates; 3,342 irrelevant papers on the basis of title/abstract; 132 reviews; 240 non-intracranial, non-neurological studies; 1,933 studies not on humans, not comparing 1.5 and 3 T or published only in abstract; and 90 non-English language papers, the remaining 275 publications were assessed in full (Fig. 1). Of these, 81 studies dealing solely with phantoms or artefacts,

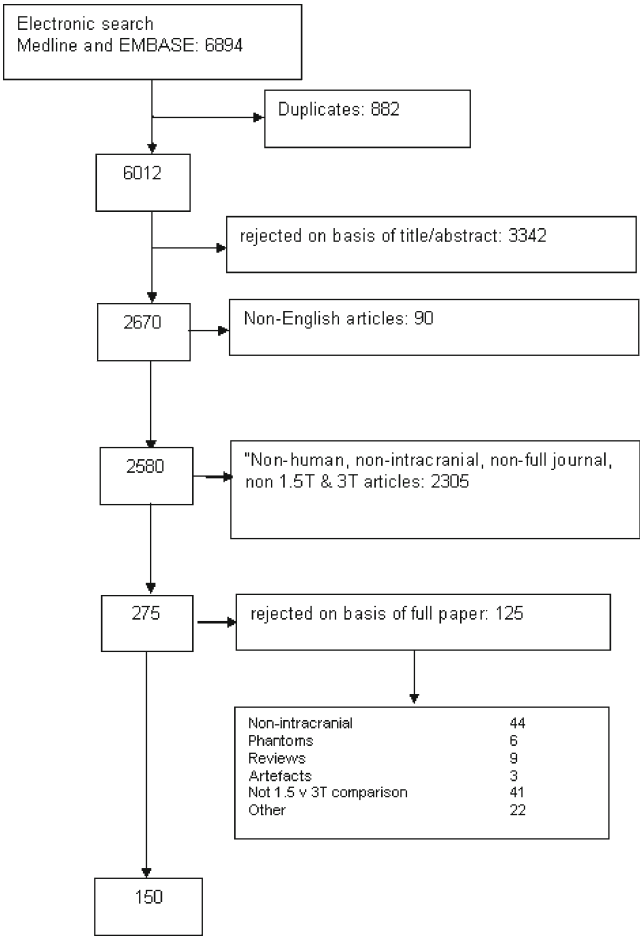


**Fig. 1** Results of the search

not comparing 1.5 and 3 T, and reviews, plus 44 on non-neurological disorders, were omitted. This left 150 studies that were included (Table 1, Electronic Supplemental Material; the full database is available at http://www.bric.ed.ac.uk).

### Study quality assessment

The 150 studies included 4,507 subjects. Only 121 studies (81 %, total $n$=1,935) definitely examined the same subjects at each field strength, 23 (15 %) examined different subjects and in 6 studies it was unclear (Table 1). A total of 147 studies gave a clear sample size (median 15, minimum 1, maximum 550, IQR 23.5), of which 36 % included 10 subjects or fewer. In the 121 studies that definitely examined the same subjects at both field strengths (median sample size 10.5, minimum 1, maximum 110, IQR 12), 49 % included 10 subjects or fewer (Fig. 1 in the Electronic Supplemental Material).

Study design was prospective in 127/150 (85 %), the rest being retrospective or unclear. Imaging at 1.5 T almost always preceded MR examinations at 3 T. The time between the two field strength acquisitions was <1 month in 53 studies, 1–6 months in 6 studies, >1 year in 9 studies and not mentioned

**Table 1** Study quality criteria

| | Yes | No | Unclear | Not reported |
|---|---|---|---|---|
| Prospective | 127 | 18 | 5 | |
| Same subjects at both field strengths | 121 | 23 | 6 | |
| Subject recruitment clear (inclusion/exclusion adequately reported) | 83 | 52 | 15 | |
| Blinding of readers to field strength | 27 | 32 | 54 | 37 |
| Blinding of readers to results at alternative field strength | | | | |
| 1.5 T | 42 | 19 | 65 | 24 |
| 3.0 T | 37 | 23 | 67 | 23 |
| Normal volunteers only | 71 | | | |
| Patients only | 42 | | | |
| Time between acquisitions < 1 month | 53 | 15 | 49 | 33 |
| Treatment given between acquisitions | 12 | 68 | 39 | 31 |
| Relevant to disease spectrum | 61 | 40 | 24 | 25 |
| Was disease verified externally? | 34 | 50 | 12 | 54 |
| Data on observer reliability/ reproducibility | 11 | 87 | 10 | 42 |
| Imaging described in replicable detail | | | | |
| 1.5 T | 122 | 21 | 6 | 1 |
| 3.0 T | 128 | 16 | 6 | 0 |
| Reporting of uninterpretable/ intermediate test results | 22 | 84 | 9 | 35 |
| Reporting of withdrawals | 14 | 106 | 8 | 22 |
| Reporting of artefacts | 46 | 85 | 16 | 3 |

Values are number of studies

in 82/150 studies (55 %). In 12 studies, a potentially disease-modifying treatment had been given between the acquisitions. Few studies reported blinding of analysis (27 studies, 18 %), subject recruitment methods, withdrawals or artefacts, or information on tolerability of either field strength.

## Subjects

Most studies (71/150; 48 %) involved only healthy volunteers, 42/150 (28 %) concerned patients with a neurological disorder, 11 (7 %) a mixed population and 26 (17 %) failed to describe the subjects. Ten papers concerned tumours, 14 multiple sclerosis (MS), 4 stroke, 13 aneurysm/arteriovenous malformation (AVM)/other vascular conditions, 9 epilepsy and the rest various miscellaneous conditions. Only 34 studies (23 %) verified any disease independently.

## Technical factors

Relatively little information was available concerning the generation or age of the MR equipment, software version or factors such as slew rates (Table 2). Roughly half the studies appeared to optimise their sequences for field strength. Forty-four studies used the same make and model series of MR equipment, 46 used the same make but not model series, 25 used different makes and 35 (23 %) did not state MR manufacturer or model.

Sixty-eight studies that provided the information showed no difference in imaging duration between 3 T (mean 15 min, 41 s) and 1.5 T (mean 15 min, 27 s) (Mann-Whitney, $P$=0.95 two-tailed). The study purpose was lesion detection (44; 29 %), technical development (43; 29 %), anatomical definition (13; 9 %), side effects assessment (1; 0.1 %) or a combination of the above (49; 33 %). Sixty-four studies concerned structural sequences, 16 diffusion tensor MRI, 27 fMRI, 13 spectroscopy, 5 perfusion imaging, 14 MRA and 24 concerned some other form of imaging (Fig. 2 in the Electronic Supplemental Material).

SNR and contrast-to-noise ratio (CNR) were both higher at 3 T, but the improvement varied with the category of imaging sequence. On angiography, SNR was higher at 3 T by a factor of 1.8–2.3 [13–16]. On spectroscopy, SNR was clearly increased, although the 100 % theoretical increase was not achieved, even in phantoms, the documented increases being 23 % [17], 28 % [18], 25–35 % [19], 23–46 % [20], or 50 % [21]. Most gain in SNR was with short echo time spectroscopy with little improvement at long echo times [18, 19, 22]. Use of a phased array or eight-channel coil instead of a quadrature coil improved the SNR to a greater extent than did higher field strength [19]. CNR was doubled on angiography [13, 16, 23], although several

**Table 2** Technical factors

| | Yes | No | Unclear | Not relevant |
|---|---|---|---|---|
| MRI manufacturer same at 1.5 and 3 T | 90 | 25 | 35 | |
| If same, is the model the same? | 44 | 46 | | |
| Slew rate stated | | | | |
| 1.5 T | 23 | | 127 | |
| 3.0 T | 25 | | 125 | |
| Age of MRI given | | | | |
| 1.5 T | 2 | | 148 | |
| 3.0 T | 2 | | 148 | |
| Same coil design at 1.5 and 3 T | 61 | 38 | 51 | |
| Sequences optimised | | | | |
| 1.5 T | 75 | 8 | 67 | |
| 3.0 T | 79 | 12 | 59 | |
| Imaging duration stated | | | | |
| 1.5 T | 64 | 86 | | |
| 3.0 T | 65 | 85 | | |

Values are number of studies

parameters affecting CNR were inconsistent at 3 and 1.5 T. CNR between enhancing tumour and non-enhancing brain was 1.2 to 2.8 times higher at 3 T on contrast-enhanced imaging with spin echo T1-weighted or magnetisation prepared rapid gradient echo (MP-RAGE) sequences [24, 25], but was four times lower in tumour vs. brain unenhanced [25] compared with 1.5 T images.

Artefacts were more common at 3 T but had little effect on diagnostic accuracy (see below). Only two studies mentioned adverse events (one at 1.5 T and one at both 1.5 and 3 T) and 14 studies specifically said there were no adverse events, but the rest did not mention adverse events. Thirteen studies reported that subjects tolerated imaging at both field strengths, the rest failing to mention tolerability.

## Diagnostic accuracy

Most studies (128; 85 %) did not provide sensitivity or specificity data or raw data for their calculation. Of the 22 studies (15 %) that did provide some sensitivity and specificity information, 7 calculated accuracy on a per patient basis [17, 26–31], 9 calculated accuracy on a per lesion basis [32–40], 4 calculated both [16, 41–43] and in two studies, sensitivity and specificity were irrelevant [44, 45], precluding a formal meta-analysis. We provide, instead, summaries of lesion detection, clinical relevance, and true and false positives and negatives (Table 3). We summarise the qualitative information to derive an overall indication of the utility of each field strength.

## Vascular abnormalities

Most studies [10, 20, 35, 43], mostly of intracranial aneurysms, reported superior image quality at 3 T, but with little effect on diagnostic accuracy [13, 14, 46]; for example, visualisation of small terminal arterial branches [13, 16, 23, 47], pathological vessels in moyamoya disease [48], intracranial aneurysms [14], residual neck in treated aneurysms [27, 49], and vascular supply of AVMs [42] was better at 3 T. Others found no difference in aneurysm diagnosis, characterisation or recurrence detection between 1.5 and 3 T despite differences in image quality [38, 47]. There was no difference in MR venography between the two field strengths [46, 50]. Artefacts were virtually all more pronounced at 3 T [13, 23, 38, 46], but, with a few exceptions [38], they had little effect on diagnosis.

**Table 3** Information on lesion detection at 1.5 and 3 T

| | At 1.5 T | At 3.0 T | Same | Mixed | Not reported/ unclear |
|---|---|---|---|---|---|
| More lesions detected | 2 | 29 | 9 | 5 | 105 |
| More positive diagnosis of disease | 2 | 16 | 11 | | 121 |
| More true positives | 1 | 11 | 7 | | 131 |
| More true negatives | 0 | 3 | 6 | | 141 |
| More false positives | 2 | 2 | 7 | | 139 |
| More false negatives | 8 | 3 | 7 | | 132 |
| Difference between 1.5 and 3.0 T[a] | 3 | 81 | 23 | 24 | 19 |

Values are number of studies

[a] From left to right: 1.5T better; 3T better; Same (ie 1.5 and 3T differed but one was not better than the other); mixed (either 1.5 or 3T better in some aspects)

## Tumours

All large solitary lesions were equally well visualised at both field strengths, although subjectively lesions were more conspicuous at 3 T [24, 25, 51]. Other differences, such as different doses of contrast medium, confounded comparisons of field strength. Triple dose contrast medium increased lesion conspicuity more than increasing field strength [24, 51]. A small proportion of very small (<5 mm diameter) metastases were only detected at 3 T, but there was no difference in the detection of lesions >5 mm [51]. Pituitary microadenomas may be more clearly defined (2/5 [30], 3/6 [52]) or located (1/5 [30]) at 3 T as there was one false positive at 1.5 T [30]. However, other substantial differences in imaging sequences (dynamic at 3 T, standard contrast-enhanced at 1.5 T [52]), or use of different machines in different imaging centres, precluded reliable comparisons. Susceptibility (pulsation and ringing) artefacts were worse at 3 T [25, 51].

## Multiple sclerosis

Differences in numbers of enhancing lesions seen at 3 vs. 1.5 T were inconsistent, including 7.5 % [33] and 21 % [29] in two studies, but only 0.5/patient overall ($P$ = NS) [53]. Few patients (2.8 % [32] to 3.6 % [53]) showed lesion enhancement only at 3 T. Suggested increases in lesion conspicuity on fluid attenuated inversion recovery (FLAIR) or T2-weighted imaging at 3 T ranged from 13 to 30 [34, 36] (average difference 3.9 %) [53]. Differences in lesion enhancement and volume were inconsistent [32, 53] and potentially confounded by differences in sequences [34]. Most lesions missed at 1.5 T were very small (1–5 mm [34]) or in the immediate periventricular tissue [34, 36]. The better tissue resolution at 3 T improved lesion detection with automated image analysis software (details not provided) [32]. Artefacts were more common at 3 T but were in areas of brain not affected by MS plaques [34, 54]. Better lesion conspicuity at 3 T could alter disease categorisation, e.g. 15 % of patients fulfilled one additional Swanton criterion and 27.5 % fulfilled one additional Barkhof criterion leading

to a diagnosis of "disseminated in space" according to McDonald criteria in 2.5 % [55]. There was no evidence that higher field strengths led to an earlier diagnosis of definite MS in patients with a clinically isolated syndrome [54] and no reliable data on whether there were more false positive diagnoses at 3 T as only 20 healthy volunteers were included [54]. There was no difference in inter-rater reliability between field strengths [53].

### Spectroscopy

Better spectral resolution at 3 T improved separation of metabolite peaks [17, 18, 20, 21, 56, 57], particularly at short echo times for closely located metabolites [18, 22, 57], although the increase in resolution was offset by an increase in line width at 3 T [18]. No significant differences were found in metabolite ratios [22] or in whole brain N-acetylaspartate concentrations [58] between 1.5 and 3 T. Coefficients of variation for metabolite ratios and between-examination and subject variability were higher at 3 T for short echo time spectroscopy, but most did not reach statistical significance in one small study [17]. Reproducibility of metabolite concentrations was better for long than short echo time spectroscopy and was independent of field strength at long echo times; coil characteristics (phased array vs. quadrature) had as much effect on metabolite quantification as field strength [19]. There was virtually no information on spectroscopy in clinical diagnosis [17, 59].

### fMRI

Most studies had eight or fewer (range 1–14) normal subjects with several duplicate publications. Where given, most papers indicated increases in the number (average 40 %; range 23–82 %) and intensity of active voxels at 3 vs. 1.5 T, mostly in visual or motor cortex. Only one paper reported effects in basal brain regions, indicating no difference in amygdaloid signal activation and increased artefactual signal dropout at 3 vs. 1.5 T [60].

## Discussion

We found only subjective evidence of increased lesion detection, finer anatomical detail or improved image resolution at 3 T and no evidence of reduced MR examination times or improved diagnostic accuracy vs. 1.5 T. Many of these differences could also be attributed to differences in generation of technology, sequences, coils or use of contrast agents. The conclusion is limited by the poor quality of the literature and numerous potential biases. Objective evidence of the advantages of 3 T only applies to short echo time spectroscopy, fMRI in non-basal brain regions and detection of small MS lesions with automated software, all of which are predominantly research tools at present [1]. The theoretical advantages of imaging at 3 T rather than 1.5 T, extensively detailed in review articles and anecdotally, which may in practice be considerable, are not supported by the current literature.

On the whole, sample sizes were extremely small (10 or fewer), with little information on blinding, time lapses between acquisitions or treatment that allowed genuine disease differences, little information on the actual breadth of disease encountered in routine practice or research, frequent use of incomparable technical factors, virtually no information on tolerability and increased artefacts and, when measured, the theoretical 100 % increase in SNR was nearer 25 % on average. The failure to achieve the expected doubling of SNR can be explained by increased magnetic field heterogeneity, susceptibility and T2 shortening effects at 3 T. There may be other unpublished work that would provide substantive evidence to fill these gaps in the published literature, but unpublished data are very difficult to find and of lesser reliability than published data as they have not been evaluated in a rigorous peer review process.

The exceptions where 3 T may outperform 1.5 T are all currently research applications: short echo time spectroscopy, fMRI of non-basal regions and possibly detection of small MS lesions with automated software. In magnetic resonance spectroscopy (MRS) and fMRI, a very subtle low amplitude signal, little stronger than background noise, is being sought. In MRS, even a 25 % increase in SNR (rather than the theoretical 100 %) may improve spectral quality and hence improve differentiation between closely located metabolites, particularly at short echo times. Perfusion imaging with arterial spin labelling may also benefit from 3 T, but we found no comparative studies on this. Improved detection of small MS lesions with automated software may arise from several factors, not just field strength. The limited fMRI evidence indicating increased activation of voxels (both number and intensity) at 3 T generally referred to non-basal brain areas. For structures near the skull base, such as amygdala, hippocampus or inferior frontal regions, 3 T conferred no advantage because the increased signal drop-out from adjacent skull base structures overwhelmed any advantage of increased voxel activation. Hence, the brain region of interest needs to be considered when deciding if an fMRI experiment should be performed at 1.5 or 3 T.

The weaknesses of this review include the difficulty of summarising this literature where the papers varied substantially in the information provided. We lacked resources to translate most non-English language papers or obtain further information from the authors, but, with a median sample size of 10.5, the likelihood of obtaining useful additional information is low. Our original aim was to perform a meta-

analysis of sensitivity and specificity, but the available data were so limited, and in such variable formats, that this was impossible. We may have included inadvertently some subjects more than once, although we made every attempt to avoid this. We included all neuroimaging studies for which we found data, i.e. we did not find any studies comparing diffusion tensor imaging at 1.5 and 3 T and we found little on spinal imaging. It was beyond the remit of this review to evaluate the literature on cardiac, musculo-skeletal, abdominal, pelvic or other MR applications, but a similar exercise would be useful for those interested in imaging these anatomical areas. Likewise, it was not possible to determine subjective or biological effects [4, 5].

The strengths include central data verification to standardise the more subjective components of this literature, our use of established standards for systematic reviews, including those specific to diagnostic imaging [11, 12], our considerable experience of systematic reviews including those of diagnostic tests, professional advice from a highly experienced professional literature searcher and many study evaluators and data extractors with considerable in-depth professional knowledge of MRI. It was therefore possible to take every advantage of the relatively sparse information provided in the papers, making the large number of reviewers a strength rather than a weakness.

Magnetic resonance imaging technology at 1.5 and 3 T is rapidly evolving. The annual incremental developments will likely lead to further improvements in image quality, artefact reduction, MR data acquisition times, better integration with ancillary devices such as electroencephalography (EEG) recording, the patient experience and hopefully diagnostic accuracy at both field strengths. Many of these future advances, as well as those in the last 10 years as suggested by several included papers, could have more effect on improving diagnostic accuracy or image quality than changes in field strength. It would be advisable for researchers and clinicians to contemplate the full range of performance features rather than just the field strength when considering which equipment to use.

Comparative studies of diagnostic imaging technologies are difficult: the technologies are large, expensive and usually immobile. However, trials of new drugs or medical devices are not particularly easy or cheap either, and perceived difficulty in performing valid comparisons should not be reinforced by an assumption of lack of need for objective comparisons. The timing of comparisons in the present review suggests that most took place when an older, usually 1.5 T MR system, was being replaced with a newer, often 3 T device, allowing for clinical disease progression and technical advances independent of field strength to confound comparison. There appeared to be few deliberate comparisons of equal generations of field strength MR systems in institutions with access to both. The consequence of not taking advantage of such situations is that the available data are from older, less technologically advanced 1.5 T systems, likely to perform poorly against newer equipment, whether the latter is 1.5 or 3 T.

It is unfortunate that diagnostic imaging is rarely subjected to the same scrutiny, or underpinned by the same evidence base, as is demanded for treatments despite being central to most modern medical care. This means that the accuracy of diagnoses underlying many treatment decisions may be suboptimal, in turn undermining treatment benefits. Diagnostic imaging is expensive, taking up increasing proportions of health care and research budgets, and may even be harmful: the largest single source of population radiation exposure in developed countries is medical diagnostic x-rays [61]. The costs are substantial [7].

To conclude, healthcare providers, government funding agencies, health departments, researchers, clinicians and patients should be more aware of the evidence deficit in diagnostic imaging. They should implement strategies to evaluate diagnostic imaging technologies rigorously, in the circumstances in which they are intended to be used in practice and thereby avoid further unnecessary inflation of already costly healthcare budgets while improving clinical outcomes.

**Acknowledgements** The SINAPSE Collaboration (Scottish Imaging Network, A Platform for Scientific Excellence, www.sinapse.ac.uk) is funded by the Scottish Funding Council and supports the salaries of many of the authors. The project did not receive any other specific funding.

## References

1. Wellcome Trust (2011) Human functional brain imaging 1990–2009: portfolio review. Wellcome Trust, London
2. Alvarez-Linera J (2008) 3 T MRI: advances in brain imaging. Eur J Radiol 67:415–426
3. Frayne R, Goodyear BG, Dickhoff P, Lauzon ML, Sevick RJ (2003) Magnetic resonance imaging at 3.0 Tesla: challenges and advantages in clinical neurological imaging. Invest Radiol 38:385–402
4. Weintraub MI, Khoury A, Cole SP (2007) Biologic effects of 3 Tesla (T) MR imaging comparing traditional 1.5 T and 0.6 T in 1023 consecutive outpatients. J Neuroimaging 17:241–245
5. de Vocht F, Stevens T, van Wendel-de-Joode B, Engels H, Kromhout H (2006) Acute neurobehavioral effects of exposure to static magnetic fields: analyses of exposure-response relations. J Magn Reson Imaging 23:291–297
6. Millennium Research Group (2011) Funding frugality: US health care facilities reduce imaging spending. http://mrg.net/News-and-Events/Medtech-Confidence-Index/Funding-Frugality.aspx. Accessed 9 Nov 2011
7. McKinsey Global Institute (2008) Accounting for the cost of U.S. health care: a new look at why Americans spend more. http://www.mckinsey.com/Insights/MGI/Research/Americas/Accounting_for_the_cost_of_US_health_care. Accessed 11 Nov 2011

8. Farah MJ (2009) A picture is worth a thousand dollars. J Cogn Neurosci 21:623–624
9. Roskies AL (2007) Are neuroimages like photographs of the brain? Philos Sci 74:860–872
10. Wardlaw JM, Schafer B (2011) What am I thinking and who has the right to know? Contributions from a workshop on the wider societal implications of neuroimaging. Cortex 47:1147–1150
11. Whiting PF, Weswood ME, Rutjes AW, Reitsma JB, Bossuyt PN, Kleijnen J (2006) Evaluation of QUADAS, a tool for the quality assessment of diagnostic accuracy studies. BMC Med Res Methodol 6:9
12. Bossuyt PM, Reitsma JB, Bruns DE et al (2003) Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. Clin Radiol 58:575–580
13. Gaa J, Weidauer S, Requardt M, Kiefer B, Lanfermann H, Zanella FE (2004) Comparison of intracranial 3D-ToF-MRA with and without parallel acquisition techniques at 1.5 T and 3.0 T: preliminary results. Acta Radiol 45:327–332
14. Bernstein MA, Huston JI, Lin C, Gibbs GF, Felmlee JP (2001) High-resolution intracranial and cervical MRA at 3.0 T: technical considerations and initial experience. Magn Reson Med 46:955–962
15. Zou Z, Ma L, Cheng L, Cai Y, Meng X (2008) Time-resolved contrast-enhanced MR angiography of intracranial lesions. J Magn Reson Imaging 27:692–699
16. Willinek WA, Born M, Simon B et al (2003) Time-of-flight MR angiography: comparison of 3.0-T imaging and 1.5-T imaging—initial experience. Radiology 229:913–920
17. Kantarci K, Reynolds G, Petersen RC et al (2003) Proton MR spectroscopy in mild cognitive impairment and Alzheimer disease: comparison of 1.5 and 3 T. AJNR Am J Neuroradiol 24:843–849
18. Barker PB, Hearshen DO, Boska MD (2001) Single-voxel proton MRS of the human brain at 1.5 T and 3.0 T. Magn Reson Med 45:765–769
19. Inglese M, Spindler M, Babb JS, Sunenshine P, Law M, Gonen O (2006) Field, coil, and echo-time influence on sensitivity and reproducibility of brain proton MR spectroscopy. AJNR Am J Neuroradiol 27:684–688
20. Li Y, Osorio JA, Ozturk-Isik E et al (2006) Considerations in applying 3D PRESS H-1 brain MRSI with an eight-channel phased-array coil at 3 T. Magn Reson Imaging 24:1295–1302
21. Sjobakk TE, Lundgren S, Kristoffersen A et al (2006) Clinical 1 H magnetic resonance spectroscopy of brain metastases at 1.5 T and 3 T. Acta Radiol 47:501–508
22. Kim JH, Chang KH, Na DG et al (2006) Comparison of 1.5 T and 3 T 1 H MR spectroscopy for human brain tumors. Korean J Radiol 7:156–161
23. Al-Kwifi O, Emery DJ, Wilman AH (2002) Vessel contrast at three Tesla in time-of-flight magnetic resonance angiography of the intracranial and carotid arteries. Magn Reson Imaging 20:181–187
24. Krautmacher C, Willinek WA, Tschampa HJ et al (2005) Brain tumors: full- and half-dose contrast-enhanced MR imaging at 3.0 T compared with 1.5 T—initial experience. Radiology 237:1014–1019
25. Nobauer-Huhmann IM, Ba-Ssalamah A, Mlynarik V et al (2002) Magnetic resonance imaging contrast enhancement of brain tumors at 3 Tesla versus 1.5 Tesla. Invest Radiol 37:114–119
26. Wolfsberger S, Ba-Ssalamah A, Pinker K et al (2004) Application of three-tesla magnetic resonance imaging for diagnosis and surgery of sellar lesions. J Neurosurg 100:278–286
27. Ramgren B, Siemund R, Cronqvist M et al (2008) Follow-up of intracranial aneurysms treated with detachable coils: comparison of 3D inflow MRA at 3 T and 1.5 T and contrast-enhanced MRA at 3 T with DSA. Neuroradiology 50:947–954
28. Knake S, Triantafyllou C, Wald LL et al (2005) 3 T phased array MRI improves the presurgical evaluation in focal epilepsies: a prospective study. Neurology 65:1026–1031
29. Sawaishi Y, Sasaki M, Yano T, Hirayama A, Akabane J, Takada G (2005) A hippocampal lesion detected by high-field 3 tesla magnetic resonance imaging in a patient with temporal lobe epilepsy. Tohoku J Exp Med 205:287–291
30. Kim LJ, Lekovic GP, White WL, Karis J (2007) Preliminary experience with 3-Tesla MRI and Cushing's disease. Skull Base 17:273–277
31. Schaafsma JD, Velthuis BK, Majoie CB et al (2010) Intracranial aneurysms treated with coil placement: test characteristics of follow-up MR angiography—multicenter study. Radiology 256:209–218
32. Sicotte NL, Voskuhl RR, Bouvier S, Klutch R, Cohen MS, Mazziotta JC (2003) Comparison of multiple sclerosis lesions at 1.5 and 3.0 Tesla. Invest Radiol 38:423–427
33. Zou KH, Greve DN, Wang M et al (2005) Reproducibility of functional MR imaging: preliminary results of prospective multi-institutional study performed by Biomedical Informatics Research Network. Radiology 237:781–789
34. Bachmann R, Reilmann R, Schwindt W, Kugel H, Heindel W, Kramer S (2006) FLAIR imaging for multiple sclerosis: a comparative MR study at 1.5 and 3.0 Tesla. Eur Radiol 16:915–921
35. Kaufmann TJ, Huston J 3rd, Cloft HJ et al (2010) A prospective trial of 3T and 1.5T time-of-flight and contrast-enhanced MR angiography in the follow-up of coiled intracranial aneurysms. AJNR Am J Neuroradiol 31:912–918
36. Wattjes MP, Lutterbey GG, Harzheim M et al (2006) Higher sensitivity in the detection of inflammatory brain lesions in patients with clinically isolated syndromes suggestive of multiple sclerosis using high field MRI: an intraindividual comparison of 1.5 T with 3.0 T. Eur Radiol 16:2067–2073
37. Simon B, Schmidt S, Lukas C et al (2010) Improved in vivo detection of cortical lesions in multiple sclerosis using double inversion recovery MR imaging at 3 Tesla. Eur Radiol 20:1675–1683
38. Buhk JH, Kallenberg K, Mohr A, Dechent P, Knauth M (2008) No advantage of time-of-flight magnetic resonance angiography at 3 Tesla compared to 1.5 Tesla in the follow-up after endovascular treatment of cerebral aneurysms. Neuroradiology 50:855–861
39. Rosso C, Drier A, Lacroix D et al (2010) Diffusion-weighted MRI in acute stroke within the first 6 hours: 1.5 or 3.0 Tesla? Neurology 74:1946–1953
40. Haacke EM, Makki M, Ge Y et al (2009) Characterizing iron deposition in multiple sclerosis lesions using susceptibility weighted imaging. J Magn Reson Imaging 29:537–544
41. Kuhl CK, Textor J, Gieseke J et al (2005) Acute and subacute ischemic stroke at high-field-strength (3.0-T) diffusion-weighted MR imaging: intraindividual comparative study. Radiology 234:509–516
42. Heidenreich JO, Schilling AM, Unterharnscheidt F et al (2007) Assessment of 3D-TOF-MRA at 3.0 Tesla in the characterization of the angioarchitecture of cerebral arteriovenous malformations: a preliminary study. Acta Radiol 48:678–686
43. Stehling C, Wersching H, Kloska SP et al (2008) Detection of asymptomatic cerebral microbleeds: a comparative study at 1.5 and 3.0 T. Acad Radiol 15:895–900
44. Sohn CH, Sevick RJ, Frayne R, Chang HW, Kim SP, Kim DK (2010) Fluid attenuated inversion recovery (FLAIR) imaging of the normal brain: comparisons between under the conditions of 3.0 Tesla and 1.5 Tesla. Korean J Radiol 11:19–24
45. van der Zwaag W, Francis S, Head K et al (2009) fMRI at 1.5, 3 and 7 T: characterising BOLD signal changes. Neuroimage 47:1425–1434
46. Hu HH, Haider CR, Campeau NG, Huston JI, Riederer SJ (2008) Intracranial contrast-enhanced magnetic resonance venography with 6.4-fold sensitivity encoding at 1.5 and 3.0 Tesla. J Magn Reson Imaging 27:653–658
47. Gibbs GF, Huston JI, Bernstein MA, Riederer SJ, Brown RD Jr (2004) Improved image quality of intracranial aneurysms: 3.0-T

versus 1.5-T time-of-flight MR angiography. AJNR Am J Neuroradiol 25:84–87
48. Fushimi Y, Miki Y, Kikuta K et al (2006) Comparison of 3.0- and 1.5-T three-dimensional time-of-flight MR angiography in moyamoya disease: preliminary experience. Radiology 239:232–237
49. Anzalone N, Scomazzoni F, Cirillo M et al (2008) Follow-up of coiled cerebral aneurysms: comparison of three-dimensional time-of-flight magnetic resonance angiography at 3 Tesla with three-dimensional time-of-flight magnetic resonance angiography and contrast-enhanced magnetic resonance angiography at 1.5 Tesla. Invest Radiol 43:559–567
50. Barth M, Nobauer-Huhmann IM, Reichenbach JR et al (2003) High-resolution three-dimensional contrast-enhanced blood oxygenation level-dependent magnetic resonance venography of brain tumors at 3 Tesla: first clinical experience and comparison with 1.5 Tesla. Invest Radiol 38:409–414
51. Ba-Ssalamah A, Nobauer-Huhmann IM, Pinker K et al (2003) Effect of contrast dose and field strength in the magnetic resonance detection of brain metastases. Invest Radiol 38:415–422
52. Agati R, Maffei M, Bacci A, Cevolani D, Battaglia S, Leonardi M (2004) 3 T MR assessment of pituitary microadenomas: a report of six cases. Riv Neuroradiol 17:890–895
53. Nielsen K, Rostrup E, Frederiksen JL et al (2006) Magnetic resonance imaging at 3.0 Tesla detects more lesions in acute optic neuritis than at 1.5 Tesla. Invest Radiol 41:76–82
54. Wattjes MP, Harzheim M, Kuhl CK et al (2006) Does high-field MR imaging have an influence on the classification of patients with clinically isolated syndromes according to current diagnostic MR imaging criteria for multiple sclerosis? AJNR Am J Neuroradiol 27:1794–1798
55. Wattjes MP, Harzheim M, Lutterbey GG et al (2008) Does high field MRI allow an earlier diagnosis of multiple sclerosis? J Neurol 255:1159–1163
56. Gonen O, Gruber S, Li BS, Mlynarik V, Moser E (2001) Multivoxel 3D proton spectroscopy in the brain at 1.5 versus 3.0 T: signal-to-noise ratio and resolution comparison. AJNR Am J Neuroradiol 22:1727–1731
57. Ethofer T, Mader I, Seeger U et al (2003) Comparison of longitudinal metabolite relaxation times in different regions of the human brain at 1.5 and 3 Tesla. Magn Reson Med 50:1296–1301
58. Benedetti B, Rigotti DJ, Liu S, Filippi M, Grossman RI, Gonen O (2007) Reproducibility of the whole-brain N-acetylaspartate level across institutions, MR scanners, and field strengths. AJNR Am J Neuroradiol 28:72–75
59. Inoue T, Ogasawara K, Kumabe T, Jokura H, Watanabe M, Ogawa A (2005) Minute glioma identified by 3.0 Tesla magnetic resonance spectroscopy—case report. Neurol Med Chir (Tokyo) 45:108–111
60. Krasnow B, Tamm L, Greicius MD et al (2003) Comparison of fMRI activation at 3 and 1.5 T during perceptual, cognitive, and affective processing. Neuroimage 18:813–826
61. Fazel R, Krumholz HM, Wang Y et al (2009) Exposure to low-dose ionizing radiation from medical imaging procedures. N Engl J Med 361:849–857