# Long-interval T2w subtraction MRI: A powerful new outcome measure in MS trials

Running Head: T2w subtraction MR imaging in phase II MS trials

*Bastiaan Moraal[1], MD*
*Ivo J. van den Elskamp[1], MD*
*Dirk L. Knol[2], PhD*
*Bernard M.J. Uitdehaag[2,6], MD, PhD*
*Jeroen J.G. Geurts[1,3], PhD*
*Hugo Vrenken[1,4], PhD*
*Petra J.W. Pouwels[4], PhD*
*Ronald A. van Schijndel[5], MSc*
*Dominik S. Meier[7], PhD*
*Charles R.G. Guttmann[7], MD*
*Frederik Barkhof[1], MD, PhD*

*[1]Department of Radiology, [2]Department of Epidemiology and Biostatistics, [3]Department of Pathology, [4]Department of Physics and Medical Technology, [5]Department of Informatics, [6]Department of Neurology, MS Center Amsterdam, VU University Medical Center, Amsterdam, The Netherlands. [7]Center for Neurological Imaging, Departments of Radiology and Neurology, Brigham and Women's Hospital, Harvard Medical School, Boston, USA*

**Corresponding author:**
Bastiaan Moraal
VU University Medical Center
Department of Radiology
De Boelelaan 1117
1081 HV Amsterdam
The Netherlands
Phone # (+31) 20 444 4596
Fax # (+31) 20 444 0397
Email: b.moraal@vumc.nl

| | |
|---|---|
| **Characters in the title (80):** | **78** |
| **Running head (50) :** | **48** |
| **Words in the abstract (250):** | **249** |
| **Words in the body (3000):** | **3227** |
| **Number of tables:** | **3** |
| **Number of figures:** | **5** |

**Abstract**

**Objective:** To compare long-interval T2-weighted subtraction (T2w-Sub) imaging with monthly gadolinium-enhanced T1-weighted (Gd-T1w) imaging for 1) detection of active lesions, 2) assessment of treatment efficacy, and 3) statistical power, in a multiple sclerosis (MS), phase II, clinical trial setting.

**Methods:** MRI data over 9 months from 120 patients (61 treatment, 59 placebo) from the oral temsirolimus trial were used. T2w-Sub images were scored for active lesions, independent of the original reading of the monthly Gd-T1w images. Treatment efficacy was evaluated using the nonparametric Mann-Whitney U test, and parametric negative binomial(NB)-regression, and power calculations were conducted.

**Results:** Datasets from 116 patients (58 treatment, 58 placebo) were evaluated. The mean number of T2w-Sub lesions in the treatment group was 3.0(±4.6) vs. 5.9(±8.8) for placebo; mean cumulative number of new Gd-T1w lesions in the treatment group was 5.5(±9.1) vs. 9.1(±17.2) for placebo. T2w-Sub imaging showed increased power to assess treatment efficacy compared with Gd-T1w imaging, when evaluated by: Mann-Whitney U test (P=.017 vs. P=.177), NB-regression without (P=.011 vs. P=.092) or with baseline adjustment (P<.001 vs. P=.002). Depending on the magnitude of the simulated treatment effect, sample size calculations showed reductions of 22%-34% in the number of patients (translating into reductions of 81%-83% in the number of MRI scans) needed to detect a significant treatment effect in favor of T2w subtraction imaging.

**Interpretation:** **C**ompared with monthly Gd-enhanced T1w imaging, long-interval T2w subtraction MRI exhibited increased power to assess treatment efficacy, and could greatly increase the cost-effectiveness of phase II MS trials, by limiting the number of patients, contrast injections and MRI scans needed.

**Introduction**

Magnetic resonance imaging (MRI) is used as the primary outcome measure in multiple sclerosis (MS), phase II, clinical trials as it represents a more sensitive and reproducible method to monitor disease activity over time, compared with clinical measures [1,2]. Gadolinium(Gd)-enhancement is thought to reflect acute inflammation in which the blood-brain barrier is disrupted [3,4], persisting on average for around one month [5,6]. Anti-inflammatory properties of new treatments are thus assessed using serial Gd-enhanced T1-weighted (T1w) imaging, which is costly because of monthly scanning and contrast injections. Furthermore, the repeated administration of gadolinium could be potentially harmful, given the possible association between the use of Gd–based contrast media and the occurrence of nephrogenic systemic fibrosis, especially in patients with (chronic) renal insufficiency [7,8].

As the scanning interval between consecutive MRI scans is lengthened, an increasing amount of disease activity is likely to be missed when assessed solely by Gd-enhanced T1w imaging. Therefore, trials of longer duration make supplementary use of T2-weighted (T2w) MR scans, performed at 3-monthly or even yearly intervals to monitor disease activity. In this setting, lesions on T2w images are thought to reflect the cumulative residual of disease activity [9]. However, the detection of active (new and enlarged) T2 lesions using serial conventional MR imaging is complicated, by repositioning errors and a background of unaltered non-active lesions [10].

T2w subtraction imaging, after image registration, provides an alternative in which the effect of repositioning errors is reduced, and the contrast between active lesions and the non-active background is enhanced as stable lesions cancel out. This method detected higher numbers of active lesions with greater interobserver agreement, compared with a standard analysis of non-registered T2w images, in both a single and multi-center clinical trial setting [11,12]. Furthermore, subtraction imaging identified higher lesion volumes of disease activity, and demonstrated increased power to evaluate treatment efficacy using direct quantification of

positive disease activity, compared with standard volumetric analyses [12,13]. Until now, long-interval T2w subtraction imaging has not been directly compared with serial Gd-enhanced T1w imaging for the identification of active lesions and assessment of treatment efficacy. If the residual disease activity identified by T2w subtraction imaging over many months exhibits at least the same power to evaluate treatment efficacy compared with serial Gd-enhanced T1w imaging, this greatly would increase the cost-effectiveness of clinical trials (by limiting the number of patients, scans and contrast injections needed), and increase patient safety and compliance. Hence the goals in this study were to explore the applicability of long-interval T2w subtraction imaging to 1) detect active lesions, 2) assess treatment efficacy, and 3) secure statistical power, in comparison with serial monthly Gd-enhanced T1w imaging, in an MS, phase II, clinical trial setting.

## Materials and Methods

*Patients and MR Image Acquisition*

Serial MR images were derived from a multicenter, randomized, double-blind, placebo-controlled, phase II, clinical trial, described in more detail in [14]. Briefly, 296 patients (with relapsing-remitting or secondary progressive MS with superimposed relapses) were randomized to 2, 4 or 8 mg temsirolimus orally every day (which selectively blocks interleukin-2 driven T-cell proliferation) or placebo for 9 months. Patients were not included in the trial if they had received systemic corticosteroid therapy within one month prior to randomization. In the present study only those patients randomized to either 8 mg temsirolimus (76 patients) or placebo (69 patients) were initially selected. Secondly, only patients with complete MRI data over all nine months were selected, which resulted in a final selection of 120 patients (61 patients on treatment and 59 on placebo). The clinical trial protocol was approved by local ethics review boards of all participating centers, and all subjects gave written informed consent.

Patients were monitored for nine months using monthly MRI scans, acquired according to published guidelines describing the use of MRI in clinical trials [15]. The MRI protocol included an oblique-axial, dual-echo, PD-weighted and T2-weighted, spin-echo or turbo/fast spin-echo (TR/TE of 2000-3000 / 15-40 & 60-100 ms) and a T1-weighted, spin-echo (TR/TE of 400-700 / 5-25 ms), before and after administration of 0.1 mmol/kg gadolinium diethylene triamine pentaacetic acid intravenously. Field of view for all examinations was 25 cm, combined with a 256 x 256 matrix, resulting in a roughly 1 x 1 mm pixel size in-plane. Images were acquired in two interleaved sets with a 3mm gap, resulting in whole brain coverage with contiguous 3mm thick slices.

*Post-processing/registration*

Subtraction images were generated by applying a series of post-processing steps. First, for both timepoints, a brain mask was created using the brain extraction tool (BET) [16], part of FMRIB's software library (FSL) [17]. All subsequent post-processing steps were calculated within the brain masks and subsequently applied to the whole scan, starting with an intra-scan intensity normalization (bias field correction) using the non-parametric method N3 [18], followed by a global scaling step to equate the average intensities of consecutive scans. Finally, images were spatially normalized (rigid body registration) to a halfway position using FMRIB's Linear Image Registration Tool (FLIRT) [19], with mutual information as cost function and sinc interpolation.

*Image Analysis*

The non-registered Gd-enhanced T1w images, had been centrally evaluated by the Image Analysis Centre (IAC), Amsterdam, The Netherlands, by expert reviewers blinded to treatment allocation. The number of new Gd-enhancing T1w lesions was assessed according to published guidelines [20] and available from the trial database. All subtraction images were evaluated by an expert reviewer (BM), also blinded to treatment allocation, using the freely available medical image processing, analysis, and visualization (MIPAV) software package developed at the National Institutes of Health, Bethesda, MD (http://mipav.cit.nih.gov/) [21]. The subtraction images comprising both the short and long echo T2w images were analyzed as one set. All subtraction images were analyzed alongside the halfway registered baseline and follow-up images, to ensure that changes identified on the subtraction images genuinely represented disease activity (rather than artifact). Lesions on subtraction images were identified according to published criteria which are described in more detail in [11,12]. Briefly, a lesion had to be a clearly visible, non-artifactual, area of change (bright or dark) against the grey background. In the subtraction image, new and enlarged lesions (positive activity) appear as bright areas, whereas resolved or shrunken lesions (negative activity) appear as dark areas. In the present study only new and enlarged lesions (positive activity) were identified, since these are used as outcome

measure in MR monitored MS treatment trials and comparable to new activity detected by Gd-enhanced T1w images.

*Statistical Analysis*

The correlation between the number of T2w subtraction lesions (month nine T2w scan minus the month one T2w scan) and the cumulative number of new Gd-enhancing T1w lesions (month two T1w scan through month nine T1w scan) was explored using Spearman's correlation coefficient. Lesion counts were expressed as means with accompanying standard deviations (SDs) for each imaging scheme, and compared using the Wilcoxon signed ranks test. By means of linear regression analysis, the predictive value of 1) the baseline number of Gd-enhancing T1w lesions, and 2) T2 lesion load (T2LL), for the development of on-study numbers of Gd-enhancing T1w lesions and T2w subtraction lesions, was determined. Statistically significant predictors were subsequently entered as covariates in the treatment efficacy analysis described below.

Treatment effects (reduction of active lesions) of oral temsirolimus as determined by the two imaging schemes was evaluated threefold; first, by means of the nonparametric Mann-Whitney U test; secondly, parametrically using a general linear model (GLM) based on the negative binomial (NB) distribution (NB-regression) *without* the inclusion of covariates; finally, again by means of NB-regression *with* the inclusion of covariates.

Subsequently, power calculations were conducted to estimate the required number of patients to determine a significant treatment effect. In addition, the number of MRI scans was calculated by multiplying the number of patients by two for T2w subtraction imaging, and eight for Gd-enhanced T1w imaging. For the sample size estimates of T2w subtraction imaging the optimal fitting distribution for describing the number of T2w subtraction lesions was selected from a selection of conceivable statistical count distributions (Poisson, NB) by comparing their maximized log-likelihood values as measure of their goodness of fit of the data. Then, the statistical model was implemented using a parametric resampling and simulation methodology

[22] which, for a given power, calculates the required sample sizes for a range of anticipated treatment effects (MATLAB version 2007a). For Gd-enhanced T1w imaging the same resampling and simulation procedure was applied, assuming the data to be NB distributed [23,24]. All statistical analyses were performed with SPSS version 15.0 (SPSS, Chicago, IL) and Stata version 10 (StataCorp, College Station, TX).

## Results

*Lesion counts and correlations*

In total 4(3.3%) of the 120 subtraction datasets were excluded from the visual analysis due to technical limitations; twice the registration procedure failed because the month one and nine T2w images were acquired with a different matrix size, and twice because extensive motion artifacts severely degraded the quality of the subtraction images. As a result, datasets from 116 patients were evaluated with both imaging schemes, of who the demographic descriptors are summarized in Table 1. The numbers of active (new and enlarged) T2w subtraction lesions and cumulative number of new Gd-enhancing T1w lesions detected were significantly correlated (rho = 0.80, P < .001), albeit that the difference between both measures increased in patients with increasing lesion numbers (Figure 1). Overall, the mean cumulative number of new Gd-enhancing T1w lesions was significantly higher (1.7-fold) than the mean number of T2w subtraction lesions (7.28 vs. 4.41, P < .001, Figures 2-4).

The baseline T2LL significantly predicted the development of on-study T2w subtraction lesions (ß = 2.464, P = .001), and stratification by the median baseline T2LL showed a significantly (2.8-fold) higher mean number of lesions developing in the "above the median" group compared with the "below the median" group (6.5 vs. 2.3, P < .001). Likewise, the baseline number of Gd-enhanced T1w lesions significantly predicted the development of on-study Gd-enhanced T1w lesions (ß = 2.189, P < .001), and stratification by the median baseline number of Gd-enhancing T1w lesions also showed a significantly (3.6-fold) higher mean number of lesions developing in the "above the median" group compared with the "below the median" group (11.9 vs. 3.3, P < .001).

*Detection of treatment effect*

The mean cumulative number of new Gd-enhancing T1w lesions identified in the treatment (temsirolimus) group was (38%) lower compared with the placebo group (5.5 vs. 9.1

respectively), but this difference was not significant when evaluated by using the Mann-Whitney U test (P = .177), nor by using NB-regression *without* covariates (P = .092, Table 2). The reduction in the cumulative number of new Gd-enhancing T1w lesions was deemed statistically significant only when either the baseline number of Gd-enhancing T1w lesions (P = .005), or baseline T2LL (P = .011), or both (P = .002), were entered as covariates in the NB-regression model.

The mean number of T2w subtraction lesions identified in the treatment group was (49%) lower compared with the placebo group (3.0 vs. 5.9, respectively). In contrast to the case using the number of Gd-enhancing lesions as outcome, this difference was significant with all the statistical methods used. In particular, the significance level was P = .017 when evaluated by using the Mann-Whitney U test, and P = .011 when using NB-regression *without* covariates (see Table 2 for more details).

*Sample size calculations*

Depending on the magnitude of the simulated treatment effect, the number of patients required per treatment arm to determine a significant treatment effect in a placebo controlled, parallel grouped clinical trial design, were 22 to 33 percent lower for T2w subtraction imaging, compared with Gd enhanced T1w imaging (Table 3, Figure 5). This translated into reductions of 81% to 83% in the number of MRI scans needed to detect a significant treatment effect (Table 3). The point estimates are based on simulations with the following NB parameters derived from the data (T2w subtraction imaging: mean ($\hat{\mu}$) = 5.9, dispersion parameter ($\hat{\theta}$) = 0.58; Gd enhanced T1w imaging: mean ($\hat{\mu}$)= 9.1 and dispersion parameter ($\hat{\theta}$) = 0.41).

**Discussion**

A previous study has shown that T2w subtraction imaging detected higher numbers of active lesions with greater interobserver agreement and demonstrated increased power to evaluate treatment efficacy, compared with a standard analysis of non-registered (raw) T2w images [12]. In this study we found that despite an expected loss in sensitivity, long-interval T2w subtraction MR imaging also exhibited increased power to assess treatment efficacy, compared with monthly serial Gd-enhanced T1w imaging. In particular, the subtraction scheme used in our study detected around half of the lesions compared with Gd-enhanced T1w imaging. Several factors contributed to this finding. First, very small (1-2mm) lesions that were identified with Gd-enhanced T1w imaging were readily missed on subtraction images. The T2w images used in this study to create the subtraction images featured an anisotropic voxel size inducing subtle misregistration artifacts and blurring in the subtraction images, which particularly hampered the detection of (very) small lesions (Figure 2). This could be ameliorated by the application of 3D sequences with intrinsically higher signal-to-noise-ratios compared with 2D sequences [25-27], allowing the acquisition of smaller isotropic voxels, leading to a potentially more accurate registration result. Secondly, flow artifacts, arising from the carotid and vertebrobasilar arteries and transverse sinuses, hampered the detection of lesions in the posterior fossa and temporal lobes on subtraction images (Figure 3). This could also be improved by the application of single-slab 3D sequences which employ non-spatially selective radio-frequency pulses [28,29], resulting in the absence of blood and cerebro-spinal fluid (CSF) flow artifacts, thereby improving the visualization of infratentorial lesions [27]. These two issues might be additional incentives to consider the implementation of single-slab 3D sequences in multicenter clinical trial design. Thirdly, some lesions only appeared as very subtle T2 hyperintensities on the month 9 T2w scan, and hence on the subtraction image, whereas their initial area of enhancement was quite extensive (Figure 4). This was most prominent for lesions identified with Gd-enhanced T1w imaging during the first

months of the study, which had the largest time-window between the initial enhancement and the month nine T2w scan.

Conversely, some lesions were detected with T2w subtraction imaging that were not detected with Gd-enhanced T1w imaging (above the dashed reference line from equation in the zoom-in of Figure 1). Gd-enhanced T1w imaging was undertaken monthly in the oral temsirolimus trial. However, some new lesions, i.e. those new lesions that appear just after a monthly scan and enhance for only a couple of weeks, will be missed. Cotton et al reported a mean duration of enhancement of new lesions of 3.07 weeks [30], which makes it likely that some new lesions will not be captured with monthly Gd-enhanced imaging. Indeed, early literature already reported that not all new or enlarging T2 lesions demonstrate enhancement on monthly scanning [5,31,32]. A future study could conduct a direct side-by-side comparison of both imaging schemes per lesion, to determine the exact properties of Gd-T1w lesions (e.g. size, location, type of enhancement) that were not seen with T2w subtraction imaging, and vice versa the properties of T2w subtraction lesions that we not seen with Gd-T1w imaging.

Using the nonparametric Mann-Whitney U test we found a significant reduction (49%, P = .017) of the mean number of T2w subtraction lesions in the treatment (temsirolimus) group compared with the placebo group, whereas we found a non-significant reduction (40%, P = .177) of the mean cumulative number of new Gd-enhancing T1w lesions. Alternatively, we used the negative binomial (NB) distribution, which has been described to adequately model lesion count data over time [23,24], allowing a statistically more powerful parametric approach (e.g. general linear regression) to evaluate treatment efficacy (i.e. the reduction in active lesions) in MS patients. Indeed, the aforementioned reductions in active lesions showed lower P values when evaluated by NB-regression compared with the non-parametric Mann-Whitney U test, which was more evident for Gd-enhanced T1w imaging (P = .092 vs. P = .177 respectively) than for T2w subtraction imaging (P = .011 vs. P = .017 respectively). Another advantage of using NB-regression is the

possibility to adjust for (baseline) covariates, which can reduce the interpatient variance, and thereby increase the statistical power of MRI-monitored trials [22,33]. We also found this in our study, especially for Gd-enhanced T1w imaging, as NB-regression adjusted for the number of baseline Gd-enhancing T1w lesions and T2LL, resulted in a lower P value (P = .002), compared with unadjusted NB-regression (P = .092). Short of stratification for the aforementioned baseline determinants, statistical analysis of MRI-monitored phase II clinical trials should employ baseline covariate adjustment to increase statistical power.

Our study had several potential limitations. Firstly, the T2w MR images were acquired after the administration of gadolinium, resulting in a slightly increased signal intensity of Gd-enhancing lesions on T2w images [34]. Consequently, lesions that had a *higher* signal intensity on the month one scan than on the month nine scan (i.e. all lesions that showed Gd-enhancement on the month one scan), appeared as negative rather than positive activity and were not identified on the subtraction images. To make a rational comparison with Gd-enhanced T1w imaging, the cumulative number of new Gd-enhancing T1w lesions was calculated from the month 2 to the month 9 scans. Significance of all outcomes was maintained when including data from the month one Gd-enhanced T1w scans (data not shown). Secondly, 4 patients were excluded from our study due to technical limitations of the subtraction scheme. However, 25 patients (15 treatment, 10 placebo) included in the original trial, were excluded from our study because Gd data was not available for all nine months of follow-up, whereas the month one and month nine T2w scans were available. Thirdly, the analysis of T2w subtraction images was undertaken by one reviewer in a research setting, whereas the Gd-enhanced T1w images were reviewed by multiple reviewers, as part of the clinical trial, in accordance with published guidelines. A previous study has reported a very high ICC (expressing interobserver agreement) of 0.98 for the detection of positive disease activity on T2w subtraction images [12]. As such, the use of a single reader in this study has probably led only to a marginal overestimation of the value of T2w subtraction imaging. Also, the reductions in the number of MRI scans needed to detected a significant

treatment effect comparing T2w subtraction with serial monthly Gd-T1w imaging might be slightly lower for phase II trials with a shorter duration (e.g. 6 months). Finally, T2w subtraction imaging could have a lower specificity compared with Gd-enhanced T1w imaging for the detection of MS activity. Age-related white matter changes, appearing as T2 hyperintensities without enhancement, also termed leukoaraiosis, are frequently found on brain MRI in elderly populations and thought to be the resultant of small vessel disease [35,36]. These lesions could be mistakenly identified as new MS lesions on T2w subtraction imaging. However, the mean age of patients in this study was around 38 years which makes it highly unlikely that the T2w subtraction lesion identified were vascular in origin.

In conclusion, despite an expected loss in sensitivity, we found long-interval T2w subtraction MR imaging to exhibit increased power to assess treatment efficacy, compared with serial monthly Gd-enhanced T1w imaging, in an MS, phase II, clinical trial setting. T2w subtraction MR imaging could greatly increase the cost-effectiveness of phase 2 MS trials, by limiting the number of patients, contrast injections, and MRI scans needed, and decrease patients' risk by obviating gadolinium administration.

**Acknowledgment**

We are grateful to the Image Analysis Centre (IAC), Amsterdam, The Netherlands and Dr. Ingrid C. Sluimer and Sjoerd B. Vos for help with image post-processing. The MS Center Amsterdam and BM, JJGG, and HV were financially supported by the Dutch MS Research Foundation (grant no. 02-358b/c). This work was also supported in part by funding provided by the US National Multiple Sclerosis Society (RG 3574-A-1), as well as the National Institutes of Health (R01 NS35142). We are also grateful to Wyeth Pharmaceuticals, Clinical R&D, Collegeville, USA and Paris, France, the sponsor of the oral temsirolimus trial, and all investigators and MS patients involved.

**References**

1. Willoughby EW, Grochowski E, Li DK, Oger J, Kastrukoff LF, Paty DW. Serial magnetic resonance scanning in multiple sclerosis: a second prospective study in relapsing patients. Ann Neurol 1989;25:43-49.

2. Thompson AJ, Kermode AG, Wicks D, et al. Major differences in the dynamics of primary and secondary progressive multiple sclerosis. Ann Neurol 1991;29:53-62.

3. Kermode AG, Thompson AJ, Tofts P, et al. Breakdown of the blood-brain barrier precedes symptoms and other MRI signs of new lesions in multiple sclerosis. Pathogenetic and clinical implications. Brain 1990;113:1477-1489.

4. Nesbit GM, Forbes GS, Scheithauer BW, Okazaki H, Rodriguez M. Multiple sclerosis: histopathologic and MR and/or CT correlation in 37 cases at biopsy and three cases at autopsy. Radiology 1991;180:467-474.

5. Harris JO, Frank JA, Patronas N, McFarlin DE, McFarland HF. Serial gadolinium-enhanced magnetic resonance imaging scans in patients with early, relapsing-remitting multiple sclerosis: implications for clinical trials and natural history. Ann Neurol 1991;29:548-555.

6. Guttmann CR, Ahn SS, Hsu L, Kikinis R, Jolesz FA. The evolution of multiple sclerosis lesions on serial MR. AJNR Am J Neuroradiol 1995;16:1481-1491.

7. Ledneva E, Karie S, Launay-Vacher V, Janus N, Deray G. Renal safety of gadolinium-based contrast media in patients with chronic renal insufficiency. Radiology 2009;250:618-628.

8. Wertman R, Altun E, Martin DR, et al. Risk of nephrogenic systemic fibrosis: evaluation of gadolinium chelate contrast agents at four American universities. Radiology 2008;248:799-806.

9. Molyneux PD, Filippi M, Barkhof F, et al. Correlations between monthly enhanced MRI lesion rate and changes in T2 lesion volume in multiple sclerosis. Ann Neurol 1998;43:332-339.

10. Goodkin DE, Vanderburg-Medendorp S, Ross J. The effect of repositioning error on serial magnetic resonance imaging scans. Arch Neurol 1993;50:569-571.

11. Tan IL, van Schijndel RA, Fazekas F, et al. Image registration and subtraction to detect active T(2) lesions in MS: an interobserver study. J Neurol 2002;249: 767-773.

12. Moraal B, Meier DS, Poppe PA, et al. Subtraction MR images in a multiple sclerosis multicenter clinical trial setting. Radiology 2009;250:506-514.

13. Duan Y, Hildenbrand PG, Sampat MP, et al. Segmentation of subtraction images for the measurement of lesion change in multiple sclerosis. AJNR Am J Neuroradiol 2008;29:340-346.

14. Kappos L, Barkhof F, Desmet A. The effect of oral Temsirolimus on new magnetic resonance imaging scan lesions, brain atrophy, and the number of relapses in multiple sclerosis: results from a randomised, controlled trial. J Neurol 2005;252:supplement 2,46.

15. Miller DH, Albert PS, Barkhof F, et al. Guidelines for the use of magnetic resonance techniques in monitoring the treatment of multiple sclerosis. US National MS Society Task Force. Ann Neurol 1996;39:6-16.

16. Smith SM. Fast robust automated brain extraction. Hum Brain Mapp 2002;17:143-155.

17. Smith SM, Jenkinson M, Woolrich MW, et al. Advances in functional and structural MR image analysis and implementation as FSL. NeuroImage 2004;23:S208-219.

18. Sled JG, Zijdenbos AP, Evans AC. A nonparametric method for automatic correction of intensity nonuniformity in MRI data. IEEE Trans Med Imaging 1998;17:87-97.

19. Jenkinson M, Smith S. A global optimisation method for robust affine registration of brain images. Medical Image Anal 2001;5:143-156.

20. Barkhof F, Filippi M, van Waesberghe JH, et al. Improving interobserver variation in reporting gadolinium-enhanced MRI lesions in multiple sclerosis. Neurology 1997;49:1682-1688.

21. McAuliffe M, Lalonde E, McGarry D, Gandler W, Csaky K, Trus B. Medical image processing, analysis and visualization in clinical research. In: Proceedings of the 14th IEEE symposium on computer-based medical systems, 2001;381–386.

22. Sormani MP, Miller DH, Comi G, et al. Clinical trials of multiple sclerosis monitored with enhanced MRI: new sample size calculations based on large data sets. J Neurol Neurosurg Psychiatry 2001;70:494-499.

23. Sormani MP, Bruzzi P, Miller DH, Gasperini C, Barkhof F, Filippi M. Modelling MRI enhancing lesion counts in multiple sclerosis using a negative binomial model: implications for clinical trials. J Neurol Sci 1999;163:74-80.

24. van den Elskamp I, Knol D, Uitdehaag B, Barkhof F. The distribution of new enhancing lesion counts in multiple sclerosis: further explorations. Mult Scler 2009;15:42-49.

25. Tubridy N, Barker GJ, Macmanus DG, Moseley IF, Miller DH. Three-dimensional fast fluid attenuated inversion recovery (3D fast FLAIR): a new MRI sequence which increases the detectable lesion load in multiple sclerosis. Br J Radiol 1998;71:840-845.

26. Tan IL, Pouwels PJ, van Schijndel RA, Adèr HJ, Manoliu RA, Barkhof F. Isotropic 3D fast FLAIR imaging of the brain in multiple sclerosis: initial experience. Eur Radiol 2002;12:559-567.

27. Moraal B, Roosendaal SD, Pouwels PJ, et al. Multi-contrast, isotropic, single-slab 3D MR imaging in multiple sclerosis. Eur Radiol 2008;18:2311-2320.

28. Mugler JP 3rd, Bao S, Mulkern RV, et al. Optimized single-slab three-dimensional spin-echo MR imaging of the brain. Radiology 2000;216:891-899.

29. Kallmes DF, Hui FK, Mugler JP 3rd. Suppression of cerebrospinal fluid and blood flow artifacts in FLAIR MR imaging with a single-slab three-dimensional pulse sequence: initial experience. Radiology 2001;221:251-255.

30. Cotton F, Weiner HL, Jolesz FA, Guttmann CR. MRI contrast uptake in new lesions in relapsing-remitting MS followed at weekly intervals. Neurology 2003;60:640-646.

31. Barkhof F, Scheltens P, Frequin ST, et al. Relapsing-remitting multiple sclerosis: sequential enhanced MR imaging vs clinical findings in determining disease activity. AJR Am J Roentgenol 1992;159:1041-1047.

32. Miller DH, Barkhof F, Nauta JJ. Gadolinium enhancement increases the sensitivity of MRI in detecting disease activity in multiple sclerosis. Brain 1993;116:1077-1094.

33. Barkhof F, Held U, Simon JH, et al. Predicting gadolinium enhancement status in MS patients eligible for randomized clinical trials. Neurology 2005;65:1447-1454.

34. Barkhof F, Valk J, Hommes OR, Scheltens P, Nauta JJ. Gadopentetate dimeglumine enhancement of multiple sclerosis lesions on long TR spin-echo images at 0.6 T. AJNR Am J Neuroradiol 1992; 13:1257-1259.

35. Basile AM, Pantoni L, Pracucci G, et al. Age, hypertension, and lacunar stroke are the major determinants of the severity of age-related white matter changes. The LADIS (Leukoaraiosis and Disability in the Elderly) Study. Cerebrovasc Dis 2006;21:315-322.

36. Vernooij MW, Ikram MA, Tanghe HL, et al. Incidental findings on brain MRI in the general population. N Engl J Med 2007;357:1821-1828.

**Figure Legends**

Figure 1.

Scatterplot visualizing the correlation (Spearman's rho 0.80, P < .001) between the cumulative number of gadolinium-enhancing T1w lesions identified (horizontal axis) versus the number of T2w subtraction lesions identified (vertical) axis for all 116 patients.

Figure 2.

Examples of Gd-enhancing T1w lesions visible as positive activity on the T2w subtraction image. MS activity on A, halfway registered month one T2w MR image; B, halfway-registered month nine T2w image; C, T2w subtraction image; D, E, F, month three, month five, and month nine non-registered Gd-enhanced T1w images respectively. Arrowheads = Gd-enhancing T1w lesions on various timepoints, also visible on the follow-up registered T2w image, and easily identified on the T2w subtraction image. Arrows = very small lesion not readily identified on the T2w subtraction image which did correspond to a Gd-enhancing T1w lesion.

Figure 3.

Example of Gd-enhancing T1w lesion in the temporal lobe which was not identified on the T2w subtraction image. A, halfway registered month one T2w MR image; B, halfway-registered month nine T2w image; C, Gd-enhanced T1w image; D, T2w subtraction image. Arrowheads = Gd-enhancing T1w lesion on month two, not identified on the T2w subtraction image, due to flow artifacts (arrows) of the basilar artery.

Figure 4.

Example of Gd-enhancing T1w lesion which was not identified on the T2w subtraction image. A, halfway registered month one T2w MR image; B, halfway-registered month nine T2w image; C, Gd-enhanced T1w image; D, T2w subtraction image. Arrowheads = Gd-enhancing T1w lesion on

month two, seen as a very subtle lesion on the month nine registered T2w image, not identified on the T2w subtraction image.

Figure 5.

Relationship between the number of patients needed in each treatment arm and various simulated treatment effects for a power of 80%, for both Gd-enhanced T1w imaging (dotted line) and T2w subtraction imaging (solid line).

**Table 1**
**Baseline descriptives**

| Characteristic | Placebo (n = 58) | Treatment (n = 58) |
|---|---|---|
| **General** | | |
| No. Female (%) | 43 (74) | 38 (66) |
| No. RRMS* (%) | 50 (86) | 53 (91) |
| Mean age (SD) | 38.3 (9.4) | 38.6 (8.5) |
| Mean disease duration (SD) | 4.7 (5.1) | 5.1 (5.3) |
| Median baseline EDSS (IQR) | 2.5 (1.5-3.5) | 2.0 (1.5-3.5) |
| **MRI** | | |
| No. ≥ 1 T1w Gd lesion (%) | 19 (33) | 35 (60) |
| Median T2 lesion volume in $cm^3$ (IQR) | 4.8 (2.1-8.6) | 4.6 (2.5-10.2) |

IQR = Interquartile range, SD = standard deviation

* remaining patients have SPMS with relapses

**Table 2**
**Analysis of treatment efficacy comparing patients on either temsirolimus or placebo, per analysis method**

| | Analysis method | |
|---|---|---|
| | **T1w Gd** | **T2w Sub** |
| Lesion Counts (Mean ± SD) | | |
| Temsirolimus (n=58) | 5.5 ± 9.1 | 3.0 ± 4.6 |
| Placebo (n=58) | 9.1 ± 17.2 | 5.9 ± 8.8 |
| | **Analysis of treatment efficacy (P value)** | |
| Mann-Whitney U test | .177 | **.017** |
| NB-regression, adjusted for: | | |
| no adjustment | .092 | **.011** |
| no. of baseline T1w Gd lesions | **.005** | **.002** |
| baseline T2 lesion load | **.011** | **< .001** |
| both | **.002** | **< .001** |

NB = negative binomial

**Table 3**

**Number of patients/MRI scans per treatment arm, necessary to perform parallel group designed trials with a statistical powers of 80%, to significantly detect a mean percentage decrease from 50% to 80% in mean number of new lesions.**

| Treatment effect | T1w Gd | | T2w Sub | | % Reduction | |
|---|---|---|---|---|---|---|
| | Patients (n) | MRI (n) | Patients (n) | MRI (n) | Patients (n) | MRI (n) |
| 50% | 150 | 1200 | 101 | 202 | 33 | 83 |
| 60% | 88 | 704 | 60 | 120 | 32 | 83 |
| 70% | 51 | 408 | 39 | 78 | 24 | 81 |
| 80% | 30 | 240 | 22 | 44 | 27 | 82 |

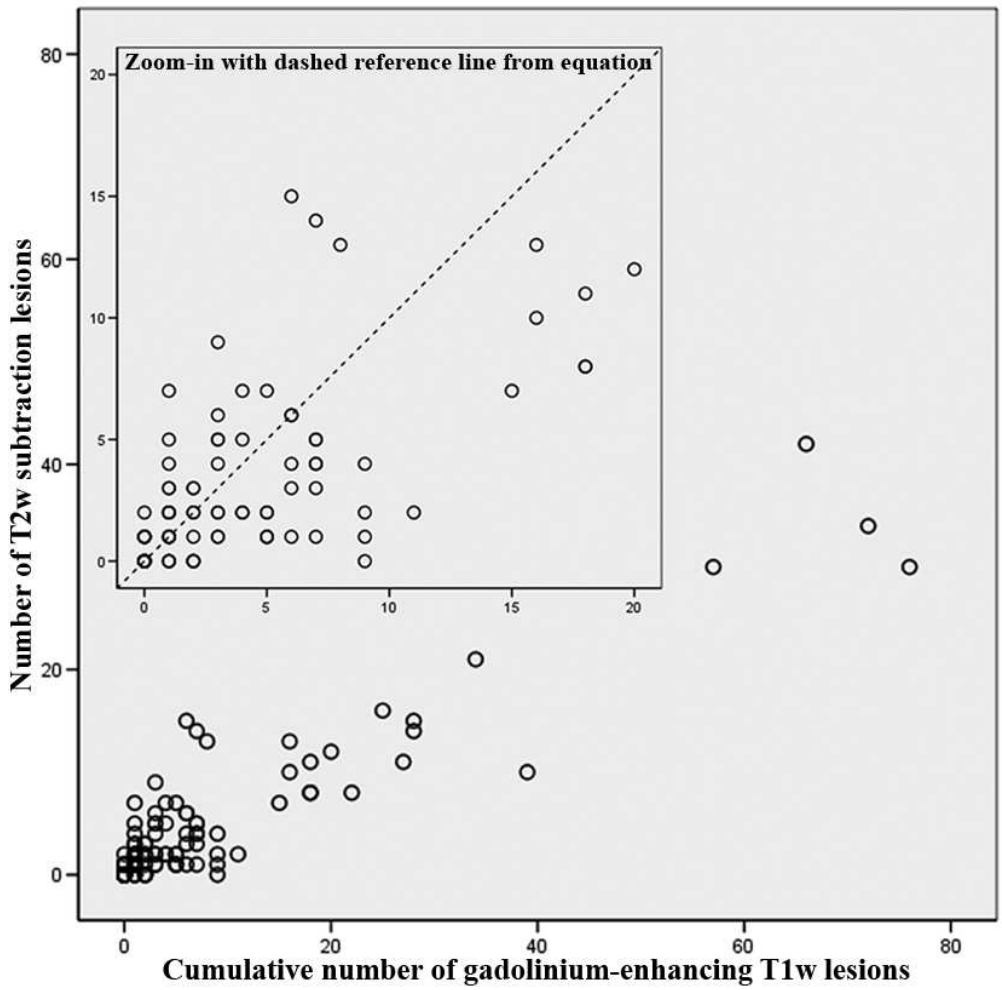


**Figure 1.** Scatterplot visualizing the correlation (Spearman's rho 0.80, P < .001) between the cumulative number of gadolinium-enhancing T1w lesions identified (horizontal axis) versus the number of T2w subtraction lesions identified (vertical) axis for all 116 patients.
70x69mm (300 x 300 DPI)

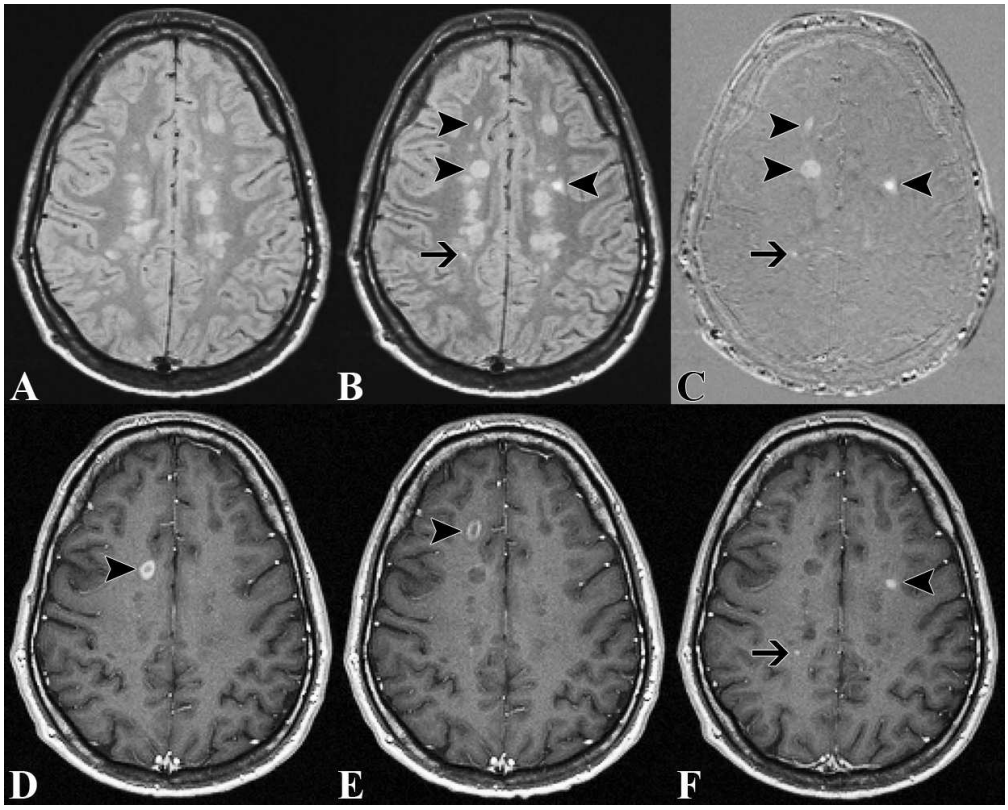


**Figure 2.** Examples of Gd-enhancing T1w lesions visible as positive activity on the T2w subtraction image. MS activity on A, halfway registered month one T2w MR image; B, halfway-registered month nine T2w image; C, T2w subtraction image; D, E, F, month three, month five, and month nine non-registered Gd-enhanced T1w images respectively. Arrowheads = Gd-enhancing T1w lesions on various timepoints, also visible on the follow-up registered T2w image, and easily identified on the T2w subtraction image. Arrows = very small lesion not readily identified on the T2w subtraction image which did correspond to a Gd-enhancing T1w lesion.
104x84mm (300 x 300 DPI)

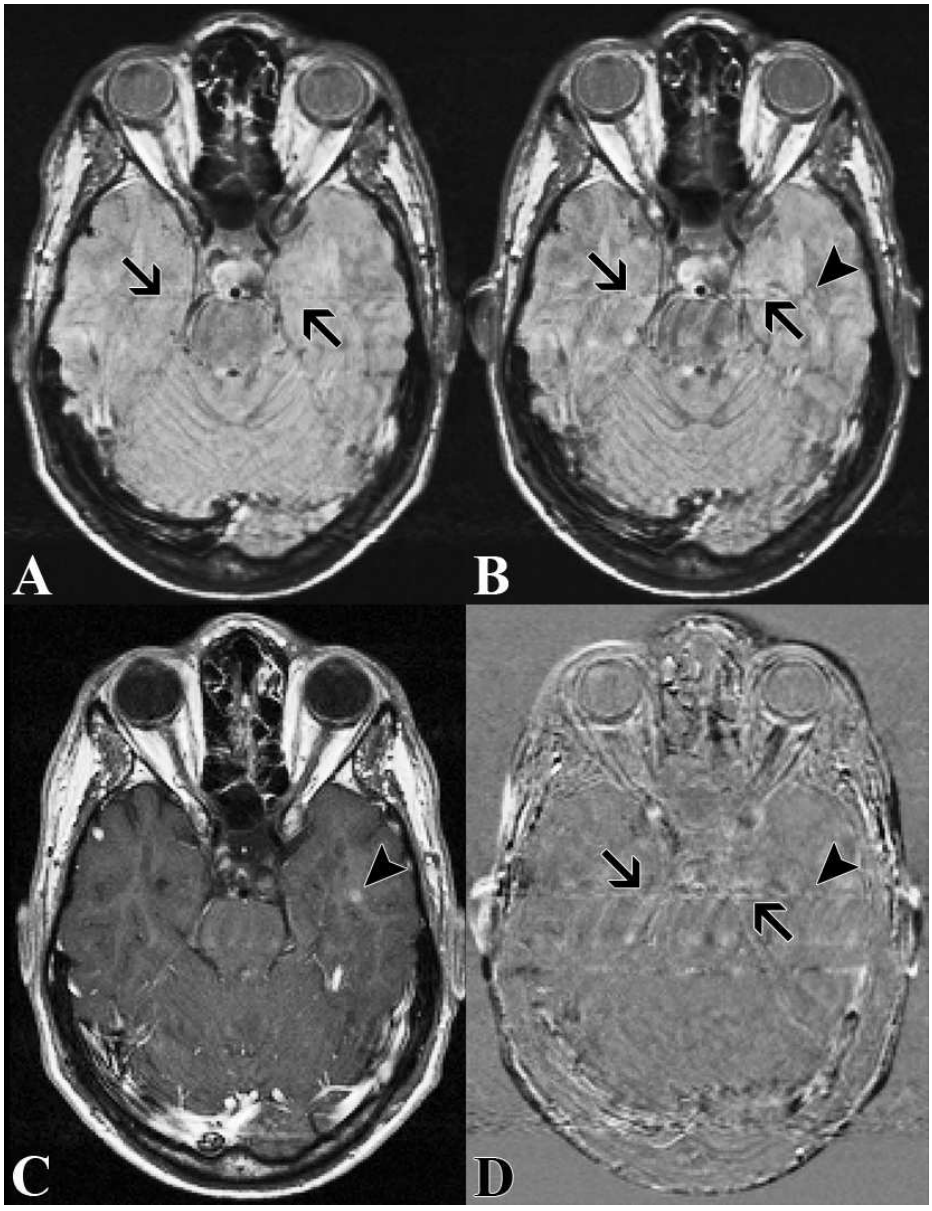


**Figure 3.** Example of Gd-enhancing T1w lesion in the temporal lobe which was not identified on the T2w subtraction image. A, halfway registered month one T2w MR image; B, halfway-registered month nine T2w image; C, Gd-enhanced T1w image; D, T2w subtraction image. Arrowheads = Gd-enhancing T1w lesion on month two, not identified on the T2w subtraction image, due to flow artifacts (arrows) of the basilar artery.
70x90mm (300 x 300 DPI)

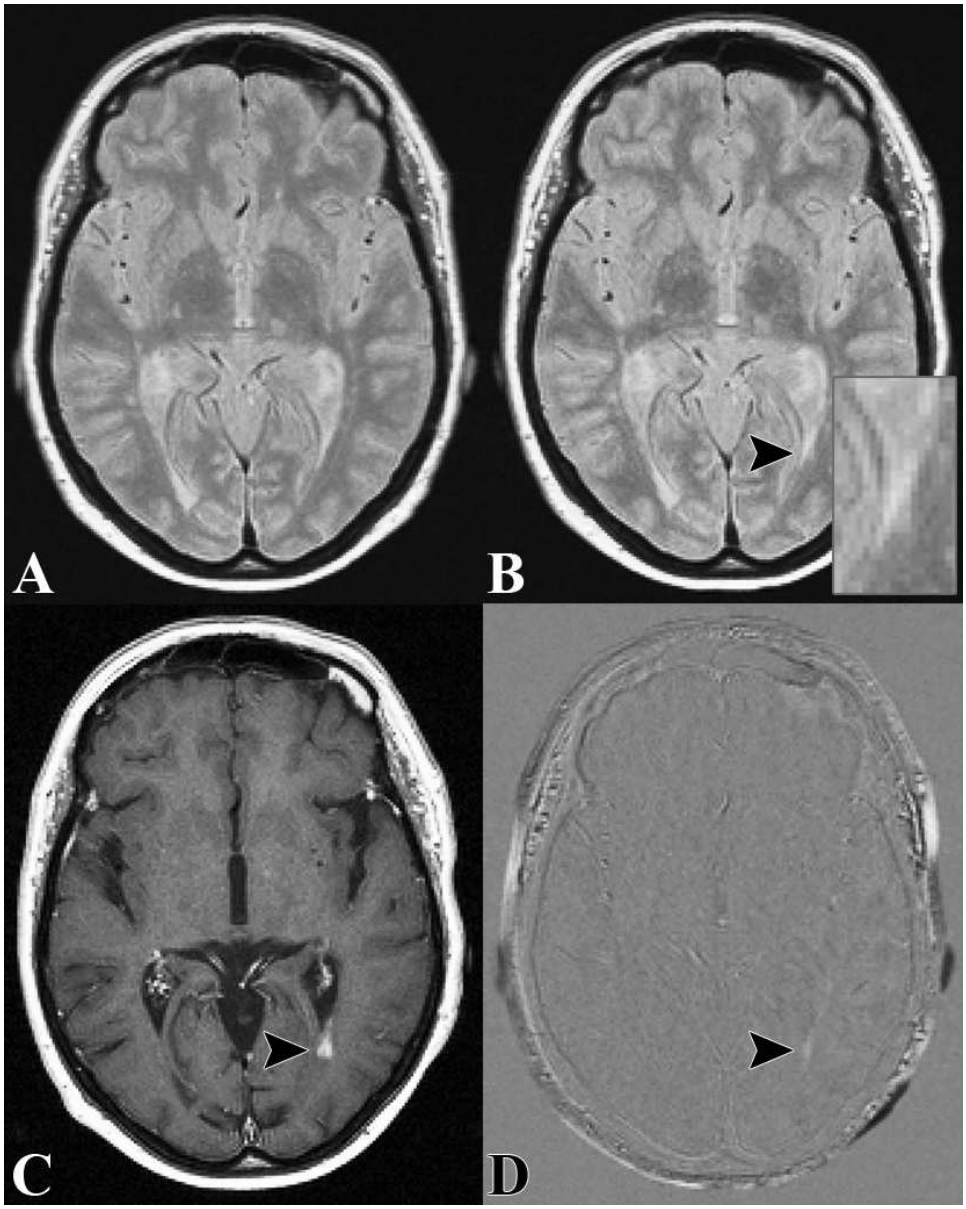


**Figure 4.** Example of Gd-enhancing T1w lesion which was not identified on the T2w subtraction image. A, halfway registered month one T2w MR image; B, halfway-registered month nine T2w image; C, Gd-enhanced T1w image; D, T2w subtraction image. Arrowheads = Gd-enhancing T1w lesion on month two, seen as a very subtle lesion on the month nine registered T2w image, not identified on the T2w subtraction image.
70x87mm (300 x 300 DPI)

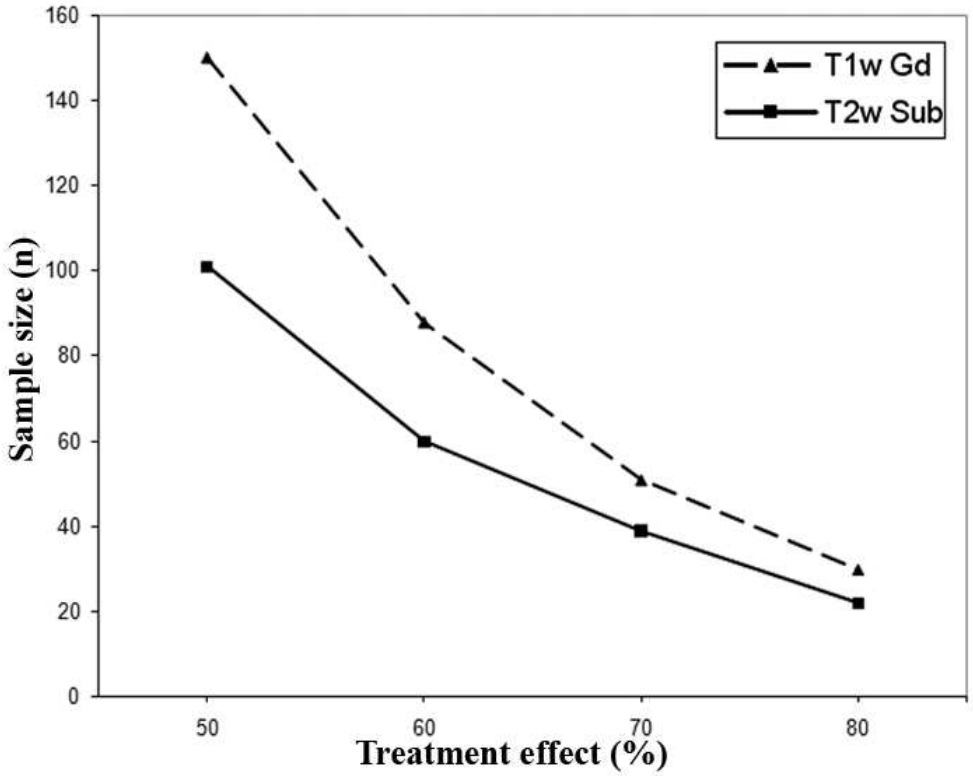


**Figure 5.** Relationship between the number of patients needed in each treatment arm and various simulated treatment effects for a power of 80%, for both Gd-enhanced T1w imaging (dotted line) and T2w subtraction imaging (solid line).
60x48mm (300 x 300 DPI)