

REPLY:

We thank Drs Lancelot, Piednoir, and Desché for their interest in our work and their comments regarding our recent article.¹ However, we disagree strongly with the critique presented in their letter. The necessity to comply with word count limitations when preparing manuscripts for publication means that the description of the statistical methodology is often too brief. We now address the points raised.

Statistical Analysis and Sample Size

The study design and analysis described in our article¹ are similar to the methodology used in many previous intraindividual comparative studies.²⁻⁹ The primary study end point was the overall diagnostic preference of each of 3 readers for one gadolinium-based contrast agent (GBCA) over the other. Other qualitative end points (determinations of lesion border delineation, definition of disease extent, visualization of lesion internal morphology, and lesion contrast enhancement) are accepted clinically relevant parameters that can directly impact patient management decisions and surgical planning, particularly for patients with glial tumors in whom macroscopically complete surgical removal is associated with improved prognosis and longer patient survival, and those with metastases, for whom determination of the precise number, size, and location of lesions can aid selection of the most appropriate treatment option.^{3,6-9} Image assessment was performed by comparing images from the 2 MR imaging examinations side-by-side, with the readers blinded to the contrast agent used and all clinical information. Each reader expressed preference for examination 1 or 2 or determined that the 2 examinations were equal. The resulting data for each reader were 1 observation per patient (ie, 1 paired sample datum with an ordinal scale). Nonparametric analysis with a 2-sided Wilcoxon signed rank test is the appropriate statistical analysis method for assessment of overall diagnostic preference. The distribution of the preference between the 2 examinations was also tested by using a 1-sample χ^2 test for equal proportions. The results obtained were similar to those from the Wilcoxon signed rank test. This analysis was not included in the article due to the word restrictions.

The results for Study Arm 1 revealed a highly significant ($P < .0001$) preference for gadobenate over gadoterate at an equivalent dose for each reader (Fig 1). Variation of measurements between readers is expected in a fully blinded read setting. Figure 1 shows that all 3 readers were in high agreement and consistent, especially concerning the very few assessments in which gadoterate was preferred over gadobenate (1.6%–3.2% of patients across 3 readers). The major reason for the 50.8% agreement among readers was the differential percentage of preferences for gadobenate across the 3 readers (49.2%, 82.3%, 69.4%). As already discussed in our article,¹ a κ value of 0.273 was due to the skewed distribution of preferences (very few preferences for gadoterate). Feinstein et al¹⁰ demonstrated clearly that a low κ can result from a substantial imbalance in marginal totals.

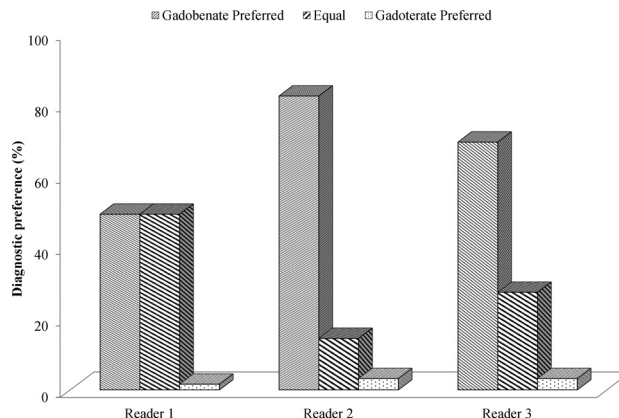


FIG 1. Distribution of overall diagnostic preference: gadobenate 0.1 mmol/kg versus gadoterate 0.1 mmol/kg.

The results for Study Arm 2 revealed no statistically significant differences between gadoterate at 0.1 mmol/kg and gadobenate at 0.05 mmol/kg of body weight.¹ In the abstract, we concluded, “No meaningful differences were recorded between 0.05 mmol/kg gadobenate and 0.1 mmol/kg gadoterate.” We understand from a statistical point of view that equivalence cannot be claimed if the test hypothesis is not prospectively defined as “noninferiority.” However, a conclusion of “no statistically significant difference” between treatments simply means that the evidence that the 2 treatments lead to different outcomes is not strong enough. As can readily be seen in Fig 2, all 3 readers determined that the images from most patients were diagnostically equal (ie, no diagnostic preference between the images with half-dose gadobenate and full-dose gadoterate).

The sample size calculation was based on the χ^2 test of specified proportions in 3 categories for paired 1-sample responses. Assumptions for the study were as follows: for Study Arm 1, an “equal” response for 50% of the patients and a ratio of superiority of either contrast agent of 4:1, with an effect size of 0.18; and for study Arm 2, an “equal” response for 50% of the patients and a ratio of superiority of either contrast agent of 3:1, with an effect size of 0.125. The sample size assumption should be based on the full distribution of the study population for the paired 1-sample data with the ordinal scale and cannot be divided in half; in addition, the hypothesis test was 2-sided and did not assume “that the preference would be in favor of gadobenate in 80% of the patients who received the full dose (Arm 1) and in 75% of those who received the half dose (Arm 2)” as stated in the comment/letter.

In summary, we believe that the statistical methods were correctly applied in line with the study objectives. The power determination and sample size consideration correctly reflected the primary analysis; the assumptions were evidence-based and reflected the information available from previous clinical trials with identical designs.^{6,8,9}

Quantitative Data

The methodology adopted for quantitative evaluation has been validated in several prior comparative studies of this type,⁵⁻⁹ and there are absolutely no surprising or biased results.

Quantitative contrast parameters are an excellent metric for

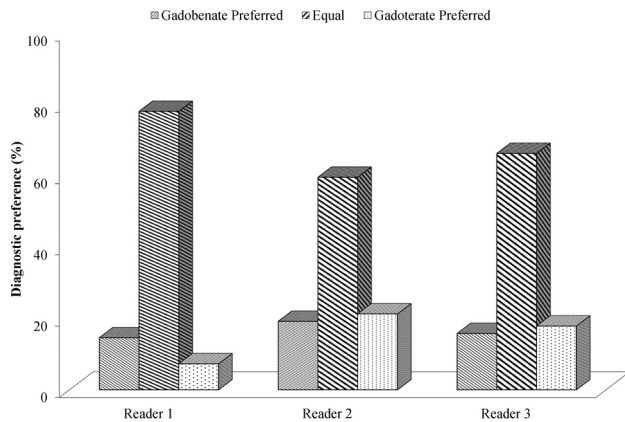


FIG 2. Distribution of overall diagnostic preference: gadobenate 0.05 mmol/kg versus gadoterate 0.1 mmol/kg.

lesion detection, which is certainly sequence-dependent. However, within-sequence intrapatient intralesion analyses were performed in this study, which eliminated any possible opportunity for biased interpretation. Criteria for measurement and selection of lesions to measure were common across readers. Moreover, training sessions were conducted with each blinded reader before the assessment of study images to ensure a consistent approach to image assessment. To standardize the size and placement of ROIs within a subject, ROIs were positioned in a paired fashion on predose T1-weighted spin-echo (SE) images and the corresponding postdose T1-weighted SE images of both examinations 1 and 2. Round or elliptical ROIs were placed on the image frame, which provided the best visualization of the lesions. ROIs were as large as possible but included only homogeneous areas. The same lesions were measured on predose and postdose T1-weighted SE/fast SE sequences for both examinations 1 and 2. This same procedure was used for the placement of ROIs on T1-weighted gradient recalled-echo (GRE) images. ROIs of the same shape and size were used for the individual measurements (lesion, normal parenchyma, and background noise) on each sequence type. For patients with multiple lesions, a maximum of 3 lesions that met the measurability criteria (ie, a homogeneous enhancing area of >5 mm, not having just very subtle rim enhancement, not being totally hemorrhagic, and not needing ROIs of <5 mm²) were considered.

Most important, in light of the issue raised by Drs Lancelot, Piednoir, and Desché, each reader individually chose the total number of lesions to be measured in a patient with multiple lesions. This approach created situations in which one reader might have measured 3 lesions while the others measured only 1 or 2 in the same patient. Likewise, readers were free to measure different numbers of lesions across different sequence types (T1-weighted SE or T1-weighted GRE). In crossover studies of this type, the intraindividual comparison (ie, the comparison within the reader of the same lesions on the same sequence type) is important. Therefore, the minimal differences in the number of lesions on the 2 different T1-weighted sequences (from 2 to 5, depending on the reader) do not bias or influence the results of the study because the analysis was performed by sequence type. Quantitative findings confirmed the predictable superiority of gadobenate at the same dose of 0.1 mmol/kg of body weight and the lack of any

meaningful difference for half-dose gadobenate compared with full-dose gadoterate. In addition to confirming the results of previous large scale intraindividual comparative studies,⁶⁻⁹ these quantitative results demonstrate once again the value of relaxivity as the only contributor to this specific outcome. The importance of relaxivity and the outcomes of previous trials that have compared gadobenate with other GBCAs^{6,8} have been recognized by regulatory agencies in Europe in section 5.1 of the current Summary of Products Characteristics.¹¹

While the study did not evaluate the impact of the diagnosis on patient management, such studies are extremely difficult to design because their interpretation presents the fundamental problem that the definition of accurate patient management based on either positive or negative test results may not be a single expected therapeutic choice and, more important, that a measured change in management does not necessarily translate into improved health outcomes. To date, there are no accepted guidelines for the design, reporting, and appraisal of patient-management studies.¹²

In conclusion, we believe that the design of this well-controlled clinical trial provides valuable information on the 2 GBCAs. First, it demonstrates that gadobenate is significantly superior to gadoterate for qualitative and quantitative enhancement of brain lesions when these agents are administered at an equivalent dose of 0.1 mmol/kg of body weight. This finding can be ascribed exclusively and unequivocally to the higher r1 relaxivity of gadobenate, which leads to superior contrast enhancement and significantly more clinically relevant morphologic information, which may be helpful for improved patient management and surgical planning. Second, it shows that there is no meaningful or relevant difference between a half dose of gadobenate and a full dose of gadoterate. The possibility of halving the amount of gadolinium administered is potentially extremely important for patients undergoing routine screening or follow-up examinations.

REFERENCES

1. Vaneckova M, Herman M, Smith MP, et al. **The benefits of high relaxivity for brain tumor imaging: results of a multicenter intraindividual crossover comparison of gadobenate dimeglumine with gadoterate meglumine (The BENEFIT study).** *AJNR Am J Neuroradiol* 2015;36:1589–98 CrossRef Medline
2. Colosimo C, Knopp MV, Barreau X, et al. **A comparison of Gd-BOPTA and Gd-DOTA for contrast-enhanced MRI of intracranial tumors.** *Neuroradiology* 2004;46:655–65 Medline
3. Knopp MV, Runge VM, Essig M, et al. **Primary and secondary brain tumors at MR imaging: bicentric intraindividual crossover comparison of gadobenate dimeglumine and gadopentetate dimeglumine.** *Radiology* 2004;230:55–64 CrossRef Medline
4. Essig M, Tartaro A, Tartaglione T, et al. **Enhancing lesions of the brain: intraindividual crossover comparison of contrast enhancement after gadobenate dimeglumine versus established gadolinium comparators.** *Acad Radiol* 2006;13:744–51 CrossRef Medline
5. Rumboldt Z, Rowley HA, Steinberg F, et al. **Multicenter, double-blind, randomised, intraindividual crossover comparison of gadobenate dimeglumine and gadopentetate dimeglumine in MRI of brain tumors at 3 Tesla.** *J Magn Reson Imaging* 2009;29:760–67 CrossRef Medline
6. Maravilla KR, Maldjian JA, Schmalfluss IM, et al. **Contrast enhancement of central nervous system lesions: multicenter intraindivi-**

- vidual crossover comparative study of two MR contrast agents. *Radiology* 2006;240:389–400 CrossRef Medline
7. Kuhn MJ, Picozzi P, Maldjian JA, et al. **Evaluation of intraaxial enhancing brain tumors on magnetic resonance imaging: intraindividual crossover comparison of gadobenate dimeglumine and gadopentetate dimeglumine for visualization and assessment, and implications for surgical intervention.** *J Neurosurg* 2007;106:557–66 CrossRef Medline
 8. Rowley HA, Scialfa G, Gao PY, et al. **Contrast-enhanced MR imaging of brain lesions: a large scale intraindividual crossover comparison of gadobenate dimeglumine versus gadodiamide.** *AJNR Am J Neuroradiol* 2008;29:1684–91 CrossRef Medline
 9. Seidl Z, Vymazal J, Mechl M, et al. **Does higher gadolinium concentration play a role in the morphologic assessment of brain tumors? Results of a multicenter intraindividual crossover comparison of gadobutrol versus gadobenate dimeglumine (the MERIT Study).** *AJNR Am J Neuroradiol* 2012;33:1050–58 CrossRef Medline
 10. Feinstein AR, Cicchetti DV. **High agreement but low kappa, I: the problems of two paradoxes.** *J Clin Epidemiol* 1990;43:543–49 CrossRef Medline
 11. MultiHance: Summary of Product Characteristics; last updated version December 4, 2015. Available at: <https://www.medicines.org.uk/emc/medicine/6132>. Accessed January 8, 2016
 12. Staub LP, Lord SJ, Simes RJ, et al. **Using patient management as a surrogate for patient health outcomes in diagnostic test evaluation.** *BMC Med Res Methodol* 2012;12:12 CrossRef Medline

A. Spinazzi

G. Pirovano

N. Shen

Global Medical and Regulatory Affairs

Bracco Diagnostics

Monroe, New Jersey

M.A. Kirchin

Global Medical and Regulatory Affairs

Bracco Imaging

Milan, Italy