
J Pruvo, V Vilgrain, C Roy, P Peretti, P Halimi, O Ernst, P Valette, C Matos, C El-Khoury

Citation

Abstract
Aim: To compare the diagnostic performance of Gd-DOTA-enhanced MRI with non-enhanced MRI in the characterization of tumoral lesions with histological (or other) corroboration.

Materials and Methods: Pooled data included 381 patients from three comparable Phase III trials of patients with abdomino-pelvic, hepatic and cerebro-spinal lesions. Each patient underwent MRI with appropriate unenhanced sequences (pre), followed by an injection of Gd-DOTA (0.1 mmol/kg) and corresponding enhanced sequences (post). Histology was used as the gold standard or, in the hepatic study, a corroborative diagnosis. Qualitative and quantitative assessments of images were done by one on-site and two independent off-site blinded readers.

Results: Technical failures were at least five times more frequent on the unenhanced sequences compared with the Gd-DOTA-enhanced sequences. Delineation of lesion borders was superior with enhanced MRI compared with unenhanced MRI, irrespective of evaluation off-site (‘post’-injection/‘pre+post’-injection: 60.9%/65.3% versus ‘pre’-injection: 35.5%, p<.0001) or on-site (post: 75.8%, versus pre: 39.9%, p<.0001). Gd-DOTA sequences improved diagnostic confidence both off-site (76.2%/83.2% versus pre: 57.6%, p<.0001) and on-site (88.7%/93.9% versus pre: 43.9%, p<.0001). Sensitivity and specificity were statistically significantly improved with Gd-DOTA for off-site (pre+post/post: 90.3%/88.8% versus pre 84.3%; and 71.1%/76.2% versus 65.3%) and on-site (95.2%/94.7% versus 71.2% and 81.9%/81.0% versus 60.8%) readings, respectively. There were no unexpected adverse events.

Conclusion: Gd-DOTA-enhanced MRI resulted in fewer technical failures, better image quality and better diagnostic performance compared with unenhanced MRI, confirming that Gd-DOTA adds clinical value and greater diagnostic confidence to the characterization of tumoral lesions.

INTRODUCTION
Magnetic resonance imaging (MRI) is a well established technique for imaging tumours. However, some tumours may appear isointense making differentiation from surrounding tissue difficult and enhancement agents are often used to provide greater delineation and characterization of these lesions (1-5).

Identification of tumours in the abdomen, pelvis, brain and spine present their own specific problems. Cross-sectional imaging of the abdomino-pelvic region offers important advantages over CT and is of particular value in the diagnosis and management of paediatric abdominal masses (6, 7).

MRI of the liver is made easier by the use of contrast agents, which improve lesion detection and characterization by increasing lesion to liver contrast, including many types of extrahepatic tumour (8, 9). Rapid breath-hold imaging techniques have been used with intravenous and intraluminal contrast media to demonstrate tumours of the solid visceral organs, the gastrointestinal tract, peritoneum, mesentery, omentum, bile ducts, lymph nodes, and osseous structures.

MRI is also considered to be the modality of choice for the evaluation of most patients presenting or suspected of having cerebral or spinal tumours in terms of both detection and differential diagnosis (10).

Gadoterate meglumine (Gd-DOTA; Magnescope in Japan; Dotarem in other countries; Guerbet, Roissy CdG Cedex, France) is a commonly used enhancement agent and its

2 of 10

safety has been well documented in clinical practice (11, 12). Efficacy has largely been judged by reader evaluation of the image, which is a subjective measure and has rarely been corroborated. Furthermore, a large variation in response has been observed between scanner types and this may impact upon the heterogeneity of overall efficacy (13).

This study evaluates a heterogeneous population by pooling data from three comparable Phase III studies. The aim of this analysis was to compare the diagnostic performance of Gd-DOTA-enhanced MRI with non-enhanced MRI in the characterization of tumoral lesions using histology or a corroborative diagnosis (hepatic patients only) as the gold standard.

The overall efficacy was to be assessed by accuracy, sensitivity and specificity of this comparative diagnosis.

MATERIAL AND METHODS

STUDY DESIGN

This analysis involved three Phase III, multicentre, open-label, fixed sequence paired studies that compared the efficacy and safety of Gd-DOTA using each patient as his/her own control. The studies evaluated: cerebral and spinal tumours (Study 1), abdominal and pelvic lesions (Study 2), and hepatic lesions (Study 3). At each centre, the study commenced after approval from the relevant Ethics Committee. Each patient provided written informed consent.

The primary objective of the analysis was to determine the intrinsic diagnostic performance of Gd-DOTA-enhanced MRI compared with non-enhanced MRI in the characterization of tumoral lesions, using histology or a corroborative diagnosis as the gold standard.

Two independent, off-site MR experts blinded to the patient's identity, clinical conditions, and histology or corroborative diagnosis, were required to read three MRI modalities: the unenhanced MRI or 'pre-injection' MRI; the Gd-DOTA-enhanced MRI or 'post-injection' MRI (post-injection images only); the 'paired' or 'pre+post'-injection MRI (unenhanced and Gd-DOTA-enhanced images had to be read jointly). These were read in a random order.

The histology was read on site by a reader who was blinded to the off-site MRI readers’ findings. MR images were also read on site; the on-site reading was unblinded to clinical conditions, but was performed without any information about the gold standard results. The readers were radiologists with a minimum of ten years experience. The centres used different equipment to perform the MRI examinations (Philips, GE or Siemens, 1 or 1.5T).

PATIENTS AND PROCEDURES

Patients could be male or female and had to be at least 18 years of age. All patients were scheduled to undergo a contrast-enhanced MRI examination to specify diagnosis or for follow-up of pre-existing lesions and were expected/scheduled to undergo either a biopsy or surgery of the lesion(s) detected. Female patients were required to be using effective contraception or to be surgically sterilised or post-menopausal (amenorrhea for at least 12 months).

Specific inclusion criteria for patients in the three individual studies were:

patients presenting or suspected of having a minimum of one cerebral or spinal tumour(s) and a maximum of three lesions detected during previous CT or MRI examination performed within 3 weeks of study MRI examination

patients suspected of having at least one complex and solid pancreatic or pelvic or renal or adrenal lesion as depicted during previous sonographic or helical CT examination performed within 3 weeks before inclusion

patients suspected of having a maximum of three complex or solid hepatic focal lesions, measuring at least 2 cm in diameter in at least one plane, depicted during previous sonographic examination performed within 2 weeks before MRI examination.

Exclusion criteria included a contraindication to MRI (such as pacemaker, aneurysm clip or severe claustrophobia), pregnancy, lactation, and known allergy to gadolinium chelates. Patients were also excluded if they were participating in another trial involving an investigational drug or if they had received a gadolinium complex within 2 days, or iron oxide nanoparticles within 7 days before the study MRI examination.

Specific exclusion criteria for patients included in the cerebral and spinal tumours study, and hepatic lesions study were: non-tumoral cerebral disease, such as diffuse pathology (e.g. multiple sclerosis, Alzheimer’s disease), or a liver cyst as unique lesion.

Patients who satisfied all eligibility requirements underwent an unenhanced MRI examination followed immediately by a
Gd-DOTA-enhanced MRI. Gd DOTA was administered as a single dose of 0.1 mmol/L/kg (0.2 mL/kg) by intravenous bolus. MRI procedures were performed under similar conditions according to the imaging laboratory's standard practice. The patient then underwent surgery and histological analysis within 1 month, or for liver tumours only, a corroborative diagnosis within 6 months after the Gd-DOTA-enhanced MRI. The stages and timings common to each of the three studies are summarized in Figure 1.

Figure 1
Figure 1. Key events and timings common to the three study schedules.

IMAGING PROTOCOL
The following sequences were performed according to anatomical investigation site.

1. Cerebral and spinal tumours:
   Before Gd-DOTA administration: T1 weighted Spin Echo, T2 weighted fast spin echo (FSE) and fluid attenuation inversion recovery (FLAIR).
   After Gd-DOTA administration: T1 weighted Spin Echo.

2. Abdominal and pelvic lesions:
   Before Gd-DOTA administration: T1 GE, T2 FSE, and Fat suppressed T1 GE.
   After Gd-DOTA administration: dynamic T1 GE (at 30, 60 and 90 seconds), T1 GE (at 180 and 300 seconds), and fat suppressed T1 GE.

3. Hepatic lesions:
   Before Gd-DOTA administration: T1 weighted breath hold gradient echo (GE) and T2 weighted FSE.
   After Gd-DOTA administration: T1 weighted breath hold GE was acquired at 20 seconds (arterial phase), 1 minute (venous phase) and from 5 to 10 minutes (delayed phase).

EFFICACY ASSESSMENTS
The off-site readers described each identified lesion in terms of location and morphology, including margin delineation, shape and appearance (homogenous/ heterogeneous). Each lesion was then characterized by the readers as ‘benign’, ‘malignant’ or ‘non assessable’. The images were also read on site.

TECHNICAL FAILURES
Technical failures of either MRI procedure were considered as diagnostic errors (non evaluable patients) relative to the gold standard procedure. The number of technical failures for each MRI procedure and the paired differences (Gd-Dota enhanced MRI minus gold standard, unenhanced MRI minus gold standard, Gd-Dota enhanced MRI minus unenhanced MRI) were calculated and compared.

LESION DELINEATION QUALITY ASSESSMENT
For each lesion, the reader had to assess the delineation quality using a 5-point scale: 1 = uncertain border between tumour and tissue that is not tumour; 2 = poorly defined border between tumour and tissue that is not tumour; 3 = fairly defined border between tumour and tissue that is not tumour; 4 = well defined border between tumour and tissue that is not tumour; 5 = very well defined border between tumour and tissue that is not tumour.

DIAGNOSTIC CONFIDENCE
For each lesion, the reader had to assess the diagnostic confidence using a 5-point scale: 1 = nil; 2 = poor; 3 = moderate; 4 = high; 5 = excellent. For the purpose of this analysis, only “high” and “excellent” diagnostic confidence scores have been considered.

DIAGNOSTIC PERFORMANCE
The patients’ diagnostic classifications (in terms of the agreement between the findings of the MRI procedures and those of the gold standard procedure) are summarized in Table 1. These classifications were True Positive (TP), True Negative (TN), False Negative (FN), and False Positive (FP).

Figure 2
Table 1. Patient’s diagnostic classification in terms of the relationship of the MRI procedures to the gold standard procedure

<table>
<thead>
<tr>
<th>MRI Results</th>
<th>Gold Standard</th>
<th>Patient’s Diagnostic Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>All benign lesions (--)</td>
<td>All benign lesions (--)</td>
<td>True Negative (TN)</td>
</tr>
<tr>
<td>All malignant lesions (++)</td>
<td>All malignant lesions (++)</td>
<td>True Positive (TP)</td>
</tr>
<tr>
<td>Non-benign lesions (++)</td>
<td>Non-benign lesions (++)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>All lesions at all (++)</td>
<td>All lesions at all (++)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>Non-benign lesions (++)</td>
<td>Non-benign lesions (++)</td>
<td>False Negative (FN)</td>
</tr>
<tr>
<td>All benign lesions (--)</td>
<td>All benign lesions (--)</td>
<td>True Negative (TN)</td>
</tr>
<tr>
<td>Non-benign lesions (++)</td>
<td>Non-benign lesions (++)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>All lesions at all (++)</td>
<td>All lesions at all (++)</td>
<td>False Positive (FP)</td>
</tr>
<tr>
<td>Non-benign lesions (++)</td>
<td>Non-benign lesions (++)</td>
<td>False Negative (FN)</td>
</tr>
</tbody>
</table>

The diagnostic performance was assessed by accuracy, sensitivity, specificity and positive and negative predictive values. These diagnostic efficacy variables were derived from matching reader evaluations resulting from both MRI and gold standard examinations. These variables were also evaluated at lesion level to investigate whether the matching of the diagnosis at this level might offer potential benefits for patient management. The matching of diagnosis at lesion level may offer a more rigorous investigation than the matching at patient level because two patient diagnoses could match even if the evaluation at the lesion level does not fit between MRI procedures.

INTRA-READER AND INTER-READER AGREEMENT

For off-site readings, intra-reader repetition on MRI results (lesions) was summarized for Reader 1 and Reader 2. Inter-reader repetition on MRI results (lesions) was also summarized for Reader 1 and Reader 2 within each MRI modality (pre-injection (unenhanced) and post-injection and pre+post).

SAFETY ASSESSMENTS

The following safety variables were reported in each of the three studies: quantity of Gd-DOTA injected; number of adverse events; number of patients experiencing at least one adverse event; number of patients treated for an adverse event; withdrawals and withdrawals due to adverse event; frequencies of each response category (intensity, outcome, pre-existence, and relationship); any changes in vital signs for up to 72 hours after the Gd-DOTA injection; and pain or discomfort at injection site.

STATISTICAL ANALYSIS

Three pooled patient populations were considered in this analysis. The all-included-patients population comprised all patients enrolled in the study who had signed the informed consent. The evaluable safety population included all patients who received at least one injection of Gd-DOTA regardless of the quantity. The evaluable efficacy population included all patients who had valid diagnostic assessments for the gold standard procedure and at least one of the two MRI procedures (Gd-DOTA-enhanced or unenhanced MRI examinations).

Two analyses were conducted on the evaluable efficacy population: a primary analysis involving assessable and all non-assessable (technical failures) MRI procedures (when paired with an assessable MRI procedure they were considered as diagnostic errors (FP, FN) and were taken into account in this primary analysis); a secondary analysis excluding these non-assessable MRI procedures (technical failures) when paired with an assessable MRI procedure.

All efficacy analyses were conducted on the evaluable population at patient-level and at lesion-level. Only off-site results were considered in primary analysis, the secondary analysis took into account on-site data.

All outputs were produced using SAS version 9.01 or a later version (SAS Institute – Cary, NC). Descriptive statistics were used for the demographic characteristics. Pre- and post-contrast MRIs were compared using McNemar’s test for the differences in accuracy, sensitivity and specificity. Statistical tests were two-sided and performed using a 5% level of significance.

Intra-reader and inter-reader agreements were both assessed using a Kappa test. Interpretation guideline for Kappa values is as follows: <0.20 poor agreement; 0.21–0.40 fair agreement; 0.41–0.60 moderate agreement; 0.61–0.80 good agreement; 0.81–1.00 very good agreement.

RESULTS

Demography, baseline characteristics and indication for examination

A total of 381 patients were enrolled in the studies. Demographic and baseline characteristics as well as indications for MRI examination for the all-included-patient population are summarized in Table 2.
Histological findings or additional evaluative corroboration of the patient’s condition was available for 381 patients (100%).

**EFFICACY EVALUATION AT PATIENT LEVEL**

**TECHNICAL FAILURES**

Technical failure rates were significantly greater for unenhanced images with both groups of readers (p<0.0001). Though the incidence of technical failures was higher for on-site readings than those made off-site and the difference between pre and post enhancement was greater on-site, with differences being seven times more frequent in pre than in pre+post injection modalities (Table 3).

**LESION DELINEATION QUALITY**

Delineation quality (percentage of lesion borders that were assessed as “well” or “very well” defined) was significantly superior on contrast-enhanced MRI modalities compared with unenhanced images, irrespective of whether readings were made off-site or on-site (Figure 2).

**Figure 2. Lesion delineation quality: percentage of lesion borders that were “well” or “very well” defined. Graph A shows results of off-site readings. Graph B shows results of on-site readings.**

Missing data at lesion level were excluded from this analysis for off-site (pre: N=47 patients, post-reading: N=10, pre+post reading: N=9), and on-site readings (pre-reading: N=92 patients, post-reading: N=18). There was no ‘pre+post’ MRI modality assessment on site.
**DIAGNOSTIC CONFIDENCE ASSESSMENT**

Diagnostic confidence (percentage of lesions for which diagnostic confidence was assessed as “high” or “excellent”) was significantly superior on contrast-enhanced MRI modalities compared with unenhanced images, irrespective of whether readings were made off-site or on-site (Figure 3). Missing data at lesion level were excluded.

**Figure 6**

Figure 3. Diagnostic confidence: percentage of lesion for which diagnostic confidence was assessed “high” or “excellent”. Graph A shows results of off-site readings. Graph B shows results of on-site readings.

The image clarity and confidence in delineation before and after Gd-DOTA administration is shown in the brain (Figures 4 and 5), the liver (Figure 6) and the kidney (Figure 7).

**Figure 7**

Figure 4. A 54-year-old woman. Same slices acquired before (a) and after (b) Gd-DOTA administration. The left parietal lesion is more well-defined and more easily assessable after Gd-DOTA administration.

**Figure 8**

Figure 5. A 81-year-old man. Same slices acquired before (a) and after (b) Gd-DOTA administration. Glioblastoma is more well-defined and more easily assessable after Gd-DOTA administration.

**Figure 9**

Figure 6. A 46-year-old woman. Same slices acquired before (a) and after (b) Gd-DOTA administration. The liver lesion is more well-defined and more easily assessable after Gd-DOTA administration.

**Figure 10**

Figure 7. A 87-year-old man. Same slices acquired before (a) and after (b) Gd-DOTA administration. The right kidney lesion is more well-defined and more easily assessable after Gd-DOTA administration.

**DIAGNOSTIC PERFORMANCE AT PATIENT**
LEVEL

Off-site readings. When off-site readings were evaluated at patient level with non-assessable data (technical failures) considered as diagnostic errors, all parameters of diagnostic performance (accuracy, sensitivity, specificity, positive predictive value [PPV], negative predictive value [NPV]) were statistically significantly higher for Gd-DOTA-enhanced MRI than for the unenhanced MRI (Table 3). When off-site readings were evaluated at patient level with technical failures, globally, all measures of diagnostic performance were higher for Gd-DOTA-enhanced MRI than for the unenhanced MRI without statistical differences between MRI modalities.

On-site readings. Similar patterns were observed when on-site readings were evaluated, although the performance differences between enhanced and unenhanced-MRI modalities were greater than those observed in the off-site situation (Table 3).

Diagnostic Performance at Lesion Level

When off-site and on-site readings were evaluated at lesion level, the findings were similar to those seen at patient level. When non-assessable data (technical failures) were taken into account, all parameters of diagnostic performance (accuracy, sensitivity, specificity, PPV, NPV) of contrast-enhanced MRI modalities were statistically significantly higher than for unenhanced MRI (Table 4). However, when technical failures were excluded globally all diagnostic performance parameters were higher for Gd-DOTA-enhanced MRI than for the unenhanced MRI without statistical differences between MRI modalities.

INTRA- AND INTER-READER AGREEMENT

Off-site readings only: The intra-reader agreement (Kappa values) was 0.702 for Reader 1 and 0.595 for Reader 2. Inter-reader agreement results (Kappa values) were 0.725 for pre-injection (unenhanced) MRI, 0.741 for post-injection and 0.710 for pre+post. These values indicate that both intra-reader and inter-reader agreement were relatively good.

SAFETY EVALUATION

The safety population comprised 379 patients who received an injection of Gd-DOTA. No new safety issue related to Gd-DOTA was identified during the course of the three trials. Thirty-seven patients experienced at least one adverse event (9.8%), representing a total of 53 adverse events. The most common adverse events were nausea, vomiting and pain. These were of mild or moderate intensity and had favourable resolution without treatment. Eighteen adverse events were reported as serious and, according to the Investigators, all were linked to the patient’s poor general condition and could not be attributed to the contrast agent administration.

DISCUSSION

In this analysis of pooled data from three tumour imaging studies, the use of Gd-DOTA-enhanced MRI resulted in fewer clinical failures, superior image quality, better diagnostic performance and greater diagnostic confidence than unenhanced MRI.
This confirms that accuracy of lesion imaging still necessitates tumour enhancement as described by Runge et al, even with the recent advances in technology which improve unenhanced image quality, including the development of high-performance imagers, improved surface coil designs, better software and changes in imaging techniques (1, 2, 3, 14). This information is particularly important where a high specificity and sensitivity is required in cancers, such as predicting patient prognosis in rectal cancer (15).

The strength of this analysis lies in the Committee for Proprietary Medicinal Products (16) guideline-based design, which uses a diagnostic gold standard as a comparator, unlike most recent efficacy studies where one enhancement agent has been compared with another (17, 18). The off-site reading of MR images, which is blinded from a methodological point of view, improves the study quality and reliability. The interpretation of the off-site images is totally objective, homogeneous between patients and without reading bias. The on-site evaluation, which takes into account a global view of the patient, may be considered to be closer to the situation in day-to-day clinical practice. Evaluation of both off-site and on-site readings were complementary and demonstrated that the use of Gd-DOTA contrast brought important improvements in diagnostic performance.

Limitations include using only one on-site reader, which may explain the higher technical failure rate for on-site readings, particularly as reader accuracy is reported to improve accordingly with the number of readers (19). The interpretation accuracy of pre-contrast images was greater for off-site readers and this may reflect their great experience as radiologists. However, post-enhancement images were described more accurately by on-site readers than off-site, which may reflect their access to detailed patient information or possibly familiarity, and hence, confidence with the particular scanner and computational presentation. Another weakness is that this study was multicentre using different types of MRI scanner and inter-centre data was not analyzed. Although the good overall diagnostic performance in this analysis suggests that this heterogeneity is not notable in terms of diagnostic impact.

The quality of delineation of lesion borders and the diagnostic confidence were superior with contrast-enhanced MRI compared with unenhanced MRI, irrespective of whether the endpoints were assessed off-site or on-site. While the results of delineation were good, it is worth noting the meta-analysis by Méndez Romero et al, whose examination of microscopic pathology found extensions beyond the tumour border identified on MRI (20).

The use of contrast-enhancement in MRI has been shown previously to improve detection, delineation, and characterization of a variety of different types of tumour (21-23). Clinical trials of Gd-DOTA-enhanced MRI performed to date have included several that have demonstrated the safety and efficacy of Gd-DOTA for imaging of the central nervous system (18, 24-26), characterising tumours of the liver (27) and abdominal and pelvic region (28-29). However, it is important to show that the efficacy of MRI contrast agents in terms of typical diagnostic performance measures (accuracy, specificity, sensitivity etc.) leads to more accurate diagnosis and optimal clinical management, with the ultimate goal of improved patient outcomes.

In conclusion, this multi-study pooled analysis demonstrates that the heterogeneity typical of day-to-day clinical practise does not impact unduly on lesion diagnosis. Even with technological advances in MRI scanning, Gd-DOTA enhanced MRI sequences provide superior accuracy, sensitivity and specificity for tumour characterization.

References

Author Information

Jean Pierre Pruvo
Neuroradiology Department, Roger Salengro Hospital

Valerie Vilgrain
Radiology Department, Beaujon Hospital

Catherine Roy
Civil Hospital Of Strasbourg

Paule Peretti
Radiology Department, Sainte-Marguerite Hospital

Philippe Halimi
Radiology Department, European George Pompidou Hospital

Olivier J. Ernst
Radiology Department, Huriez Hospital

Pierre-Jean Valette
Digestive Radiology Department, Edouard Herriot Hospital

Celso Matos
Radiology And Medical Imaging Department, Erasme Hospital

Carl El-Khoury
Radiology Department, Curie Institut