ORIGINAL REPORT

Standardization of the quantification of iron concentration in the liver by magnetic resonance imaging

J.M. Alústiza\textsuperscript{a,*}, J.I. Emparanza\textsuperscript{b}, P. Aldazábal\textsuperscript{c}, A. Garrido\textsuperscript{d}, N. García\textsuperscript{c}, E. Salvador\textsuperscript{a}, M. San Vicente\textsuperscript{a}, Grupo Burnia\textsuperscript{b}

\textsuperscript{a} Osatek-Donostia, Spain
\textsuperscript{b} Servicio de Epidemiología Clínica, Hospital Donostia, Spain
\textsuperscript{c} Unidad Experimental, Hospital Donostia, Spain
\textsuperscript{d} Laboratorio de Bioquímica, Hospital Donostia, Spain

Received 20 September 2010; accepted 16 December 2010

KEYWORDS
Hemochromatosis; Iron overload; Magnetic resonance imaging

Abstract
Objective: To calibrate 1.5 T magnetic resonance (MR) scanners for the quantification of the concentration of iron in the liver.

Material and methods: We analyzed twenty-eight 1.5 T MR scanners using a phantom with four tubes containing different concentrations of iron (III) chloride and one tube without iron. The phantom represented two typical patients: one with moderate iron overload and one with high iron overload. We measured the signal intensity (SI) ratio between each iron-containing tube and the tube without iron; then we calculated the theoretical levels of iron concentration in each scanner according to the model for the two levels of overload. We compared the results of each scanner with those of the reference scanner in which the model and the phantom had been designed, and we calculated the percentage of difference between the two scanners.

Results: The mean difference in the ratios compared to the reference center was 11\% (0.3--39). The mean concentration of iron was 71 μmol Fe/g for moderate overload and 193 μmol Fe/g for high overload. The mean difference was 6\% (1.2--7\%) and 3.4\% (0--16\%), respectively.

In two scanners, we applied a correction factor so that the difference was below 25\% in all cases.

Conclusion: We calibrated twenty-eight 1.5 T scanners for the concentration of iron in the liver and achieved variability less than 25\%.

© 2010 SERAM. Published by Elsevier España, S.L. All rights reserved.

\textsuperscript{*} Please cite this article as: Alústiza JM, et al. Estandarización de la cuantificación de la concentración de hierro en el hígado por resonancia magnética. Radiología. 2012;54:149--54.
\textsuperscript{*} Corresponding author.
\textsuperscript{E-mail address: jmalustiza@osatek.es (J.M. Alústiza).
\textsuperscript{†} See Annex 1 for the members of Burnia Group.
Introduction

Iron overload is harmful because the body is unable to eliminate the excess iron and this accumulates in different organs, particularly liver, heart and pancreas, which can result in irreversible damage if not appropriately treated. Indirect indicators of iron overload (elevated plasma ferritin values and/or transferrin saturation index) are highly sensitive but not very specific, making it difficult to determine which patients with elevated values actually have tissue iron overload. Elevated measurements appear in conditions as different as hemochromatosis, metabolic syndrome, resistance to insulin, chronic alcoholism or hepatopathy.

Other analytical and clinical data may help in the diagnosis of iron overload, but many cases remain not sufficiently defined, making necessary in some instances the direct measurement of body iron store through quantification of liver iron concentration (LIC), which typically entails performing a liver biopsy.

Since the 1990s, MR has proved its ability to accurately measure LIC and, as a non-invasive technique, it is bound to play a significant part in the diagnosis of patients with abnormal iron metabolism. This requires standardization of the method with the use of acquisition techniques that are widely available and allow reproducibility of results.

At our institution, between 1999 and 2002 we compared quantitative measurements of LIC obtained by biopsy with MR measurements in 122 patients and developed a calculation model for LIC quantification by MR, yielding a high correlation (r = 0.937) with measures provided by liver biopsy. The model uses the liver/muscle ratio of two gradient echo (GE) sequences following the protocol proposed by Dr. Gandon, from Rennes University, for LIC quantification by MR imaging. Since 2002, we have performed more than 1000 studies of liver iron quantification using MR imaging, contributing to the diagnosis of abnormalities in iron metabolism in the clinical practice. As a result, other institutions have shown an interest in having this technique and the calculation model, thus raising the problem of model reproducibility in different MR systems.

Some authors have expressed their interest in having dedicated phantoms in order to investigate this reproducibility and some trials have already been published in the literature.

Between 2002 and 2007, our working group constructed and validated a phantom consisting of four vials containing different concentrations of iron III chloride and one vial containing no iron. This phantom simulates two average patients: one with moderate iron overload and the other with high iron overload according to the calculation model described elsewhere (FIS Project 2004-2006 P041652 and Patent P200601279). As a result of that study, twelve 1.5 T systems in the Basque Country were calibrated. Mean variation between the different institutions was 13% and 8% for moderate and high iron overload, respectively, which is similar to that obtained in 9 real patients with different iron concentrations that were also studied in different institutions.

The objective of this study was to calibrate a number of 1.5 T scanners in Spain in order to carry out studies to quantify LIC using a calculation model and a phantom that were previously validated.

Material and methods

Participating centers

Twenty-eight Spanish centers with 1.5 T scanners were included in this study. The centers voluntarily registered for participation at the SEDIA (Spanish Society of Abdominal Imaging) website (www.sedia.es) in February–March 2007.
Figure 1  Phantom with the different concentrations of iron III chloride. (A) Photograph of the prototype. (B) MR image of the PD sequence (120/4/20°) (TR/TE/Flip angle). (C) MR image of the T2 sequence (120/14/20°). (1) Solution containing no iron. (2) Solution containing 0.3 mg Fe/mL. (3) Solution containing 0.5 mg Fe/mL. (4) Solution containing 0.6 mg Fe/mL. (5) Solution containing 1.2 mg Fe/mL. (a) Copper sulfate CuSO₄²⁻. (b) Water.

The signal from the tube containing no iron (solution 1) is used as the reference to calculate the ratio in a way similar to that of the muscle used in real patients. The signal intensity ratio of solutions 2 and 3 is similar to the average ratios of patients with moderate iron overload for the PD and T2 sequences. The ratios of solutions 4 and 5 are similar to the average ratios of patients with high iron overload. Coppers sulfate provides great signal intensity to the phantom, stable, facilitating its handling in the MRI systems. The water surrounding the tubes helps prevent paramagnetic susceptibility artifacts.

Table 1  List of the Spanish centers with 1.5 T scanners that participated in the study.

<table>
<thead>
<tr>
<th>Center</th>
<th>City</th>
<th>Scanner brand</th>
</tr>
</thead>
<tbody>
<tr>
<td>Clínica Universitaria de Navarra</td>
<td>Pamplona</td>
<td>Siemens</td>
</tr>
<tr>
<td>Complejo Hospitalario Universitario de Vigo</td>
<td>Vigo</td>
<td>General Electric</td>
</tr>
<tr>
<td>Centro Hospitalario Universitario Juan Canalejo</td>
<td>Coruña</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital Universitario Virgen del Rocio</td>
<td>Seville</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital Miguel Servet-I</td>
<td>Zaragoza</td>
<td>General Electric</td>
</tr>
<tr>
<td>Hospital Miguel Servet-II</td>
<td>Zaragoza</td>
<td>General Electric</td>
</tr>
<tr>
<td>Institut de Diagnóstic per la Imatge (IDI) Girona (Hospital Dr. Josep Trueta)</td>
<td>Girona</td>
<td>General Electric</td>
</tr>
<tr>
<td>Clínica Girona</td>
<td>Girona</td>
<td>General Electric</td>
</tr>
<tr>
<td>Hospital Clinic</td>
<td>Barcelona</td>
<td>Siemens</td>
</tr>
<tr>
<td>Corporació Sanitaria (CRC) – Sagrat Cor</td>
<td>Barcelona</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital Son Dureta</td>
<td>Palma de Mallorca</td>
<td>Siemens</td>
</tr>
<tr>
<td>Hospital Universitario Dr. Peset</td>
<td>Valencia</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital Quirón Valencia</td>
<td>Valencia</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital de la Ribera</td>
<td>Alzira</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital General Universitario Gregorio Marañón</td>
<td>Madrid</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital de León</td>
<td>León</td>
<td>Siemens</td>
</tr>
<tr>
<td>Clínica La Milagrosa</td>
<td>Madrid</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital Universitario 12 de Octubre</td>
<td>Madrid</td>
<td>Philips</td>
</tr>
<tr>
<td>Sanatorio San Francisco de Arés</td>
<td>Madrid</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital Universitario Gran Canaria Dr. Negrín</td>
<td>Las Palmas</td>
<td>Siemens</td>
</tr>
<tr>
<td>Hospital Santa María del Rosell</td>
<td>Cartagena</td>
<td>General Electric</td>
</tr>
<tr>
<td>UDIAT</td>
<td>Sabadell</td>
<td>Siemens</td>
</tr>
<tr>
<td>Hospital de Madrid – Montepríncipe</td>
<td>Boadilla del Monte</td>
<td>Philips</td>
</tr>
<tr>
<td>Centro Integral de Diagnóstico y Tratamiento “Francisco Díaz”</td>
<td>Alcalá de Henares</td>
<td>General Electric</td>
</tr>
<tr>
<td>Hospital Universitario Príncipe de Asturias</td>
<td>Alcalá de Henares</td>
<td>Philips</td>
</tr>
<tr>
<td>Hospital Universitario Reina Sofia</td>
<td>Córdoba</td>
<td>General Electric</td>
</tr>
<tr>
<td>Hospital Mesa del Castillo</td>
<td>Murcia</td>
<td>General Electric</td>
</tr>
<tr>
<td>Hospital General Universitario “Moraes Meseguer”</td>
<td>Murcia</td>
<td>General Electric</td>
</tr>
<tr>
<td>Osatek Donostia</td>
<td>San Sebastián</td>
<td>Philips</td>
</tr>
</tbody>
</table>
(Table 1). The group included 5 Siemens, 10 General Electric and 13 Philips MRI units.

Optimization of the technique

Between May and October 2007, information was exchanged with the different centers regarding the optimization parameters for the sequences and the image media format that would be subsequently used to submit the measurements of SI in order to verify the correct implementation of all technical aspects. Some centers were required to modify the bandwidth to adjust all the parameters.

Studies on phantom models

Four copies of the phantom were constructed on a support safe for transportation to the different participating centers. On the reference scanner, in Osatek-San Sebastian, we verified that the measurements from the four phantoms were equivalent (percentage deviation <4.5% and <3.2% for moderate and high iron overload, respectively). Between October and December 2007, each center received any of the four copies of the phantoms. In each center the two sequences described in the quantification model, called PD (proton density) and T2 (Table 2), were obtained from the phantom, with 5 slices perpendicular to the vials and recording the SI values from the 3 central slices. The results of the SI measurements were sent to the Osatek Unit in San Sebastian using the image media format previously mentioned.

Statistical analysis

(1) For each tube, the SI ratio was calculated with respect to the tube containing no iron in the two sequences. Therefore, 8 ratios were obtained from each scanner. These ratios were compared to the ratios of the reference scanner, calculating the percentage differences (Ratio X – Ratio Ref/Ratio Ref × 100) (Ratio X: ratio of the MRI unit to be calibrated, Ratio Ref: ratio of the scanner used as reference).

(2) For each center, the theoretical LIC of the two virtual patients simulated by the phantom (with moderate and high iron overload) were also estimated using the corresponding ratios of the calculation model. The results were compared with the values from the reference scanner. The percentage difference was also calculated and the clinical relevance of the observed results was assessed.

The mean, confidence interval (CI) at the 95% confidence level, and range (minimum value–maximum value) were estimated for the two cases.

Table 2 Parameters for the two gradient echo sequences used in the model for iron liver determination.

<table>
<thead>
<tr>
<th>Sequence</th>
<th>TR (ms)</th>
<th>TE (ms)</th>
<th>Flip angle (°)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PD</td>
<td>120</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>T2</td>
<td>120</td>
<td>14</td>
<td>20</td>
</tr>
</tbody>
</table>

Results

The mean percentage difference for the 8 ratios estimated for each center compared to the reference center was 11% (range 0.3–39) (95% CI: 7–14). In order to obtain these results, two of the centers had to modify the original sequences, even after optimization of parameters, because the GE technique, they used, led to very high percentage differences.

Mean theoretical values of LIC of the two virtual patients simulated by the phantom estimated for all participating centers were 71 μmol Fe/g (range 50–93) (95% CI: 67–74.5) for moderate overload, and 193 μmol Fe/g (range 165–217) (95% CI: 188–197) for high overload. Mean percentage differences for these LIC values compared to the reference MRI unit were 6% for moderate iron overload (range 1.2–37%) and 3.4% for high iron overload (range 0–16%).

For the virtual patient with moderate overload, only two centers showed an overestimation >30% (22 and 25 μmol Fe/g, respectively, in absolute terms). None of the scanners showed differences >30% for high iron overload. Only those two scanners required an ad hoc correction. The correction factor involved subtracting 12 and 21 μmol Fe/g, respectively, to the results estimated by the calculation model, so that the variability with respect to the reference scanner was <25%, for both moderate and high iron overload.

Discussion

Reproducibility level between the different scanners, for both the SI ratios between the tubes and the estimated LIC for the two levels of iron overload simulated by the phantom, was very high, within the range of the validation study carried out in the Basque Country between 2004 and 2007, on phantom and on real patients (Table 3).

A cut-off of 30% was used as standard of quality similar to that observed in the variation of LIC obtained by liver biopsy, which ranges between 24% and 41%. Reproducibility of the two centers the maximum variation was >30% after directly applying the model. Once the correction factor was applied in these two centers, maximum variation was reduced to 25%. For the virtual patient with high overload, maximum variation was 16%, and it was therefore not necessary to apply a correction factor in any of the centers.

A significant aspect of this study is the fact that two of the participating centers had to modify their GE technique. One of them previously used this technique, obtaining errors that went therefore unnoticed in the clinical practice when trying to estimate LIC using Dr. Gandon’s protocol from Rennes University. Both scanners were older than 9 years and used GE techniques with low paramagnetic susceptibility, known as “f2D-we” and “f2D×10s”, where the SI did not change significantly between the two sequences of the calculation model for the same tube.

Our results demonstrate the reproducibility of the technique but this reproducibility should not be taken for granted and we recommend that each center carry out a specific verification.

Taking into consideration the current need for standardization of the technique for quantification of LIC using MR imaging, we believe that this study represents a significant
contribution given the fact that 28 scanners have been actually calibrated. Other studies have used phantoms for scanner calibration with similar results to those obtained in this study, but they have not calibrated such a large number of scanners. In addition, the phantoms are available from our team free of charge upon request.

The LIC can be estimated using the quantification model available free of charge at the SEDIA website (www.sedia.es). The final report of this study is also available there.

A major limitation of this study results from the fact that the calculation model was developed from patients with iron overloads <400 μmol Fe/g, for this reason, the conclusions cannot be applied to patients with higher overload, common in patients with secondary hemochromatosis.

Since this phantom ensures reproducibility of results, its use is very interesting in multicenter studies involving patients with abnormalities in iron metabolism of different etiologies because it offers higher statistical accuracy in the evaluation of diagnostic protocols and the possibility of assessing new iron-chelating therapies in patients with secondary hemochromatosis.

**Conflict of interest**

The authors declare not having any conflict of interest.

**Funding**

The present study has been fully funded by SEDIA-Investiga (www.sedia.es).

**Authorship**

1. Responsible for the integrity of the study: JMA.
2. Conception of the study: JMA, JIE, PA.
3. Design of the study: JMA, JIE, PA.
4. Acquisition of data: JMA, JIE, PA, AG, NG, ES, MSV and Burnia Group.
5. Analysis and interpretation of data: JMA, JIE, PA, AG, NG.
6. Statistical analysis: JIE, NG, ES.
7. Bibliographic search: JMA, PA, AG, ES, MSV.
8. Drafting of paper: JMA, JIE, PA, AG, NG, ES, MSV and Burnia Group.
9. Critical review of the manuscript with intellectually relevant contributions: JMA, JIE, PA, AG, NG, ES, MSV and Burnia Group.
10. Approval of the final version: JMA, JIE, PA, AG, NG, ES, MSV and Burnia Group.

**Acknowledgements**

We would like to thank Eusebio Azcue from the Osatek Unit-Donostia for coordinating the transfer of the phantoms between centers. Our thanks also to Dr. Luis Martí Bonnati and Ángel Alberich for their support from SEDIA-Investiga.

**Annex 1.**

Members of the Burnia Group:

- Dr. José M. Alústiza, Dra. Marta Aduna, Dr. José A Larena, Dr. Óscar Casado, Dra. Blanca Martínez de Guereñu, Dra. Amaia Merino, Dr. Guillermo Fernández, Dr. Manuel San Vicente, Cristina de la Calle, Dra. Emma Salvador (Osatek),
- Dra. Arantxa Moyua (Policlínica Gipuzkoa, San Sebastián),
- Dr. Xabier Arozena, Dr. Miguel Zarco (Clínica Quirón, San Sebastián),
- Dra. Lourdes Jáuregui, Ohiana Vicente (Instituto Oncológico, San Sebastián),
- Dra. Rosa M. Rodrigo (Clínica San Francisco Javier, Bilbao),
- Dra. Elena Astigarraga (Clínica Quirón-Preiteimagen, Bilbao),
- Dr. Fernando García Delgado (Clínica Dr. San Sebastián, Bilbao),
- Dr. Jon Egidazu (Clínica Virgen Blanca, Bilbao),
- Dr. Javier Grande (Hospital Basurto, Bilbao),
- Dra. Begoña Canteli (Hospital de Cruces, Bilbao),
- Dra. Ana Gabilondo (Manterola SA, Elbar),
- Dr. Agustín Castiella, Eva Zapata (Hospital Mendaro, Mendaro),
- Dr. José I. Empananz, Dr. Adolfo Garrido, Dr. Pablo Aldazábal, Nerea García, Dra. Ana B. Asensio (Hospital Donostia, San Sebastián),
- Dr. Alfonso Casado (Universidad del País Vasco, San Sebastián).

**Table 3** Mean (minimum and maximum) of liver iron concentration values, estimated for ‘‘average patients’’ simulated by the phantom, obtained in the two multicenter studies mentioned.

<table>
<thead>
<tr>
<th>No.</th>
<th>Moderate overload</th>
<th>High overload</th>
</tr>
</thead>
<tbody>
<tr>
<td>Validation Basque Country 2004–2007</td>
<td>12</td>
<td>64 (48–85)</td>
</tr>
</tbody>
</table>

No.: number of calibrated scanners.

* Theoretical LIC (μmol Fe/g).

**References**