Liver iron overload assessment by MRI R2* relaxometry in highly transfused pediatric patients: An agreement and reproducibility study

S. Verlhac a, *, M. Morel a, F. Bernaudin b, S. Béchet c, C. Jung d, M. Vasile a

a Service d'imagerie médicale, centre de référence de la drépanocytose, centre hospitalier intercommunal de Créteil, 40, avenue de Verdun, 94010 Créteil cedex, France
b Service de pédiatrie, centre de référence pédiatrique de la drépanocytose, université Paris Est Créteil, centre hospitalier intercommunal de Créteil, 40, avenue de Verdun, 94010 Créteil cedex, France
c Association clinique et thérapeutique infantile du Val-de-Marne (ACTIV), 94100 Saint-Maur des Fossés, France
d Centre de recherche clinique, centre hospitalier intercommunal de Créteil, 40, avenue de Verdun, 94010 Créteil cedex, France

KEYWORDS
Iron overload;
Liver MRI;
Sickle cell disease

Abstract
Aims: Perform an agreement and reproducibility study of the estimation of iron overload in highly transfused pediatric patients comparing R2* relaxometry (R2* = 1000/T2*) to the reference technique liver/muscle signal intensity ratio (SIR).

Patients and methods: Ninety-two MRI were performed in 68 children who were mainly transfused for sickle cell disease, mean age 9.9 years old. The examination included six sequences for the SIR protocol and a single multiecho T2* sequence. R2* relaxometry was measured by two radiologists independently, either by a region of interest (ROI) in the right liver, or an outline of the whole liver. Hepatic iron load was determined by the Wood formula (Fe mg/g = R2* × 0.0254 + 0.202). The validity of R2* relaxometry compared to SIR was evaluated by the coefficient of variation and the quadratic weighted Kappa value.

Results: The correlation between R2* relaxometry and SIR was very good with a Pearson coefficient of 0.89 and a coefficient of variation of 17.3%. The inter- and intraobserver reproducibility of the measurement of R2* relaxometry by ROI and whole liver mapping was excellent.
Iron overload is a risk factor of mortality in highly transfused patients. It is the main prognostic factor in beta-thalassemia major, but it also affects patients with sickle cell disease [1] because of increasingly broad indications for long-term blood transfusion, mainly to prevent cerebral vasculopathy in children and in case of severe complications in adults (pulmonary hypertension, cardiac or renal failure and stroke). Assessment of iron overload has become essential for the therapeutic management of multitransfused patients whether it is for beta-thalassemia major, sickle cell disease, or myelodysplastic syndrome because it makes it possible to begin and adapt chelation treatment.

It is well-known that hepatic iron content (HIC) is a reliable reflection of overall iron content in the organism [2] and MRI has become the reference technique for assessing HIC, replacing liver biopsy because it is noninvasive. The paramagnetic properties of the iron in hepatocytes and Kupffer cells result in local changes in the magnetic field causing a drop of the liver signal. This effect is especially visible on T2-weighted sequences, which are the most sensitive for detecting iron. Two main protocols are available, and each of them has been validated in comparison to the histological results of iron measurements obtained from liver biopsies. In France, the reference technique is liver/muscle signal intensity ratio (SIR). Developed by Gandon [3] in 2004, and completed by Rose [4], it is based on the comparison of the liver signal to that of the paravertebral muscles, which serve as a reference because they do not store excess iron. To perform this technique, five or six sequences must be obtained, the signal of the liver and muscles must be measured on each of the sequences and the data must transferred to a calculator that is available free of charge on internet. SIR has been validated in comparison to HIC measurements based on biopsy samples in 139 adult patients with primary hemochromatosis [3]. The second method is based on the measurement of transversal relaxation time T2* of the liver on a single multiecho gradient-echo T2 sequence, a sequence which is also used to evaluate myocardial iron. As T2* has a negative relationship to HIC, T2* is converted into its reciprocal $R^2 = 1000/T2^*$, i.e. the relaxation rate which is directly proportional to HIC. This relaxometry method has been validated in several studies including children with post-transfusion iron overload. In a series of 102 thalassemic or sickle cell patients and 13 controls, all of whom were evaluated by 1.5 T MRI and 23 of whom underwent liver biopsy, Wood [5] found an excellent correlation between R2* in the liver and the HIC measured on biopsy specimens, with a variation coefficient of 0.97, resulting in a calibration curve between R2* and HIC which is now commonly used: HIC mg/g dry liver = R2* × 0.0254 + 0.202. The same correlation between R2* and liver biopsies was found by Hankins [6] in 43 patients including 32 with sickle cell disease between 7 and 35 years old (correlation coefficient 0.98, P<0.001). A more recent study [7] comparing T2* and liver biopsy in 25 patients, mean age 43, showed that a decrease in relaxation time $T2^* < 10.07$ ms was a good predictor of hepatic iron overload with excellent sensitivity 84% and specificity 100%. With a cut-off of 15.47 ms, sensitivity was 89.5% and specificity 83%. The goal of our study was to evaluate the reproducibility and agreement of liver iron measurements by $R^2^* = 1000/T2^*$ in children using liver/muscle signal intensity ratio as the reference standard.

**Patients and methods**

**Patients**

This single center study included all children who underwent liver MRI to evaluate hepatic iron content between August 2009 and March 2013. Ninety-two MRI were performed in 68 patients of mean age 9.9 years old (2.3–24). Forty-six patients underwent one examination (67.65%), 21 underwent 2 (30.88%), and 1 underwent 4 (1.47%). The diseases causing iron overload were sickle cell disease (n=60), thalassemia (n=6) and Blackfan Diamond anemia (n=2). Chelation therapy at the moment of the examination was deferasirox (Exjade®) in 46 cases, deferiprone (Ferriprox®) in 17 cases, the combination of deferoxamine (Desferal®)/deferiprone (Ferriprox®) in one case, deferoxamine (Desferal®)/deferoxamine (Exjade®) in one case, and no treatment in 16 cases. Ten patients had undergone hematopoietic stem cell transplantation. Clinical and biological data were obtained from each patient’s file. The median total number of transfusions or exchange transfusions received by each patient was 39 (7–207). For the entire cohort, median ferritinemia was 2001 ng/mL (94–16927), AST 49 U/L (18–158), ALT 24 U/L (12–196), total bilirubin 29 μmol/L (4–157), iron 28 μmol/L (6–63) and the saturation coefficient was 59% (10.8–101.6). Median hemoglobin was 9.35 g/dL (7–13.6).

**MRI**

MRI examinations were performed on a 1.5 Tesla, General Electric Healthcare, Milwaukee, WI, USA. Each MRI included six sequences according to the Rennes/Lille protocol described on the website [8] (SIR method) using a body coil, and a single axial mid-hepatic slice fast gradient-echo multiecho T2* sequence called MFGRE/20 using a torso-phased array coil positioned over the heart and liver. The following parameters were used: axial plane

---

However, we observed a common positive variation of one class between SIR and R2* relaxometry, with higher hepatic iron content values with SIR than with R2* relaxometry.

**Conclusion:** Hepatic iron content can be rapidly and precisely estimated on MRI by multiecho gradient-echo sequences.

© 2014 Éditions françaises de radiologie. Published by Elsevier Masson SAS. All rights reserved.
breath-hold acquisition (without breath-holding in young children), flip angle 20°, FOV 30, slice thickness 10 mm, 16 regularly spaced echo times from 1.2 ms–25.8 ms, TR 200 ms, 1 excitation, matrix 96/96, bandwidth 83 kHz, acquisition time 21 s (Figs. 1 and 2). Post-processing of the SIR technique included positioning three regions of interest (ROI) in the liver and one in each paravertebral muscle to measure the signal, then entering the data in the calculator on the website and converting the results from μmol Fe/g dry liver into mg Fe/g dry liver on excel using the formula μmol/g/17.8 = mg/g performed by a senior reader. For relaxometry, R2* = 1000/T2* was calculated on the Windows 4.4 workstation with the manufacturer’s R2* software by 2 methods. First, a small

Figure 1. Hepatic MRI, MFGRE sequence. Up: images of the 5 initial echos. Down: signal decay curve, R2* mapping with a small ROI and whole liver mapping. Slow signal decay can be noted in the liver. The R2* is 47 Hz, which corresponds to an HIC of 1.39 mg Fe/g dry liver, i.e. 24 μmol Fe/g dry liver, indicating an absence of iron overload.

Figure 2. Hepatic MRI, MFGRE sequence, Images of the 5 initial echos. There is rapid decay in the signal in the liver after the second TE. The R2* is 700 Hz, which corresponds to 18 mg Fe/g dry liver, i.e. 320 μmol Fe/g dry liver, indicating severe iron overload.
10 cm² region of interest (ROI) was drawn in the right lobe on the R2* map, secondly, a map of the whole liver was obtained by outlining the entire cross-sectional area of the liver, excluding vessels (Figs. 1 and 2). Measurements were made independently by two radiologists, a junior reader (MM) after a short training session and a senior reader (SV) who analyzed the images twice, on the day of the examination and during reviewing sessions. It should be noted that the R2* maps were determined by measuring the R2* pixel by pixel using a reconstruction algorithm with a double exponential to prevent underestimating high iron overload which induces a sharp increase in the R2* relaxation rate curve followed by a plateau, for which only the initial R2* values should be taken into account. Iron content was determined using Wood’s formula [5] (Fe mg/g = R2* × 0.0254 + 0.202). Four classes of hepatic iron content (HIC) were defined for each examination according to the classification described by Olivieri [9] in thalassemic patients: class 1: normal HIC < 3 mg Fe/g dry liver, i.e. < 50 μmol/g, R2* < 88 Hz; class 2: mild overload HIC 3—7 mg Fe/g dry liver, i.e. 50—125 μmol/g, R2* 88—263 Hz; class 3: moderate HIC overload 7—15 mg Fe/g dry liver, i.e. 125—270 μmol/g, R2* 263—555 Hz; class 4: severe HIC overload > 15 mg Fe/g dry liver, i.e. > 270 μmol/g, R2* > 555 Hz.

Myocardial T2* measurements were also obtained for each patient from the interventricular septum by multiecho sequence using a short axis view of the heart [10].

Statistical calculations

The validity of R2* was evaluated in relation to SIR by the Bland Altman plot and the coefficient of variation for quantitative values, and by the quadratic weighted Kappa value for the classes of iron overload. To study agreement, SIR values ≥ 300 μmol/g (16.8 mg/g) were not taken into account because it is known that sequences from the Gandon protocol [3] are saturated at concentrations above this threshold. Intra- and interobserver reproducibility was evaluated by the Pearson correlation coefficient, the Bland Altman plot and calculation of the intra-class correlation coefficient ICC. Statistical analyses were performed with Stata 12 software.

This study was approved by the ethics committee.

Table 1 Classes of hepatic iron content: SIR versus R2* relaxometry. A common overestimation of one class is noted by SIR. For example 12 patients defined as class 1 (no iron overload) by R2* were defined as class 2 (mild overload) by SIR.

<table>
<thead>
<tr>
<th>SIR</th>
<th>R2*</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>6</td>
</tr>
<tr>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>0</td>
</tr>
</tbody>
</table>

Results

Hepatic R2* (median, range) measured by the senior reader was 234 Hz (49—857), corresponding to a HIC of 6.15 mg Fe/g dry liver (1—22), or 109 μmol/g Fe/g dry liver (17.8—391). The HIC (median, range) measured by the SIR method was 200 μmol Fe/g dry liver (5—798), or 11.23 mg Fe/g dry liver (0.28—44.8). The correlation between R2* and SIR was very good (P < 0.001) with a Pearson coefficient of 0.89; the coefficient of variation was also very good (17.3%) (Fig. 3). The comparative study of the classes of iron overload showed good agreement for the classes between the two techniques with a weighted Kappa value of 0.63 with a common positive variation of one class between SIR and R2* relaxometry (Table 1). Of the 20 tests that were classified as normal by R2*, 12 (13% of the total) were classified as having mild overload by SIR. The median cardiac T2* was 33 ms (15—78). There was a correlation between HIC and serum ferritin (r = 0.553, P < 0.001), and between HIC and the number of transfusions or exchange blood transfusions (r = 0.350, P = 0.001). On the other hand, we did not find any correlation between HIC and total bilirubin, liver enzymes (ALT, AST), serum iron saturation coefficient, or between HIC and myocardial T2*. The correlation between the two drawing techniques (ROI and mapping) was excellent for both the senior (Pearson = 0.97, ICC = 0.96) and junior reader (Pearson = 0.98, ICC = 0.97) (Fig. 4). Inter and intraobserver reproducibilities of R2*

Figure 3. a: Correlation between SIR measurements and R2* relaxometry (r = 0.89, P < 0.001); b: Agreement of SIR/R2* relaxometry measurements according to the Bland and Altman plot. Means of measurements on the abscissa (x-axis) and the differences of the measurements (SIR-R2* relaxometry) on the ordinate (y-axis). Slight overestimation of iron overload by SIR compared to iron overload measured by R2* relaxometry.
by small ROI (respective ICC 0.97 and 0.96) and interobserver reproducibility by whole liver mapping (ICC = 0.97) were excellent.

Discussion

As reported in the literature, we found excellent intra- and inter-reproducibility of the R2* liver measurement [11–13]. Chandara [11] reported an excellent intra-patient and inter-device reproducibility for the T2* measurement with a coefficient of variation of less than 5%. In an international study [13] conducted in 5 countries, the intercenter reproducibility was excellent with a variation coefficient of 5.9% for the heart and 5.8% for the liver. Moreover, the interobserver reproducibility was 5.4% for the heart and 4.4% for the liver. As reported by McCarville [14], intra and inter-reviewer agreement was excellent for both small and whole liver ROI methods. Our study showed very good agreement between SIR and R2* methods for measuring HIC of less than 16 mg/g. It should be remembered that we chose to exclude SIR values of ≥ 300 μmol/g, which were obtained from the additional sequence designed by Rose [4] from the study on agreement, because our initial calculations resulted in a marked dispersion of values above this threshold, probably due to the limitations of the two techniques (no possibility for our machine to set a TE of less than 1.2 ms and limitation of the sequence designed by Rose which was not validated by other teams). The SIR technique has the advantage of being easy to use. The sequences are described on the website and there is free access to the calculator which automatically selects the most well adapted sequence to determine the patient’s iron content. However, SIR technique has several limitations. Despite extensive validation of the technique, it has mainly been assessed in adult patients with genetic hemochromatosis. It requires several [5,6] breath-holding acquisitions. Breath-holding is often impossible in children and there is a risk of variation between the sites of measure in youngest patients. Nevertheless, it should be noted that the SIR technique could also be performed by multi-echo. SIR is rarely used outside of France, which makes it difficult to participate in international studies. Finally, we found that Gandon’s algorithm tended to overestimate HIC. This has already been reported by a Spanish team, which proposes another algorithm [15]. This group compared the SIR technique with iron measurements in biopsy specimens in 171 patients, mean age 46.5, and showed that the specificity of SIR for the absence of iron overload was only 57%: 43% of normal patients were classified with SIR as having iron overload; 44.7% of the patients with moderate iron overload were classified as having severe overload. This could be due to the preparation of the biopsy specimens that were used for calibration of the Gandon technique. Numerous studies have shown that T2* measurement was a rapid and accurate MRI method to determine HIC [5–7,10–12,14,15]. Thanks to modernization of MRI devices, multi-echo gradient sequences are widely available and could replace multi-sequences protocols. In pediatrics, the importance of a single sequence is obvious. Besides reducing the examination time, this can prevent imprecise measurements due to spatial variations of multiple sequences in children who have trouble holding their breath. When setting parameters, to avoid underestimating high iron loads, it is important to use the shortest initial TE, less than 1 ms if possible. For the moment, the main limitation of relaxometry is that the availability of image processing software varies among manufacturers. Ideally calculation software packages should be available on the manufacturer’s workstation. However, several commercial software packages are available and it is also possible to perform your own calculations with Excel® or MATLAB® [16]. Agreement of results is good as long as the algorithm used eliminates late echos corresponding to background noise when the drop in signal is early and rapid, either by truncating the monoexponential signal decay curve or by using a double exponential. Finally, R2* relaxometry allows simultaneous evaluation of myocardial iron overload, with the coil placed over the heart and the liver. Several studies have shown that there is no correlation between iron overload in the liver and heart, which probably results from different iron accumulation and chelation kinetics in the two organs [10,17], justifying simultaneous assessment of both organs.

The main limitation of our study is that no correlation has been performed between MRI R2* and histological iron measurements, because liver biopsy is no longer indicated in France in these cases. For this reason, we compared two MRI techniques that are commonly used in daily practice and have already been validated in comparison to biopsies by different teams.
Conclusion

In conclusion, our study indicates that R2* relaxometry is highly correlated to the SiR method, provided that initial TE of the multiecho sequence is short. R2* relaxometry may be preferable in patients with thalassemia and sickle cell disease, especially in children, because it allows simultaneous quantification of liver and myocardial iron overload.

Disclosure of interest

The authors declare that they have no conflicts of interest concerning this article.

References