MRI for the measurement of liver iron content, and for the diagnosis and follow-up of iron overload disorders

Anita Paisingt, Gaspard d’Assignies, Elise Bannier, Edouard Bardou-Jacquet, Yves Gandon

Available online: 11 November 2017

CHU de Rennes, service de radiologie, 35033 Rennes, France

Correspondence:
Yves Gandon, CHU de Rennes, service de radiologie, 35033 Rennes, France.
yves.gandon@chu-rennes.fr

Summary

MRI is now the reference method for detecting and quantifying hepatic and extrahepatic iron overload, regardless of its cause. The decrease of the hepatic signal is proportional to the amount of iron in the tissues. It is more pronounced with T2*-weighted gradient echo sequences. It increases proportionally with the strength of the magnetic field. Thus a 3-T MRI is more sensitive and probably more accurate to detect a slight iron overload, as seen in dysmetabolic hepatosiderosis. Conversely, a 1.5-T MRI better estimates a high overload. Quantification can be done with the calculation of T2* (or R2*) or by using the liver to muscle signal intensity ratio (SIR). Today with a single multi-echo gradient-echo sequence, obtained in a unique apnea, the two methods can be used simultaneously. An associated quantification of steatosis is also obtained. This same type of sequence is proposed for quantification of iron in other tissues and in particular for the myocardium.

Introduction

Iron overload is a common clinical situation, due either to an intestinal absorption increase as for genetic hemochromatosis, or to repeated transfusion (thalassemia, drepanocytosis, myelodysplasia…), or even to dysmetabolic iron overload syndrome (DIOS). Quantification of iron is also an emerging topic in other fields such as hemodialysis follow-up [1]. An easier and earlier diagnostic, added to a greatest facility to evaluate severity of the disease, the follow-up and the treatment improved patients’ survival.

Regardless to the etiology of iron overload, hepatosiderosis is constant feature and its quantification is the best indicator to evaluate severity. Liver biopsy with biochemical iron quantification was the historical gold standard to evaluate liver iron concentration (LIC). The pathological threshold was 2 mg Fe/g dry-weight (36 μmol Fe/g). However, liver biopsy was an invasive method with some risks.

Many studies were about having a non-invasive method for LIC evaluation. Nowadays, MRI is the only method in routine clinical use. It is widely usable on all devices and results are sufficient enough for clinical answers: the presence or not of iron overload, quantification, extra-hepatic iron overload.
deposition and treatment follow-up. Today, MRI has become the gold standard for LIC evaluation and biopsies are no more done in this indication [2,3].

**Detection and quantification**

MRI evaluates, in an intense magnetic field, the exponential decrease of liver signal after a stimulation wave. The decrease depends on T1 and T2 relaxation times. Iron atoms have a paramagnetic effect inducing a T1 and T2 decrease of the nearby hydrogen nuclei. But their grouping in clusters of ferritin or hemosiderin crystals increases this effect, which becomes superparamagnetic, leading mainly to a reduction of the T2 of the adjacent protons in the spin-echo sequence. The gradient-echo sequences, which are faster and compatible with a single apnea, are more sensitive to the heterogeneities of the magnetic field. The decrease of the signal then takes place more rapidly according to the parameter T2* whose reduction is accelerated in the presence of iron crystals. It is therefore possible to quantify the hepatic iron concentration from the measured signal [4]. In order to reduce the influence of T1, a small flip angle is used in gradient echo. The shorter the TR, the smaller the flip angle should be.

To quantify this effect and thus the iron concentration, two types of methods have been proposed, either by relaxometry by calculating the relaxation times of T2 in spin-echo [5] or T2* in gradient-echo [6–10], or an algorithm based on the Signal Intensity Ratio (SIR) between the liver and the muscle [11,12]. This magnetic susceptibility increases proportionally with the magnetic field.

**Relaxometry**

T2 and T2* are expressed in milliseconds and their frequency equivalents R2 and R2* are expressed in Hertz or sec⁻¹. The conversion is thus done as R2 = 1000/T2 and R2* = 1000/T2*. Most publications concerning hepatic concentration use R2 or R2* whereas for the analysis of iron overload of the myocardium, references are expressed in T2*.

The decay of the signal is exponential and more important with the longer echo times and in the case of shorter T2 or T2*, according to the formula:

\[ \text{Signal} = S_0 e^{(-TE/T2)} \]

S0 corresponds to the starting liver signal, when TE = 0 ms.

**T2 or R2**

At 1.5 T, St Pierre et al. have proposed quantification based on the calculation of T2 from spin-echo sequences [5]. Their “Ferriscan” method has been FDA approved and is commercially available. It first requires calibration of the instrument and an external reference via a test object. The MRI acquisition sequence is rather long (about 10 minutes) incompatible with apnea. Data must be transferred to the Australian company for quality control and analysis. The validation of the method is based on a series of 105 confrontations demonstrating a curvilinear relationship between the R2 and the LIC [5]. This method is particularly useful in case of high iron overload such as in β-thalassemia disease. In a large international study (ESCALATOR), recruiting 252 heavily iron-overloaded patients treated by an oral chelator, the 95% confidence interval of the Ferriscan™ method was between 71 and 74% [13].

**T2* or R2***

The determination of R2* from gradient echo sequence is more widely used. Advantages are undeniable. First, the acquisition time is very short, all the information being obtained with a single sequence of approximately 15 seconds, therefore in a single apnea. This removes the kinetic blur and greatly improves the quality of the image. Then this type of sequence is more sensitive to the presence of iron and allows a better quantification of low overloads. The corollary is that the quantification of very high overloads can become difficult due to the very fast decrease of the signal, which requires the use of a very short first TE around 1 ms at 1.5 T [2]. One of the initial limitations of this method was that few devices could perform a multi-echo sequence with such a short first echo. Wood’s first publication in 2005 had bypassed the problem by repeating the same sequence with a gradual increase of the TE without recalibrating between sequences [7]. Nowadays, multi-echo 2D gradient echo sequences are proposed by most constructors, either in the basic package or as an option. Data must then be processed by software calculating the T2 or T2*.

T2* calculation can be simply done by the determination of the exponential curve corresponding to the decay of the hepatic signal measured on the increasing echo times. The method used to calculate T2* must also take into account the presence of background noise. The formula can be modified by adding an offset:

\[ \text{Signal} = S_0 e^{(-TE/T2*)} + \text{offset} \]

Usually, this has little involvement in the analysis of the image because the signal-to-noise ratio is high, especially by using a surface coil with lots of sensors. However, in the event of the hepatic signal collapses in case of high iron overload, the measurements obtained from the longest echoes merge with the background noise signal preventing a correct calculation of the T2*. He et al. studied the effects of noise on the T2* signal decay and evaluated different curve fitting models (monoeponential, baseline subtraction, offset, truncation) [14]. At maximum, in case of total collapse, all measurements are at the level of background noise and the value of T2* becomes unreliable or even aberrant. This underlines the importance of using a first echo time as short as possible.

Several studies have compared the R2* to the LIC obtained by biochemical analysis of the biopsy sample. The main ones are detailed in table 1. They validated the existence of a linear
relation between the R2* and the LIC, which simplifies the conversion in comparison with the curvilinear relation observed with the R2. Despite differences in equipment and echo time, the conversion formulas between published series with biochemical confrontation are quite similar (Figure 1).

By pooling the main results Henninger determined an average relationship [10]:

\[
\text{LIC (mg/g)} = R2^* \times 0.02886 + 0.08
\]

Or in \(\mu\text{mol/g}:

\[
\text{LIC (\mu\text{mol/g})} = R2^* \times 0.519 + 1.44
\]

By simplifying, it can therefore be considered, at 1.5 T, that the LIC expressed in \(\mu\text{mol/g} \text{ dry weight} is approximately equal to the \(R2^*\) divided by two. Thus, using this equation, the maximum normal limit of \(R2^*\) would be of the order of \(65 \text{ s}^{-1}\) and the normal minimum limit of \(R2^*\) of 15 ms. The mean of liver \(R2^*\) on healthy controls value was, as expected, below this limit, between 20 and 25 ms with a normal minimum limit of \(R2^*\) of 18 or even 20 ms [15, 16]. However these series have no correlation to biopsy and the he LIC in the normal population is usually low, far from the upper limit, particularly in women.

Two studies showed that the method was transferable between different devices from different centers. Tanner demonstrated that inter-center reproducibility of \(R2^*\) in heart and liver was 5.0% and 7.1%, with mean absolute differences in \(R2^*\) of 1.3 ms and 0.45 ms, respectively [17]. Using the same software to calculate \(R2^*\) all along its study, Kirk found a similar reproducibility, with no significant difference between local and intersite reproducibility [18].

But, variations are probably more related to the methodology of calculation of the \(R2^*\) or to the analyzed population than to the MRI device used. According to the pathologies, the intrahepatic distribution of the overload may be different. The liver may also have a fat overload.

The influence of steatosis was neglected in most studies validating the method because of the recruitment of patients that was essentially haematological, free of fat overload. For haematological recruitment patients and especially in case of suspected DIOS, the steatosis frequency is higher. The problem is that steatosis introduces signal oscillation with a maximum for the in-phase TE, multiples of 4.4 ms at 1.5 T and 2.2 ms at 3 T. The maximum signal of a fatty liver is a little higher than the one of a healthy liver, due to the T1 intensity of fat. Except for in-phase TE, the signal of the liver is lowered and can even collapse in case of strong steatosis for a TE in opposed-phase (2.2 ms at 1.5 T). To avoid errors related to steatosis, a manufacturer proposes to use a multi-echo sequence with saturation of the fat. The other manufacturers offer as an option a dedicated 3D sequence that integrates the simultaneous calculation of the \(T2^*\) or \(R2^*\) but also the fat concentration. The calculation of \(T2^*\) is then made by integrating the cyclic variation of the signal related to the steatosis, the fat concentration becoming then a second unknown. Calculations are done pixel by pixel resulting in maps of fat concentration and \(R2^*\) or \(T2^*\). Results are correct but in the case of a high overloads, obtained results must always

---

**Table 1**

<table>
<thead>
<tr>
<th>Patients (n)</th>
<th>TE range (ms)</th>
<th>R2* - LIC in mg/g</th>
<th>MR-LIC in (\mu\text{mol/g})</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Slope (mg/g)</td>
<td>Intercept (mg/g)</td>
</tr>
<tr>
<td>Anderson [5]</td>
<td>27</td>
<td>2.2-20.1</td>
<td>0.0146</td>
</tr>
<tr>
<td>Wood [6]</td>
<td>22</td>
<td>0.8-4.8</td>
<td>0.0254</td>
</tr>
<tr>
<td>Hankins [7]</td>
<td>43</td>
<td>1.1-17.3</td>
<td>0.028</td>
</tr>
<tr>
<td>Garbowskii [8]</td>
<td>50</td>
<td>0.93-16</td>
<td>0.032</td>
</tr>
<tr>
<td>Henninger [9]</td>
<td>17</td>
<td>0.99-16.5</td>
<td>0.025</td>
</tr>
</tbody>
</table>

---

**Figure 1**

Comparison of \(R2^*\) LIC calibration models based on results published by Anderson [5] LIC = 0.0146(R2*) − 0.27, Wood [6] LIC = 0.0254(R2*) + 0.202, Hankins [7] LIC = 0.028(R2*) − 0.45, Garbowskii [8] LIC = 0.032(R2*) − 0.14 and Henninger [9] LIC = 0.025(R2*) + 0.277.
be correlated to the visual aspect because the complex calculation on small voxels can lead to major errors as in the example figure 2. It should also be noted that, for now, at 1.5 T, there is no study demonstrating the performance of estimating hepatic iron concentration by T2* in a population with mixed overloads. In parallel, it should be known that the quantification of hepatic fat by MRI is no longer possible when the signal decays too fast. At 3 T, the double magnetic susceptibility and the same iron concentration is responsible for reduction of the signal twice as fast [19]. Our work, with comparison with the LIC from biopsies, confirms this hypothesis with a minimum normal limit of T2* at 11 ms [20]. The acquisition protocols must therefore be adapted to 3 T by minimizing the echo times.

At 3 T, we propose this equation [20]:

\[
\text{LIC (\mu mol/g)} = R^2 \times 0.314 - 0.96
\]

The oscillation due to fat is similarly shortened and the first TE in phase is no longer 4.4 ms but at 2.2 ms. In theory, to obtain a correct calculation of the T2* during heavy overloading it would be necessary to obtain a first echo shorter than 0.5 ms which is not technically possible with the sequences proposed by the manufacturers. Thus, it is clearly preferable to use a 3 T MRI to quantify slight overloads as in DIOS and a 1.5 T MRI to evaluate high overloads such as in β-thalassemia. But it is also possible to use in parallel the second “SIR” method which analyzes the liver to muscle ratio which is less limited than the calculation of the T2*, with comparable TE, for the evaluation of the high overloads [20].

**Signal intensity ratio (SIR)**

Here, instead of taking into account, for each echo, the absolute value of the signal in order to calculate its decay rate, the relative value between the signal of the liver and that of the paravertebral muscles will be used for each echo. Indeed, under normal conditions, with a low flip-flop angle reducing T1 effect, their signal is quite close, the muscle being a little more intense than the liver. From these ratios, the concentration of iron can be obtained by two different algorithms but with similar acquisition rules [11,12]. First of all, the most important, is the use a body coil and not to a multi-channel surface coil, because the signal would then decrease from the surface to the depth, thus prohibiting a signal comparison at different depths. Next, a multi-echo 2D gradient echo sequence with a 120 ms TR, a 20° flip angle and a dozen multiple echoes of 1.2 ms must be used to have both opposed and in-phase echoes. Examples for each type of device are available at [http://www.mrquantif.org](http://www.mrquantif.org). If this sequence is not available on the used device, it can be replaced by several sequences with one or two echoes. The required echo times depend on the magnetic field and the calculation algorithm (table II). The acquisition time of each sequence is 15 seconds, compatible with apnea. Measurements are then made by placing on an image, regions of interest on the most homogeneous areas of the liver and paravertebral muscles, avoiding all artefacts, such as the decrease of muscle signal on the upper sections, close to the pleura (figure 3). At 1.5 T, we proposed to use at least 3 echo times of 4, 9 and 14 ms. From a series of 174 patients the comparison of the different SIR to the LIC obtained from the biopsy fragment

**Figure 2**

Genetic hemochromatosis with high iron overload. a. Major decrease of the liver signal intensity including on the shortest TE. T2* is dropped off. b. Miscalculation of T2* on the pixel wise map reconstructed by the MR software. T2* calculation can be done by an external software or by the SIR method

**Table II**

<table>
<thead>
<tr>
<th>Echo times proposed for SIR method at 1.5 and 3 T</th>
</tr>
</thead>
<tbody>
<tr>
<td>TE (ms)</td>
</tr>
<tr>
<td>-------------------</td>
</tr>
<tr>
<td>Echo 1</td>
</tr>
<tr>
<td>Echo 2</td>
</tr>
<tr>
<td>Echo 3</td>
</tr>
<tr>
<td>Echo 4</td>
</tr>
</tbody>
</table>

© 2019 Elsevier Masson SAS. All rights reserved. - Document downloaded on 23/07/2019 It is forbidden and illegal to distribute this document.
allowed us to establish an algorithm [11]. It allows to estimate the overloads to a maximum of 350 μmol/g. Beyond that, the signal of the liver was collapsed including on the shortest TE at 4 ms. The online calculation based on this algorithm is always possible on the website https://imagemed.univ-rennes1.fr/mrquantif but we propose since then to evolve towards a combined T2* and SIR quantification using the MRQuantif free-ware (see infra).

Another study found the same results and proposed to add a shorter TE of 1.8 ms to quantify overloads greater than 350 μmol/g using the formula available on the website http://oernst.f5lvg.free.fr/liver/iron .html [21]:

\[
\text{LIC} (\mu \text{mol/g}) = 937 - (5.37 \times \text{SIR})
\]

On a series of 112 patients, Alustiza confirmed the results of this SIR method but using only two echos, one with a TE of 4 ms producing an intermediate weighting (IW) and the second with a TE of 14 ms producing a T2 weighting (T2). The calculation of the liver iron concentration based on SIR obtained can be done with the formula:

\[
\text{HIC} = e^{-[S \times 0.888 \times \text{SIR}_{T2} - (1.518 \times \text{SIR}_{TE})]}
\]

Comparing the two methods on a series of 171 patients with biochemical confrontation Castiella observed an overestimation with our algorithm for low and medium values, but the 9 msec TE, that is important for these levels of overloads, was not realized on the large majority of patients [22]. The inter-device variability is around 10% [23]. In general, the SIR method is probably less accurate for low overloads than the T2* method, when the data for T2* calculation are of good quality. On the other hand, SIR method is probably easier to implement and its accuracy is quite sufficient in clinical routine, provided that no coil error is made.

Indeed, the risk of major error of the SIR method is to use, alone or in association with the body coil, a surface coil integrated into the bed of the device. This gives a signal decreasing from the back to the front resulting to liver to muscle ratio falsely lowered as in the example figure 4. To reduce the risk of error we recommend that you no longer use the online calculation but to download the DICOM software available at http://www.mrquantif.org. This software checks the acquisition parameters, in particular the selected coil. It also allows the comparison of LIC by the two previously described methods, but also provides, in case of a multi-echo sequence, the R2* calculation and the LIC based on. Hepatic fat can also be evaluated according to different methods, which can become very important in the case of DIOS. At 3 T, we validated the SIR method on the basis of echo times divided by two (see table) [24]. This method allows quantification of all levels of overload. Indeed, the very fast decay of the
Extra-hepatic overload
In case of iron overload, hepatic involvement is constant but depending on the mechanisms and level of involvement, there may be an overload of other organs such as the spleen, pancreas or heart. It may be necessary to investigate and quantify these other disorders in order to guide the etiologic diagnosis and, in case of cardiac overload, prevent a major complication such as heart failure.

Splenic iron overload
There is no iron overload in spleen in HFE hemochromatosis because hepcidin deficiency leads to an increase in intestinal absorption with iron storage in the hepatocytes. However, a reduction of the splenic signal can be demonstrated in the rare genetic hemochromatosis type IV, which is accompanied by a deficiency in ferroportin and especially in hemosiderosis linked to hemolytic diseases with repeated transfusions (β-thalassemia, sickle cell anemia) (figure 5). To a lesser degree there is also sometimes a slight overload of splenic iron in DIOS. Obviously the quantification of iron in the spleen can only be done by analogy to the liver because there is no study that allows to calibrate the splenic MRI results with biochemical assays. However, as in normal patients the signal of the spleen is similar to that of the liver on the sequences used for the quantification of iron and that there is no reason for the excess of iron to give a different signal decay, transposition of hepatic methods is therefore acceptable.

Pancreatic iron overload
It is observed only in major and prolonged forms of iron overload. In genetic hemochromatosis it is therefore understood that its presence has been linked to the existence of cirrhosis with

Figure 4
Two examples of coil selection errors: a: in the first case, only the coil integrated in the MR system bed was selected. Then, error was obvious: the signal of the paraspinous muscle is clearly stronger with a wrong 0.1 liver to muscle ratio; b: in the second example, the usual phased-array surface coil was selected instead the body coil with a relatively slighter increase of the paraspinous muscle signal. This error was more difficult to detect visually. That justify the use of an automatic analysis of acquisition parameters

signal in the presence of iron affects the high overloads quantifications by T2* calculation. Several echoes with a signal over the background noise are needed to get a reliable calculation. With the SIR method, the first echo, in opposed-phase, can be used alone to estimate the iron concentration, as proposed at 1.5 T [21]. Of course it is necessary to use this echo only when the others do not give signal any more, in other not to mistake a steatosis. The algorithm we have defined is integrated in the software mentioned above.

Figure 5
Hepatic and splenic iron overload due to repeated transfusions. Decrease of the signal of the liver and the spleen in comparison with the paravertebral muscles
portal hypertension [25]. However, it can also occur in β-thalassemia, especially after splenectomy [26,27]. It is correlated with the presence of diabetes mellitus. There was a slight correlation between the an excess of iron in the pancreas and in the myocardium but without a clear correlation with the level of hepatic overload [28]. The analysis of the pancreas must therefore be systematic in high overloads and quantification by analogy to the liver can also be done for this organ.

**Myocardial iron overload**

Myocardial overload is a serious complication because it can lead to potentially lethal heart failure. In genetic hemochromatosis it is described in major and extended forms of iron overload. Early detection and efficacy of depletion by bleeding considerably limit the frequency. In β-thalassemias it appears when hepatic iron overload reaches a threshold of 350 μmol/g [29]. However, the correlation with mainly in patients treated with chelation [30-33]. The kinetics of capturing and releasing myocardial iron, which is slower than that of hepatocyte iron, may explain these discrepancies.

The diagnosis and quantification of myocardial overload is also based on MRI with gradient echo sequences allowing the calculation of the T2* of the myocardium. The protocol is based on a single, multiecho gradient echo, midventricular short-axis slice acquired in presystole or late diastole to minimize artifacts from blood flow and motion. A T2* multiecho black-blood sequence obtained after a double inversion recovery preparation pulse is preferable [34] (Figure 6).

Anderson found that a T2* less than 20 ms at 1.5 T detected all patients with impaired left ventricular function [6]. To optimize the calculation of the T2*, it is therefore necessary to acquire TE longer than those defined for the liver, for example between 2 and 20 ms.

A comparison between MRI data obtained in patients with heart failure and post-mortem biochemical analysis revealed a correlation between R2* and myocardial iron concentration (MIC) [35]:

$$\text{MIC} = 0.00958 \times (R2^*)^{-1.22}$$

However, in practice, unlike the liver, T2* is used as a biomarker for myocardium without translating it into iron concentration. It has been demonstrated that the 1-year risk of development of heart failure increase with T2* reduction and is 14% for T2* between 8 and 10 milliseconds at 1.5 T, 30% for T2* between 6 and 8 milliseconds, and 50% for T2* below 6 milliseconds [36]. At 3 T the observed T2* limit values are divided by 2 [19].

**Follow-up under treatment**

Although the reproducibility is good [17,18,23,37], one should not think too quickly the quantification of iron by MRI is completely standardized. For the same method as the calculation of T2* we saw that there were different formulas with different results. And even if exactly the same formulas used the conversion of the R2* to the LIC can be done with a different reference.

One of the major challenges is therefore to reduce the variability of all these factors. One possible way to standardize the measurement would be to apply the same sequence to all devices and to use identical calculation software, independent of the manufacturer, such as the commercial software CMR Tools or the one we offer online.

Pending this standardization it is strongly recommended that, if a patient or group of patients is to be monitored, the same radiology structure should be used and the same method should be used for each patient. We can thus judge the progressive decrease of the iron overload.

The repetition of the examinations is not useful for the follow-up of the bleedings because the impact of this treatment is totally known. Thus, in genetic hemochromatosis, follow-up by ferritinemia...
is often sufficient, except in the case of poor tolerance of bleeding in discordance with the supposed overload. On the other hand, for patients treated with chelator, the effectiveness of the treatment is not systematic and may have variations from one patient to another. This is why MRI monitoring is warranted.

**Conclusion**

The quantification of the iron overload must be done by a gradient echo sequence with about ten echoes, with very short first echoes, preferably in a body antenna, in order to also have the sIR method, according to a protocol transposable on all devices (https://imagemed.univ-rennes1.fr/en/mrquantif).

It is thus possible to obtain the calculation of the concentration of hepatic iron with the various methods, favoring the calculation of the T2* controlled by the sIR method in case of very high overload at 3 T.

Steatosis, common in DIOs, can also be estimated jointly, as well as other potential pancreatic and splenic overloads.

A similar protocol, with a cardiac antenna, preferentially choosing a black blood sequence, makes it possible to look for a myocardial overload.

**Disclosure of interest:** the authors declare that they have no competing interest.

## References


