Methodologies and tools used today for measuring iron load

Evangelos Alexiou

CT-MRI Department, General Hospital of Larissa, Larissa, Greece

Abstract

Iron overload is a matter of an extreme clinical importance, in the overall management of Thalassaemia. Magnetic Resonance Imaging (MRI), has evolved in a novel tool for iron quantification during the last decade and it is considered as a validated, accurate and noninvasive method with worldwide distribution. The MRI scanner exploits the intrinsic magnetic properties of the hydrogen nuclei in order to discriminate the tissue characteristics. The presence of iron in a tissue causes a faster dephasing of the protons and a reduction in T2* and T2. R2 and R2* represent the reciprocal of T2 and T2*. In order to measure the signal intensity and quantify iron concentration the Gradient Echo (GRE) T2* and the Spin Echo (SE) T2 sequence are used. There are two broad groups of techniques to quantify the iron. The signal intensity ratio (SIR) methods and the relaxometry methods. The later are sub grouped in the R2 (T2) relaxometry methods with the predominant of this category being the FerriScan® and the R2* (T2*) methods. CMR Gradient Echo T2* pulse sequence is the preferred technique for the quantification of iron in the heart. The R2 and R2* methodologies are both very accurate in predicting the true LIC with high levels of sensitivity and specificity in the range of clinically important LIC thresholds and can be both used over a wide clinical range, individually.

Introduction

Iron overload is a matter of an extreme clinical importance, in the overall management of Thalassaemia. Due to its toxicity in many tissues and in order to provide the appropriate chelation therapy and thus avoid severe clinical complications, close monitoring is vital for the thalassaemic patients.

Magnetic Resonance Imaging (MRI), has evolved in a novel tool for iron quantification during the last decade and it is considered as a validated, accurate and noninvasive method with worldwide distribution.

Many argue that the MRI scanner is one of the most complicated devices in the world and the cooperation of a whole team of scientists, such as radiologists, medical physicists, radiology technicians, is mandatory to exploit all of its useful benefits.

Introduction to methodology

One of the most basic and fundamental topics in MRI imaging is that the hydrogen nuclei have intrinsic magnetic properties and we can say that they behave like tiny magnets. Hydrogen is a very abundant element in our body and we can find it mainly in the form of water as 70% of our body consists of water and secondarily as fat combined with carbon and oxygen. The MRI scanner exploits magnetic properties in order to discriminate the tissue characteristics.

Under the earth’s magnetic field, the hydrogen nuclei are orientated randomly in our body canceling each other out and producing no magnetic field. When the patient is placed into the scanner, which has a strong magnetic field, usually from 1.5 to 3 Tesla, known as B0, the hydrogen protons are forced to align towards the direction of the applied field. That alignment produces a second magnetic field parallel to B0 known as B1 or net magnetization.

By disturbing the above equilibrium state with the application of a radiofrequency pulse (RF), energy is transmitted to the nuclei exciting them at a higher-energy level and forcing them to align perpendicular to B0 usually at 90 degrees, an effect called “resonance”. Immediately after the pulse stops, the nuclei start to return to the previous low-energy state, a procedure that is called “relaxation”. The relaxation has two components: the T1 or longitudinal relaxation which is the time needed for spinning protons to re-align with the external magnetic field and the T2 or transverse relaxation which is the time needed for the loss of phase coherence among spins due to interactions between the spins. In practice there is a T1 recovery and a T2 decay, as during the relaxation procedure the longitudinal magnetization increases and the transverse relaxation decrease. These procedures occur simultaneously. Finally, at a chosen time, the energy is collected back in the form of a signal with an RF coil. Because the T1 and the T2 time constants are different for each tissue they are a powerful contrast mechanism.

There are two types of T2 Decay. The first one is the T2* Relaxation or Free Induction Decay (FID) which occurs due to the interaction between the neighboring protons and due to inhomogeneities of the magnetic field the latter of which can be described by a constant called T2’ prime. The second component, the T2, can be measured after the application of specific techniques that eliminate the T2 prime contribution to the signal and it is an extremely powerful contrast mechanism as it occurs in different degrees in different tissues.

The presence of iron in a tissue causes a faster dephasing of the protons and a reduction in T2* and T2. The signal that is received by the coils is weaker and the signal intensity (SI) reduction is propor-
Liver iron quantification

The liver iron concentration (LIC or HIC) is considered as the best marker for the estimation of total body iron stores as it accounts for more than 70% of it. The accurate estimation of LIC faces endogenous difficulties, due to facts that have to do with the heterogeneous distribution of iron, the heavy burden of the iron load, the presence of hepatitis and fibrosis and errors in biopsy sampling.

There are two broad groups of techniques to quantify the iron. The signal intensity ratio (SIR) methods and the relaxometry methods. The later are sub grouped in the R2 (T2) relaxometry methods with the predominant of this category being the FerriScan® and the R2* (T2*) methods.

In the SIR method described in the paper of Ganton et al. (2004), a single image is acquired with the use of five Gradient Recalled pulse sequences at a particular echo time. Following this, the SI in a liver region of interest (ROI) and a non-iron overloaded tissue, such as paraspinal muscle, are measured. The liver SI is then compared with the SI of the paraspinal muscle and the ratio between the two signals can be used, with the help of a calibration curve formed from 174 liver biopsies, to measure the iron concentration in the liver. The heavier the iron load is in the liver the smaller the signal ratio will be. The use of the above method is limited to LIC between 60-375 µmol/g, limiting its wide use in the clinical field.

The most widely known and commercial R2 relaxometry method is the FerriScan®, which has international regulatory clearances (USA, Europe, Canada, Australasia). The images are generated with the use of a Spin Echo T2 pulse sequence and are obtained locally from an MRI scanner where the appropriate software is installed and the data analysis is being conducted centrally (ResonanceHealth, Australia). In general, the maximal liver cross section slice is selected from the multi-slice image sequence. The signal intensity as well as the R2 value is calculated for each voxel. In this way, a liver iron map is generated where light areas indicate high iron concentrations, while dark areas show low iron concentrations. Furthermore, a histogram with the R2 value distribution is generated and the mean R2 value is considered as the liver R2 value. The first calibration curve was published in 2005 by T. G. St Pierre et al. (2005). Biopsy LIC values were measured for 105 patients with hepatitis C, hereditary hemochromatosis, b-thalassaemia/haemoglobinopathy-H disease and thalassaemia major. The patients were scanned with the R2 relaxometry method and the R2 values were calculated in the right lobe of the liver near the biopsy site. Consequently, the values were plotted against the liver iron concentration as measured by a liver biopsy. A highly significant correlation was found between biopsy LIC and liver R2 measurements for the region of interest in the right lobe of the liver. Although a general increase in R2 variability with increasing biopsy LIC was noticed, for the clinically important LIC thresholds of up to 15.0 mg Fe/g dry tissue, high levels of sensitivity and specificity were demonstrated. In 2005, Wood et al. (2005) plotted the HIC estimated by a biopsy and R2*against the R2 relaxometry and reproduced with consistency the calibration curve published by St. Pierre, proving the robustness of the method.

The R2* relaxometry methods make use of a Gradient Echo T2* sequence and in general are faster and easier to perform. The image acquisition is being performed locally by the majority of the 1.5T MRI scanners and the data analysis is being conducted by the use of commercial or free software, locally. Initially, a single echo sequence was used due to hardware limitations, but the modern scanners are able to capture complete R2* information for multiple slices in a single breath-hold sequence. There are two commonly used approaches to LIC quantification using MR R2* relaxometry. The first is the pixel-wise method (R2* mapping) where the SI is calculated for each pixel, an R2* map is produced and finally averaged the mapped R2* in the ROI to produce the final R2* value. The second method is to average the signal in the ROI of a large area in the right liver lobe excluding large vessels. The first T2* results, were published by Anderson et al. (2001) and a near-linear relationship between liver T2* and biopsy iron was demonstrated in 24 patients, despite the hardware limitations of a long first Echo Time unsuitable for LIC measurements and a non-constant TR which increased the T1 effect. Wood et al. (2005) in a following study in hepatitis C negative patients, demonstrated a stronger linear relationship between liver R2* and liver iron. In the same paper, using a combination of iron estimates by R2* and by liver biopsy, the authors were able to closely replicate the St Pierre calibration curve. Despite significant differences in signal acquisition and processing, an internal consistency among the R2 and R2* techniques was hypothesized. Hankins et al. (2009) plotted the data obtained from liver biopsies of 47 iron overloaded patients versus R2*-MRI measurements and the calibration equation produced, was in excellent agreement with the one reported by Wood et al. but with a different slope compared to that of Anderson et al.

The R2* relaxometry, despite not being fully standardized with external quality control, performs greatly in terms of interscanner reproducibility. A mismatch of R2* values in higher levels of iron burden has been reported. It is observed that the relaxation decay is decomposed of two components: a fast decay component and a slow one. This phenomenon is mild in low and moderate iron concentrations but is quite obvious in heavily iron overloaded tissues such as the liver. In the later case, the signal decays very fast at early TEs due to the presence of iron and later on, a “plateau” is produced in the slope from iron-poor tissues. Two solutions have been proposed to deal with the issue and improve the
R2* measurements. The first one is the truncation model, in which the later points that form the plateau, are removed and the slope is fitted in the remaining points. The second one is the offset-correction, in which a constant is added to represent the slowly decaying components. The offset model tends to overestimate R2* due to its inaccurate compensation mechanism. The truncation model generally produces accurate R2* measurements by removing the major source of errors, but underestimates very high R2* values due to the limited acquisitions of low TE images. He et al. (2008), compared the two models in vivo and found better agreement and more reproducible results with the truncation technique. In general the truncation model is considered more reproducible in both heart and liver R2* measurements.

Cardiac iron quantification

The iron quantification in the heart, faces some endogenous difficulties. The heart is a fast moving organ and the acquisition procedure must be very quick. The blood flow, the air in the lungs and the intestine and the coexistence of heavily liver iron overload are all potential factors of artifacts. Moreover, the formation of a calibration curve is very difficult due to ethical reasons. CMR Gradient Echo T2* pulse sequence is the preferred technique rather than a spin-echo T2 sequence because it is very fast and has a greater sensitivity to iron deposition and lower sensitivity to motion. The first described method by Anderson et al. (2001) due to hardware limitations, multiple separate image acquisitions were performed, each of which required a breath hold. This was time consuming and prone to artifacts that made it difficult to assess the exact myocardial borders with longer echo times, and it created problems with image registration between the images. In modern scanners, a multi-echo sequence is possible, allowing the acquisition of a single short-axis mid-ventricular slice at multiple echo times in a single breath hold and has also the advantage of minimizing the T1 effect because of the constant repetition time. The most recent technical improvement in the T2* sequence has been the development of the black-blood sequence. This sequence greatly reduces the blood signal, which significantly reduces the blood artifact propagating in the phase-encoding direction that typically spreads across the interventricular septum and reduces measurement variability. A multi-slice, multi-echo T2* sequence has been published for the global and segmental measurement of cardiac iron overload with reported multi-center validation, but its clinical advantage of this more complex protocol remains to be demonstrated.

The analysis of the results can be done locally, as some MR vendors have included specific software in their workstations for the T2* calculation. Alternatively, commercial applications are available (CMR Tools Thalassaemia Tool Kit, London-UK) or even cost free options such as Microsoft Excel spreadsheet.

The iron calibration and distribution in human heart was published in 2011 by Carpenter et al. (2011). Twelve postmortem heart of transfusion-dependent patients were scanned by CMR and R2* measurements were related to tissue iron concentration. The calibration equation was found to be: Fe=45•(T2*)-1.22 (Fe in mg/g d.w.). Despite that iron levels were enhanced when the two methods are compared to one another. The latest developments that have been achieved during the recent years. The American Heart Association (2013), released a consensus statement about the cardiovascular function and treatment in beta-Thalassaemia major. Concerning the CMR measurements of cardiac iron, the preferred method was stated to be the Gradient-Echo T2* instead of Spin-Echo T2 due to its increased sensitivity to the presence of iron and its decreased sensitivity to motion artifacts. The proposed protocol should include a single breathhold, multiecho, black-blood T2* sequence, with a gating delay of 0 ms after the R wave in order to acquire a single short-axis mid-ventricular slice at multiple Echo Times. In this way, the exam is time effective with reduced blood, motion and registration artifacts together with minimum T1 effect given the constant repetition time (TR). S.T. Pierre et al. (2013), published a paper in which the originally described calibration equation for the R2-Ferriscan® method, was recalibrated using spin-density projection-assisted (SDPA) R2-MRI measurements and liver biopsy samples that were obtained from a subgroup of 233 b-thalassaemia major patients of the ESCALATOR trial. The study, also investigated the effect of scanner type, fibrosis stage, necroinflammation grade, age and chelation therapy. The results suggested that the initial calibration curve, which had already been proved to be applicable in a range of iron-loaded patients, is applicable across all stages of liver fibrosis and is unaffected by the use of different types of 1.5T MRI scanners, the necroinflammation grade, the patient age and the iron chelation therapy with deferasirox.

Carpenter et al. (2014), published the calibration curve of myocardial T1 and T2 relaxation times against iron concentration, by examining twelve postmortem hearts from transfusion-dependent patients. Although the limitation of the effect of formalin fixation, myocardial T2 correlated well with tissue iron and the logarithmic correlation (lnR2 and lnFe) was even stronger giving promises that the method could serve additional information to T2* for patients with cardiac siderosis. The results from T1 measurements were not reliable, possibly due to the greater impact of formalin.

Wood et al. (2014) published a longitudinal study comparing R2* and R2-Ferriscan® techniques. The study indicated that well-performed and logarithmic transformed LIC-R2* and LIC-R2 measurements, individually, yield comparable longitudinal performance suitable for clinical trials and for clinical practice. Though, the methods cannot be interchanged with one another in the same patient due to discrepancies that are related to patient-specific differences in iron distribution and which are enhanced when the two methods are compared to one another.

Garbowski et al. (2014), published a paper where an optimized T2* pulse sequence was used in order to recalibrate the T2* method against liver iron concentration and 50 liver biopsy samples were used on 25 with haemosiderosis. Furthermore, the novel T2*-LIC was compared with R2-LIC Ferriscan. The results revealed a strong linear correlation between lnT2* and lnLIC and the relationship between the original proof of concept T2* method was corrected by 2.2 times. On the other hand, there was a poor agreement between T2*-LIC and R2-LIC, possibly due to different sensitivity of the two methods, to iron distribution or to patient-specific differences.

Latest developments

During the last fourteen years, a strenuous scientific work has been done in the field of non-invasive iron quantification by the use of MRI; and it is a ongoing procedure.

Instead of conclusions, it would be of great interest to mention the latest developments that have been achieved during the recent years.

The American Heart Association (2013), released a consensus statement about the cardiovascular function and treatment in b-Thalassaemia major. Concerning the CMR measurements of cardiac iron, the preferred method was stated to be the Gradient-Echo T2* instead of Spin-Echo T2 due to its increased sensitivity to the presence of iron and its decreased sensitivity to motion artifacts. The proposed protocol should include a single breathhold, multiecho, black-blood T2* sequence, with a gating delay of 0 ms after the R wave in order to acquire a single short-axis mid-ventricular slice at multiple Echo Times. In this way, the exam is time effective with reduced blood, motion and registration artifacts together with minimum T1 effect given the constant repetition time (TR). S.T. Pierre et al. (2013), published a paper in which the originally described calibration equation for the R2-Ferriscan® method, was recalibrated using spin-density projection-assisted (SDPA) R2-MRI measurements and liver biopsy samples that were obtained from a subgroup of 233 b-thalassaemia major patients of the ESCALATOR trial. The study, also investigated the effect of scanner type, fibrosis stage, necroinflammation grade, age and chelation therapy. The results suggested that the initial calibration curve, which had already been proved to be applicable in a range of iron-loaded patients, is applicable across all stages of liver fibrosis and is unaffected by the use of different types of 1.5T MRI scanners, the necroinflammation grade, the patient age and the iron chelation therapy with deferasirox.

Carpenter et al. (2014), published the calibration curve of myocardial T1 and T2 relaxation times against iron concentration, by examining twelve postmortem hearts from transfusion-dependent patients. Although the limitation of the effect of formalin fixation, myocardial T2 correlated well with tissue iron and the logarithmic correlation (lnR2 and lnFe) was even stronger giving promises that the method could serve additional information to T2* for patients with cardiac siderosis. The results from T1 measurements were not reliable, possibly due to the greater impact of formalin.

Wood et al. (2014) published a longitudinal study comparing R2* and R2-Ferriscan® techniques. The study indicated that well-performed and logarithmic transformed LIC-R2* and LIC-R2 measurements, individually, yield comparable longitudinal performance suitable for clinical trials and for clinical practice. Though, the methods cannot be interchanged with one another in the same patient due to discrepancies that are related to patient-specific differences in iron distribution and which are enhanced when the two methods are compared to one another.

Garbowski et al. (2014), published a paper where an optimized T2* pulse sequence was used in order to recalibrate the T2* method against liver iron concentration and 50 liver biopsy samples were used on 25 with haemosiderosis. Furthermore, the novel T2*-LIC was compared with R2-LIC Ferriscan. The results revealed a strong linear correlation between lnT2* and lnLIC and the relationship between the original proof of concept T2* method was corrected by 2.2 times. On the other hand, there was a poor agreement between T2*-LIC and R2-LIC, possibly due to different sensitivity of the two methods, to iron distribution or to patient-specific differences.
Technical considerations

The use of the above described techniques is quoted as “very precise”, by a variety of papers in the medical literature. Though, there are some technical issues that need to be taken into account.

Both the R2 and R2* methodologies face reduced sensitivity and specificity for the prediction of biopsy LIC at higher LICs. This is obvious as a broadening of the scatter in the curves above the level of 15 mg Fe/gr dw for R2 (Ferriscan®) an the level of 10 mg Fe/gr dw for R2*.

When the two techniques are compared to one another, there is an increased scatter in their data above the limit of 10 mg Fe/gr dw limiting their overall cross-sectional agreement up to ±50% in a way that cannot be used interchangeably at any given time-point.

Another observation is that there are differences in the slopes of the published calibration curves of the R2* method.

The proposed explanations for the above can be categorized to calibration-related, biopsy-related and patient-related issues.

Concerning the calibration related issues, there are differences in the use of correction models to compensate for the bi-component relaxation decay consisting of fast and slow protons and also the correction of the background noise. The R2-Ferriscan uses a bi-exponential function to model the signal decay in each voxel and uses a specific phantom to measure the noise while the R2* technique proposes a monoexponential function and make use of correction models such as the truncation and the offset model.

Concerning the biopsy related issues, it has been reported that the coefficient variation in biopsy LIC measurements varies from 19% for disease-free liver to greater than 40% with the presence of increased iron loading, fibrosis and cirrhosis. Other points to be mentioned are the specimen size and specimen treatment (formalin and paraffin use), laboratorv variabilities in the ratio estimation of wet-to-dry tissue and the liver quantification.

Concerning the patient related issues, it is proposed that R2 and R2* methods are in different ways sensitive to noise, particle size, pools of iron and patient-specific patterns of iron deposition in the liver which cannot be corrected neither with pre-processing nor with post-processing methodologies, as if each patient has their own calibration curve.

Conclusions

From all the aforementioned, the MRI is a non-invasive, safe and accurate tool for the monitoring of iron overload.

The T2* technique is considered as the gold standard for the cardiac iron measurement.

Concerning the liver, the commercial R2 (Ferriscan®) is the only method that has clearances in USA, Europe, Canada and Australia, employs external quality assessment and validation, but also has an additional cost per scan for the iron concentration analysis.

R2* has been proven to be a fast and accurate technique which can be performed locally, may well combine liver and cardiac measurements and results can be extracted on-site.

In an attempt to answer the clinical question of which of the two methods should be used for the liver iron quantification, both logistical and scientific considerations should be incorporated. The R2 and R2* methodologies are both very accurate in predicting the true LIC with high levels of sensitivity and specificity in the range of clinically important LIC thresholds and can be both used over a wide clinical range, individually. Concerning the logistics, the total cost, the expected range of iron burden, the availability of well trained personnel and quality control and the potential one-stop measurement (both liver and cardiac) should be taken into account.

References


