Magnetic resonance imaging measurement of iron overload

John C. Wood
Divisions of Pediatric Cardiology and Radiology, Children’s Hospital Los Angeles, Los Angeles, California, USA

Abstract

Purpose of review—To highlight recent advances in magnetic resonance imaging estimation of somatic iron overload. This review will discuss the need and principles of magnetic resonance imaging-based iron measurements, the validation of liver and cardiac iron measurements, and the key institutional requirements for implementation.

Recent findings—Magnetic resonance imaging assessment of liver and cardiac iron has achieved critical levels of availability, utility, and validity to serve as the primary endpoint of clinical trials. Calibration curves for the magnetic resonance imaging parameters R2 and R2* (or their reciprocals, T2 and T2*) have been developed for the liver and the heart. Interscanner variability for these techniques has proven to be on the order of 5–7%.

Summary—Magnetic resonance imaging assessment of tissue iron is becoming increasingly important in the management of transfusional iron load because it is noninvasive, relatively widely available and offers a window into presymptomatic organ dysfunction. The techniques are highly reproducible within and across machines and have been chemically validated in the liver and the heart. These techniques will become the standard of care as industry begins to support the acquisition and postprocessing software.

Keywords
heart; iron overload; liver; magnetic resonance imaging; thalassemia

Introduction

Significant progress in magnetic resonance imaging (MRI) quantitation has occurred since a National Institute of Diabetes and Digestive and Kidney Diseases (NIDDK) workshop in 1991, but the consensus statement from that meeting serves as a good reference for prior work [1]. The present article focuses on the use of MRI to quantitate organ iron burden in systemic iron overload disorders. Supporting validation studies for liver and cardiac iron calibration are presented as well as examples demonstrating integration of the new techniques into clinical practice.

Utility of liver iron estimates

Patients with thalassemia major and other transfusion-dependent anemias receive roughly 0.4 mg/kg/day of heme iron, nearly 50 times the physiologic rate of iron absorption [2,3]. Without aggressive iron chelation therapy, these patients die from endocrine and cardiac complications in the second decade of life [4]. Chelation therapy is life saving, but requires close monitoring of iron balance [5]. Trends in serum ferritin are useful in tracking chelator responsiveness, are
relatively inexpensive and are widely available [6–8]. Ferritin values, however, can be confounded by inflammatory state and may give wildly inaccurate estimates of total body iron in selected patients [9,10].

As a result, some thalassemia and sickle cell disease centers have used liver iron concentration by biopsy as their gold standard for chelation titration [11,12]. The liver is the dominant iron storage organ and liver iron concentration correlates closely with the total iron balance [13, 14]. Elevated liver iron also prospectively predicts poor endocrine and cardiovascular outcomes in patients with thalassemia major [14,15].

Unfortunately, liver biopsy is invasive, expensive and subject to sampling error [16–18]. Although the risk of ultrasound-guided liver biopsy is relatively low, hemorrhage requiring prolonged hospitalization occurs in approximately 0.5% of cases [19,20]. Proper sedation and local anesthesia can minimize the discomfort, however postprocedure pain decreases patient acceptance, increasing the ‘effective’ interval between studies. Lastly, the reproducibility of liver iron measurements is poor in diseased livers, limiting its utility for serial measurement [21,22].

To counter the shortcomings of liver biopsy, a number of noninvasive techniques have been applied to liver iron estimation including the superconducting quantum interference device (SQUID) [23,24], quantitative computed tomography (qCT) [25–27], and MRI [28–30]. MRI has clearly emerged as the dominant technique because of its sensitivity, reproducibility [31, 32], availability [33•], and ability to image multiple organs in the body during a single imaging session [34,35]. MRI has been integrated into the standard of care at centers where it is available and it is providing new windows into the pathophysiology of iron overload [36].

### Principles of magnetic resonance imaging-based iron measurements

MRI operates like many imaging modalities in that it transmits a signal into the body and creates an image from the signal returning from the body after it has interacted with the microenvironment. With MRI, the transmitted signal is a microwave which excites water protons in the body to higher magnetic energy states. As these water protons relax back to the unexcited state, they emit microwaves that are received and interpreted by the scanner; these waves reflect the magnetic milieu near the protons. In noniron overloaded tissues, the magnetic environment is fairly magnetically homogeneous. This means that the signals received from different areas in the tissue remain coherent with one another and the signals last for a long duration (bright images without much contrast). Iron deposits, however, act like little magnets when placed in a strong magnetic field; protons diffusing along different paths experience wildly different magnetic profiles, disrupting coherence among the protons and darkening the image more quickly [37,38].

This process is illustrated in Fig. 1. MRI has the ability to ‘refocus’ the radio waves from the tissues at specific time intervals known as echo times. The longer the echo time (moving from left to right in Fig. 1), the more discordant the proton signals become and the darker the image. Iron overloaded liver (Fig. 1, top panels) simply darkens more quickly with echo time. In fact, this darkening process behaves similarly to radioactive material and can be described by a ‘half life’. The MRI scanner can refocus the returning signal either using a special radiofrequency pulse (forming a so-called spin echo), or by using special small magnets known as gradients (forming a so-called gradient echo). The time constant for a spin echo is known as T2 and for a gradient echo is known as T2*. The greater the tissue iron, the shorter the signal half lives, and the smaller the T2 and T2* become.
Some investigators prefer to report rates of signal decay, $R_2$ or $R_2^*$, instead of the half lives $T_2$ or $T_2^*$. These distinctions are purely computational and not related to the imaging itself. These rates of signal decay are simply the reciprocals of $T_2$ and $T_2^*$:

$$R_2 = \frac{1000}{T_2}$$

(1)

$$R_2^* = \frac{1000}{T_2^*}$$

(2)

The factor of 1000 is used because $T_2$ and $T_2^*$ are usually reported in ms and the units of $R_2$ and $R_2^*$ are Hertz or s$^{-1}$. So a $T_2^*$ of 20 ms is equal to an $R_2^*$ of 50 Hz and vice versa. The advantage of $R_2$ and $R_2^*$ notation is that these parameters are directly proportional to iron, rather than inversely proportional to iron; results in the liver are typically reported as $R_2$ and $R_2^*$ values [28,30] whereas $T_2$ and $T_2^*$ reporting is more common in the heart [33•,34,35].

**Validation of liver iron measurements**

Liver $R_2$ and $R_2^*$ measurements were the first to be calibrated to tissue iron because liver biopsy was routinely obtained in clinical practice. In a study of over 100 patients, St Pierre et al. [30] demonstrated a curvilinear relationship between liver iron estimated by $R_2$ and by biopsy (Fig. 2), having a correlation coefficient of 0.98. The prediction error was comparable to the intrinsic variability of liver biopsy. This work followed a number of supporting manuscripts validating the methodology [39–43]. The first $R_2^*$ results, published by Anderson et al. [34], demonstrated a near-linear relationship between liver $R_2^*$ and biopsy iron in 24 patients. Unfortunately, the study was confounded by a high incidence of liver fibrosis which dramatically increased liver iron concentration variability. Our laboratory followed with a study in hepatitis C negative patients that demonstrated a stronger linear relationship between liver $R_2^*$ and liver iron [Fig. 3(a)] [28]. Furthermore, using a combination of iron estimates by $R_2^*$ and by liver biopsy, we were able to closely replicate the St Pierre calibration curve. Figure 3(b) demonstrates an updated figure reflecting 384 simultaneous $R_2$ and $R_2^*$ determinations in iron overloaded patients. Despite significant differences in signal acquisition and processing there is strong internal consistency among the $R_2$ and $R_2^*$ techniques.

These data all represent single-echo MRI acquisitions, that is, only one echo time is collected per tissue excitation. Newer MRI scanners can speed $R_2$ and $R_2^*$ acquisition by acquiring multiple echoes (either gradient echoes or spin echoes) with each excitation. Fortunately, $R_2^*$ values are nearly identical whether acquired by either technique, allowing comparison of results from old and new scanners [32]. Multiple-echo spin-echo acquisitions exhibit more complicated behavior [37,38]. The calibration curve derived by St Pierre does not apply and observed $R_2$ values are lower than predicted for single-echo measurements [44•,45,46].

$R_2$ and $R_2^*$ values also scale linearly with magnetic field strength. Although 1.5 T magnets remain the most popular imaging platform, 3 T magnets are becoming increasingly popular. Storey et al. [47] demonstrated that cardiac and liver $R_2^*$ values measured at 3 T are almost exactly double those observed at 1.5 T. High-field $R_2$ measurements have not been similarly validated in humans, although there are excellent data from a marmoset hemosiderosis model showing similar effects [48].

High field imaging may have advantages when tissue iron levels are low and high image resolution is imperative, such as for brain iron quantitation. There are two important disadvantages, however. First, magnetic susceptibility artifacts cause more problems at higher field strength; this is particularly important when organs are near lung, bowel gas, or sinuses.

*Curr Opin Hematol. Author manuscript; available in PMC 2010 June 28.*
Second, heavy hepatic iron burden can cause R2 and R2* to increase beyond the limits of detection. Therefore, higher field strengths for cardiac and liver iron estimation should only be used if there is no practical alternative.

**Need for target organ assessment**

While the liver iron is a good surrogate for total body iron flux, the majority of iron toxicities occur in tissues carrying a trivial proportion of the total somatic iron burden. More importantly, these iron-sensitive ‘target’ organs have different mechanisms and kinetics of iron uptake/clearance [49–51]. Specifically, endocrine tissue and heart take up circulating labile iron species that are not bound to transferrin (so-called NTBI), while liver iron uptake is predominantly mediated via transferrin [49,50]. As a result, serum ferritin and liver iron values have almost no predictive value for cardiac iron deposition when evaluated on a cross-sectional basis [34,35,52•]. High liver iron values still convey a negative prospective risk [14,15,53], however low liver iron values are not necessarily reassuring. Patients may silently harbor or even accumulate cardiac iron despite apparently adequate chelation as judged by liver iron and serum ferritin values [54,55].

The ability of MRI to detect and monitor the otherwise silent iron infiltration of cardiac tissue has truly revolutionized the management of transfusional iron overload. Figure 4 contrasts two patients with discordant cardiac and liver iron burden. One patient has heavy liver iron load (>60 mg/g dry weight) with no cardiac iron deposition. The second patient exhibits moderate cardiac iron burden with normal hepatic iron levels. The first patient required high-dose chelation therapy to prevent subsequent cardiac iron accumulation. The second patient required continuous deferoxamine therapy to facilitate cardiac iron removal but at low dose to avoid chelator toxicity.

Better insight into the dynamics of liver and heart iron loading can be obtained by serially plotting scattergrams of the two variables. Figure 5(a) demonstrates iron trajectories of five patients. Arrows indicate increasing time and temporal sampling ranging from 6 months to 2 years. The heart loaded quickly during periods of heavy hepatic siderosis but cleared substantially only when liver iron levels were quite low. The trajectories of these patients trace a complete oval, destroying the cross-sectional relationship between liver and cardiac iron.

The apparent hysteresis exhibited by cardiac iron burden suggests that there is a critical liver burden for cardiac iron loading. Figure 5(b), however, demonstrates two important counter examples. Both patients developed cardiac iron loading despite apparently ‘safe’ levels of liver iron. Similar findings have been observed by others [54,55]. Therefore, while high liver iron and ferritin values are always undesirable, low values do not guarantee cardiac protection.

**Validation of cardiac iron measurements**

Cardiac T2* was initially received with considerable skepticism [56–58]. Since cardiac biopsy is more variable and more dangerous than hepatic biopsy, direct tissue validation of cardiac T2* was a challenge [59,60]. Animal models represent a logical first step. Our laboratory demonstrated that cardiac R2* (1/T2*) rose linearly with cardiac iron in a gerbil model. Interestingly, liver and cardiac iron calibration curves were quite similar, on a wet weight basis [61]. As a subsequent ‘thought experiment’, we then extrapolated our human liver R2*–iron calibration curves to the heart T2* data and found that predicted iron levels were consistent with values derived from autopsy. Our studies were followed by gerbil work validating the utility of R2 and magnetic susceptibility measurements [62].

Since iron storage and distribution likely vary between humans and gerbils, we extended our validation studies to include fresh postmortem tissue [63•]. This study exploited the patchy
nature of cardiac iron deposition to correlate variations in cardiac $R_2$ and $R_2^*$ with variations in tissue iron concentration (Fig. 6). Although it was not possible to characterize the MRI–iron calibration over the entire pathophysiologic range, we were able to demonstrate that $R_2$ and $R_2^*$ increases (T2 and T2* shortening) reflect cardiac iron deposition in human tissue. Further autopsy studies are ongoing in nearly a dozen specimens to place tighter bounds on the absolute calibration.

Clinically, however, determination of absolute cardiac iron levels is unnecessary; functional correlates suffice. Figure 7 demonstrates the interplay between left ventricular ejection fraction and cardiac T2* in thalassemia major patients. Important points may be summarized as follows: all patients having a normal T2* (>20 ms) had normal ejection fraction; the prevalence of left ventricular dysfunction progressively increased as T2* decreased into the abnormal range; many patients with detectable cardiac iron (T2*<20 ms) nonetheless had normal cardiac function. This implies that abnormal cardiac T2* represents a preclinical degree of cardiac iron loading [34]. Figure 7(b) demonstrates the increasing prevalence of cardiac dysfunction at low T2* in a retrospective analysis of 972 cardiac T2* examinations [64]. Although the data represent a cross-sectional sampling rather than a prospective risk assessment, the implied prospective risk for a patient with low T2* is obvious.

From a practical perspective, we divide patients into three categories based upon their cardiac T2* values. Patients with T2*>20 reside in the ‘green’ zone and iron chelation is guided by trends in their estimated liver iron. Patients with T2* between 10 and 20 reside in the ‘yellow’ zone where cardiac iron deposition has occurred but there is little immediate risk of cardiac decompensation. Patients with T2*<10 ms occupy the ‘red zone’, representing sufficiently increased risk of cardiac decompensation to require immediate review and intensification of chelation therapy.

Other target organ assessments

MRI has also been used to characterize tissue iron deposits in the brain, pituitary gland, bone marrow, kidney and pancreas [65–71]. Although less well validated and clinically exploited, these approaches are likely to become increasingly important in the assessment of transfusional iron overload. The merits and limitations of these investigations are beyond the scope of this review, but several key references are included for interested readers.

Magnetic resonance imaging examinations

MRI has the remarkable ability to evaluate anatomic, physiologic, and chemical data in patients. All iron examinations should measure cardiac and liver iron as well as cardiac function. This portion of the examination can be performed in a half hour imaging slot (15–20 min of imaging time). Our practice is to utilize a full 1 h imaging slot (40 min of imaging time) to characterize liver iron by both $R_2$ and $R_2^*$ methods, liver volume, pancreatic and kidney $R_2^*$ as well as cardiac output and diastolic function. Children under 7 years old are typically sedated using propofol infusion. Older patients take no medications unless a benzodiazepine is needed to treat claustrophobia. Since cardiac function is an established, billable procedure, insurance reimbursement has been satisfactory.

Requirements to perform magnetic resonance imaging assessment of liver and cardiac iron

MRI scanners run computer programs called ‘pulse sequences’ necessary to form T2 and T2* images. State-of-the-art, fast iron imaging sequences are available on all major MRI platforms for sites having research agreements. Sites lacking such agreements can also usually obtain
satisfactory iron imaging, but it takes more effort to validate the approach. In fact, the lack of local radiologic experience or incentive is the single greatest barrier to iron quantitation by MRI.

Once proper images are acquired, images must be computer processed to generate the R2 or R2* values. Most sites have developed their own local expertise by contracting physicists or programmers to write the reconstruction algorithms; algorithmic details may be found in several publications [40,42,72•]. Other sites have used commercial software (http://www.cmrttools.com) or services (http://www.ferriscan.com) to generate their iron estimates.

Conclusion

MRI combines accurate, reproducible and platform-robust estimates of organ iron concentration with structural and functional correlations in patients with iron overload. Further work is ongoing to better understand interpatient and interorgan differences in iron loading and unloading. These techniques will become the standard of care as industry begins to support the technical demands of their implementation.

Acknowledgments

This work has been supported by the NIH 5 RO1 HL075592, the General Clinical Research Center (GCRC, RR00043-43), Novartis Pharma, Department of Pediatrics at Children’s Hospital Los Angeles, Wright Foundation, and Gunther Foundation.

Abbreviation

MRI magnetic resonance imaging

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

• of special interest

•• of outstanding interest

Additional references related to this topic can also be found in the Current World Literature section in this issue (p. 291).


55. Kolnagou A, Economides C, Eracleous E, Kontoghiorghes GJ. Low serum ferritin levels are misleading for detecting cardiac iron overload and increase the risk of cardiomyopathy in thalassemia patients: the importance of cardiac iron overload monitoring using magnetic resonance imaging T2 and T2*. Hemoglobin 2006;30:219–227. [PubMed: 16798647]


Figure 1. Gradient echo images of liver collected at four different echo times

The top four images were collected from a patient having a liver iron of 6 mg/g. The bottom four images were collected from a normal volunteer. All images darken as the echo time (TE) lengthens, but the iron-heavy tissue darkens faster. The half life of this process is called T2* and the rate is called R2* (R2*=1000/T2*).
Figure 2. Plot of liver R2 versus biopsied liver iron concentration in 104 patients
Data reproduced with permission from St Pierre [30].
Figure 3. Plots of liver R2 versus biopsied liver iron concentration

(a) Plot of liver R2* versus biopsied liver iron concentration (HIC) in 22 patients. There is a strong linear correlation ($r = 0.97$). (b) Plot of R2 versus liver iron predicted by biopsy (+ signs) or by simultaneously measured R2* HIC estimates ($n = 384$). Bold line represents the Tim St Pierre calibration curve, supporting strong internal consistency between R2 and R2* measurements. This research was originally published in Wood et al. [28]. Copyright American Society of Hematology.
Figure 4. Gradient echo (T2*) imaging illustrating discordant iron loading of the liver and the heart
(a) Heavy liver iron loading (dark tissue) with heart sparing. (b) Heavy cardiac iron loading
with no liver iron deposition.
Figure 5. Iron trajectories of five patients
(a) Scattergram demonstrating cardiac iron versus liver iron (HIC) in three patients. Temporal trajectories trace out an oval, demonstrating the lack of cross-sectional correlation between liver and heart iron. (b) Same scattergram with two additional patients who developed cardiac iron loading despite having apparently adequate iron chelation.
Figure 6. Correlation of variations in cardiac R2 and R2* with variations in tissue iron concentration
(a) R2 and R2* ‘maps’ measured in vitro from a fresh postmortem specimen. The grid represents approximate divisions for iron sampling. (b) Rise of R2 and R2* as a function of assayed tissue iron concentration. Graphs redrawn with permission from [63•].
Figure 7. Interplay between left ventricular ejection fraction (LVEF) and cardiac T2* in thalassemia major patients
(a) Plot of LVEF versus cardiac T2* value. Data reproduced with permission from [34]. (b) Prevalence histogram of left ventricular dysfunction (LVEF <56%) as a function of cardiac T2*. Data reproduced with permission from [64].