HHS Public Access

Author manuscript

Am J Hematol. Author manuscript; available in PMC 2016 September 01.

Published in final edited form as:

Am J Hematol. 2015 September ; 90(9): 806–810. doi:10.1002/ajh.24089.

Liver Iron Concentration Measurements by MRI in Chronically Transfused Children with Sickle Cell Anemia: Baseline Results from the TWiTCH Trial

John C. Wood[1], **Sara Pressel**[2], **Zora R. Rogers**[3], **Isaac Odame**[4], **Janet L. Kwiatkowski**[5], **Margaret T. Lee**[6], **William C. Owen**[7], **Alan R. Cohen**[8], **Timothy St. Pierre**[8], **Matthew M. Heeney**[9], **William H. Schultz**[10], **Barry R. Davis**[2], **Russell E. Ware**[10], and **for the TWiTCH Investigators**

[1]Children's Hospital Los Angeles, Los Angeles, CA

[2]The University of Texas Health Science Center, Houston TX

[3]University of Texas Southwestern Medical Center, Dallas, TX

[4]University of Toronto, Division of Haematology/Oncology, The Hospital for Sick Children, Toronto, Canada

[5]Children's Hospital of Philadelphia, Philadelphia, PA

[6]Columbia University, New York, NY

[7]Children's Hospital of the King's Daughters, Norfolk VA

[8]School of Physics, University of Western Australia, Crawley, Australia

[9]Boston Children's Hospital, Boston, MA

[10]Cincinnati Children's Hospital Medical Center, Cincinnati, OH

Abstract

Non-invasive, quantitative, and accurate assessment of liver iron concentration (LIC) by MRI is useful for patients receiving transfusions, but R2 and R2* MRI techniques have not been systematically compared in sickle cell anemia (SCA). We report baseline LIC results from the TWiTCH trial, which compares hydroxyurea with blood transfusion treatment for primary stroke prophylaxis assessed by transcranial Doppler sonography in pediatric SCA patients. Liver R2 was collected and processed using a FDA-approved commercial process (FerriScan®), while liver R2* quality control and processing were performed by a Core Laboratory blinded to clinical site and patient data. Baseline LIC studies using both MRI techniques were available for 120 participants.

Correspondence: John C. Wood MD PhD, Division of Cardiology, Mailstop 34, Children's Hospital Los Angeles, 4650 Sunset Blvd, Los Angeles, CA, 90027, 323 361 5470, jwood@chla.usc.edu.

AUTHORSHIP

Designed research (JCW, REW), performed research (ZRR, IO, JLK, MTL, WCO, ARC, MMH, WHS, REW), contributed analytical tools (JCW, TSP), analyzed data (JCW, SP, BRD), and wrote the paper (JCW, SP, ZRR, IO, JLK, MTL, WCO, ARC, TSP, MMH, WHS, BRD, REW)

CONFLICT OF INTEREST DISCLOSURE

Sara Pressel, William Owen, Barry Davis, and Russell Ware have no conflicts to disclose.

LIC_{R2*} and LIC_{R2} results were highly correlated ($r^2 = 0.93$). A proportional bias of LIC(R2*)/LIC(R2), decreasing with average LIC, was observed. Systematic differences between LIC_{R2*} and LIC_{R2} were also observed by MRI manufacturer. Importantly, LIC_{R2*} and LIC_{R2} estimates had broad 95% limits of agreement with respect to each other. We recommend LIC_{R2} and LIC_{R2*} not be used interchangeably in SCA patients to follow individual patient trends in iron burden.

INTRODUCTION

Stroke remains one of the most feared complications among children with sickle cell anemia (SCA) because of its medical, neurocognitive, psychosocial, and financial burdens to patients and their families [1–3]. The implementation of universal transcranial Doppler (TCD) screening in all children with SCA, with subsequent chronic transfusion therapy in high risk patients with abnormal TCD velocities, has markedly decreased primary stroke incidence in pediatric patients [4, 5]. However, chronic transfusion therapy inevitably causes iron overload, necessitating long-term iron chelation that requires monitoring and management of treatment-related toxicities [6].

Serum ferritin typically guides the treatment of transfusional iron overload, but in patients with SCA, ferritin is a notoriously unreliable marker of iron overload [7]. In contrast, liver iron concentration (LIC) is an excellent marker of total body iron balance [8]. While liver biopsy was traditionally used for this purpose, it is invasive, subject to sampling error [9], unpopular with patients, and is being replaced by magnetic resonance imaging (MRI) methods [10]. Tissue iron disrupts the magnetic field in different organs, causing the images to darken more quickly than normal [11]. MRI estimates tissue iron by measuring the rate of darkening which is called R2 or R2*, depending on how the images are collected. These R2 and R2* LIC measurements (LIC_{R2} and LIC_{R2*}) are becoming standard of care at major sickle cell centers [10] and have even become accepted surrogates for clinical trials [12]. In fact, one LIC_{R2} measurement technique has been FDA approved for monitoring of chelation therapy for certain indications (FerriScan®, Resonance Health, Nedlands, Western Australia). However, concordance between FerriScan® LIC_{R2} and LIC_{R2*}, particularly in the context of a multicenter clinical trial, has never been reported for patients with SCA.

TCD With Transfusions Changing to Hydroxyurea (TWiTCH) is a multi-centered, open label, randomized clinical trial, comparing 24 months of open label hydroxyurea to chronic transfusion therapy for primary stroke prophylaxis. Pediatric patients who received at least 12 months of regular blood transfusions for abnormal TCD velocities were randomized to either continue transfusion therapy or transition to hydroxyurea therapy at maximum tolerated dose. In the transfusion arm, iron overload is managed with chelation, while in the hydroxyurea arm, iron overload is controlled with monthly phlebotomy. The primary outcome variable in TWiTCH is TCD velocity. However, the LIC represents a critical secondary outcome variable, with LIC_{R2} and LIC_{R2*} being measured at baseline, midpoint (1 year), and at study exit. This manuscript reports the baseline measurements of LIC, directly comparing LIC_{R2} and LIC_{R2*} results from two specific MR LIC measurement techniques, and examining the relationships with regard to study participant and MRI scanner characteristics.

METHODS

The TWiTCH study is registered at ClinicalTrials.gov (NCT01425307). The study is being performed at 26 clinical sites and approved by Institutional Review Boards at each institution. Informed consent was obtained for all patients. Although LIC measurements represent a secondary outcome, efforts were made to collect LIC_{R2} and LIC_{R2*} measurements in as many institutions as possible. LIC_{R2} was performed at 24 out of 26 centers. LIC_{R2*} had more stringent hardware constraints and was performed at 21 centers. All imaging was performed at 1.5 Tesla using phased array torso or cardiac coils. MRI manufacturers included Siemens (67 exams), General Electric Medical Systems (31 exams), and Philips Healthcare (15 exams). A total of 159 patients enrolled in the study and initiated screening; 138 patients underwent at least one baseline LIC measurement, and 120 patients successfully completed both LIC_{R2} and LIC_{R2*} baseline assessments.

LIC_{R2} was collected using the FerriScan® protocol [13, 14]. Free-breathing single spin echo images were collected with the following parameters: repetition time 1000 ms, echo times 6, 9, 12, 15, and 18 ms; slice thickness 5 mm; matrix 192 × 256; and a ¾ phase field of view. All images were sent to an independent commercial imaging laboratory (Resonance Health Ltd, Western Australia) for processing using proprietary software [14]. Signal decay was fitted with a two-component model on a pixel-by-pixel basis [15] and the mean value was transformed to the LIC value using an established [13] and independently verified [12, 16] calibration curve.

LIC_{R2*} was collected using multiple echo, gradient echo sequences. Hardware differences among the 3 manufacturers precluded completely controlling the imaging parameters. Supplemental Table 1 summarizes the “targeted” and achieved imaging parameters for all 120 studies as well as a cohort restricted to fully compliant protocols (N=87). All images were transferred to an independent core laboratory at Children’s Hospital Los Angeles and processed using a previously published algorithm [17]. Signal decay was fitted with an exponential with a constant [18], using adaptive binning to group pixels of similar decay rates [17]. Median R2* was translated to LIC using an established [16] and independently verified calibration curve [19].

LIC_{R2} and LIC_{R2*} results were compared using Bland-Altman analysis. Log transformation was needed because differences scaled proportionally to average LIC. Log differences were then compared to both biological and methodological predictors, to identify potential contributors of bias by analysis of variance. Biological predictors included patient age, body mass index (BMI), years of transfusion and chelation therapy, and a variety of enrollment laboratory parameters including hemoglobin concentration, serum ferritin, absolute reticulocyte count, lactate dehydrogenase (LDH), serum iron studies, alanine aminotransferase (ALT), and C-reactive protein (CRP) values. Methodological predictors included the MRI manufacturer, initial echo time, second echo time, final echo time, echo spacing, number of echoes, slice thickness, frequency resolution, phase resolution, percent phase sampling, percent frequency sampling, bandwidth, and two relative signal-to-noise (SNR) metrics. Univariate and multivariate regression were initially performed using all variables, followed by stepwise forward regression to identify the most robust predictors of

the Bland Altman agreement. All statistics were performed by the TWiTCH Data Coordinating Center at the University of Texas School of Public Health.

RESULTS

Patient demographics for the baseline assessments are summarized in Supplemental Table 2. The enrolled subjects had laboratory evidence of transfusional iron overload, with an average chronic transfusional exposure of 4.1 years and mean serum ferritin of 2708 ng/mL. More than 95% were regularly chelated with either deferasirox or deferoxamine.

Figure 1 (upper panel) demonstrates a scattergram of LIC_{R2*} versus LIC_{R2} values for all 120 TWiTCH subjects who had both baseline assessments. Data are displayed on logarithmic axes to stabilize the variance. Linear regression performed on the log-transformed LIC values yielded an r^2 of 0.93 with approximation to the line of identity. Since some sites had difficulty achieving the ideal imaging parameters (Supplemental Table 1) for liver R2* imaging, the lower panel of Figure 1 includes only examinations that were fully compliant with core lab recommendations (N=87). Results are nearly identical to the complete dataset ($r^2 = 0.92$), suggesting that variation in imaging parameters are not a major contributor to the broad Bland-Altman limits of agreement observed. Figure 2 illustrates the Bland Altman relationship between the log-transformed LIC metrics on the smaller data set. Data are plotted on logarithmic axes for clarity, resulting in a graph of the ratio of LIC_{R2*} to LIC_{R2} versus their geometric mean. There was a significant negative trend ($r^2 = 0.195$, $p < 0.0001$), with LIC_{R2*} greater than LIC_{R2} at lower iron concentrations while the converse was true at higher iron concentrations. 95% confidence intervals were 0.60 – 1.48.

To identify possible biological or methodological predictors of bias and variance, we next performed univariate and multivariate linear regression. For clarity, only variables that exhibited a univariate regression p-value less than 0.1 are shown in Table 1. Separate columns are shown for the complete and restricted datasets. For both cohorts, serum ferritin and average LIC were the strongest biological predictors of measurement bias between LIC_{R2*} and LIC_{R2}, consistent with downward slope of the Bland-Altman relationship. Serum LDH was also positively correlated with a larger difference between LIC_{R2*} and LIC_{R2} on univariate analysis.

Several technologic variables were also significant on univariate analysis including MRI manufacturer, SNR metric 1, number of echoes, Interecho spacing, repetition time, slice thickness (full cohort only), points in frequency direction, and frequency direction sampling percentage. In backwards stepwise regression, the following variables were significantly associated with the ratio R2*/R2 in the restricted cohort: average LIC, R2, minimum echo time, signal to noise ratio metrics 1 and 2, LDH, and vendor in the restricted cohort.

Figure 3 demonstrates an analysis of variance of the LIC_{R2*} to LIC_{R2} ratio by MRI manufacturer. Analysis of variance according to MRI manufacturer corrected for LIC was highly significant (p=0.0004). Mean LIC ratios were greater than one for Philips and General Electric magnets, while they were less than one for Siemens magnets. Vendor differences were retained in the final multivariate model (Supplemental Table 3).

DISCUSSION

The accurate measurement of LIC is critical for the management of transfusional iron burden among patients receiving chronic transfusion therapy. Historically much of the clinical focus was on children and young adults with transfusion-dependent thalassemia. More recently, substantial iron burdens have been identified in chronically transfused patients with SCA, but these patients have highly variable clinical manifestations and potentially different pathophysiology [20, 21] than patients with transfusion-dependent thalassemia. The availability of non-invasive, quantitative, and accurate assessment of LIC using MRI technology represents an important opportunity to improve the clinical care for such chronically transfused patients, especially as many children with SCA are receiving transfusion therapy for primary stroke prophylaxis[4, 10].

In this analysis of baseline iron burden in a large cohort of children with SCA and transfusional iron overload, systematic bias was identified between LIC_{R2} and LIC_{R2*} measurements. Although there was generally excellent correlation between the two measures (Figure 1, $r^2 = 0.93$), the LIC_{R2*} results exceeded LIC_{R2} for values <5 mg/gm dry weight liver, but were lower than LIC_{R2} for higher iron concentrations (Figure 2). The Bland-Altman relationship cannot distinguish whether this bias is arising from measurements of LIC_{R2}, LIC_{R2*}, or both. The most likely source of this bias is that the R2 and R2* calibration curves were not calibrated to an identical biopsy standard; differences in biopsy processing can have profound systematic effects on estimated LIC [22]. Furthermore, published calibration curves for both LIC-R2 and LIC-R2* are both imperfect because of liver biopsy sampling variability, MRI measurement errors, and intersubject variability in iron storage[12, 13, 16, 19, 23] Similar results have been described in a large cohort of thalassemia major patients, suggesting that our observations may be independent of the disease phenotype[24].

The contribution of LDH to the observed bias is unclear. LDH levels are influence by many physiologic variables including hemolytic rate, vascular stress, as well as transfusion intensity and duration. These physiologic parameters, in turn, could produce patient-specific differences in tissue iron distribution affecting the relationships between R2*, R2, and liver iron concentration [25].

The source of LIC-R2/LIC-R2* bias also appears to have a technical component, with greater Bland-Altman differences identified on exams using Siemens scanners than on those collected from Philips or General Electric magnets (Figure 3). Siemens R2* imaging protocols often use a technique to decrease signal from fat, an approach known to lower observed R2* values[26]. However, we were not able to document that this was responsible for the bias. We found that minimum echo time was the most important parameter [18], but that relative signal to noise estimates also contributed weakly to observed differences [27]. While one prior paper demonstrated that LIC_{R2} measured by FerriScan® is robust across different manufacturers (Siemens, GE, and Philips Healthcare)[12], this comparison was performed using older models than those employed in TWiTCH. Previous studies with LIC_{R2*} have not suggested systematic inter-manufacturer bias, but may have been underpowered to detect effects of this magnitude [28–30]. However, our observed

manufacturer bias is sufficiently small, compared with the 95% limits of agreement of the Bland-Altman relationship, that it is unlikely to impact patient management. High quality data, yielding reproducible LIC_{R2} and LIC_{R2*} values, can be obtained on all three manufacturers.

While the biases between the LIC_{R2}, and LIC_{R2*} measurements used in this study were notable and interesting, they were easy to correct once characterized. However, even after compensating for systematic differences, LIC_{R2} and LIC_{R2*} also demonstrated significant intrasubject variability, precluding use of these metrics interchangeably in clinical practice. The source of these errors is controversial, since both LIC_{R2} and LIC_{R2*} are extremely reproducible, with inter-study variation of 5–8% [14, 16]. Furthermore, both LIC_{R2} and LIC_{R2*} have also been separately calibrated against liver biopsy, with 95% limits of agreement ranging from ±42% [19] to ±72%[12]LIC_{R2} and LIC_{R2*} measurements sample tissue volumes of 100 – 300 grams, nearly five orders of magnitude greater than that collected by liver biopsy, eliminating sampling variability as a major contributor to the observed differences. However, LIC_{R2} and LIC_{R2*} are both indirect metrics of tissue iron, quantifying the interaction between diffusing water protons and disturbances in the magnetic field produced by iron [25]. Each patient has a unique magnetic signature that introduces significant, patient-specific errors in LIC_{R2} and LIC_{R2*} values [12, 25] that are additive when LIC_{R2} and LIC_{R2*} are compared to one another.

To the practicing hematologist, these patient-specific LIC errors are not large enough to change a patient's clinical risk and they vary slowly over time. MRI manufacturer/model bias is also likely to be static over time. As a result, serial LIC measurements by either of the two methods employed in this study accurately represent changes in total body iron [24]. In fact, detailed simulation work suggests that carefully-performed LIC_{R2} and LIC_{R2*} behave so well longitudinally that they both are superior to liver biopsy for the determination of iron balance [31]. Thus, clinicians can use either approach on any imaging platform with confidence, but should maintain consistency using the same methodology and machines for serial patient evaluation as much as practically possible.

In summary, we provide a large cohort data analysis that demonstrates both LIC_{R2} and LIC_{R2*} quantitative measurements of liver iron burden exhibit a systematic bias ranging from +0.2 mg/g at LIC values of 1 mg/g to −9.1 mg/g at LIC values of 40 mg/g. The observed bias varies among scanner manufacturers, with the Siemens platform yielding larger differences. More importantly, we demonstrate that comparisons of LIC_{R2} and LIC_{R2*} have broad limits of agreement. Accordingly, we recommend that different methods of MR LIC measurement not be used interchangeably in the same patient to follow trends in iron loading or unloading. Longitudinal data from the TWiTCH trial will further test this recommendation by measuring interval changes in LIC_{R2} and LIC_{R2*} and comparing differences between techniques and across time. Exit data from TWiTCH will extend prior work regarding iron unloading in thalassemia major to patients with SCA[8]. Furthermore, in TWiTCH half of the subjects receive serial phlebotomy, and these quantitative changes in iron will serve as independent validation of the changes in total body iron predicted by LIC_{R2} and LIC_{R2*}. Serial analyses will also offer insights into inter-machine differences that could not be appreciated in smaller trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

ACKNOWLEDGEMENTS

This work was supported by the National Heart Lung and Blood Institute (R01HL095647-05 and 5 R01 HL095511-05). We would like to credit all participating sites and investigators, as outlined in the appendix. The authors would also like to thank Resonance Health for their help setting up and interpreting the FerriScan® analyses for this study.

JCW has served as a paid consultant to Biomed Informatics, Shire, and Pfizer. He has research funding from NIH and Shire and has served on Medical Advisory Panels for ApoPharma. ZR has served as a consultant for ApoPharma, Baxter, Roche, Glaxo Smith Kline, and the State of Texas. She has served as a paid DSMB member (NIH) and is an elected Board Member for ASPHO and the executive committee for the SOHO section of AAP. IO has received a Grand Challenge award from Canada and has received honoraria from Novartis for a speaking engagement and expert testimony. JLK has served as a consultant for Sideris and Shire and received research funding from Resonance Health. ML has research funding from the FDA and NIH. Alan Cohen has served as a consultant to Novartis. Tim St Pierre has served as a consultant for, owns stock in, and has served on the Board of Directors for Resonance Health. Matthew Heeney has received research funding from Novartis and Eli Lilly.

APPENDIX

Participating sites and clinical investigators are listed as follows: Texas Children’s Hospital (Alex George), Children’s Hospital Boston (Matt Heeney), Cincinnati Children’s Hospital Medical Center (Theodosia Kalfa), Children’s Hospital of Minneapolis (Steve Nelson), Children’s Healthcare of Atlanta, Eggleston (Clark Brown), Hughes Spaulding (Beatrice Gee), Scottish Rite (Clark Brown), Children’s Hospital of Philadelphia (Janet Kwiatkowski), Toronto Sick Kids Hospital (Issac Odame), Children’s National Medical Center (Lori Luchtman-Jones), Columbia University (Margaret Lee), Rainbow Babies and Children’s Hospital (Connie Piccone), University of South Alabama (Hamayun Imran), Medical University of South Carolina, Charlston (Sherron Jackson), Cohen Children’s Medical Center, New Hyde Park (Banu Aygun), St. Jude Children’s Research Hospital, Memphis (Kerry Nottage), State University of New York Downstate, Brooklyn (Scott Miller), University of Alabama, Birmingham (Lee Hilliard), University of Miami (Ofelia Alvarez), University of Mississippi, Jackson (Margaret Smith), University of Texas Southwestern, Dallas (Zora Rogers), Wayne State University, Detroit (Ingrid Sarniak), Lurie Children’s Hospital of Chicago (Alexis Thompson), Children’s Hospital of the King’s Daughter, Norfolk (William Owen), Nemour’s Children’s Hospital, Jacksonville (Cynthia Gauger), University of South Carolina, Columbia (Carla Roberts), Duke University Medical Center, Durham (Jennifer Rothman), and Eastern Carolina University, Greenville (Beng Fuh).

REFERENCES

1. Ohene-Frempong K, Weiner SJ, Sleeper LA, et al. Cerebrovascular accidents in sickle cell disease: rates and risk factors. Blood. 1998; 91:288–294. [PubMed: 9414296]
2. Scantlebury N, Mabbott D, Janzen L, et al. White matter integrity and core cognitive function in children diagnosed with sickle cell disease. J Pediatr Hematol Oncol. 2011; 33:163–171. [PubMed: 21325970]
3. Steen RG, Miles MA, Helton KJ, et al. Cognitive impairment in children with hemoglobin SS sickle cell disease: relationship to MR imaging findings and hematocrit. AJNR Am J Neuroradiol. 2003; 24:382–389. [PubMed: 12637286]

4. Bernaudin F, Verlhac S, Arnaud C, et al. Impact of early transcranial Doppler screening and intensive therapy on cerebral vasculopathy outcome in a newborn sickle cell anemia cohort. Blood. 2011; 117:1130–1140. quiz 1436. [PubMed: 21068435]
5. Enninful-Eghan H, Moore RH, Ichord R, et al. Transcranial Doppler ultrasonography and prophylactic transfusion program is effective in preventing overt stroke in children with sickle cell disease. J Pediatr. 2010; 157:479–484. [PubMed: 20434165]
6. Vichinsky EP. Current issues with blood transfusions in sickle cell disease. Semin Hematol. 2001; 38:14–22. [PubMed: 11206956]
7. Karam LB, Disco D, Jackson SM, et al. Liver biopsy results in patients with sickle cell disease on chronic transfusions: poor correlation with ferritin levels. Pediatr Blood Cancer. 2008; 50:62–65. [PubMed: 17457853]
8. Angelucci E, Brittenham GM, McLaren CE, et al. Hepatic iron concentration and total body iron stores in thalassemia major. New England Journal of Medicine. 2000; 343:327–331. [PubMed: 10922422]
9. Aygun B, Mortier N, Kesler K, et al. Therapeutic Phlebotomy is Safe in Children with Sickle Cell Anaemia and can be Effective Treatment for Transfusional Iron Overload. Br J Haematol. 2015 In Press.
10. Kwiatkowski JL, Kim HY, Thompson AA, et al. Chelation use and iron burden in North American and British thalassemia patients: a report from the Thalassemia Longitudinal Cohort. Blood. 2012; 119:2746–2753. [PubMed: 22279056]
11. Wood JC. Magnetic resonance imaging measurement of iron overload. Curr Opin Hematol. 2007; 14:183–190. [PubMed: 17414205]
12. St Pierre TG, El-Beshlawy A, Elalfy M, et al. Multicenter validation of spin-density projection-assisted R2-MRI for the noninvasive measurement of liver iron concentration. Magn Reson Med. 2014; 71:2215–2223. [PubMed: 23821350]
13. St Pierre TG, Clark PR, Chua-anusorn W, et al. Noninvasive measurement and imaging of liver iron concentrations using proton magnetic resonance. Blood. 2005; 105:855–861. [PubMed: 15256427]
14. St Pierre TG, Clark PR, Chua-Anusorn W. Single spin-echo proton transverse relaxometry of iron-loaded liver. NMR Biomed. 2004; 17:446–458. [PubMed: 15523601]
15. Clark PR, Chua-anusorn W, St Pierre TG. Bi-exponential proton transverse relaxation rate (R2) image analysis using RF field intensity-weighted spin density projection: potential for R2 measurement of iron-loaded liver. Magn Reson Imaging. 2003; 21:519–530. [PubMed: 12878262]
16. Wood JC, Enriquez C, Ghugre N, et al. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. Blood. 2005; 106:1460–1465. [PubMed: 15860670]
17. Meloni A, Zmyewski H, Rienhoff HY Jr, et al. Fast approximation to pixelwise relaxivity maps: validation in iron overloaded subjects. Magn Reson Imaging. 2013; 31:1074–1080. [PubMed: 23773621]
18. Ghugre NR, Enriquez CM, Coates TD, et al. Improved R2* measurements in myocardial iron overload. Journal of Magnetic Resonance Imaging. 2006; 23:9–16. [PubMed: 16329085]
19. Hankins JS, McCarville MB, Loeffler RB, et al. R2* magnetic resonance imaging of the liver in patients with iron overload. Blood. 2009; 113:4853–4855. [PubMed: 19264677]
20. Fung EB, Harmatz PR, Lee PD, et al. Increased prevalence of iron-overload associated endocrinopathy in thalassaemia versus sickle-cell disease. Br J Haematol. 2006; 135:574–582. [PubMed: 17054676]
21. Jenkins ZA, Hagar W, Bowlus CL, et al. Iron homeostasis during transfusional iron overload in beta-thalassemia and sickle cell disease: changes in iron regulatory protein, hepcidin, and ferritin expression. Pediatr Hematol Oncol. 2007; 24:237–243. [PubMed: 17613866]
22. Butensky E, Fischer R, Hudes M, et al. Variability in hepatic iron concentration in percutaneous needle biopsy specimens from patients with transfusional hemosiderosis. Am J Clin Pathol. 2005; 123:146–152. [PubMed: 15762291]

23. Garbowski MW, Carpenter JP, Smith G, et al. Biopsy-based calibration of T2* magnetic resonance for estimation of liver iron concentration and comparison with R2 Ferriscan. J Cardiovasc Magn Reson. 2014; 16:40. [PubMed: 24915987]
24. Wood JC, Zhang P, Rienhoff H, et al. R2 and R2* are equally effective in evaluating chronic response to iron chelation. Am J Hematol. 2014; 89:505–508. [PubMed: 24452753]
25. Ghugre NR, Wood JC. Relaxivity-iron calibration in hepatic iron overload: probing underlying biophysical mechanisms using a Monte Carlo model. Magn Reson Med. 2011; 65:837–847. [PubMed: 21337413]
26. Meloni A, Tyszka JM, Pepe A, et al. Effect of inversion recovery fat suppression on hepatic r2* quantitation in transfusional siderosis. AJR Am J Roentgenol. 2015; 204:625–629. [PubMed: 25714295]
27. He T, Gatehouse PD, Kirk P, et al. Myocardial T(*)2 measurement in iron-overloaded thalassemia: An ex vivo study to investigate optimal methods of quantification. Magn Reson Med. 2008; 60:350–356. [PubMed: 18666131]
28. Westwood MA, Anderson LJ, Firmin DN, et al. Interscanner reproducibility of cardiovascular magnetic resonance T2* measurements of tissue iron in thalassemia. J Magn Reson Imaging. 2003; 18:616–620. [PubMed: 14579406]
29. Westwood MA, Firmin DN, Gildo M, et al. Intercentre reproducibility of magnetic resonance T2* measurements of myocardial iron in thalassaemia. Int J Cardiovasc Imaging. 2005; 21:531–538. [PubMed: 16175443]
30. Kirk P, He T, Anderson LJ, et al. International reproducibility of single breathhold T2* MR for cardiac and liver iron assessment among five thalassemia centers. J Magn Reson Imaging. 2010; 32:315–319. [PubMed: 20677256]
31. Wood JC, Zhang P, Rienhoff H, et al. Liver MRI is more precise than liver biopsy for assessing total body iron balance: A comparison of MRI relaxometry with simulated liver biopsy results. Magn Reson Imaging. 2015

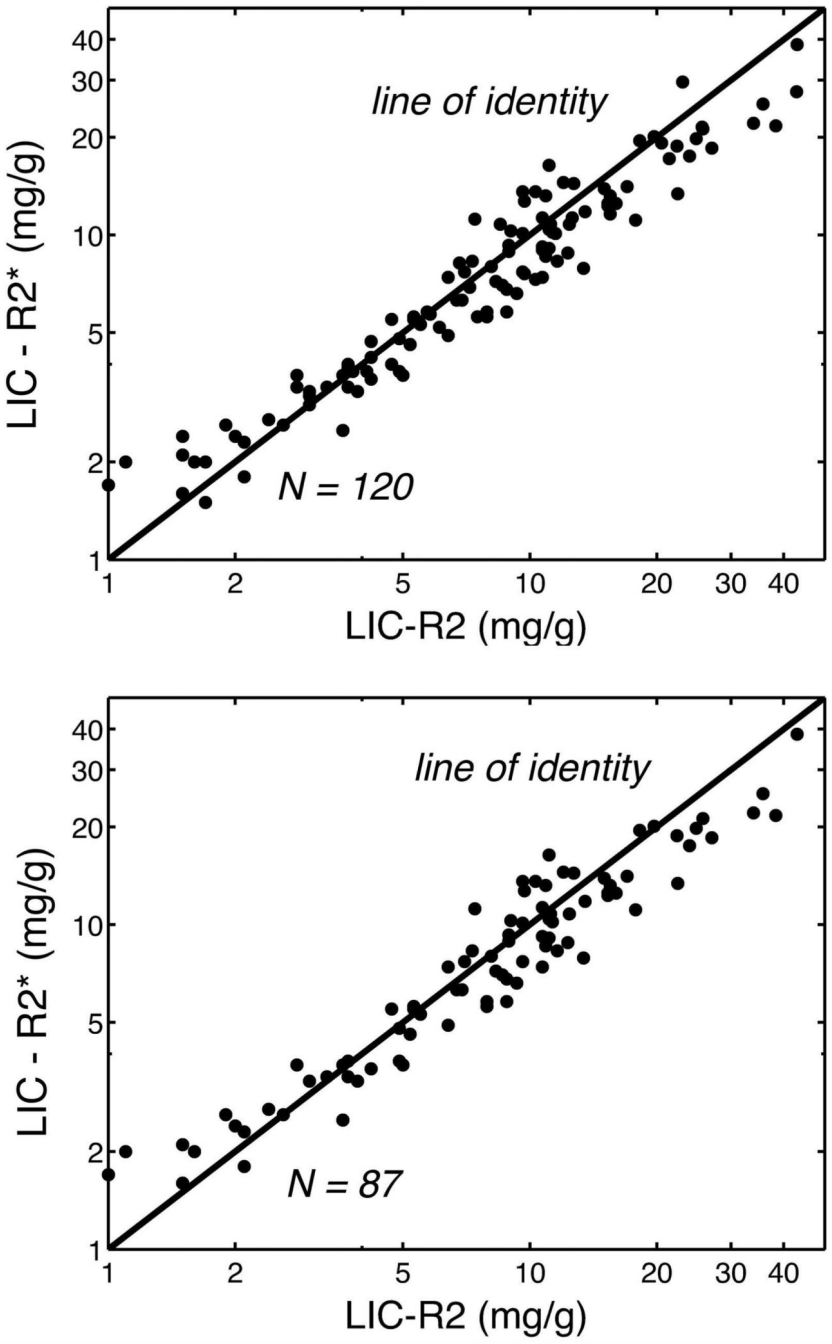

Figure 1.

Comparison of LIC_{R2*} and LIC_{R2} techniques using full cohort (top, N = 120) and restricted data set (bottom, N = 87). Both horizontal and vertical axes are logarithmic.

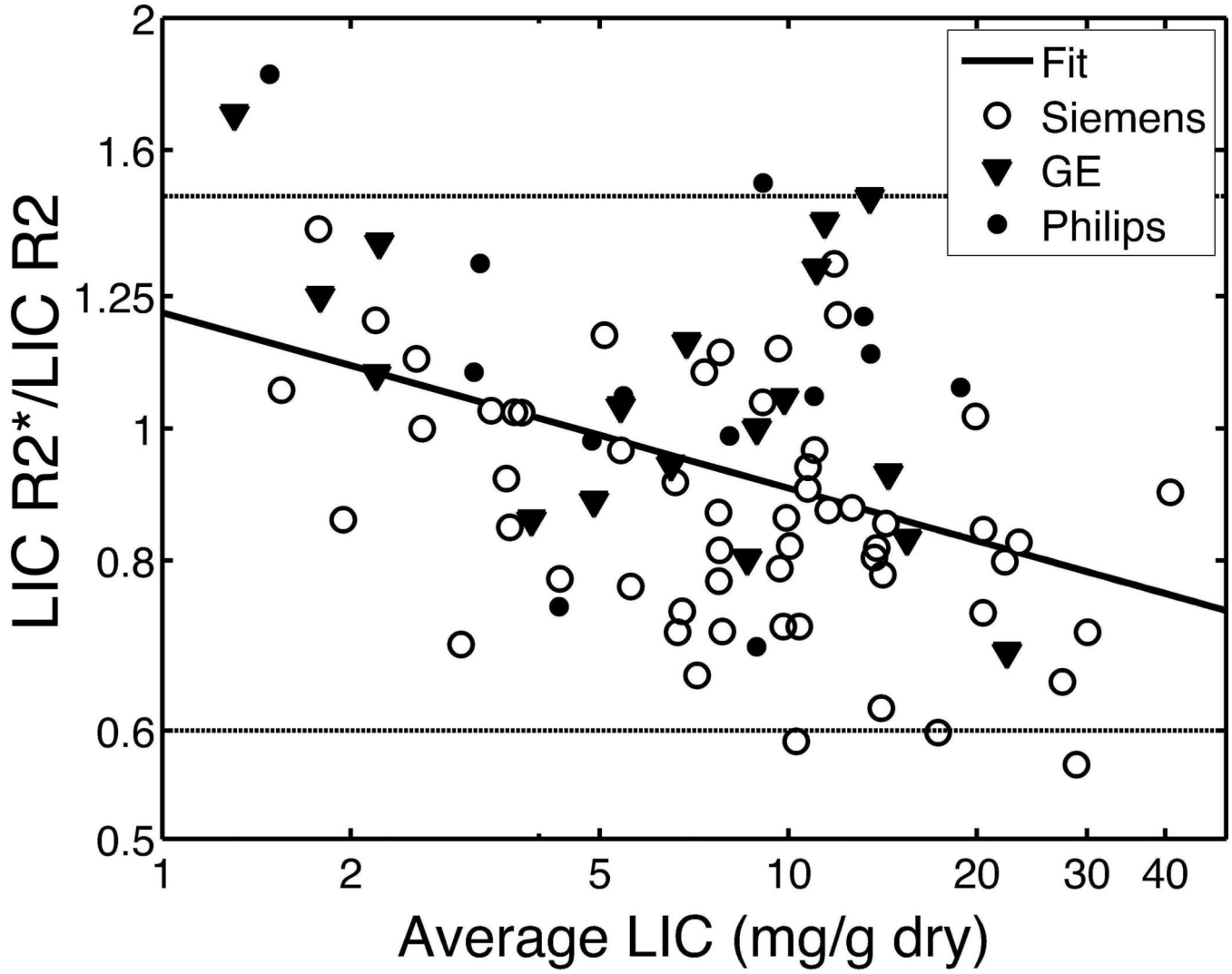

Figure 2.

Bland-Altman comparison between LIC_{R2*} and LIC_{R2} techniques on reduced data set (N = 87). Data are represented using logarithmic axes, corresponding to the ratio of LIC_{R2*} to LIC_{R2} on the vertical axis plotted against the geometric mean of LIC_{R2*} and LIC_{R2} on the horizontal axis. The three magnet manufacturers are indicated by different symbols. Horizontal lines represent the Bland-Altman 95% confidence intervals uncorrected for bias.

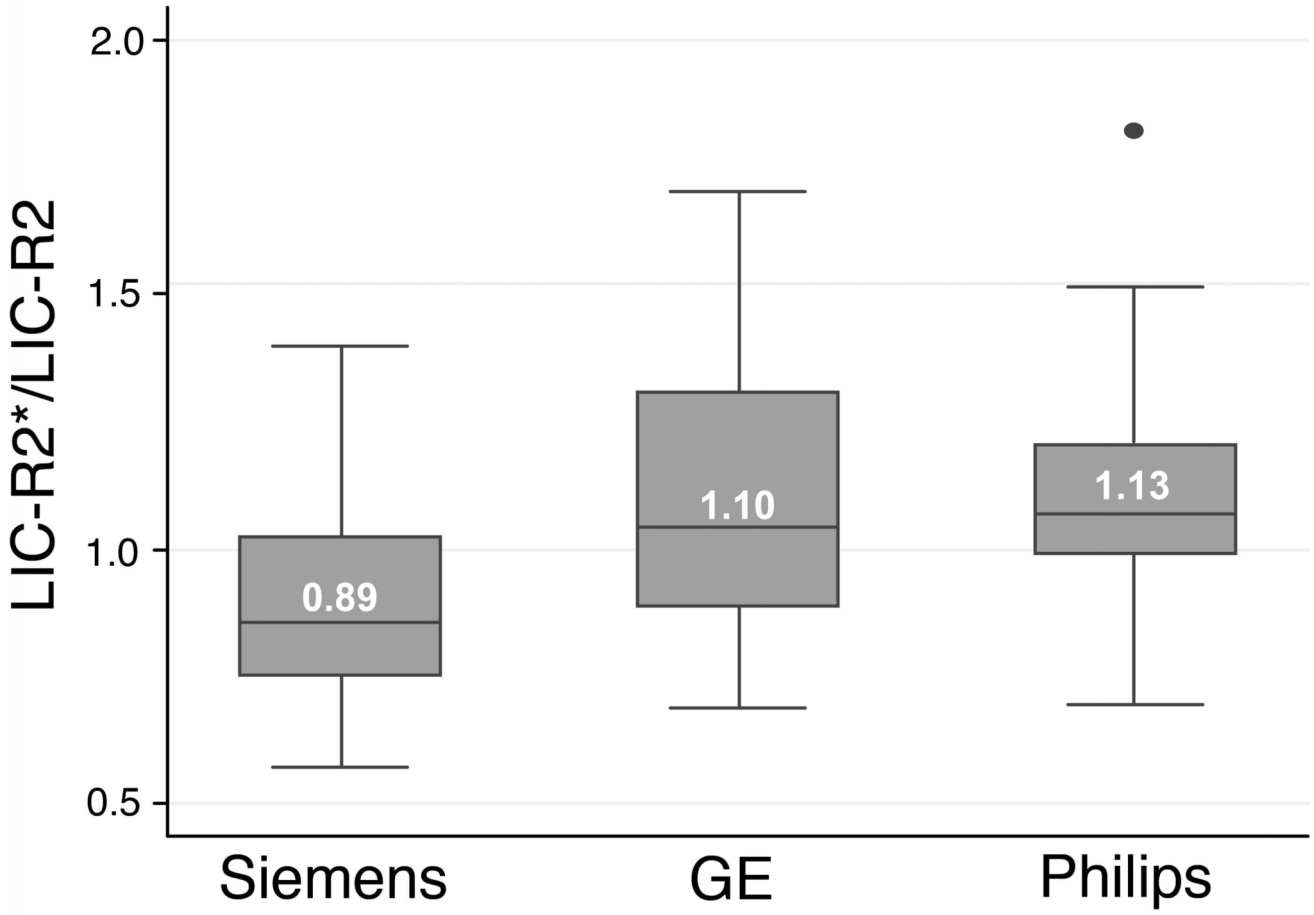

Figure 3.

Ratio of LIC_{R2*} to LIC_{R2}, by manufacturer, for the restricted cohort. Asterisk indicates a ratio that is statistically different from unity ($p<0.05$). Horizontal line and white text denote median values, the box represents the interquartile range, while the whiskers represent the upper and lower adjacent values.

Table 1

Univariate regression for the difference in the natural logarithm of LIC_{R2*} and LIC_{R2}

	Full cohort (N=120)		Restricted (N=87) Cohort	
Biological Parameters	**Coefficient**	**p-value**	**Coefficient**	**p-value**
Average(ln LIC_{R2}), ln(LIC_{R2*}))	−0.129	**<0.001**	−0.140	**<0.001**
LIC median	−0.095	**0.001**	−0.102	**0.005**
Sex (male)	0.073	0.089	0.110	**0.038**
Deferoxamine in the past year	−0.224	**0.020**	−0.215	0.055
Serum ferritin (ng/mL)	−0.00005	**<0.001**	−0.000056	**<0.001**
LDH (U/L)	0.0002	**0.035**	0.00030	**0.015**
Technical Parameters				
Scanner manufacturer: Siemens (reference) GE Phillips	 0.163 0.242	 **<0.001** **< 0.001**	 0.210 0.234	 **0.001** **0.001**
Relative SNR metric 1	0.019	0.136	0.034	**0.029**
Relative SNR metric 2	0.061	0.209	0.116	0.051
Number of echoes	0.024	**0.002**	**0.041**	**<0.001**
Interecho spacing (ms)	−0.099	0.151	−0.288	**0.012**
Repetition time (ms)	−0.002	**0.049**	−0.0030	**0.034**
Slice thickness (mm)	−0.014	0.439	*	*
Points in frequency direction	−0.001	0.138	−0.004	**0.014**
Frequency direction sampling	0.004	**0.014**	0.004	**0.021**
Minimum echo time	0.128	0.251	0.510	**0.043**

Bold text depicts p<0.05 on univariate analysis

*Slice thicknesses were the same in all exams for the restricted data set.