A multicenter study on the quantification of liver iron concentration in thalassemia patients by means of the MRI T_2* technique

Fengming Xu[1], Yuzhao Peng[1], Hanhong Xie[1], Bumin Liang[2], Gaohui Yang[3], Fanyu Zhao[4], Yu Liu[5] and Peng Peng[1,6]*

[1]Department of Radiology, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China, [2]School of International Education, Guangxi Medical University, Nanning, Guangxi, China, [3]Department of Hematology, The First Affiliated Hospital of Guangxi Medical University, Nanning, Guangxi, China, [4]Department of Radiology, People's Hospital of Guangxi Zhuang Autonomous Region, Nanning, Guangxi, China, [5]Department of Radiology, The Affiliated Tumor Hospital of Guangxi Medical University, Nanning, Guangxi, China, [6]NHC Key Laboratory of Thalassemia Medicine, Guangxi Medical University, Nanning, Guangxi, China

Objective: To investigate the feasibility and accuracy of quantifying liver iron concentration (LIC) in patients with thalassemia (TM) using 1.5T and 3T T_2* MRI.

Methods: 1.5T MRI T_2* values were measured in 391 TM patients from three medical centers: the T_2* values of the test group were combined with the LIC (LIC_F) provided by FerriScan to construct the curve equation. In addition, the liver 3T MRI liver T_2* data of 55 TM patients were measured as the 3T group: the curve equation of 3T T_2* value and LIC_F was constructed.

Results: Based on the test group LIC_F (0.6−43 mg/g dw) and the corresponding 1.5T T_2* value, the equation was $LIC_F = 37.393T_2*^{\wedge}(-1.22)$ ($R^2 = 0.971$; $P < 0.001$). There was no significant difference between $LIC_{e-1.5T}$ and LIC_F in each validation group ($Z = -1.269, -0.977, -1.197$; $P = 0.204, 0.328, 0.231$). There was significant consistency (Kendall's $W = 0.991, 0.985, 0.980$; all $P < 0.001$) and high correlation ($r_s = 0.983, 0.971, 0.960$; all $P < 0.001$) between the two methods. There was no significant difference between the clinical grading results of $LIC_{e-1.5T}$ and LIC_F in each validation group ($\chi^2 = 3.0, 4.0, 2.0$; $P = 0.083, 0.135, 0.157$), and there was significant consistency between the clinical grading results (Kappa's K = 0.943, 0.891, 0.953; $P < 0.001$). There was no statistical correlation between the LIC_F (≥ 14 mg/g dw) and the 3T T_2* value of severe iron overload ($P = 0.085$). The LIC_F (2−14 mg/g dw) in mild and moderate iron overload was significantly correlated with the corresponding T_2* value ($r_s = -0.940$; $P < 0.001$). The curve equation constructed from LIC_F and corresponding 3T T_2* values in this range is $LIC_F = 18.463T_2*^{\wedge}(-1.142)$ ($R^2 = 0.889$; $P < 0.001$). There was no significant difference between LIC_F and LIC_{e-3T} in the mild to moderate range ($Z = -0.523$; $P = 0.601$), and there was a significant correlation ($r_s = 0.940$; $P < 0.001$) and significant consistency (Kendall's $W = 0.970$; $P = 0.008$) between them. LIC_{e-3T} had high diagnostic efficiency in the diagnosis of severe, moderate, and mild liver iron overload (specificity = 1.000, 0.909; sensitivity = 0.972, 1.000).

Conclusion: The liver iron concentration can be accurately quantified based on the 1.5T T_2* value of the liver and the specific LIC-T_2* curve equation. 3T T_2* technology can accurately quantify mild-to-moderate LIC, but it is not recommended to use 3T T_2* technology to quantify higher iron concentrations.

KEYWORDS

thalassemia, magnetic resonance imaging, iron overload, liver iron concentration, equation

1. Introduction

The liver is one of the major iron storage organs; LIC reflects the total iron load, which is an important clinical indicator for clinical monitoring, evaluation, and treatment of iron overload (1). Although the actual liver iron concentration provided by liver biopsy serves as the "gold standard" for clinical indicators, most scholars and medical centers prefer to use non-invasive MRI technology for LIC monitoring because biopsy provides only small samples and has the disadvantages of invasiveness and poor repeatability (2). The LIC (LIC_F), based on the MRI T_2/R_2 ($1000/T_2$) technique and reported by FerriScan (Resonance Health Limited, Burswood, WA, Australia), has been certified by the Food and Drug Administration (FDA) of the United States and has high reliability (3). However, there are many limitations to this technique (4): it requires patient-related MRI T_2/R_2 data to be sent to FerriScan for off-site post-processing and analysis. The off-site sending of patient data requires approval from the relevant center and involves time costs that will prolong the time to obtain LIC results. The additional cost of the analysis will also increase the cost of LIC monitoring at one time, which results in the use of FerriScan technology for liver iron quantification being limited to a few large medical centers, and the possibility of regular or long-term LIC monitoring in patients being greatly reduced.

The T_2^* technique, based on the gradient recalled echo (GRE) imaging sequence of MRI has been established as a non-invasive standard for quantifying tissue iron levels (5–7). Many centers have been using the T_2^* relaxation method and corresponding software technology to measure organ relaxation parameters, such as T_2^* and R_2^* ($1000/T_2^*$) values, to indirectly obtain the estimated value of organ iron concentration (8). Some studies have explored the relationship between liver 1.5T T_2^*/R_2^* and LIC in patients with iron overload and constructed the corresponding LIC-T_2^*/R_2^* curve equation (8–13). However, the equations constructed in these studies, which were partly based on liver iron concentrations obtained from small biopsy samples, have not been validated in multiple centers, and their reliability needs to be verified. Moreover, most of the current studies in this area are based on 1.5T MRI, while the studies based on 3T MRI are few and limited. many studies only discuss the correlation between 3T T_2^*/R_2^* and LIC and whether the diagnosis is liver iron overload, but there is less analysis of the cutoff value of clinical classification of mild, moderate, and severe iron overload (14, 15).

The aim of this study was to investigate the relationship between liver 1.5T, 3T T_2^* values and LIC_F in thalassemia (TM) patients based on large sample size and multicenter data and also to investigate the feasibility, reliability, and accuracy of 1.5T and 3T MRI T_2^* techniques in quantifying LIC in TM patients.

2. Materials and methods

2.1. Research data

Liver 1.5T MRI T_2^* data of thalassemia patients from three medical centers from January 2014 to June 2022 were retrospectively analyzed: 273 patients from the First Affiliated Hospital of Guangxi Medical University (Center 1), 54 from the Guangxi Zhuang Autonomous Region Ethnic Hospital (Center 2), and 64 from the Guangxi Medical University Affiliated Tumor Hospital (Center 3). In total, 13 patients underwent 1.5T MRI T_2^* liver scans at three centers (within 24 h of the same patient being scanned at different centers). In addition, 3T liver MRI T_2^* data of 55 TM patients from center 1 were collected. The inclusion criteria were: (1) patients diagnosed with thalassemia by genetic diagnosis, with a history of regular or irregular blood transfusion; (2) age ≥ 9 years old; (3) MRI for thalassemia was performed with liver T2 (as required by FerriScan) and T_2^* sequences. Exclusion criteria were: (1) the image data artifacts were large and did not meet the measurement requirements; (2) patients were complicated with other chronic liver diseases or neoplastic diseases. This study was conducted in accordance with the principles of the Declaration of Helsinki and approved by the Ethics Committee of the First Affiliated Hospital of Guangxi Medical University (2022-E457-01).
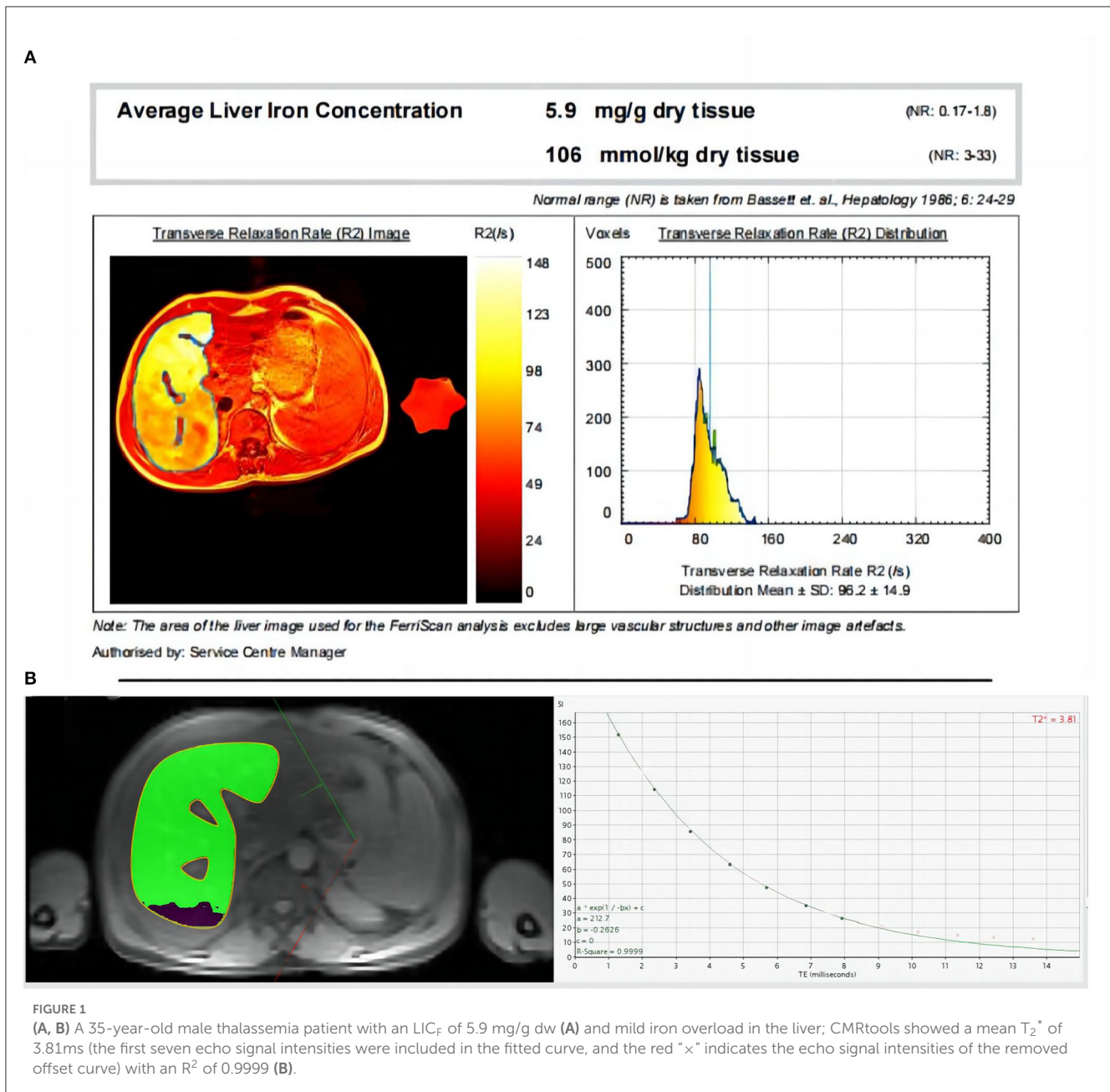
2.2. MR scanning methods

1.5T data: a Siemens 1.5T MRI scanner (MAGNETOM Avanto Fit & Altea, Siemens Healthcare, Erlangen, Germany), a Philips 1.5T MRI scanner (Achieva, Philips Medical Systems, Best, Netherlands), and the body coil were used. The images reported by FerriScan use a free-breathing two-dimensional multilayer spin-echo pulse sequence: rollover angle = 90°, repeat time (TR) = 1,000 ms, echo time (TE) = 6.0, 9.0, 12.0, 15.0, 18.0 ms, matrix = 256 × 256 mm^2, layer thickness = 5 mm, FOV = 400 mm × 400 mm. The scan time was 15 min. GRE scan sequence: the same level above the hepatic hilum was scanned with one breath hold at the end of exhalation. rotation angle = 20°, TR = 200.00 ms, TE = 1.29, 2.35, 3.43, 4.6, 5.68, 6.85, 7.93, 9.1, 10.18, 11.35, 12.43, 13.6 ms, matrix = 256 × 256 mm^2, layer thickness = 10 mm, FOV = 400 mm × 400 mm.

3T data: a Siemens 3T MRI scanner was used (Verio, Siemens Healthcare, Erlangen, Germany). TR = 200 ms, TE = 0.97, 2.38, 3.79, 5.20, 6.61, 8.02, 9.43, 10.84, 12.25, 13.66, 15.07, 16.48 ms; Rotation angle = 20°, matrix = 64 × 128 mm^2, layer thickness = 10 mm, FOV = 200 mm × 400 mm.

2.3. Data measurement and analysis

The 1.5T T2 image data was sent to FerriScan for post-processing and analysis. The LIC_F used in the study was obtained from the final FerriScan report (Figure 1A). The overall technical procedure of this study is shown in Figure 2.

All T_2^* image data were measured using CMRtools (CMRtools/Thalassemia Tools 2014, Cardiovascular Imaging Solutions, London, UK). Measurement procedure: T_2^* image data were exported from the PACS system and imported into a personal computer with CMRtools software installed. The "Thalassemia" function of CMRtools was used to draw a roughly similar ROI according to the range of liver levels measured by FerriScan, avoiding the visible intrahepatic vessels and bile ducts at the same liver level. The delineated ROIs and the fitted T_2^* values were then displayed in the post-processing software. The truncation method

FIGURE 1
(A, B) A 35-year-old male thalassemia patient with an LIC_F of 5.9 mg/g dw **(A)** and mild iron overload in the liver; CMRtools showed a mean $T_2{}^*$ of 3.81ms (the first seven echo signal intensities were included in the fitted curve, and the red "×" indicates the echo signal intensities of the removed offset curve) with an R^2 of 0.9999 **(B)**.

(14) was used to reject signal intensity values (SI) that deviated from the fitted curve one by one, and the $T_2{}^*$ value was recorded when the goodness of fit (R^2) ≥ 0.98 (Figure 1B).

2.4. Statistical methods

SPSS 26.0 statistical software package was used for statistical analysis. All test results were statistically analyzed according to the test significance α = 0.05. MedCalc 19.8.0 statistical software was used to analyze the consistency of $T_2{}^*$ values measured by different MRI scanners.

The 1.5T data of 273 patients from center 1 were divided into 191 patients in the test group and 82 patients in the validation group 1 according to the ratio of 7:3 by random number method. The 1.5T data of 54 patients from Center 2 were used as validation group 2. The 1.5T data of 64 patients from Center 3 were used as validation group 3. The 3T data of 55 cases from Center 1 were taken as the 3T group (3T data were only self-verified due to a small amount of data and were not grouped).

The intraclass correlation coefficient (ICC) was used, according to the ratio of 7:3. The 1.5T $T_2{}^*$ image data of 50 randomly selected patients from the test group ($n = 35$) and validation group 1 ($n = 15$) were measured to evaluate intra-observer and inter-observer agreement (independently performed by two radiologists with 5 years of experience in abdominal radiology diagnosis). The intra-observer ICC was calculated by comparing the $T_2{}^*$ values measured by observer A twice. The inter-observer ICC was calculated by comparing the $T_2{}^*$ values measured by observer B with the $T_2{}^*$ values measured by observer A. The ICC between the different

FIGURE 2
The overall technology roadmap.
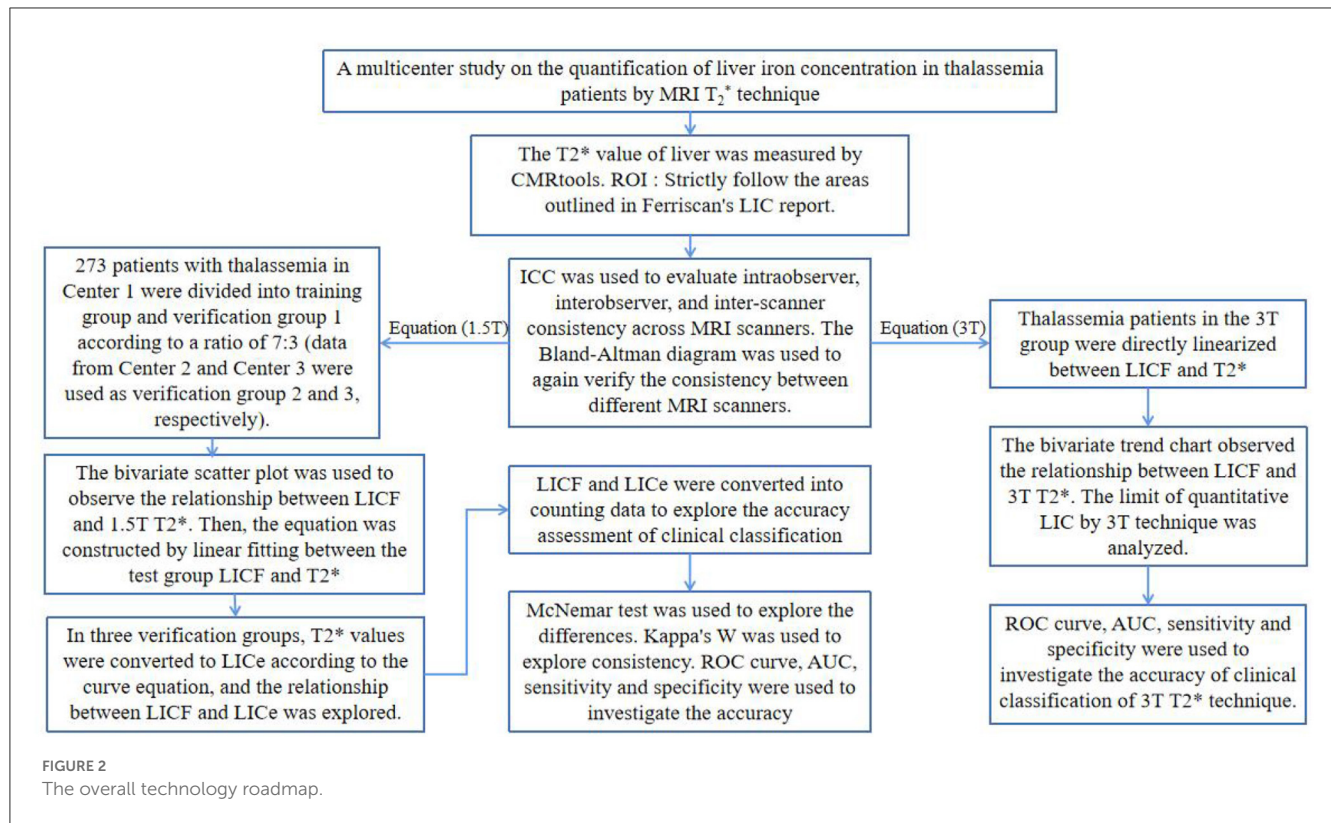
MRI scanners was calculated by comparing the 1.5T T_2^* values of 13 patients at three centers. "Two-way random" was selected for "model" and "absolutely consistent" was selected for "type." When ICC > 0.75 and $P < 0.05$, the measured T_2^* values were considered to be highly consistent. The remaining T_2^* value measurements were performed by observer A. Bland–Altman plots were used to analyze the consistency of the 1.5T T_2^* values of 13 patients measured by different MRI scanners.

The age, LIC_F, and the measured liver T_2^* values of each group did not follow the normal distribution by the normality test ($P < 0.05$). The interquartile range ($P_{25\%}$, $P_{75\%}$) and median (M) were used as statistical descriptors. By curve fitting, the calibration curve equation was constructed between the T_2^* value of the test group and the LIC_F of the 3T group. The T_2^* values in each validation group and the 3T group were converted into $LIC_{e-1.5T}$ and LIC_{e-3T} by the constructed curve equation. Both $LIC_{e-1.5T}$ and LIC_{e-3T} did not conform to the normal distribution ($P < 0.05$). The Wilcoxon signed-rank test was used to examine the difference, and $P > 0.05$ was considered as no significant difference. Kendall's W coefficient was used to examine consistency. If the consistency coefficient $W > 0.75$ and $P < 0.05$, it indicated a high degree of consistency. Spearman rank correlation analysis was used to examine the correlation; if the correlation coefficient of $|r_s| > 0.75$ and $P < 0.05$, it indicated a high correlation.

LIC_e and LIC_F were graded according to the severity of clinical liver iron overload, which was divided into normal (<1.8 mg/g dw), mild (1.8–7.0 mg/g dw), moderate (7.0–14.0 mg/g dw), and severe (>14.0 mg/g dw) liver iron overload (12). The McNemar test was used to examine the difference in clinical grading results. When $P > 0.05$, there was no significant difference between the two. Kappa's coefficient (Kappa's K) was used to examine the

consistency of clinical grading results. If K > 0.75 and $P < 0.05$, the two were highly consistent. With the LIC_F grading results as the reference standard, the area under the curve (AUC), sensitivity, and specificity of the receiver operating characteristic curve (ROC) was used to evaluate the accuracy of the $LIC_{e-1.5T}$ and LIC_{e-3T} clinical classification results in each validation group. The Youden index was used to evaluate the authenticity of the clinical classification results, and the cutoff values of different clinical classifications were recorded when the Youden index was maximum.

3. Results

3.1. Basic data information

In Center 1, there were 152 male subjects (55.68%) and 121 female subjects (44.32%), aged from 9 to 49 years (M = 19.00, $P_{25\%}$ = 12.00, $P_{75\%}$ = 27.00). In Center 2, there were 19 male subjects (35.19%) and 35 female subjects (64.81%), aged from 9 to 47 years (M = 13.00, $P_{25\%}$ = 12.00, $P_{75\%}$ = 21.25). In Center 3, there were 42 male subjects (65.63%) and 22 female subjects (34.37%), aged from 10 to 63 years (M = 15.50, $P_{25\%}$ = 11.00, $P_{75\%}$ = 25.00). In the 3T group, there were 33 males (60.00%) and 22 females (40.00%), aged from 9 to 25 years (M = 13.00, $P_{25\%}$ = 11.00, $P_{75\%}$ = 14.00).

3.2. Consistency analysis

The intra-observer ICC calculated based on the two measurements of observer A was 0.996 (95% confidence interval (CI) = 0.992–0.998) and $P < 0.001$. The inter-observer ICC between observers A and B was 0.978 (95%CI = 0.940–0.990), and

FIGURE 3
Bland–Altman plots of T_2^* measured between different MRI scanners. There is no data point exceeding 95%CI in **(A, B)**, or panel **(C)**. The arithmetic mean between the T_2^* values measured by the different MRI scanners was 0.06769 ($P = 0.4323$; 95%CI $= -0.1138$ to 0.2492), 0.009231 ($P = 0.9264$; 95%CI $= -0.2040$ to 0.2225), and 0.07692 ($P = 0.1832$; 95%CI $= -0.04173$ to 0.1956), respectively.
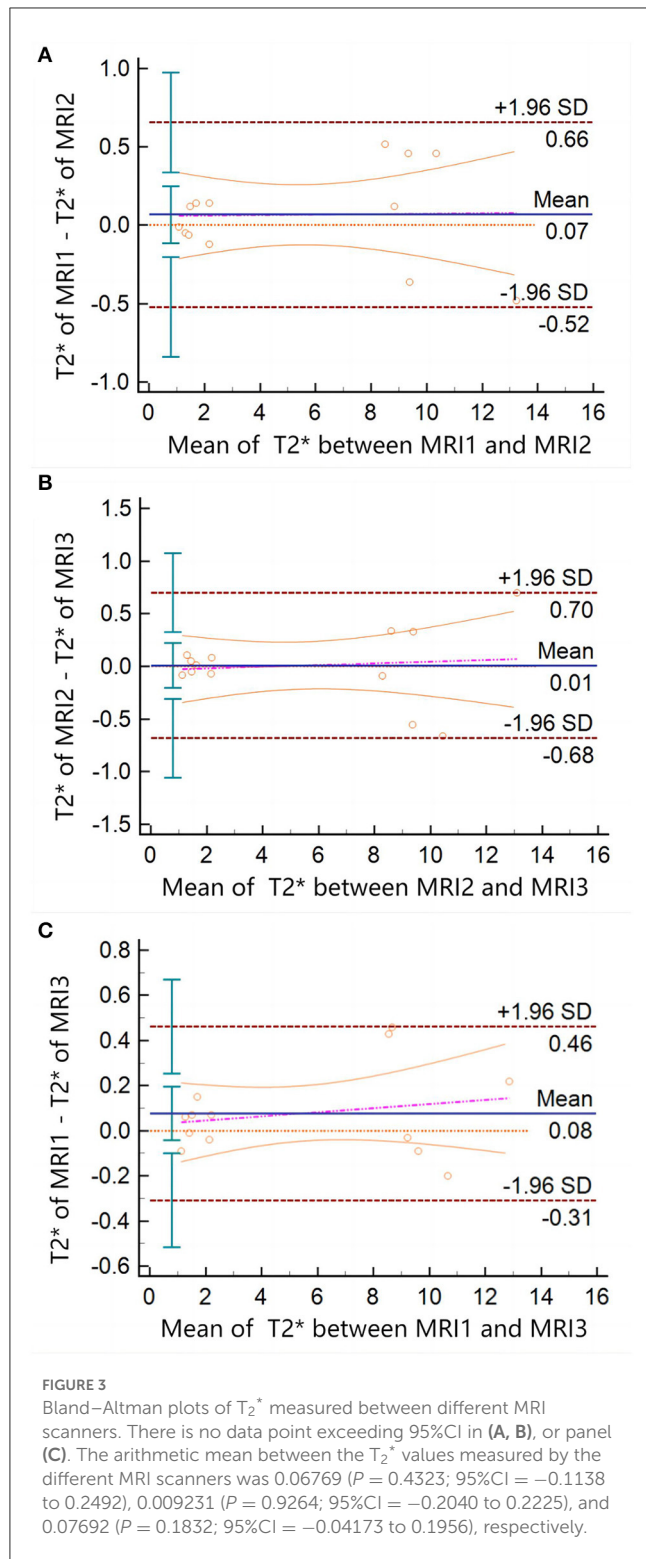
$P < 0.001$. The ICC between different MRI scanners was 0.999 (95%CI $= 0.998$–1) and $P < 0.001$. The results showed that the T_2^* values obtained by different MRI scanners were highly consistent with intra-observer and inter-observer. The Bland–Altman plots also showed significant consistency in the T_2^* values measured by different MRI scanners (Figure 3).

3.3. Formula construction

The trends of the relationships between LIC_F and T_2^* and LIC_F and LIC_e in each group are shown in Figure 4. According to the T_2^* value of the test group and LIC_F, the curve equation constructed was $LIC_F = 37.393 T_2^{*\wedge}(-1.22)$ ($R^2 = 0.971$, $P < 0.001$) (Figure 4A). According to the T_2^* value of the 3T group and LIC_F (<14 mg/g dw) of mild-to-moderate iron overload, the curve equation was $LIC_F = 18.463 T_2^{*\wedge}(-1.142)$ ($R^2 = 0.889$, $P < 0.001$) (Figure 4B). The liver T_2^*, LIC_F, $LIC_{e-1.5T}$, and LIC_{e-3T} values, and statistical description indicators of thalassemia patients in each group are shown in Table 1.
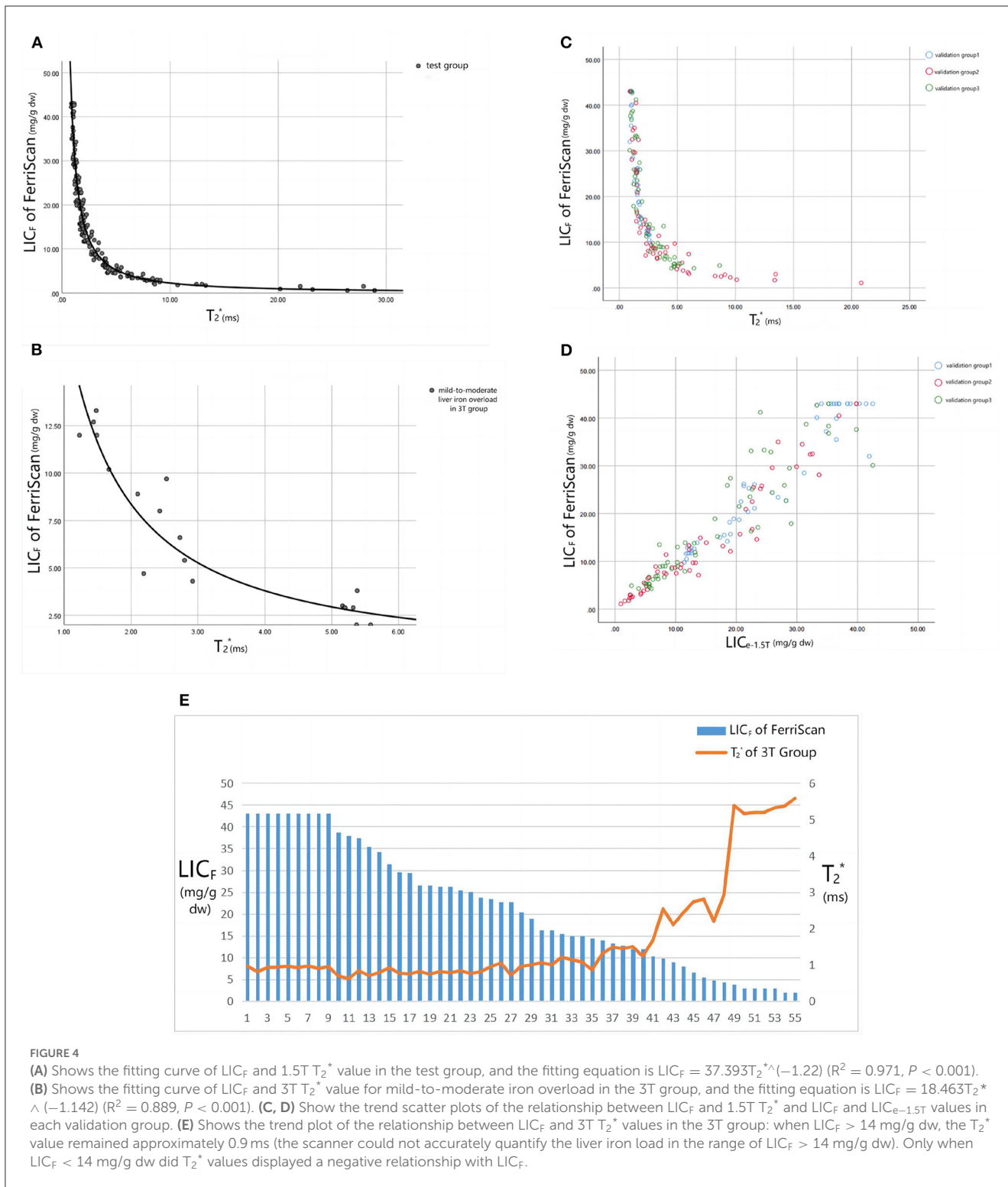
3.4. Formula verification

There was no significant difference between $LIC_{e-1.5T}$ and LIC_F in verification groups 1, 2, and 3 ($Z = -1.269$, -0.977, -1.197; $P = 0.204$, 0.328, 0.231). There was significant consistency between them (Kendall's $W = 0.991$, 0.985, 0.980; $P < 0.001$) and high correlation ($r_s = 0.983$, 0.971, 0.960; $P < 0.001$). There was no statistical difference between the clinical grading results of $LIC_{e-1.5T}$ and LIC_F in verification groups 1, 2, and 3 ($\chi^2 = 3.0$, 4.0, 2.0; $P = 0.083$, 0.135, 0.157), and the grading results are shown in Figure 5. There was significant consistency among the clinical grading results (Kappa's $K = 0.943$, 0.891, 0.953; $P < 0.001$). The accuracy indexes and corresponding cutoff values of the clinical classification results of $LIC_{e-1.5T}$ for the test group and each validation group are shown in Table 2.

In the 3T group, LIC_F in the mild-to-moderate range was significantly correlated with the corresponding T_2^* value ($r_s = -0.940$, $P < 0.001$). There was no statistically significant correlation between LIC_F and T_2^* values in the severe range ($P = 0.085$). Of the 36 patients with severe iron overload diagnosed by LIC_F, 1 (2.78%) was diagnosed with moderate iron overload by LIC_{e-3T}. Of the eight patients with moderate iron overload diagnosed by LIC_F, two (25%) were diagnosed with mild iron overload by LIC_{e-3T}. Of the 11 patients with mild iron overload diagnosed by LIC_F, one (9.10%) was diagnosed with moderate iron overload by LIC_{e-3T}. There was no significant difference between LIC_F and LIC_{e-3T} in the mild to moderate range ($Z = -0.523$, $P = 0.601$). There was a significant correlation between them ($r_s = 0.940$, $P < 0.001$) and significant consistency (Kendall's $W = 0.970$, $P = 0.008$). There was no significant difference between the clinical grading results of full range LIC_F ($\chi^2 = 1.333$, $P = 0.513$) and LIC_{e-3T} ($\chi^2 = 1.333$, $P = 0.513$). There was significant consistency among the clinical grading results (Kappa's $K = 0.860$, $P < 0.001$). The ROC curve of LIC_{e-3T} clinical grading results is shown in Figure 6, and the evaluation indices of diagnostic accuracy are shown in Table 2.

4. Discussion

As mentioned, magnetic resonance imaging has been widely considered the primary method for the non-invasive determination of liver iron concentration (6). FerriScan based on R_2 technology can generate reports including liver iron concentration, but it cannot be widely applied due to a variety of limiting factors,

FIGURE 4
(A) Shows the fitting curve of LIC_F and 1.5T T_2^* value in the test group, and the fitting equation is $LIC_F = 37.393 T_2^{*\wedge}(-1.22)$ ($R^2 = 0.971$, $P < 0.001$).
(B) Shows the fitting curve of LIC_F and 3T T_2^* value for mild-to-moderate iron overload in the 3T group, and the fitting equation is $LIC_F = 18.463 T_2^* \wedge (-1.142)$ ($R^2 = 0.889$, $P < 0.001$). **(C, D)** Show the trend scatter plots of the relationship between LIC_F and 1.5T T_2^* and LIC_F and $LIC_{e-1.5T}$ values in each validation group. **(E)** Shows the trend plot of the relationship between LIC_F and 3T T_2^* values in the 3T group: when $LIC_F > 14$ mg/g dw, the T_2^* value remained approximately 0.9 ms (the scanner could not accurately quantify the liver iron load in the range of $LIC_F > 14$ mg/g dw). Only when $LIC_F < 14$ mg/g dw did T_2^* values displayed a negative relationship with LIC_F.

especially for long-term and timely quantitative monitoring of LIC in patients (16). T_2^*/R_2^* image-based relaxometry and related measurement techniques have been developed in many centers. After years of research, many scholars have verified that the T_2^*/R_2^* value has an obvious linear relationship with LIC and partially constructed the curve equation of the relationship between them (8–13).

In the first quantitative study of liver iron overload using R_2^* relaxation measurement by Henninger et al. (8), the relevant parameters set were repetition time (TR) = 200 ms and initial echo time (TE) = 0.99 ms. Liver biopsy and MRI were performed on 17 patients with clinical suspicion of liver iron overload. The final regression model between R_2^* and LIC was constructed as follows: $LIC = 0.024 R_2^* + 0.277$, correlation coefficient = 0.926,

TABLE 1 Statistical descriptive indicators of liver T_2^*, LIC_F, $LIC_{e-1.5T}$, and LIC_{e-3T} values of thalassemia patients in each group.

Group	Variable name	Number (N)	Min. ~ maximum value	Quartile ($P_{25\%}$)	Quartile ($P_{75\%}$)	Median (M)
Test group	T_2^* (ms)	191	0.86–28.92	1.19	4.25	2
	LIC_F (mg/g dw)	191	0.60–43.00	5.7	29.5	14.9
Validation group1	T_2^* (ms)	82	0.90–25.44	1.0875	4.4375	1.85
	LIC_F (mg/g dw)	82	0.90–43.00	6.025	35.925	14.6
	$LIC_{e-1.5T}$ (mg/g dw)	82	0.72–42.52	6.08	33.755	17.6
Validation group2	T_2^* (ms)	54	0.95–20.83	1.5075	5.155	2.625
	LIC_F (mg/g dw)	54	1.10–43.00	5.325	23.175	9.55
	$LIC_{e-1.5T}$ (mg/g dw)	54	0.92–39.81	5.2751	22.9435	12.1968
Validation group3	T_2^* (ms)	64	0.90–28.92	1.4725	4.76	2.885
	LIC_F (mg/g dw)	64	0.60–43.00	5.225	24.925	9.8
	$LIC_{e-1.5T}$ (mg/gdw)	64	0.62–42.52	5.2751	23.3285	10.2663
3T group (mild-to-moderate range)	T_2^* (ms)	19	1.23–5.58	1.67	5.2	2.73
	LIC_F (mg/g dw)	19	2.00–13.30	2.9	10.2	5.4
	LIC_{e-3T} (mg/g dw)	19	2.59–14.58	2.81	10.28	5.86
3T group (moderate-to-severe range)	T_2^* (ms)	46	0.61–2.80	0.8	1.2075	0.94
	LIC_F (mg/g dw)	46	5.40–43.00	14.375	37.55	24.4
	LIC_{e-3T} (mg/g dw)	46	5.70–32.47	14.8875	23.82	19.81

FIGURE 5
Cluster plot of the constituent ratio of clinical grades of liver iron overload in LIC_F and $LIC_{e-1.5T}$ for each validation group. The results showed that the overall distribution of the two methods was almost the same, and there were slight differences in some distributions: in validation group 1, three patients (3.66%) with liver iron overload were classified as mild according to LIC_F classification, and as normal according to LIC_e classification. In validation group 2, two patients (3.70%) had liver iron overload classified as mild according to LIC_F and moderate according to LIC_e. In total, two patients (3.70%) had moderate liver iron overload according to LIC_F and severe liver iron overload according to LIC_e. In validation group 3, two patients (3.13%) had mild liver iron overload according to LIC_F classification and moderate liver iron overload according to LIC_e classification.
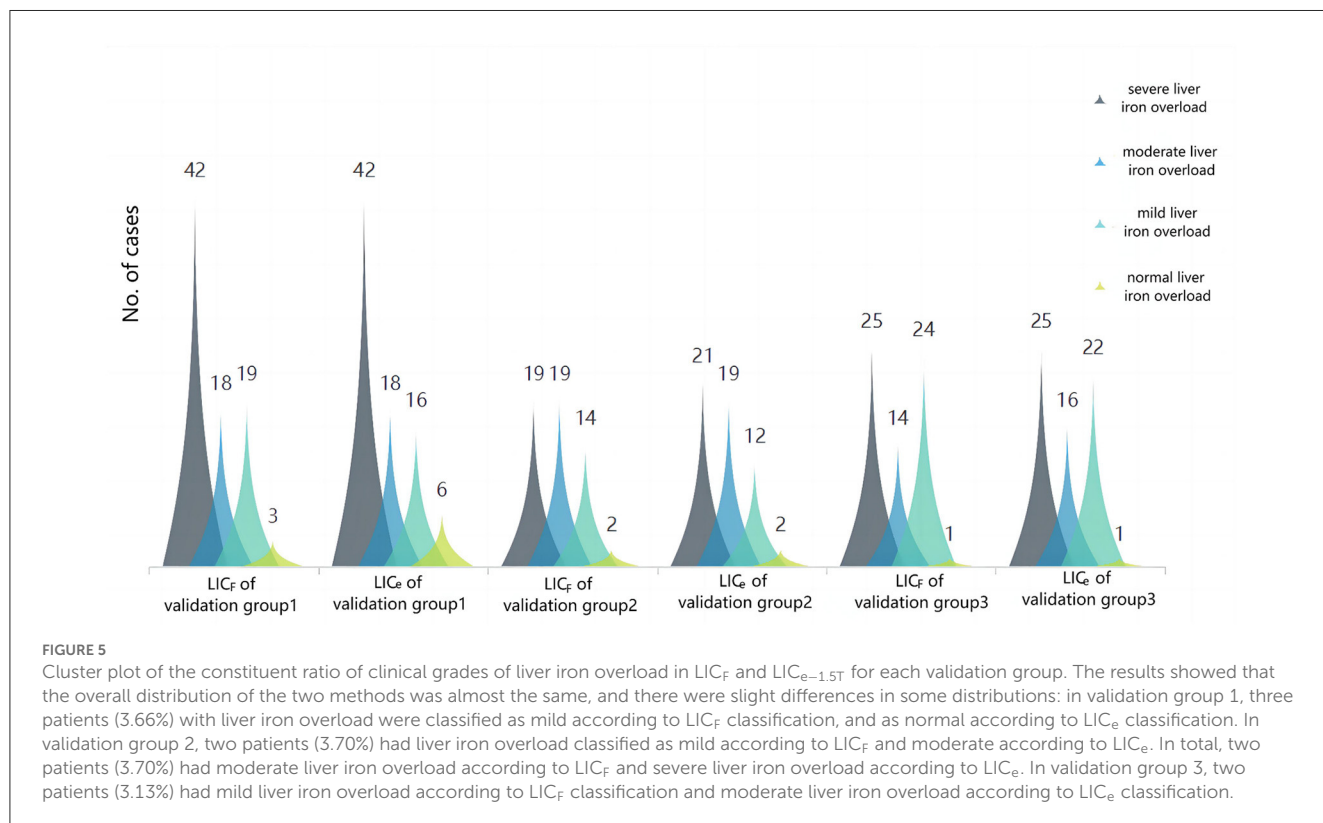
TABLE 2 Accuracy indicators and clinical classification cutoff values of $LIC_{e-1.5T}$ clinical classification results in the validation group.

Group	Liver iron overload grading	AUC	95%CI	P	Specificity	Sensitivity	Youden index	Cut-off values (mg/g dw)
Test group-$LIC_{1.5T}$	Severe vs. moderate groups	0.991	0.983–0.999	<0.0001	0.967	0.94	0.907	16.005
	Moderate vs. mild groups	0.942	0.904–0.980	<0.0001	0.929	1	0.929	7.49
	Mild vs. normal group	0.999	0.995–1.000	<0.0001	0.989	1	0.989	2.3
Validation group 1-$LIC_{1.5T}$	Severe vs. moderate groups	1	1.000–1.000	<0.0001	1	1	1	15.425
	Moderate vs. mild groups	0.969	0.927–1.000	<0.0001	0.937	1	0.937	7.485
	Mild vs. normal group	0.988	0.963–1.000	0.019	0.987	1	0.987	1.225
Validation group 2-$LIC_{1.5T}$	Severe vs. moderate groups	0.995	0.985–1.000	<0.0001	1	0.947	0.947	19.814
	Moderate vs. mild groups	0.95	0.985–1.000	<0.0001	0.95	1	0.95	6.123
	Mild vs. normal group	1	0.963–1.000	0.019	0.981	1	0.981	2.302
Validation group 3-$LIC_{1.5T}$	Severe vs. moderate groups	1	1.000–1.000	<0.001	0.974	1	0.974	13.22
	Moderate vs. mild groups	0.972	0.923–1.000	<0.001	0.975	0.958	0.933	7.199
	Mild vs. normal group	1	1.000–1.000	0.008	0.984	1	0.984	2.182
3T group-LIC_{3T}	Severe vs. moderate groups	0.997	0.988–1.000	<0.001	1	0.972	0.972	14.785
	Moderate vs. mild groups	0.977	0.920–1.000	<0.001	0.909	1	0.909	6.13

slope = 0.024 (mg/g) [95% CI = 0.013–0.024], intercept = 0.277 (mg/g) [95% CI = 0.328–2.49]. In an early study by Wood et al. (9), the set TE was increased from the initial 0.8 ms to 4.8 ms at 0.25 ms intervals in a breath hold, and TR = 25 ms. After MRI evaluation of 102 patients with liver iron overload (the biopsy-measured LIC was evenly distributed between 1.3 mg/gdw and 32.9 mg/gdw, and one patient had a HIC of 57.8 mg/gdw), the final LIC-R_2* regression equation was constructed as follows: the correlation coefficient was 0.97, the slope was 37.4 Hz/mg/gdw, and the y-intercept was 23.7 Hz. In an early study by Hankins et al. (10), TE = 1.1–17.3 ms (20 echoes) was set, and 43 patients (32 with sickle cell anemia, six with major β-thalassemia, five patients with bone marrow failure) underwent MRI examination and liver biopsy (LIC range = 0.6 mg Fe/g to 27.6 mg Fe/g). The final LIC-R_2* regression model was constructed as follows: the intercept was −454.85, the slope was 28.02 (P < 0.001), the R^2 was 0.72, and the correlation coefficient was 0.98. In an early study by Christoforidis et al. (11), MRI was performed on 94 patients with β-thalassemia

major with TE = 2.24–20.13 ms and TR = 200 ms. The relationship between liver–muscle ratio (MRI-LIC = 5–350 μmol/g) and R_2* (27.03–1,298.70 s⁻¹) was compared. The final LIC-R_2* regression model was $R_2* = 0.851$(MR-LIC) − 2.137 (correlation coefficient = 0.851). In the study by Garbowski et al. (12), TE was set as 0.93–16.0 ms. A total of 54 patients (36 cases of thalassemia major, seven cases of sickle cell anemia, four cases of myelodysplastic syndrome, three cases of Diamond-Blackfan anemia, two cases of red cell aplasia, two cases of pyruvate kinase deficiency anemia), and 31 healthy volunteers underwent liver biopsy (LIC = 1.7–42.3 mg/g dw) and MRI examination (R_2* range = 28.7–54.4 s⁻¹). The final regression models of LIC (biopsy) -T_2* and LIC (biopsy)-R_2* were constructed: (1) LIC = $31.94(T_2*)^{-1.014}$, 95%CI of coefficient = 27.8–36.7 (87–115%), 95%CI of index = −1.118–0.91(110–90%). (2) LIC = $0.029(R_2*)^{1.014}$, 95%CI of coefficient = 0.016–0.054 (55–186%), 95%CI of index = 0.910–1.118 (90–110%). Garbowski et al. (12) also constructed the correction relationship between LIC (Ferriscan)-R_2* and LIC-T_2*: (1) R_2-LIC = $0.83T_2*$ − $LIC^{1.04}$,
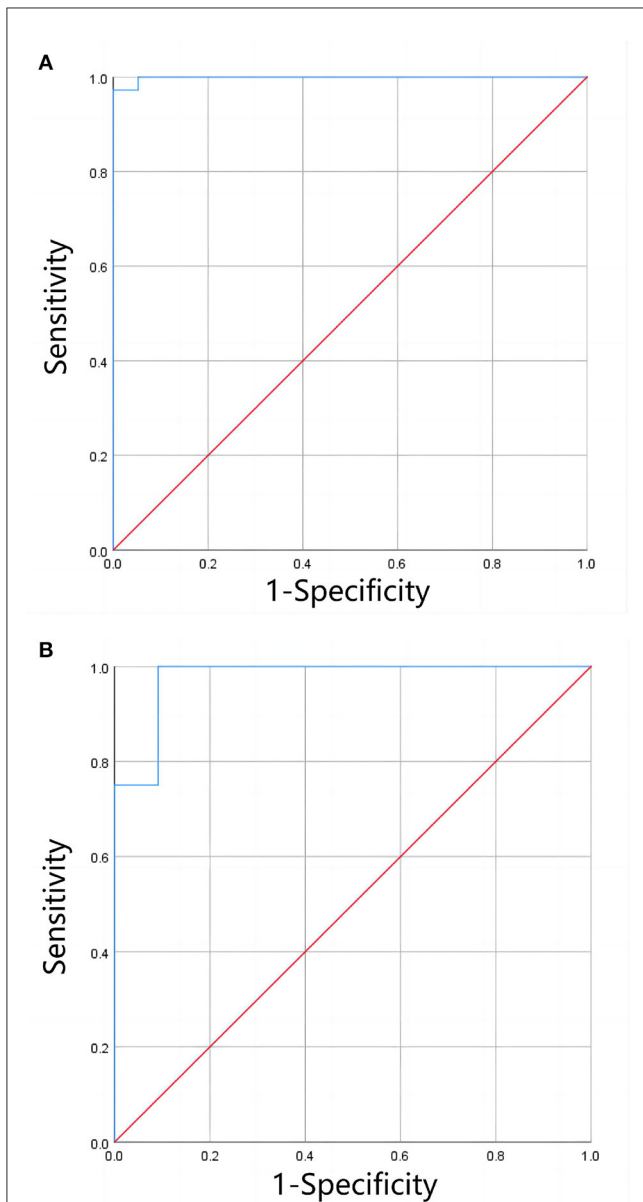
FIGURE 6
(A, B) Shows the ROC curves for clinical grading of moderate and severe liver iron overload, and of mild and moderate liver iron overload by LIC_{e-3T}, respectively. The results showed that LIC_{e-3T} had a high diagnostic efficiency for clinical grading of liver iron overload when only clinical grading diagnosis was performed (without precise quantification of LIC).

95%CI of coefficient = 0.96 ~ 1.11, 95%CI of index = 0.55 ~ 1.29. (2) R_2-LIC = $0.87R_2^*$ − LIC − 0.55, 95%CI of slope = 0.74–0.99, 95%CI of intercept = −0.01–1.19.

In this study, based on large sample size, multicenter validation, and complete statistical analysis, the equation $LIC_F = 37.393T_2^{*\wedge}$ (−1.22) was proposed to quantify LIC from 1.5T MRI T_2^*. For the 3T MRI quantification of liver iron overload, the relationship $LIC_F = 18.463T_2^{*\wedge}(-1.142)$ was proposed to quantify LIC in patients with mild-to-moderate iron overload. However, it is still not possible to accurately quantify LIC in patients with severe liver iron overload at 3T field strength. This also suggests that

the 3T T_2^* technique should be avoided for the quantification of LIC in patients with severe iron overload, and 1.5T or other methods should be used instead. This conclusion is similar to that of d'Assignies (15). It is worth noting that this study found that when 3T T_2^* technology was used to quantify liver iron overload, although the T_2^* value of patients with severe liver iron overload was almost maintained at 0.9 ms, it was impossible to further accurately quantify the LIC value; however, if only the clinical classification of liver iron overload was performed, that is, only the classification of mild-to-moderate and severe iron overload was performed, the classification of liver iron overload would have high diagnostic efficacy.

The slope of the calibration curve proposed by different studies is different, and for the LIC calculated by the earlier calibration curve, Garbowski et al. (12) also proposed further calibration coefficients to calibrate the final LIC. The specific reasons for the differences are analyzed as follows: (1) the previous T_2^*/R_2^*-LIC calibration curve equation was based on liver biopsies, such as the study by Henninger et al. (8), Wood et al. (9), Hankins et al. (10), Christoforidis et al. (11), and Garbowski et al. (12). Although LIC provided by liver biopsy has been used as the "gold standard" for a long time, the materials and methods used in the process of liver biopsy, and the heterogeneity of iron in the liver will lead to differences between different studies; (2) The sample size used in some studies is small. A small sample size will not only increase the sampling error, but also limit the range of LIC used, and the final fitted calibration curve equation cannot be extended to quantify a wider range of LIC. For example, the LIC of 17 patients collected by Henninger et al. (8) through liver biopsy ranged from 0.917 mg/g to 11.646 mg/g. The authors believe that because the range of LIC studied is small, it is not appropriate to use the corresponding calibration curve to quantify a wider range of LIC; (3) Different models used to measure T_2^*/R_2^* will directly lead to differences in the final LIC. For example, Wood et al. (9) used an offset model, while Garbowski et al. (12) used a truncated model to measure R_2^*. Garbowski et al. (12) proposed that the R_2^* value measured by the offset model was high, while the R_2^* value measured by the truncated model was low, and emphasized the importance of using the appropriate measurement model to quantify T_2^*/R_2^* and the appropriate analysis techniques to construct the curve equation in clinical practice; (4) The high iron concentration corresponds to a very low T_2^* value, thus it is necessary to set a very short minimum echo time for more accurate measurement. However, due to the differences in the technique and scanning sequence used by different research centers, the different minimum echo time set obviously limits the lowest T_2^* value, which is the maximum value of LIC measured by the center. Some studies have shown that LIC with severe iron overload in the liver should be measured in combination with the signal intensity ratio between the liver and the paravertebral muscles (SIR) (17).

The shortcomings of this experiment are as follows: (1) the definition of the ROI. In this study, the delineated T_2^* image ROI was as close as possible to the T_2 image ROI delineated by FerriScan, but the artificial delineation of ROI was susceptible to various subjective and objective factors, and measurement error was inevitable; (2) In this study, the proportion of patients with moderate or severe liver iron overload was relatively large, and the proportion of patients with mild iron concentration was relatively

small, which had a certain bias. However, in general, this study was analyzed with large sample size and was validated in multiple centers, which makes the results reliable; (3) In this retrospective study, LIC_F was used as the reference standard, in other words, the T_2^*-LIC calibration equation was constructed under the assumption that FerriScan based on T_2/R_2 technique was very reliable. Therefore, the equations obtained should not be extended to other techniques or organs for calculating iron concentration; (4) The sample size of 3T data is small, and it is difficult to perform grouping verification; (5) Accurate quantification of severe liver iron load by 3T T_2^* technique still cannot be achieved due to the technical limitation of the 3T scanning sequence in this study.

5. Conclusion

This study explored the relationship between liver T_2^* value and LIC_F provided by FerriScan in patients with thalassemia, and the related curve equation was constructed. After measuring the liver 1.5T T_2^* value, the liver iron concentration could be accurately quantified, and liver iron concentration in patients with iron overload could be better monitored. An important reference for timely and better formulation of appropriate diagnosis and treatment plans can be made. 3T T_2^* can be used to quantify liver iron concentration in patients with mild-to-moderate liver iron overload, while 1.5T T_2^* or other methods are recommended for patients with severe liver iron overload.

Data availability statement

The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author.

Ethics statement

The studies involving human participants were reviewed and approved by the Ethics Committee of the First Affiliated Hospital of Guangxi Medical University (2022-E457-01). Written informed consent from the participants' legal guardian/next of kin was not required to participate in this study in accordance with the national legislation and the institutional requirements.

Author contributions

Funding

Acknowledgments

Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

1. Labranche R, Gilbert G, Cerny M, Vu KN, Soulières D, Olivié D, et al. Liver iron quantification with MR imaging: a primer for radiologists. *Radiographics.* (2018) 38:392–412. doi: 10.1148/rg.2018170079

2. Xu F, Yi J, Liang B, Tang C, Feng Q, Peng P. Comparative study on the measurement of liver LICdw between Ferriscan and T_2^* based licdw obtained by different software's. *Mediterr J Hematol Infect Dis.* (2022) 14:e2022072. doi: 10.4084/MJHID.2022.072

3. Jhaveri KS, Kannengiesser SAR, Ward R, Kuo K, Sussman MS. Prospective Evaluation of an R_2^* Method for Assessing Liver Iron Concentration (LIC) against ferriscan: derivation of the calibration curve and characterization of the nature and source of uncertainty in the relationship. *J Magn Reson Imag.* (2019) 49:1467–74. doi: 10.1002/jmri.26313

4. Scott B. Reeder DH. *System and Method for Assessing Susceptibility of Tissue Using Magnetic Resonance Imaging.* USA: Wisconsin Alumni Research Foundation. (2018).

5. Henninger B, Plaikner M, Zoller H, Viveiros A, Kannengiesser S, Jaschke W, et al. Performance of different Dixon-based methods for MR liver iron assessment in comparison to a biopsy-validated R_2^* relaxometry method. *Eur Radiol.* (2021) 31:2252–62. doi: 10.1007/s00330-020-07291-w

6. Khadivi Heris H, Nejati B, Rezazadeh K, Sate H, Dolatkhah R, Ghoreishi Z, et al. Evaluation of iron overload by cardiac and liver T_2^* in β-thalassemia: Correlation with serum ferritin, heart function and liver enzymes. *J Cardiov Thor Res.* (2021) 13:54–60. doi: 10.34172/jcvtr.2021.18

7. Henninger B, Alustiza J, Garbowski M, Gandon Y. Practical guide to quantification of hepatic iron with MRI. *Eur Radiol.* (2020) 30:383–93. doi: 10.1007/s00330-019-06380-9

8. Henninger B, Zoller H, Rauch S, Finkenstedt A, Schocke M, Jaschke W, et al. R_2^* relaxometry for the quantification of hepatic iron overload: biopsy-based calibration and comparison with the literature. *RoFo.* (2015) 187:472–9. doi: 10.1055/s-0034-1399318

9. Wood JC, Enriquez C, Ghugre N, Tyzka JM, Carson S, Nelson MD, et al. MRI R2 and R2* mapping accurately estimates hepatic iron concentration in transfusion-dependent thalassemia and sickle cell disease patients. *Blood.* (2005) 106:1460–5. doi: 10.1182/blood-2004-10-3982

10. Hankins JS, McCarville MB, Loeffler RB, Smeltzer MP, Onciu M, Hoffer FA, et al. R_2^* magnetic resonance imaging of the liver in patients with iron overload. *Blood.* (2009) 113:4853–5. doi: 10.1182/blood-2008-12-191643

11. Christoforidis A, Perifanis V, Spanos G, Vlachaki E, Economou M, Tsatra I, et al. MRI assessment of liver iron content in thalassamic patients with three different protocols: comparisons and correlations. *Eur J Haematol.* (2009) 82:388–92. doi: 10.1111/j.1600-0609.2009.01223.x

12. Garbowski MW, Carpenter JP, Smith G, Roughton M, Alam MH, He T, et al. Biopsy-based calibration of T_2^* magnetic resonance for estimation of liver iron concentration and comparison with R2 Ferriscan. *J Cardiov Magn Reson.* (2014) 16:40. doi: 10.1186/1532-429X-16-40

13. Ouederni M, Ben Khaled M, Mellouli F, Ben Fraj E, Dhouib N, Yakoub IB, et al. Myocardial and liver iron overload, assessed using T_2^* magnetic resonance imaging with an excel spreadsheet for post processing in Tunisian thalassemia major patients. *Ann Hematol.* (2017) 96:133–9. doi: 10.1007/s00277-016-2841-5

14. Anwar M, Wood J, Manwani D, Taragin B, Oyeku SO, Peng Q. Corrigendum to "Hepatic iron quantification on 3 tesla (3 T) magnetic resonance (MR): technical challenges and solutions". *Radiol Res Pract.* (2018) 2018:8142478. doi: 10.1155/2018/8142478

15. d'Assignies G, Paisant A, Bardou-Jacquet E, Boulic A, Bannier E, Lainé F, et al. Non-invasive measurement of liver iron concentration using 3-Tesla magnetic resonance imaging: validation against biopsy. *Eur Radiol.* (2018) 28:2022–2030. doi: 10.1007/s00330-017-5106-3

16. Healy GM, Kannengiesser SAR, Espin-Garcia O, Ward R, Kuo KHM, Jhaveri KS. Comparison of Inline R_2^* MRI versus FerriScan for liver iron quantification in patients on chelation therapy for iron overload: preliminary results. *Eur Radiol.* (2021) 31:9296–305. doi: 10.1007/s00330-021-08019-0

17. Fernandes JL, Fioravante LAB, Verissimo MP, Loggetto SR. A free software for the calculation of T_2^* values for iron overload assessment. *Acta radiol.* (2017) 58:698–701. doi: 10.1177/0284185116666416