The use of appropriate calibration curves can correct the systematic differences between softwares in hepatic R2* estimation

Antonella Meloni, Hugh Young Rienhoff, Amber Jones, Aessia Pepe, Massimo Lombardi, and John C Wood

1CMR Unit, Fondazione G. Monasterio CNR-Regione Toscana and Institute of Clinical Physiology, Pisa, Italy, 2FerroKin BioSciences, Inc, San Carlo, California, United States, 3Division of Cardiology, Children’s Hospital Los Angeles, Los Angeles, California, United States

Introduction. Liver R2* can be used as a surrogate for liver iron concentration (LIC) in iron overloaded subjects. Two different signal decay models, truncated exponential [2,3] and exponential plus constant [1,4], have been validated for R2* estimation and calibrated to liver biopsy [5]. However, reported calibration curves for these two analysis methods differ by 15%.

Our aim was to evaluate if the different fitting models yielded significantly different R2* estimates and if these differences disappeared once R2* estimates were converted to LIC units using method-appropriate calibration curves.

Materials and methods. A single-center (N=45) and a multi-center cohort (N=47) of patients were used. Gradient echo images optimized for R2* estimation were collected at each site according to local clinical practice. R2* values were generated using the CMRTools introduced by the Pennel’s group (truncated exponential model; R2*_{Pennell}) and custom Matlab code (exponential plus constant model; R2*_{Wood}). R2* values were converted to dry weight liver iron concentration using calibrations published by Garbowski (equation 1) [5] and Wood (equation 2) [1], respectively:

\[ \text{LIC}_{\text{Pennell}} = 0.03 \times \text{R2*}_{\text{Pennell}} + 0.7 \quad (\text{equation 1}) \]
\[ \text{LIC}_{\text{Wood}} = 0.0254 \times \text{R2*}_{\text{Wood}} + 0.2 \quad (\text{equation 2}) \]

Bland Altman analysis was performed with respect to both R2* and LIC estimates.

Results. For the single-center cohort the R2*_{Pennell} values ranged from 28.1 to 1219.5 s^{-1}, with a mean value of 367.5 ± 380.6 s^{-1}. The R2*_{Wood} values ranged from 29.7 to 1344.9 s^{-1}, with a mean value of 422.3 ± 445.6 s^{-1}. Figure 1a shows R2*_{Wood} values as a function of R2*_{Pennell} values. The line of best fit had a slope of 1.160 ± 0.024, significantly different from 1 (P<0.0001), an intercept of -3.992 ± 12.723 s^{-1}, and an R-squared value of 0.982. Figure 1b is the Bland-Altman plot. Results were unbiased for R2* < 300 s^{-1}, but R2* values obtained using exponential plus constant were systematically larger at higher R2* and the difference increased with increasing values. The mean difference was 54.7 ± 85.7 s^{-1} (95% confidence intervals of the difference: lower 28.9 and upper: 80.5 s^{-1}), corresponding to a percentage difference in R2* values of 9.1 ± 11.8%. The bias was eliminated following conversion to LIC units. The LIC_{Pennell} values ranged from 1.5 to 37.3 mg/g dry, with a mean value of 11.7 ± 11.4 mg/g dry. The LIC_{Wood} values ranged from 0.95 to 34.4 mg/g dry, with a mean value of 10.9 ± 11.3 mg/g dry. The line of best fit had a slope of 0.982 ± 0.020, not significant different from the unity (P=0.382), an intercept of -0.589 ± 0.334 mg/g dry (Figure 2a). Figure 2b is the Bland-Altman plot. LIC_{Pennell} values were systematically higher for LIC’s up to 10 mg/g and the two estimates were unbiased thereafter. The mean difference was -0.8 ± 1.5 mg/g dry (95% confidence intervals of the mean difference: lower -1.3 and upper: -0.3 mg/g dry). 95% confidence intervals of the individual LIC estimates were -3.8 – 2.2 mg/g dry weight

Similar differences in R2* estimation were found in the multi-center cohort and the conversion of R2* values to LIC units again removed the disparity.

Conclusion. R2* values vary with post-processing method but yield statistically identical LIC values when technique-appropriate calibration curves are used. LIC, rather than R2* values, should be reported and compared across studies.

References.