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

Antonella Meloni<sup>1</sup>, Hugh Young Rienhoff<sup>2</sup>, Amber Jones<sup>2</sup>, Aessia Pepe<sup>1</sup>, Massimo Lombardi<sup>1</sup>, and John C Wood<sup>3</sup>

<sup>1</sup>CMR Unit, Fondazione G. Monasterio CNR-Regione Toscana and Institute of Clinical Physiology, Pisa, Italy, <sup>2</sup>FerroKin BioSciences, Inc, San Carlo, California, United States, <sup>3</sup>Division 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 [1]. 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.

<u>Materials and methods.</u> 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:

$$LIC_{Pennel} = 0.03 * R2 *_{Pennell} + 0.7$$
 (equation 1)  $LIC_{Wood} =$ 

$$LIC_{Wood} = 0.0254 * R2 *_{Wood} + 0.2$$
 (equation 2)

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

Results. For the single-center cohort the R2\*<sub>Pennell</sub> values ranged from 28.1 to 1219.5 s<sup>-1</sup>, with a mean value of 367.5  $\pm$  380.6 s<sup>-1</sup>. The R2\*<sub>Wood</sub> 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  $1.160 \pm 0.024$ , significantly different from 1 (P<0.0001), an intercept of -3.992 ± 12.723 s<sub>-1</sub>, and an R-squared value of 0.982. Figure 1b is the Bland-Altman plot. Results were unbiased for  $R2^* <$ 300 s<sup>-1</sup>, 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  $\pm$  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  $9.1 \pm 11.8\%$ . The bias was eliminated following conversion to LIC units. The LIC<sub>Pennell</sub> values ranged from 1.5 to 37.3 mg/g dry, with a mean value of  $11.7 \pm 11.4$  mg/g dry. The LIC<sub>Wood</sub> values ranged from 0.95 to 34.4 mg/g dry, with a mean value of  $10.9 \pm 11.3$  mg/g dry. The line of best fit



had a slope of  $0.982 \pm 0.020$ , not significant different from the unity (P=0.382), an intercept of  $-0.589 \pm 0.334$  mg/g dry (Figure 2a). Figure 4b is the Bland-Altman plot. LICPennell values were systematically higher for LIC's up to 10 mg/g and the two estimates were unbiased thereafter. The mean difference was  $-0.8 \pm 1.5$  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.

<u>Conclusion</u>. 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.**

[1] Wood JC et al. Blood 2005;106(4):1460-1465. [2] Tanner MA et al. Haematologica 2006;91(10):1388-1391. [3] Kirk P et al. Circulation 2009;120(20):1961-1968. [4] Meloni A et al. J Magn Reson Imaging 2011;33(2):348-355. [5] Garbowski MJ et al. Blood 2009.