Electronic Supplementary Material

Artificial intelligence analysis of magnetic particle imaging for islet transplantation in a mouse model

Hasaan Hayat[1,2#], Aixia Sun[1,3#], Hanaan Hayat[2,4], Sihai Liu[1,3,5], Nazanin Talebloo[1,6], Cody Pinger[4], Jack Owen Bishop[1,7], Mithil Gudi[1,2], Bennett Francis Dwan[1,8], Xiaohong Ma[1,3,9], Yanfeng Zhao[1,3,9], Anna Moore[1,3], Ping Wang[1,3*]

[1] Precision Health Program, Michigan State University, East Lansing, MI, USA.

[2] Lyman Briggs College, Michigan State University, East Lansing, MI, USA.

[3] Department of Radiology, College of Human Medicine, Michigan State University, East Lansing, MI, USA.

[4] Institute for Quantitative Health Science and Engineering, Department of Biomedical Engineering, Michigan State University, East Lansing, MI, USA

[5] Department of Orthopedics, Beijing Charity Hospital, Capital Medical University, Beijing, China.

[6] Department of Chemistry, College of Natural Science, Michigan State University, East Lansing, MI, USA.

[7] Department of Neuroscience, College of Natural Science, Michigan State University, East Lansing, MI, USA.

[8] College of Natural Science, Michigan State University, East Lansing, MI

[9] Department of Radiology, National Cancer Center/Cancer Hospital, Chinese Academy of Medical Sciences and Peking Union Medical College, Beijing, China.

[#] Contributed equally to the study.

*Address correspondence to: Ping Wang, Precision Health Program, Michigan State University, 766 Service Road, Rm. 2020, East Lansing, MI 48823, tel. (517)-353-3817, email: wangpin4@msu.edu

Supplementary Methods

Ratio model

As a result, we obtained the sum of pixel intensities of the main ROI and reference fiducial marker. We also have the known TIV of the reference fiducial marker which was calculated and recorded by hand during preparation of the solution and phantom sample for imaging. To calculate the unknown TIV of a main ROI we used the following equation:

$$Ratio: \frac{\sum Pixel\ Intensity\ of\ Fiducial\ Marker}{TIV\ Fiducial\ Marker}$$

$$TIV\ Main\ ROI = \frac{\sum Pixel\ Intensity\ of\ Main\ Blob}{Ratio}$$

In order to predict the TIV of the main ROI using the aforementioned ratio method, we calculated the ratio of the sum of pixel intensities in fiducial marker to its TIV in the fiducial marker region, and then we used this ratio to predict the TIV of the main ROI using the formula above, because we extracted the total pixel sum of the ROI from the *K-means++* segmentation algorithm and solved for the unknown TIV through this ratio. Although the ratio method is the

current manual approach to predicting TIV of ROI within an MPI scan by imaging specialists, there are limitations to this approach including its inability to operate with a high degree of specificity and accuracy since the calculated ratio is only relative to a single fiducial marker region's pixel sum and known TIV, and therefore does not account for variations in region size and TIV fluctuations. Therefore, this method may not render a standardized approach to TIV prediction of an ROI within an MPI scan and limits the viability and extent of analysis of the *K-means++* segmentation results (Supplemental Fig. 1 A).

Standard curve model

Due to limitations in the Ratio model for predicting the unknown TIV of the main ROI within an MPI scan, we propose another approach using the *K-means++* segmentation. In order to create a more robust, standardized model for TIV prediction, we generated MPI Images with four reference fiducial markers with known TIV of increasing value instead of just one, along with the main ROI in the center (Supplemental Fig. 1). Using these four reference markers, we created a *linear regression* model and predicted the TIV of the main ROI using this model. The linear regression model was created using the sum of pixel intensities of fiducial markers as the independent variable and total iron value of these blobs as the dependent variable. The algorithm generated a standard curve (SC) from these values. From this, using the known total pixel sum of the main ROI, we used the standard curve to predict the corresponding TIV of the main ROI. This approach certainly posed segmentation challenges. Instead of the previous two ROIs in the ratio method, we now had to identify five regions of interest. To accomplish this, we devised an entirely region-based *K-means++* approach by exploiting the center position of the

image. Since reference blobs are always located in the corners and main blob is near the center, as per our standard setup for MPI scans in this study, we found our five regions of interest and their locations. Then, we calculated the distance from the center of the image. Based on this distance we easily assigned numbers to our regions, such as closest to the center was the main blob, left and below of center was reference blob one, left and above of center was reference two, right and above was reference three and right and below was reference four (Supplemental Fig. 1 B). Using this numbering system, we input the known TIV for each reference blob in the standard curve model, with their corresponding segmented total pixel sum values. Here, the *K-means++* segmentation worked the same as in the ratio method, except that it segmented all four of the markers and the main ROI pixel cluster instead of just two such blobs. An overview of the pipeline flow implemented is shown in Supplemental Fig. 2.

Supplementary Results

Ratio model vs. standard curve model for TIV prediction of phantom

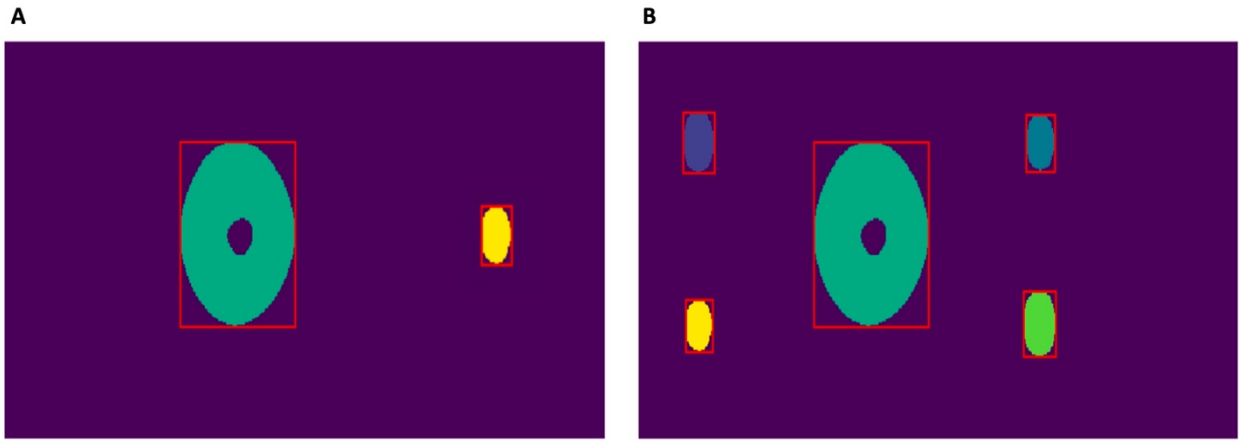
Upon optimization of the *K-means++* algorithm for segmentation and extraction of the ROI, the efficacy of the two proposed models for TIV prediction (the Ratio model and the Standard curve model) was assessed using the two 3D printed phantoms generated for this study: the 'S' and 'Circle' phantom. First, the ratio model was employed for TIV prediction (Supplemental Fig. 3). The *K-means++* algorithm was used for segmentation of the reference and main ROI, and total pixel sum of the regions was extracted. A ratio of the total pixel sum of the fiducial marker to

the known TIV of the marker was calculated. This ratio was then used to estimate the TIV of the main ROI. For segmentation, both the “S” and the “circle” were segmented with minimal inclusion of pixels from the desired ROI, as seen in the increasing concentrations of the ‘S and ‘Circle phantom (Supplemental Fig. 3A). Noticeably, the ‘Circle’ phantom segmentation included little relative noise, and therefore the final output ROI was rather crescent shaped than a full circle (Supplemental Fig. 3B). This is critical to proper analysis of the ROI and potential TIV prediction, especially due to the results produced by board-certified radiologists that included noise and bleeding of signals in the final ROI analysis. As segmented ROIs increased in total pixel sum as the injected amount of iron oxide increased in the phantom constructs, the predicted TIV value from the applied ratio model algorithm also increased in a linear fashion (Supplemental Fig. 3C). However, it is apparent that variations from the actual injected TIV and that predicted by the algorithm were rather wide with the applied ratio method. From the calculated ICC score, the data purports that for single measures, the ratio method did not have a great degree of inter-rater reliability, although the algorithm performed adequately in regards to the segmentation of the ROI with *K-means++* (Supplemental Table 1A, B). For TIV prediction using this method, the ICC score was slightly greater, but not considered above the .9 threshold, a value that would indicate excellent inter-rater reliability (Supplemental Table 1C, D). This is likely due to variations in ROI inclusion of pixel values during segmentation, including instances of false-positive signal analysis by the human rater. However, due to the indicated linear correlation between the ROIs total pixel sum and the predicted TIV, these findings warranted further study in which a linear regression-based SC model was applied to extract more accurate TIV predictions from the algorithm.

When the *K-means++* algorithm was applied to the standard curve (SC) model for TIV prediction, a clear correlation between the increasing amount of injected iron, indicating increasing signal intensity of total pixel sum, and the TIV prediction was observed (Supplemental Fig. 4). Furthermore, for the S shaped phantom, the segmentation output preformed with a high degree of specificity for the true signals of nanoparticles and excluded existing noise from the nanoparticle signal. The final output showed islands of pixel values where the signals originated and were the strongest (Supplemental Fig. 4A). This is an important feature of this unsupervised machine learning segmentation algorithm, as board-certified radiologists can rarely account for situations where bleeding of the original nanoparticle signal is often overlooked or difficult to detect. In these situations, the ROI chosen by the individual often includes false positive/negative signals which are incorrectly quantified in the manually detected ROI and distorts subsequent analysis of TIV and other parameters associated with the ROI. Furthermore, this phenomenon is also inherent in the *K-means++* segmentation of the signals from the circle phantom, in which the final output ROI is not merely a complete circle, but instead only includes the true signals from the nanoparticles without including bleeding of signal noise in the segmentation result (Supplemental Fig. 4B). From these segmentation results, the algorithm then preformed TIV prediction from the total pixel intensity sum of the segmented ROI's of multiple MPI scans of both the 'S' and 'Circle' phantom (Supplemental Fig. 4C). The standard curve generated by the algorithm permitted subsequent prediction of the TIV of the segmented ROI, and a linear trend between the total pixel intensity sum of the ROI and its TIV was observed for the increasing injected total iron in the phantom constructs. This was apparent in both the 'S' and 'Circle' phantom. Furthermore, due to the

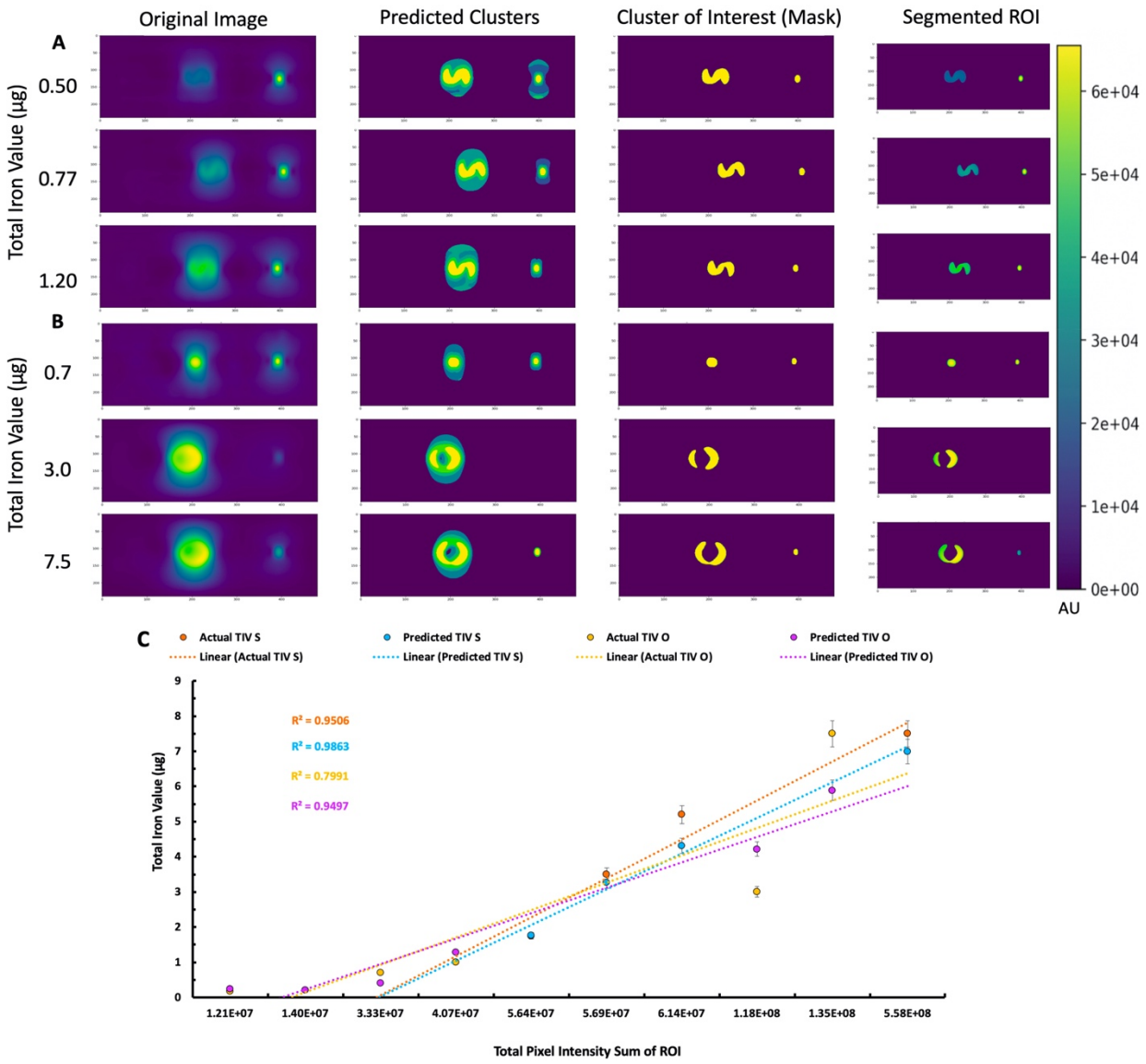
greater degree of accuracy in TIV prediction by the SC model in comparison to the Ratio model, as indicated by the greater ICC and degree of correlation between the actual and predicted TIV for both phantom shapes, the SC-*K-means++* based model was selected as the unsupervised machine learning algorithm of precedence throughout the rest of the study and was used in *in vitro*, *in vivo*, and *ex vivo* experiments (Supplemental Table 2A-D).

Supplemental Figures

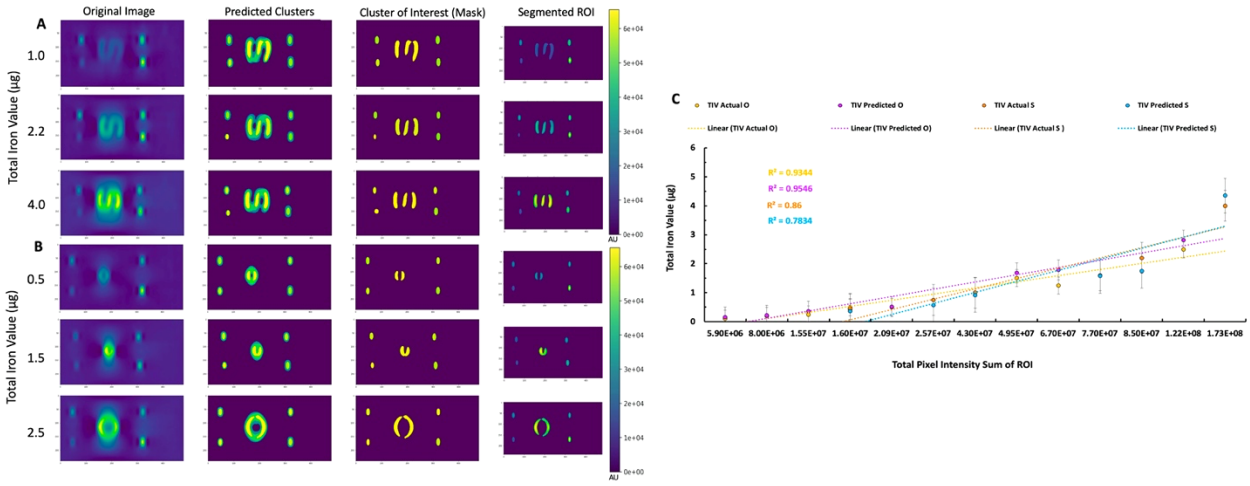


Supplemental Figure 1. Layout of different fiducial marker and main ROI regions for use in the ratio *K-means++* model (A) and standard curve *K-means++* model (B). The blobs in the periphery indicate fiducial reference markers of known TIV, and the larger blob in the middle indicates a sample main ROI.

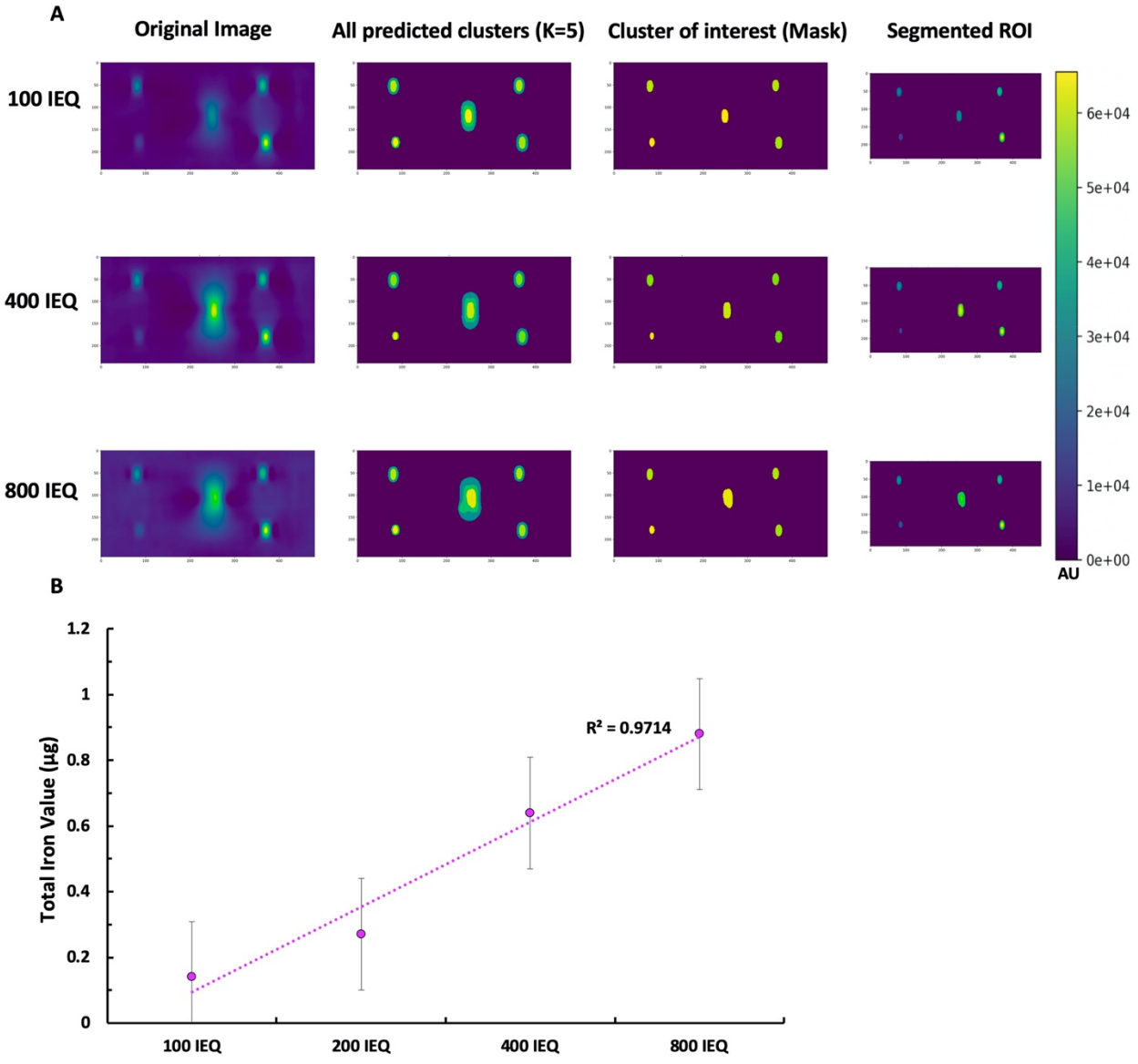
Supplemental Figure 2. Flow diagram of the pipeline for the ratio and standard curve *K-means++* machine learning algorithms.



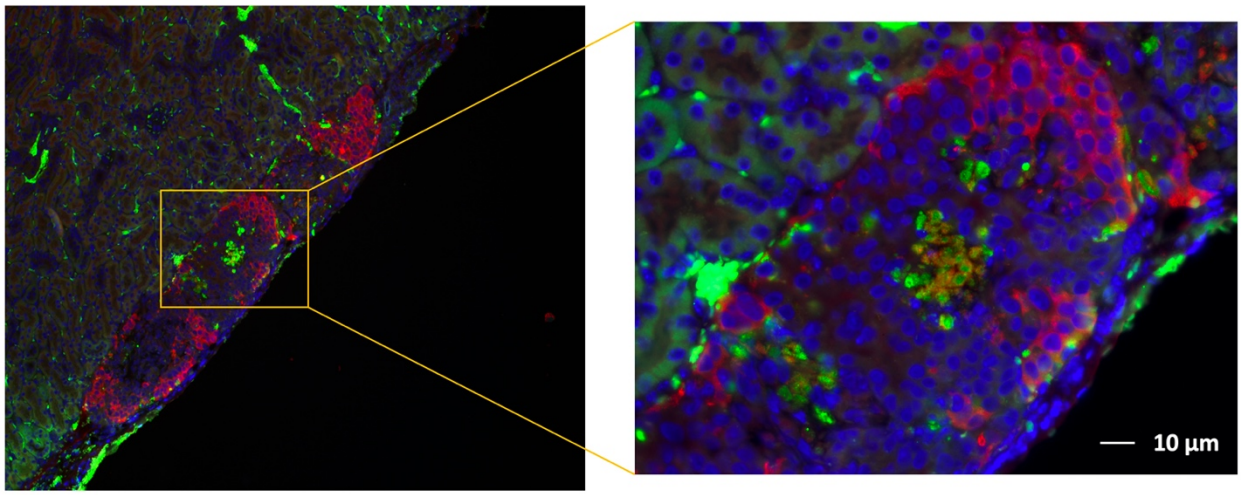
Supplemental Figure 3. *K-means++* Segmentation and Ratio TIV prediction for MPI phantom images. **A.** *K-means++* Segmentation and Ratio TIV prediction of 'S' phantom with 0.5 µg iron, 'S' phantom with 0.77 µg iron, 'S' phantom with 1.2 µg iron; The size of the 'S' phantom was 15mm. **B.** *K-means++* Segmentation and Ratio TIV prediction of 'Circle' phantom with 0.7 µg iron, 'Circle' phantom with 3.0 µg iron, 'Circle' phantom with 7.5 µg iron; **C.** TIV prediction of algorithm from total pixel intensity sum of ROI via ratio model. The size of the 'Circle' phantom was 10mm.



Supplemental Figure 4. *K-means++* segmentation and standard curve model TIV prediction for MPI phantom images. **A.** *K-means++* segmentation and standard curve model TIV prediction of 'S' phantom with 1.0 µg iron, 'S' phantom with 2.2 µg iron, and 'S' phantom with 4.0 µg iron; The size of the 'S' phantom was 30mm.**B.** *K-means++* segmentation and standard curve model TIV prediction of 'Circle' phantom with 0.5 µg iron, 'Circle' phantom with 1.5 µg iron, and 'Circle' phantom with 2.5 µg iron; **C.** TIV prediction of algorithm from total pixel intensity sum of ROI via SC model. The size of the 'Circle' phantom was 20mm.



Supplemental Figure 5. *K-means++* segmentation and SC model TIV prediction of ex vivo MPI of excised kidneys. **A.** *K-means++* segmentation and SC model TIV prediction of kidney with 100 IEQ, 400 IEQ, and 800 IEQ; **B.**TIV prediction of extracted ROIs for 100, 200, 400 and 800 IEQ's using SC model.



Supplemental Figure 6**.** *Ex vivo* immunostaining of a kidney section containing transplanted human islets labeled with VivoTrax. Green = VivoTrax, Red = Insulin, Blue = DAPI staining of cell nucleus. Bar=10µm.

Supplemental Tables

Supplementary Tables 1: Phantom Ratio Method ICC's.

A: Actual vs. Predicted ROI for "S" phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value
Single measures	0.629	0.080-0.891	9	0.03
Average Measures	0.921	0.148-0.942	9	0.03

B: Actual vs. Predicted ROI for “O” phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value
Single measures	0.797	0.375-0.945	9	0.02
Average Measures	0.887	0.545-0.972	9	0.02

C: Actual vs. Predicted TIV for “S” phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value
Single measures	0.851	0.488-0.968	7	0.02
Average Measures	0.919	0.619-0.984	7	0.02

D: Actual vs. Predicted TIV for “O” phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value
Single measures	0.820	0.353-0.961	7	0.04
Average Measures	0.901	0.522-0.980	7	0.04

Supplementary Tables 2: Phantom SC Method ICC’s.

A: Actual vs. Predicted ROI for “S” phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0 df	F test with true value 0 sig
Single measures	0.891	0.693-0.965	12	0.01
Average Measures	0.942	0.818-0.982	12	0.01

B: Actual vs. Predicted ROI for “O” phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0 df	F test with true value 0 sig
Single measures	0.916	0.753-0.973	12	0.01
Average Measures	0.956	0.859-0.987	12	0.01

C: Actual vs. Predicted TIV for “S” phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value
Single measures	0.921	0.711-0.970	10	0.01
Average Measures	0.959	0.831-0.989	10	0.01

D: Actual vs. Predicted TIV for “O” phantoms

Type	Interclass correlation	95% confidence interval	F test with true value 0	F test with true value 0

			Degree of freedom	P value
Single measures	0.905	0.699-0.973	10	0.01
Average Measures	0.950	0.823-0.986	10	0.01

Supplementary Tables 3: In Vitro Method ICC.

Algorithm vs. Raters

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value
Single measures	0.812	0.316-0.959	7	0.02
Average Measures	0.896	0.480-0.979	7	0.02

Supplementary Tables 4: In Vivo Method ICC.

Algorithm vs. Raters

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value
Single measures	0.892	0.621-0.972	9	0.01
Average Measures	0.943	0.011-0.986	9	0.01

Supplementary Tables 5: Ex Vivo Method ICC.

Algorithm vs. Raters

Type	Interclass correlation	95% confidence interval	F test with true value 0 Degree of freedom	F test with true value 0 P value

Single measures	0.828	0.341-0.963	7	0.04
Average Measures	0.906	0.509-0.981	7	0.04