Infinite Science Publishing

Research Article

MPI region of interest (ROI) analysis and quantification of iron in different volumes

Olivia C. Sehl[a,*] · Brice Tiret[b] · Maryam A. Berih[a] · Ashley V. Makela[c] · Patrick W. Goodwill[b] · Paula J. Foster[a]

[a] Robarts Research Institute, Western University, London, Canada
[b] Magnetic Insight Inc., Alameda, USA
[c] The Institute for Quantitative Health Science & Engineering, Michigan State University, East Lansing, USA
*Corresponding author, email: osehl@uwo.ca

Abstract

MPI directly detects superparamagnetic iron oxides (SPIONs), which should enable precise, accurate, and linear quantification. However, selecting a region of interest (ROI) has strong effects on MPI quantification results. Ideally, ROI selection should be simple, user-independent, and widely applicable. In this work, we describe and compare four MPI ROI selection methods and assess their performance *in vitro* and *in vivo*. To explore the effect of ROI selection, ten ferucarbotran phantoms were imaged, each contained the same amount of iron but varied in volume. Three users tested the accuracy of the ROI methods for quantification of these samples. Lastly, the four ROI methods were applied to quantify ferucarbotran *in vivo* after intravenous, intramuscular, and subcutaneous injections in mice. We discuss the strengths and limitations of each ROI method, such as the ability to capture MPI signals of custom shapes (*i.e.* size of the ROI), degree of user variability, speed of analysis, and quantification accuracy of SPIONs in different volumes.

I. Introduction

Magnetic particle imaging (MPI) is a quantitative modality that directly detects superparamagnetic iron oxide nanoparticles (SPIONs). MPI quantification of SPIONs provides unique opportunities in tracking cell therapies, imaging inflammation, and magnetic hyperthermia. MPI can be used to track and quantify adoptive cells as they move within the body by first labeling them with SPIONs *in vitro* [1–4]. Here, quantification is necessary to understand the success of adoptive cell therapies which rely on a sufficient number of cells arriving at a target site while minimizing off-target accumulation of cells. Development of SPIONs for imaging inflammatory sites is another area of research requiring robust quantification [5–7]. SPIONs are administered intravenously

and are internalized by phagocytic cells at inflammatory sites and in the reticuloendothelial system (liver, spleen, bone marrow, and lymph nodes) [5, 6, 8, 9]. Quantification of inflammatory cells within these sites can be used to assess the degree of inflammation, the response to immunomodulatory therapies [9, 10] and the density of tumor-associated macrophages (TAMs) [11–13]. MPI-guided magnetic hyperthermia involves exciting SPIONs with alternating magnetic fields, within a region defined by stronger gradient fields. Precise localization and quantification of SPIONs in the specified region is necessary to apply a specific heating prescription, as described in [14–16].

Our development of cell tracking techniques [3] has identified a demand for standardized, reproducible, and accurate quantification methods for MPI. *In vivo,*

SPIONs can be dispersed in larger volumes (*e.g.* a tumor, 1 cm^3), whereas MPI quantification is conducted using calibration samples, of known amounts of iron, typically in small fluid volumes (~10 µL). Careful and rigorous choice of the region of interest (ROI) is thus essential for quantification. Previous work by Hayat *et al.* [17] used a K-means++ machine learning algorithm for segmentation of ROIs and demonstrated accurate quantification of transplanted ferucarbotran-labeled cells *in vivo*. However, for some MPI users without such computational background, more straightforward methods would be of value.

Positron emission tomography (PET) is another imaging modality that provides hotspot images with direct quantification. Quantification of PET signals are typically conducted by reporting the maximum standard uptake value (SUV). There is concern that the maximum value does not represent the signal(s) of interest and this is discussed particularly in the application of heterogenous tumors with distributed PET signals. There are similar concerns for MPI as the distribution of SPIONs can influence the maximum value of the point spread function (PSF). Additional methods for PET quantification involve ROIs defined by manual delineations [18, 19], anatomical scans (whole organ segmentations) [20], fixed diameter circular ROI [18], or an ROI threshold defined as a percentage of the maximum signal [18, 19], however these are not routine clinical practices. Importantly, the quantitative outcome for PET is dependent on the method of ROI selection, involving the size, shape, and placement of the ROI. This has been a concern for several decades [21].

Many of these ROI selection methods used for PET have been used for MPI. A common method for ROI definition with MPI has been to select pixels which are a percentage of the maximum value (*e.g.* 50 % maximum) [22–27]. Other groups have used manual delineations of MPI signal [5, 28, 29], fixed circular ROIs [7, 30–33], or standard thresholding [17, 34, 35]. These ROIs have been applied to MPI images with both field free line (FFL) and field free point (FFP) acquisitions, with X-space reconstruction and other methods. For quantification of stenosis there are unique methods described [36]. For MPI, the PSF extends beyond the borders of the SPION source and the extent of this depends on many factors, including the SPION resolution, amount of SPION, and the acquisition and reconstruction methods. In this paper, we hope to expand the toolbox for MPI quantification, by describing and directly comparing four methods of ROI selection, including both custom and 'one-size-fits-all' approaches.

I.I. Study objectives

We first developed four unique ROI selection methods and assessed their linearity of quantification using *in vitro* ferucarbotran phantoms. We then evaluated the inter-user reproducibility of these ROI methods. Lastly,

these methods were applied to quantify SPIONs that were administered locally or systemically in mice.

II. Materials and methods

II.I. *In vitro* sample preparation and imaging

Ten ferucarbotran phantoms (Vivotrax™, 5.5 µg Fe/µL, Magnetic Insight Inc., Alameda, CA, USA) were prepared, each containing the identical amount of iron (34.4 µg) mixed with increasing volumes of saline, up to 1.2 mL (see Table 1). These samples were imaged in the base of a 15 mL Falcon tube which is approximately 5 mm in diameter at the base. The tube widens to a diameter of 15 mm, therefore as saline is added to the sample the volume expands in all directions.

For MPI signal calibration, a separate series of calibration samples containing different, known amounts of ferucarbotran in the same fluid volume (range 0.34 – 55 µg Fe in 10 µL) were prepared. An additional five samples of ferucarbotran (range 10 – 150 µg) were imaged for calibration of *in vivo* images (refer to Section II.V for imaging protocol).

Each ferucarbotran sample was imaged separately using a MOMENTUM™ MPI system (Magnetic Insight Inc.). Images were acquired in 2D with a 6 cm x 6 cm field of view, 5.7 T/m selection field gradient, and drive field strengths of 20 mT and 26 mT in the X and Z axes, respectively. The images were reconstructed using x-space methods described in [37]. Images from the X and Z planes were combined using signal averaging. The image pixels are 250 x 250 µm and the resolution of a ferucarbotran point source using a 5.7 T/m gradient is expected to be 1.73 mm, based on MP relaxometry measurements [38].

Table 1: Ten samples created by dilution of ferucarbotran in different volumes of saline.

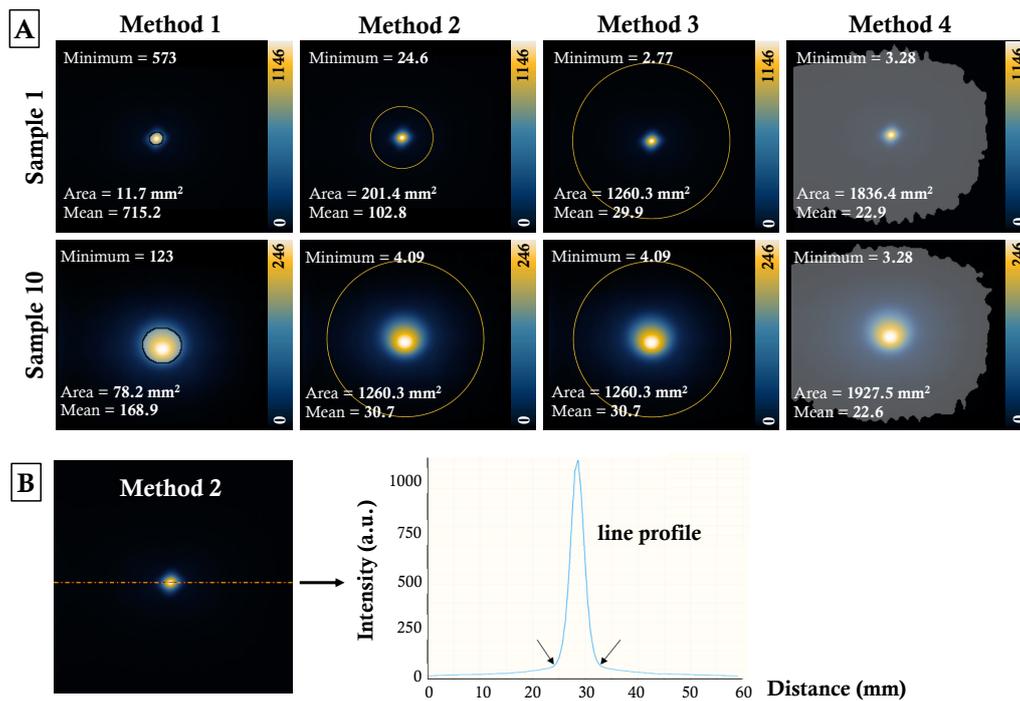| #	Total sample volume (µL)	Volume ratio (Vivotrax: saline)	Iron Concentration (µg/µL)
1	6.25	1:0	5.500
2	12.5	1:1	2.750
3	25.0	1:3	1.375
4	50.0	1:7	0.688
5	100.0	1:15	0.344
6	200.0	1:31	0.172
7	400.0	1:63	0.086
8	600.0	1:95	0.057
9	800.0	1:127	0.042
10	1200.0	1:191	0.029

Figure 1: (A) A demonstration of ROI methods 1–4 on two samples of ferucarbotran (Samples 1 and 10). The minimum value (a.u.), area (mm^2), and mean value (a.u.) is recorded for each ROI. (B) A depiction of ROI selection method 2.

II.II. ROI selection methods

All of the 2D images obtained were analyzed using four ROI selection methods using Horos imaging software (Annapolis, MD, USA). Each method is shown visually in Figure 1A and described in the sections below. For each ROI, the MPI signal (A.U.) was calculated as a product of average pixel value (A.U./mm^2) and the ROI volume (mm^2).

Method 1: ROI threshold set by the maximum signal.

An ROI was drawn at pixel value $> a \cdot s_{max}$, where a is a scaling factor < 1 and s_{max} (A.U./mm^2) is the maximum value for the signal of interest. We used $a = 0.1$, 0.5, and 0.7 for *in vitro* analysis and $a = 0.5$ *in vivo*.

Method 2: Circular ROI with the diameter determined visually, with scaling factor.

A line profile was drawn through the maximum value of the signal of interest, and the distance between the two points of maximum curvature was estimated (Δx) (see Figure 1B). This spatial distance was used to define a circular ROI with diameter $d = b\Delta x$, where b is a scaling factor > 1 that increases the ROI diameter to account for user variability and capture more of the signal from the wide PSF. We used a scaling factor $b = 2$, and the peak signal of the ROI was visually centered on the maximum value of the signal of interest.

Method 3: Large circular ROI.

ROI diameters were measured for all images in the dataset using Method 2. Then, all images were quantified with the same circular ROI with diameter $d = b\Delta x_{max}$, where Δx_{max} is the largest ROI diameter in the dataset.

Method 4: ROI threshold set by image noise characteristics.

The standard deviation of system noise, S.D., was measured by imaging an empty sample holder using the same image settings. A mask was generated that selected pixels with pixel value $> c \cdot SD$, where c is a multiplier chosen to select signal produced by the sample while rejecting noise produced by the imaging system. We used $c = 5$ to select signals with Signal to Noise Ratios (SNR) > 5 according to the Rose Criterion [39] for *in vitro* and *in vivo* datasets. Higher thresholds were also tested with $c = 10$ and $c = 25$ for *in vitro* ferucarbotran samples. This method makes the reasonable assumption that the system noise is similar in each image in the dataset.

II.III. MPI signal calibration

After the MPI signal associated with calibration samples was measured, calibration lines were formed to determine the relationship between known iron mass (µg) and MPI signal. A linear equation was obtained for each ROI selection method (MPI signal = slope · iron mass) and the intercept of the line was set at (0,0). These equations were rearranged as:

$$\text{Estimated iron content (µg)} = \frac{\text{MPI signal (A.U.)}}{\text{Calibration slope}}$$

This equation was used to estimate the iron mass from ferucarbotran phantoms and ferucarbotran *in vivo* using each ROI method.

II.IV. Inter-user reproducibility

The ROI analysis (Section II.II and Section II.III) was repeated by 2 additional users on the same set of images. The pairwise differences in the user's measurements of iron mass were calculated, and for each ROI method, the average of these differences was determined. Bland-Altman plots were used to display the difference between user's measurements of iron mass compared to their average measurements. The absolute inter-user variability was calculated as the standard error of measurement (SEM):

$$\text{SEM} = SD(\text{value 1, value 2, value 3})$$

SEM was calculated for each of the ten samples of ferucarbotran and the average SEM for each ROI method was reported. As an additional measurement, the coefficient of variation (CoV) was calculated as:

$$\text{CoV (\%)} = \frac{\text{SEM}}{\text{Average (value 1,value 2,value 3)}} \cdot 100$$

and this value represents the relative interobserver variability. When the user's measurements are in perfect correspondence, SEM and CoV are equal to 0 [19].

II.V. ROI selection for *in vivo* quantification

Two Nu/Nu mice were obtained and cared for per the standards of the Canadian Council on Animal Care, under an approved protocol by the Animal Use Subcommittee of Western University's Council on Animal Care. The first mouse received 25 µg ferucarbotran by intramuscular point injection in the left hind limb. This mouse was imaged immediately before and after a second injection of 25 µg subcutaneously. The second mouse received 100 µg ferucarbotran intravenously and was imaged 20 hours later. Ferucarbotran is expected to accumulate predominately in the liver, and to a lesser extent the spleen [7, 23,
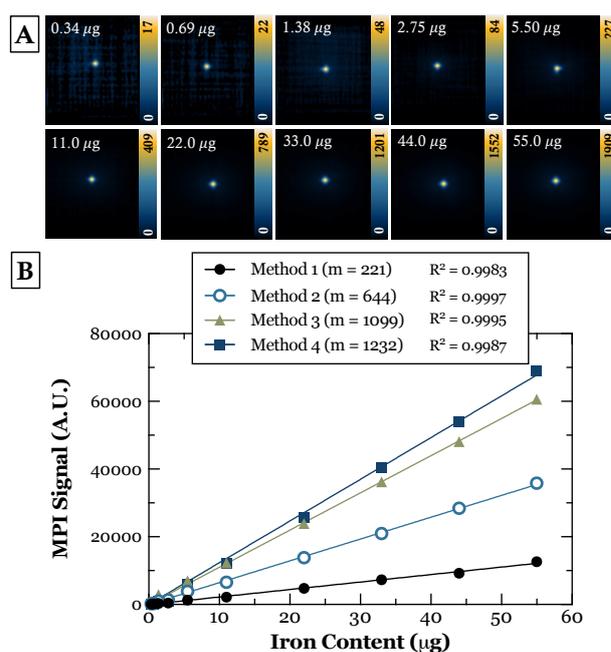
Figure 2: (A) MPI calibration performed by imaging 10 samples of known ferucarbotran mass (0.34–55 µg). (B) Using each ROI method, a linear relationship between iron content and MPI signal was established as MPI signal (A.U.) = $m \cdot$ iron content (µg), where m is the slope of each line.

40–42]. Isotropic 2D MPI images were acquired using a 3.0 T/m selection field gradient, a larger field of view (12 x 6 cm), and the same parameters listed in Section II.I For a 3.0 T/m gradient, the image resolution for a ferucarbotran point source is expected to be 3.29 mm [38]. Separate calibration with ferucarbotran (10 – 150 µg) was conducted for these parameters and applied to *in vivo* images to quantify SPION mass, using the same methods described in Section II.III.

III. Results

III.I. MPI signal calibration

The MPI images of the calibration samples are visually similar in shape and primarily differ in signal amplitude (Figure 2A). This is due to all samples having the same volume but with different iron quantities. Quantification of calibration points from all methods resulted in a linear ($R^2 > 0.998$) relationship between iron mass and measured MPI signal (Figure 2B). More MPI signal (steeper slope, m) was collected for methods with larger ROIs (*i.e.* Method 3 and 4). The linear equations for each quantification method are used to estimate the iron contents of the *in vitro* and *in vivo* samples.
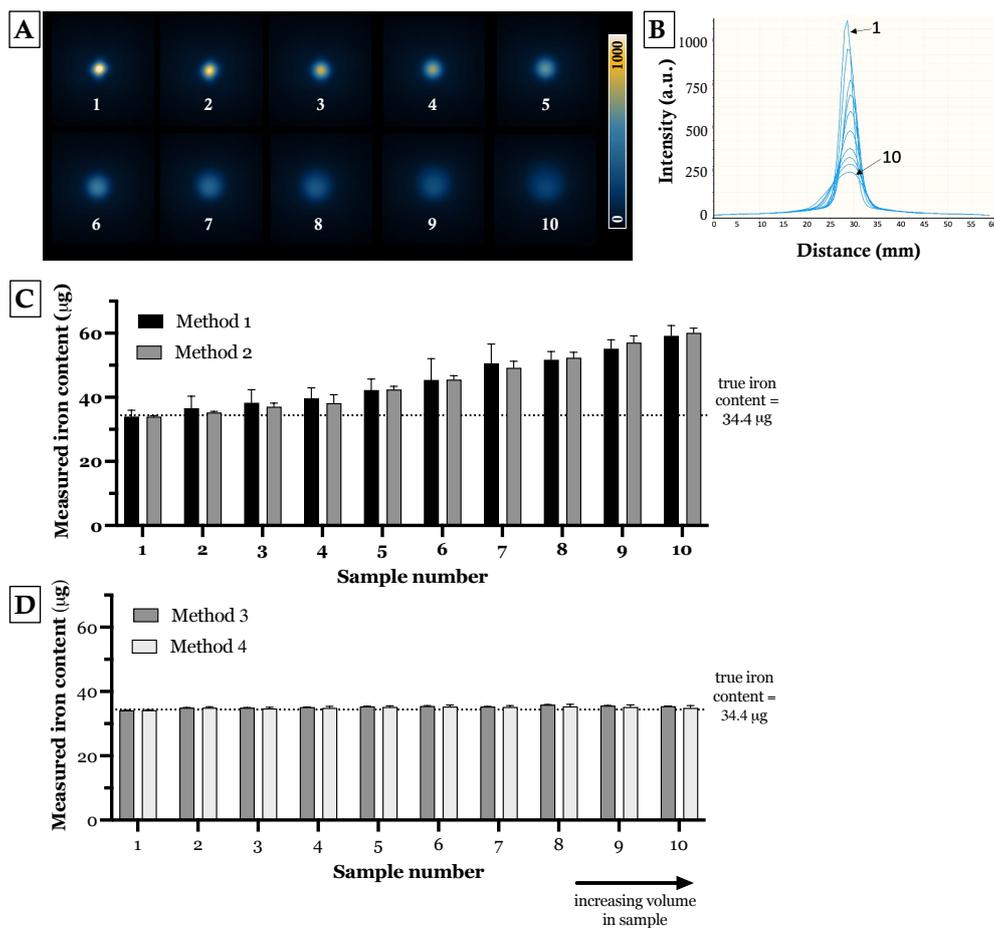
Figure 3: (A) Projection images showing the same amount of ferucarbotran (34.4 µg) in increasing volumes of saline. Image values displayed are 0-1000 A.U. (B) Line profiles showing differences in signal intensity and resolution from Samples 1-10. Undiluted ferucarbotran (Sample 1) produces higher signal intensity as the sample is concentrated, while the point source spreads out more as the sample is diluted. Iron content (µg) measured from samples 1-10 using methods 1 and 2 (C), and methods 3 and 4 (D).

III.II. Quantification of *in vitro* phantoms

MPI images of Samples 1-10 are shown in Figure 3A and line profiles through these images are shown in Figure 3B. As the fixed mass of ferucarbotran tracer is diluted with increasing amounts of saline, the volume of the sample expands. Thus, the MPI signal is localized over a larger region and the effective amount of iron per imaging voxel is reduced, leading to a diminished peak signal intensity and broadened width of the PSF.

As shown in Figure 3C, ROI methods 1 and 2 accurately quantified the iron mass for samples with volumes similar to the calibration samples (*i.e.* 10 µL). Unfortunately, as the sample is diluted, methods 1 and 2 overestimated iron mass by up to 70 %. For method 1, this pattern persists with different threshold factors (Figure 4A). ROI methods 3 and 4 are larger ROIs and adequately capture

the broader extents of the MPI signal. This leads to a more accurate estimation (<5 % error) of iron mass, regardless of sample volume (Figure 3D). Method 4 was accurate when thresholding with higher values of c (Figure 4B). For ROIs defined by $c = 10$, the estimation of iron mass in phantoms was 35.22 ± 0.44 µg. For $c = 25$, the iron mass was estimated at 33.69 ± 0.50 µg, therefore a slight systematic underestimation was observed.

III.III. Inter-user reproducibility

All three users estimated less accurately and precisely using methods 1 and 2 and more accurately and precisely using methods 3 and 4. Box plots of the iron mass mean and range estimated by each user for each method are shown in Figure 5A. Bland-Altman plots show that method 4 has the highest accuracy with an average 0.08 µg difference between the estimated iron con-
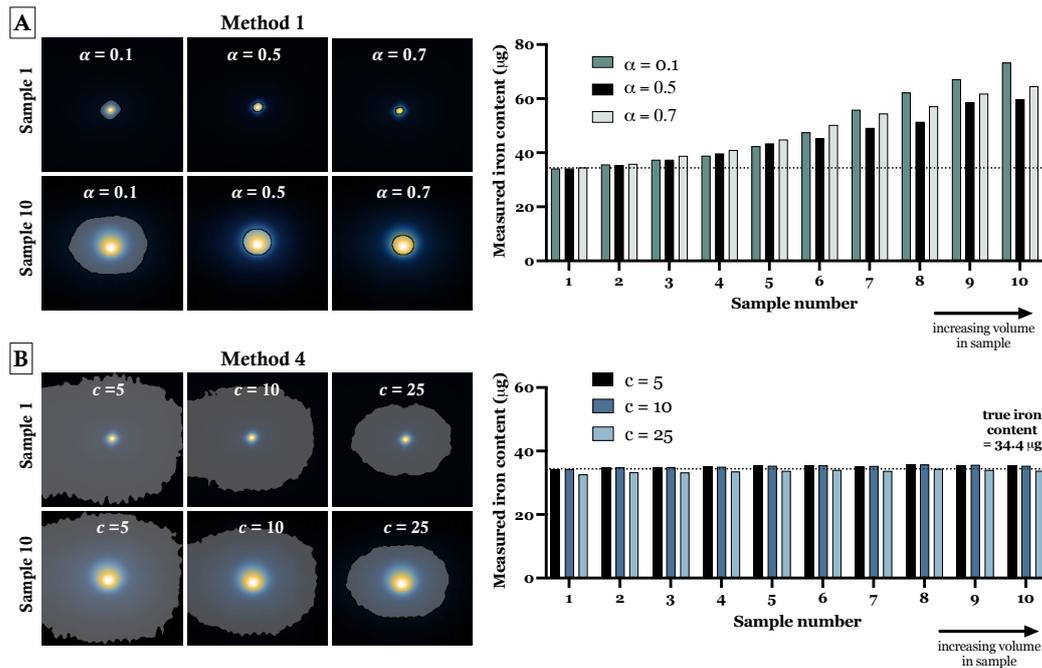
Figure 4: (A) Various threshold factors are applied to the maximum value (a = 0.1, 0.5, and 0.7) to quantify Samples 1-10. For ROI method 1, quantification of iron is overestimated for dilute samples regardless of the scaling factor. (B) Various threshold factors are applied to the background SD (c = 5, 10, 25). Using method 4, iron mass is consistent across scaling factors and accurate with dilution of ferucarbotran.

Table 2: Inter-user reproducibility measures for ROI methods 1-4. The average difference for each user pair, standard error of measurement (SEM) and coefficient of variance are reported.

ROI method	1	2	3	4
Average difference (A.U.)	1.75	4.84	0.53	0.08
SEM (A.U.)	1.389	3.770	0.433	0.061
CoV (%)	3.02	8.49	1.23	0.17

tent (Figure 5B). Method 2 shows the largest variation between users, averaging a difference of 4.84 µg iron and with a maximum difference of 13.25 µg.

Table 2 summarizes the agreement parameters: average difference, SEM, and CoV. The SEM and CoV are lowest for method 4, indicating that this method has the smallest inter-user variability.

III.IV. *In vivo* quantification

The application of ROI methods 1-4 for *in vivo* quantification of iron is demonstrated in Figure 6. Each ROI selection method is shown for the three *in vivo* images in Figure 6A-C. The iron mass estimated from each ROI was calculated using the slope from the corresponding calibration curves (Figure 6D).

Figure 6A shows 2D MPI of a mouse administered 100 µg ferucarbotran intravenously. From the bloodstream, ferucarbotran accumulates in the phagocytic cells and is primarily distributed throughout the liver and spleen. Method 1 uses the smallest ROI and while method 2 ROI is customized, the circular ROI is applied to a non-circular MPI shape, complicating the aim of including more MPI signal. The use of these ROI methods resulted in overestimation of iron mass measured from the mouse liver and spleen (method 1: 144 µg, method 2: 123.5 µg, Figure 6E). Note that ROI method 2 and 3 are the same for this image because disperse signal in the mouse liver and spleen resulted in the largest PSF in the dataset. Next, we applied higher values of α with method 1, in attempt to increase the amount of signal included in the ROI while maintaining a custom ROI shape (Figure 7A). However, this did not improve the accuracy of this method for quantifying the SPION mass in the mouse liver and spleen (153.35 µg).

Compared to the other ROI methods, method 4 with $c = 5$ creates the largest ROI with signal spread occupying most of the MP image (Figure 6A). Thus, the liver and spleen signal is adequately sampled and results in accurate iron quantification (101.4 µg, Figure 6E). A smaller ROI with a customized shape can be achieved in this application by using a higher threshold value, *i.e.* setting $c = 100$ with Method 4. As shown in Figure 7B, the ROI is
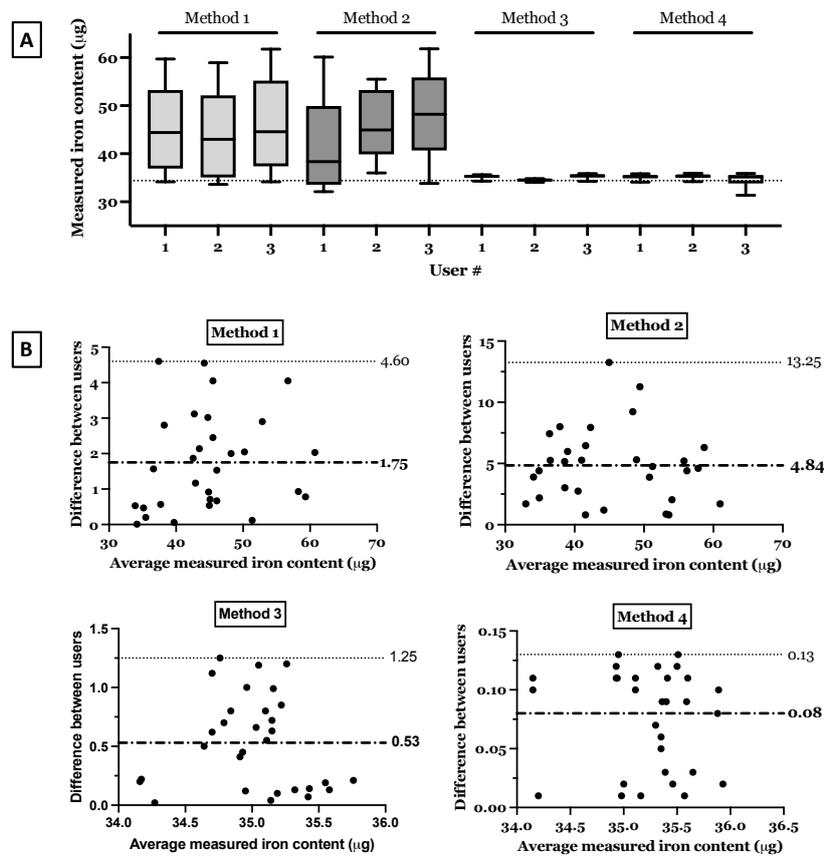
Figure 5: Analysis of inter-user reproducibility for ROI methods 1-4. (A) For each method, the range of iron content measured by each user for samples 1-10 is shown. The dotted line indicates the true iron mass (34.4 µg). (B) Bland-Altman plots comparing absolute difference between users with average measured iron content (µg). The dotted lines indicate the maximum difference and the average difference (bolded line) between users. Method 4 had the lowest user differences (mean = 0.08 µg).

confined to the liver and spleen signal and quantitation accuracy is maintained (99.47 µg).

For point injections *i.e.* intramuscular injection (Figure 6B) and subcutaneous injection (Figure 6C), the MPI signal originates from ferucarbotran in a smaller volume. For quantification of ferucarbotran point sources, all ROI methods return accurate estimates of iron mass (Figure 6F,G; ground truth = 25 µg). However, as shown in Figure 6C, it is challenging to separate two sources of MPI signal when using ROI selection method 4 with $c = 5$ since the ROI boundaries for each source overlap. Next, using method 4 we applied higher values of c in attempt to distinguish these two signals. At $c = 35$, these two point sources of SPION could be separately quantified (Figure 7C). However, since a smaller ROI is applied, the quantitative accuracy is compromised; the subcutaneous injection of 25 µg was estimated to have 22.8 µg ferucarbotran and the intramuscular injection to have 23.3 µg ferucarbotran.

IV. Discussion

MPI directly detects SPIONs, which should enable precise, accurate, linear quantification. However, selection of an ROI has strong effects on MPI quantification results. Ideally ROI selection should be simple, user-independent, and applicable to *in vitro* and *in vivo* situations. This manuscript describes four MPI ROI selection methods and assesses their performance *in vitro* and *in vivo*. Methods 1-4 use different sized ROIs to quantify MPI signal and each approach has suitable applications. These are discussed below and summarized in Table 3. The methods were first tested on 10 ferucarbotran phantoms with identical iron mass but varying sample volumes (6.25 – 1200 µL). The resulting images presented a quantification challenge since the point source spreads out as the ferucarbotran samples is diluted and the overall MPI signal intensity is reduced. To assess user-reproducibility, the samples were anlysed by three users. Last, these methods were applied to *in vivo* quantifica-
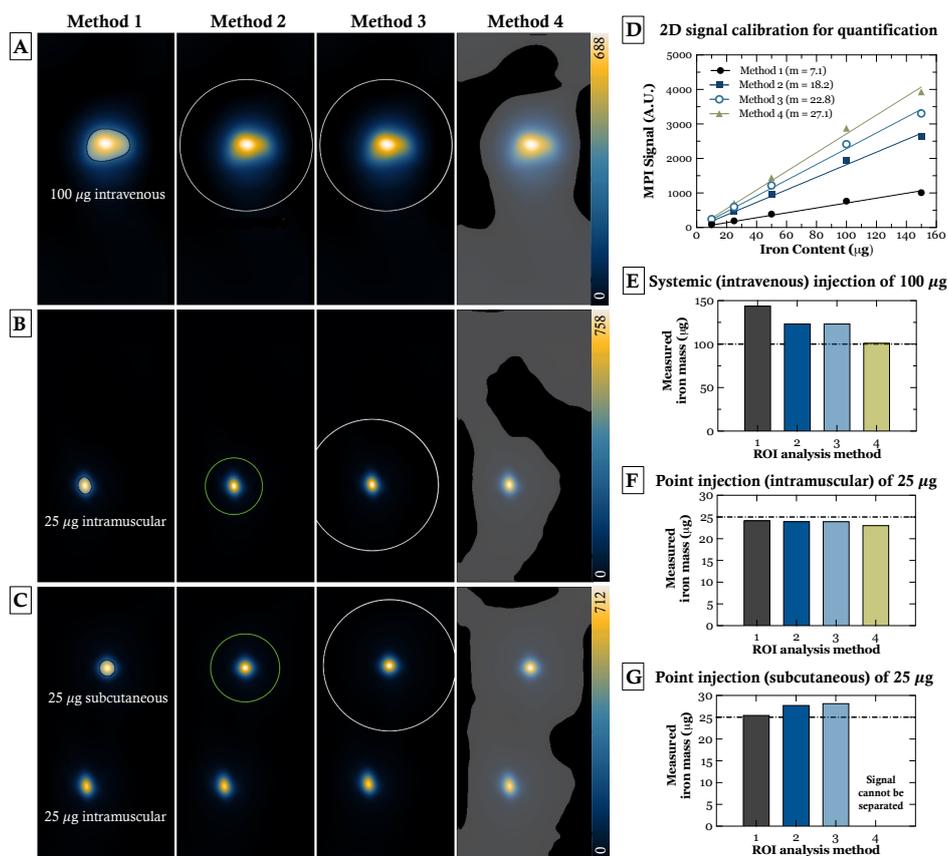
Figure 6: Demonstration of *in vivo* quantification using methods 1-4 following the (A) intravenous administration of 100 µg ferucarbotran or an (B) intramuscular injection of 25 µg ferucarbotran, followed by a (C) subcutaneous injection of 25 µg ferucarbotran. (D) Signal calibration to determine the relationship between iron mass (ferucarbotran) and MPI signal using 3.0 T/m gradients (the same parameters as *in vivo* images). MPI signal of calibration samples was measured using all 4 methods. (E-G) Iron mass measured from respective *in vivo* images A-C, comparing ROI methods 1-4.

Table 3: Comparison of ROI methods 1-4 in terms of size of the ROI, time spent on analysis, user variability, the ability to create custom shaped ROIs, and the quantification accuracy of iron in different volumes.

ROI method	1	2	3	4
Size of ROI	Small	Medium	Large	Large
Speed	Medium	Slow	Medium	Fast
User variability	Medium	High	Medium	Low
Custom shapes	Yes	No	No	Yes
Different volume samples	Poor	Poor	Good	Good

tion of ferucarbotran after intravenous administration or point injection (intramuscular and subcutaneous) to mice.

Method 1 selects an ROI based on the peak signal value within the target volume. Pixels are included in the ROI by selecting a threshold, *e.g.* $0.5 \cdot s_{max}$, and including all pixels with values larger than the threshold. This approach is beneficial to generate small ROIs, which helps achieve separation of multiple adjacent signals present in a single image. For example, Method 1 can separately quantify the MPI signal from ferucarbotran administered to multiple regions of a mouse with accuracy (Figure 6C). Ultimately this method is best suited for high-SNR signals from samples in a similar volume. We saw that samples of ferucarbotran in small volumes lead to sharp signals with high intensities relative to larger volumes, which produce broader signals. Since dilute samples have broader signal with lower peak intensity, the $a \cdot s_{max}$ threshold is lower and includes more area in
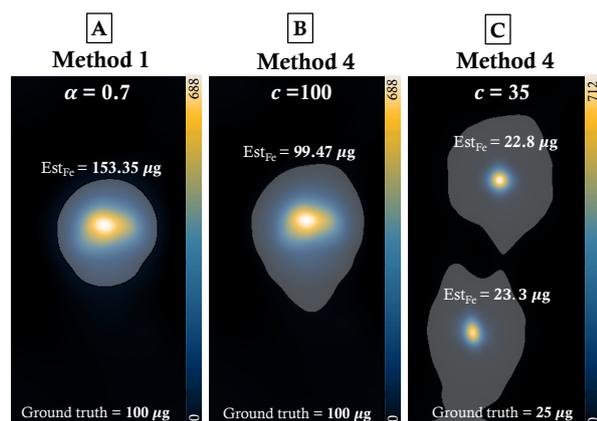
Figure 7: Application of alternative threshold values using methods 1 and 4 for *in vivo* analysis. (A) ROI method 1 with $a = 0.7$ was applied to the signal associated with ferucarbotran in the mouse liver and spleen. This creates a larger ROI than $c = 0.5$, but did not return an accurate estimate of SPION mass (153.3 µg). (B) ROI method 4 with high value of c (= 100) applied to the same mouse results in accurate quantification (99.47 µg). (C) With ROI method 4, increasing c (=35) allows for separate quantification of two iron sources *in vivo* but quantitative accuracy is reduced.

the ROI (Figure 1A and Figure 4A). Thus, an overestimation of iron mass is expected for dilute samples, which is confirmed *in vitro* (Figure 3C and Figure 4A). This is also true for *in vivo* quantification of ferucarbotran that is administered intravenously, which disperses throughout the mouse liver and spleen, regardless of the value chosen for α (Figure 6E and Figure 7A). This suggests that when using method 1, the fluid volume in the calibration samples must be carefully controlled to match that of the signals that are being quantified.

Method 2 selects a circular ROI that is manually placed on the peak signal, with the ROI diameter selected from the line profile. These ROIs are larger than method 1 and include more of the PSF; however, the quantitative accuracy of dilute *in vitro* samples of ferucarbotran did not improve. Likewise, the quantification of dilute ferucarbotran in the mouse liver and spleen after intravenous administration was overestimated. For *in vivo* analysis, line profiles tend to be less sharp and consequently the determination of the PSF width is challenging. For non-circular objects, automatic image thresholding in the presence of high SNR (*e.g.* Otsu's method) could help to identify the initial ROI. This method is quantitative only for high-SNR and high-resolution signals and breaks down for more complex images. Method 2 was accurate for quantifying *in vitro* phantoms of small volumes (Figure 3C) and point injections *in vivo* (Figure 6F,G).

Method 3 uses ROIs with fixed dimensions on all images in a dataset, using the largest-sized ROI from method 2. This method is optimal for image datasets

that assume the same physical layout of objects and can be used for quantification across a large SNR range. With the ROI diameter tied to the largest signal spread, this method guarantees sampling of signals at lower concentrations, and the amount of noise introduced to the ROI is consistent for each image. Method 3 provided precise and accurate iron mass measurement *in vitro* (Figure 3D) and this approach is faster than method 2.

Method 4 uses an SNR threshold-based segmentation system (*e.g.* SNR ≥ 3) and has many advantages and few disadvantages (Table 3). This method is simple to perform and can be conducted quickly. The size of the ROI depends on the standard deviation of the background signal, the threshold factor (c), and the size of the PSF. Method 4 applies the same lower threshold *i.e.* $c \cdot SD$ to all images in the dataset (including calibration samples), regardless of MPI signal strength and dimensions. For low SNR thresholds (*e.g.* SNR ~3), this technique produces large ROIs compared to methods 1 and 2, and similar to method 3 (shown in Figure 1A). This method illustrates that the MPI signal is spreading in nature, with image pixels being affected despite their distance from the source. Our results show that method 4 returns accurate quantification of ferucarbotran in varying volumes both *in vitro* (Figure 3D) and *in vivo* (Figure 6E,F). For higher threshold values (*e.g.* $25 \cdot SD$) the method results in smaller ROIs without compromising accuracy if a consistent threshold is used (Figure 4B and Figure 7B). However, this technique can struggle to quantify multiple overlapping signal sources *in vivo*. As shown in Figure 6C, two ferucarbotran sources separated by 5.3 cm could not be separately quantified using a $5 \cdot SD$ threshold. We observed that a minimum threshold of $35 \cdot SD$ was necessary to separate two sources of 2D MPI signal *in vivo* (Figure 7C) although this led to loss in quantitative accuracy. At higher threshold values ($c > 5$), more MPI signal is neglected at which point the quantification accuracy may be degraded. In our experience, overlapping signals is less of an issue in 3D data sets. Overall, there exists an important tradeoff between ROI area and quantification result, and this is dependent on many factors including MPI acquisition, SPION mass (and therefore, the intensity of the MPI signal), and SPION type.

All three MPI users analysed the identical set of images, therefore user differences reported in Section III.III are attributed solely to image analysis. Overall, we observed that method 4 has the highest reproducibility as shown by the smallest user differences, and method 2 had the greatest variability. We would expect methods 1 and 2 to be more error prone and show increased user differences because a different, unique ROI threshold is applied to all ten ferucarbotran phantoms and all ten calibration samples. In contrast, method 3 and 4 would be expected to have lower user differences because they apply a uniform ROI or threshold value to all 20 images. ROI method 1 is a semi-automatic method. When calculating

the threshold value ($a \cdot s_{max}$), we would expect there to be rounding errors depending on how many decimal places the user included. Method 2 is the most manual and error-prone method since it requires on 3-user inputs per image. The first is a subjective measurement of the PSF edges, and though a multiplication factor (b) is applied to mitigate it, inherent user variability is expected. Finally, the ROI is placed manually over the peak of the signal, which may also contribute to user differences. Method 3 has reduced user variability compared to method 2 since there is only a single manual bias. Lastly, method 4 may show differences resulting from the measurement of the background standard deviation from the same blank MPI image. There could be slight differences in the selected region that would subsequently lead to a different threshold value ($c \cdot SD$).

The applicability of these methods presented is transferrable to other reconstructions than X-space. For reconstructions that have more localized signal, the different ROIs resulting from the methods presented would look more similar, as the majority of the signal is confined in the same region and the transition at peak curvature would be more pronounced. For such a dataset, with relatively lower background signal, method 4 would be recommended as it would be the easiest and most reliable to implement. Furthermore, a lower scaling factor would be needed to capture an ROI close to the object imaged. Note that using methods like deconvolution lowers the peak SNR and caution is recommended where sensitivity may be compromised. With reconstructions that provide more confined MPI signals, other groups have used ROIs defined by the sample position or anatomy for quantification [43, 44].

Methods 3 and 4 use large ROIs for quantifying SPIONs. We have considered an additional approach, where the entire image serves as the ROI. Interestingly, this method was accurate for quantification of ferucarbotran phantoms (Section III.II) and ferucarbotran *in vivo* (Section III.IV). This approach led to estimation of 34.68 ± 0.61 µg ferucarbotran in Samples 1-10 (truth = 34.4 µg). For ferucarbotran administered intravenously, this ROI method yielded an estimate of 100.1 µg from the entire image (truth = 100 µg). Lastly, for ferucarbotran administered intramuscularly, this ROI yielded an estimate of 22.2 µg (truth = 25 µg). Overall, this approach could serve as a quick, crude analysis technique for quantifying images with a single signal source.

In this study we observed that the dispersion of SPIONs reduces the quality of MPI quantification when using ROI selection methods 1 and 2. There are additional factors which influence reliable quantification of SPION with MPI, including SPION immobilization [3, 45–48] and degradation [30, 48]. Immobilization of SPION occurs with protein binding or cellular internalization and is expected to increase Brownian relaxation times, leading to MPI signal reduction and blurring. As SPIONs undergo

degradation over time, their MPI sensitivity can be variable [30]. Importantly, ferucarbotran in this study was vortexed before use and stored at appropriate conditions to avoid clustering and degradation. *In vivo* imaging was conducted immediately or within 1 day of iron administration, therefore SPION degradation is expected to be negligible.

V. Conclusion

There is an unmet need for standardized, reproducible, simple, and accurate ROI methods for quantification of MPI signal. Here we described and tested four ROI selection method approaches. All four methods showed a strong linear relationship between ferucarbotran mass and quantified MPI signal for iron samples of the same fluid volume.

Each ROI method has advantages. Method 1 relies on segmentation at a fraction of the maximum MPI signal and forms the smallest ROIs. This method provided accurate quantification of ferucarbotran in small volumes and ferucarbotran administered by point injection *in vivo*. This method is beneficial for separately quantifying multiple sources of iron present in a single image. Ultimately ROI method 1 was not accurate for quantification of ferucarbotran in larger fluid volumes and after systemic administration *in vivo*. Method 2 applies a custom size ROI to each MPI signal by assessing the PSF. While this method includes more MPI signal than method 1, the quantitative accuracy for dilute ferucarbotran *in vitro* and *in vivo* did not improve. It can be challenging to assess the PSF of low-SNR signals and dilute iron, particularly *in vivo*. Method 3 addresses this by applying the largest ROI established from method 2 to all images in a dataset, regardless of MPI signal strength and dimensions. Lastly, method 4 uses a uniform threshold-based segmentation at a factor of the background SD ($c \cdot SD$) which results in larger ROIs. This method provided accurate and precise estimates of ferucarbotran mass for *in vitro* phantoms, including samples which were in larger fluid volumes. Likewise, *in vivo* quantification of ferucarbotran administered systemically was accurate with method 4. The limitation of this method is the use of large ROIs which can be challenging when attempting to quantify multiple MPI signals in a single image. ROI method 4 overall has many advantages, including quick analysis, simplicity, and high reproducibility between users.

No single method meets all desired criteria (Table 3), therefore a careful choice of ROI selection method must be made for analysing each dataset. This paper was focused on quantification of ferucarbotran but these ROI methods are applicable for other SPIONs. Our hope is these ROI selection methods will be widely adopted by

MPI users to improve the accuracy and consistency of SPION quantification.

References

[1] A. Martín-Fontecha, S. Sebastiani, U. E. Höpken, M. Lipp, A. Lanzavecchia, and F. Sallusto. Regulation of Dendritic Cell Migration to the Draining Lymph Node. *Journal of Experimental Medicine*, 198(4):615–621, 2003, doi:10.1084/jem.20030448.

[2] F. Chapelin, C. M. Capitini, and E. T. Ahrens. Fluorine-19 MRI for detection and quantification of immune cell therapy for cancer. *Journal for ImmunoTherapy of Cancer*, 6(1):105, 2018, doi:10.1186/s40425-018-0416-9.

[3] O. C. Sehl, J. J. Gevaert, K. P. Melo, N. N. Knier, and P. J. Foster. A Perspective on Cell Tracking with Magnetic Particle Imaging. *Tomography*, 6(4):315–324, 2020, doi:10.18383/j.tom.2020.00043.

[4] A. Rivera-Rodriguez, L. B. Hoang-Minh, A. Chiu-Lam, N. Sarna, L. Marrero-Morales, D. A. Mitchell, and C. M. Rinaldi-Ramos. Tracking adoptive T cell immunotherapy using magnetic particle imaging. *Nanotheranostics*, 5(4):431–444, 2021, doi:10.7150/ntno.55165.

[5] A. V. Makela, J. M. Gaudet, M. A. Schott, O. C. Sehl, C. H. Contag, and P. J. Foster. Magnetic Particle Imaging of Macrophages Associated with Cancer: Filling the Voids Left by Iron-Based Magnetic Resonance Imaging. *Molecular Imaging and Biology*, 22(4):958–968, 2020, doi:10.1007/s11307-020-01473-0.

[6] P. Chandrasekharan, K. B. Fung, X. Y. Zhou, W. Cui, C. Colson, D. Mai, K. Jeffris, Q. Huynh, C. Saayujya, L. Kabuli, B. Fellows, Y. Lu, E. Yu, Z. W. Tay, B. Zheng, L. Fong, and S. M. Conolly. Non-radioactive and sensitive tracking of neutrophils towards inflammation using antibody functionalized magnetic particle imaging tracers. *Nanotheranostics*, 5(2):240–255, 2021, doi:10.7150/ntno.50721.

[7] S. Liu, A. Chiu-Lam, A. Rivera-Rodriguez, R. DeGroff, S. Savliwala, N. Sarna, and C. M. Rinaldi-Ramos. Long circulating tracer tailored for magnetic particle imaging. *Nanotheranostics*, 5(3):348–361, 2021, doi:10.7150/ntno.58548.

[8] V. Dousset, C. Delalande, L. Ballarino, B. Quesson, D. Seilhan, M. Coussemacq, E. Thiaudiére, B. Brochet, P. Canioni, and J.-M. Caillé. In vivo macrophage activity imaging in the central nervous system detected by magnetic resonance. *Magnetic Resonance in Medicine*, 41(2):329–333, 1999, doi:10.1002/(SICI)1522-2594(199902)41:2<329::AID-MRM17>3.0.CO;2-Z.

[9] H. Nejadnik, J. Tseng, and H. Daldrup-Link. Magnetic resonance imaging of stem cell–macrophage interactions with ferumoxytol and ferumoxytol-derived nanoparticles. *WIREs Nanomedicine and Nanobiotechnology*, 11(4):1–13, 2019, doi:10.1002/wnan.1552.

[10] G. Ren, J. M. Gaudet, M. Gerosa, Y. Zhang, J. Mansfield, M. Wintermark, and P. Goodwill. Abstract 2770: Imaging cancer immunology: Monitoring CD47 mAb treatment in vivo by magnetic particle imaging. *Cancer Research*, 80(16_Supplement):2770–2770, 2020, doi:10.1158/1538-7445.AM2020-2770.

[11] M. Iv, P. Samghabadi, S. Holdsworth, A. Gentles, P. Rezaii, G. Harsh, G. Li, R. Thomas, M. Moseley, H. E. Daldrup-Link, H. Vogel, M. Wintermark, S. Cheshier, and K. W. Yeom. Quantification of Macrophages in High-Grade Gliomas by Using Ferumoxytol-enhanced MRI: A Pilot Study. *Radiology*, 290(1):198–206, 2019, doi:10.1148/radiol.2018181204.

[12] M. Aghighi, A. J. Theruvath, A. Pareek, L. L. Pisani, R. Alford, A. M. Muehe, T. K. Sethi, S. J. Holdsworth, F. K. Hazard, D. Gratzinger, S. Luna-Fineman, R. Advani, S. L. Spunt, and H. E. Daldrup-Link. Magnetic Resonance Imaging of Tumor-Associated Macrophages: Clinical Translation. *Clinical Cancer Research*, 24(17):4110–4118, 2018, doi:10.1158/1078-0432.CCR-18-0673.

[13] A. V. Makela, J. M. Gaudet, and P. J. Foster. Quantifying tumor associated macrophages in breast cancer: a comparison of iron and fluorine-based MRI cell tracking. *Scientific Reports*, 7(1):42109, 2017, doi:10.1038/srep42109.

[14] D. Hensley, Z. W. Tay, R. Dhavalikar, B. Zheng, P. Goodwill, C. Rinaldi, and S. Conolly. Combining magnetic particle imaging and magnetic fluid hyperthermia in a theranostic platform. *Physics in Medicine and Biology*, 62(9):3483–3500, 2017, doi:10.1088/1361-6560/aa5601.

[15] Z. W. Tay, P. Chandrasekharan, A. Chiu-Lam, D. W. Hensley, R. Dhavalikar, X. Y. Zhou, E. Y. Yu, P. W. Goodwill, B. Zheng, C. Rinaldi, and S. M. Conolly. Magnetic Particle Imaging-Guided Heating in Vivo Using Gradient Fields for Arbitrary Localization of Magnetic Hyperthermia Therapy. *ACS Nano*, 12(4):3699–3713, 2018, doi:10.1021/acsnano.8b00893.

[16] S. Healy, A. F. Bakuzis, P. W. Goodwill, A. Attaluri, J. W. M. Bulte, and R. Ivkov. Clinical magnetic hyperthermia requires integrated magnetic particle imaging. *WIREs Nanomedicine and Nanobiotechnology*, 14(3):1–20, 2022, doi:10.1002/wnan.1779.

[17] H. Hayat, A. Sun, H. Hayat, S. Liu, N. Talebloo, C. Pinger, J. O. Bishop, M. Gudi, B. F. Dwan, X. Ma, Y. Zhao, A. Moore, and P. Wang. Artificial Intelligence Analysis of Magnetic Particle Imaging for Islet Transplantation in a Mouse Model. *Molecular Imaging and Biology*, 23(1):18–29, 2021, doi:10.1007/s11307-020-01533-5.

[18] N. C. Krak, R. Boellaard, O. S. Hoekstra, J. W. R. Twisk, C. J. Hoekstra, and A. A. Lammertsma. Effects of ROI definition and reconstruction method on quantitative outcome and applicability in a response monitoring trial. *European Journal of Nuclear Medicine and Molecular Imaging*, 32(3):294–301, 2005, doi:10.1007/s00259-004-1566-1.

[19] Y. W. S. Jauw, F. Bensch, A. H. Brouwers, O. S. Hoekstra, J. M. Zijlstra, S. Pieplenbosch, C. P. Schröder, S. Zweegman, G. A. M. S. van Dongen, C. W. Menke-van der Houven van Oordt, E. G. E. de Vries, H. C. W. de Vet, R. Boellaard, and M. C. Huisman. Interobserver reproducibility of tumor uptake quantification with 89Zr-immuno-PET: a multicenter analysis. *European Journal of Nuclear Medicine and Molecular Imaging*, 46(9):1840–1849, 2019, doi:10.1007/s00259-019-04377-6.

[20] N. Sato, K. Stringaris, J. K. Davidson-Moncada, R. Reger, S. S. Adler, C. Dunbar, P. L. Choyke, and R. W. Childs. In Vivo Tracking of Adoptively Transferred Natural Killer Cells in Rhesus Macaques Using 89Zirconium-Oxine Cell Labeling and PET Imaging. *Clinical Cancer Research*, 26(11):2573–2581, 2020, doi:10.1158/1078-0432.CCR-19-2897.

[21] J. W. Keyes. SUV: standard uptake or silly useless value? *Journal of nuclear medicine: official publication, Society of Nuclear Medicine*, 36(10):1836–1839, 1995. URL: http://www.ncbi.nlm.nih.gov/pubmed/7562051.

[22] N. S. Sarna, L. Marrero-Morales, R. DeGroff, A. Rivera-Rodriguez, S. Liu, A. Chiu-Lam, H. J. Good, and C. M. Rinaldi-Ramos. An anatomically correct 3D-printed mouse phantom for magnetic particle imaging studies. *Bioengineering & Translational Medicine*, 2022, doi:10.1002/btm2.10299.

[23] P. Keselman, E. Y. Yu, X. Y. Zhou, P. W. Goodwill, P. Chandrasekharan, R. M. Ferguson, A. P. Khandhar, S. J. Kemp, K. M. Krishnan, B. Zheng, and S. M. Conolly. Tracking short-term biodistribution and long-term clearance of SPIO tracers in magnetic particle imaging. *Physics in Medicine and Biology*, 62(9):3440–3453, 2017, doi:10.1088/1361-6560/aa5f48.

[24] F. Griese, T. Knopp, R. Werner, A. Schlaefer, and M. Möddel. Submillimeter-Accurate Marker Localization within Low Gradient Magnetic Particle Imaging Tomograms. *International Journal on Magnetic Particle Imaging*, 3(1), 2017, doi:10.18416/IJMPI.2017.1703011.

[25] L. Wöckel, J. Wells, O. Kosch, S. Lyer, C. Alexiou, C. Grüttner, F. Wiekhorst, and S. Dutz. Long-term stable measurement phantoms for magnetic particle imaging. *Journal of Magnetism and Magnetic Materials*, 471:1–7, 2019, doi:10.1016/j.jmmm.2018.09.012.

[26] J. Zhong, M. Schilling, and F. Ludwig. Simultaneous Imaging of Magnetic Nanoparticle Concentration, Temperature,

and Viscosity. *Physical Review Applied*, 16(5):054005, 2021, doi:10.1103/PhysRevApplied.16.054005.

[27] P. Szwargulski, M. Wilmes, E. Javidi, F. Thieben, M. Graeser, M. Koch, C. Gruettner, G. Adam, C. Gerloff, T. Magnus, T. Knopp, and P. Ludewig. Monitoring Intracranial Cerebral Hemorrhage Using Multicontrast Real-Time Magnetic Particle Imaging. *ACS Nano*, 14(10):13913–13923, 2020, doi:10.1021/acsnano.0c06326.

[28] P. Wang, P. W. Goodwill, P. Pandit, J. Gaudet, A. Ross, J. Wang, E. Yu, D. W. Hensley, T. C. Doyle, C. H. Contag, S. Conolly, and A. Moore. Magnetic particle imaging of islet transplantation in the liver and under the kidney capsule in mouse models. *Quantitative Imaging in Medicine and Surgery*, 8(2):114–122, 2018, doi:10.21037/qims.2018.02.06.

[29] J. Haegele, J. Rahmer, B. Gleich, J. Borgert, H. Wojtczyk, N. Panagiotopoulos, T. M. Buzug, J. Barkhausen, and F. M. Vogt. Magnetic Particle Imaging: Visualization of Instruments for Cardiovascular Intervention. *Radiology*, 265(3):933–938, 2012, doi:10.1148/radiol.12120424.

[30] J. Guzy, S. Chakravarty, F. J. Buchanan, H. Chen, J. M. Gaudet, J. M. L. Hix, C. L. Mallett, and E. M. Shapiro. Complex Relationship between Iron Oxide Nanoparticle Degradation and the Signal Intensity in Magnetic Particle Imaging. *ACS Applied Nano Materials*, 3(5):3991–3999, 2020, doi:10.1021/acsanm.0c00779.

[31] H. Paysen, N. Loewa, A. Stach, J. Wells, O. Kosch, S. Twamley, M. R. Makowski, T. Schaeffter, A. Ludwig, and F. Wiekhorst. Cellular uptake of magnetic nanoparticles imaged and quantified by magnetic particle imaging. *Scientific Reports*, 10(1):1922, 2020, doi:10.1038/s41598-020-58853-3.

[32] M. G. Kaul, T. Mummert, M. Graeser, J. Salamon, C. Jung, E. Tahir, H. Ittrich, G. Adam, and K. Peldschus. Pulmonary blood volume estimation in mice by magnetic particle imaging and magnetic resonance imaging. *Scientific Reports*, 11(1):4848, 2021, doi:10.1038/s41598-021-84276-9.

[33] A. V. Makela, M. A. Schott, C. S. Madsen, E. M. Greeson, and C. H. Contag. Magnetic Particle Imaging of Magnetotactic Bacteria as Living Contrast Agents Is Improved by Altering Magnetosome Arrangement. *Nano Letters*, 22(12):4630–4639, 2022, doi:10.1021/acs.nanolett.1c05042.

[34] O. C. Sehl and P. J. Foster. The sensitivity of magnetic particle imaging and fluorine-19 magnetic resonance imaging for cell tracking. *Scientific Reports*, 11(1):22198, 2021, doi:10.1038/s41598-021-01642-3.

[35] J. Dieckhoff, M. G. Kaul, T. Mummert, C. Jung, J. Salamon, G. Adam, T. Knopp, F. Ludwig, C. Balceris, and H. Ittrich. In vivo liver visualizations with magnetic particle imaging based on the calibration measurement approach. *Physics in Medicine and Biology*, 62(9):3470–3482, 2017, doi:10.1088/1361-6560/aa562d.

[36] S. Herz, P. Vogel, T. Kampf, M. A. Ruckert, S. Veldhoen, V. C. Behr, and T. A. Bley. Magnetic Particle Imaging for Quantification of Vascular Stenoses: A Phantom Study. *IEEE Transactions on Medical Imaging*, 37(1):61–67, 2018, doi:10.1109/TMI.2017.2717958.

[37] J. J. Konkle, P. W. Goodwill, D. W. Hensley, R. D. Orendorff, M. Lustig, and S. M. Conolly. A Convex Formulation for Magnetic Particle Imaging X-Space Reconstruction. *PLOS ONE*, 10(10):e0140137J. Najbauer, Ed., 2015, doi:10.1371/journal.pone.0140137.

[38] J. J. Gevaert, C. Fink, J. D. Dikeakos, G. A. Dekaban, and P. J. Foster. Magnetic Particle Imaging Is a Sensitive In Vivo Imaging Modality for the Detection of Dendritic Cell Migration. *Molecular Imaging and Biology*, 2022, doi:10.1007/s11307-022-01738-w.

[39] A. Rose. The Sensitivity Performance of the Human Eye on an Absolute Scale. *Journal of the Optical Society of America*, 38(2):196, 1948, doi:10.1364/JOSA.38.000196.

[40] G. M. Siegers, S. Krishnamoorthy, L. E. Gonzalez-Lara, C. McFadden, Y. Chen, and P. J. Foster. Pre-Labeling of Immune Cells in Normal Bone Marrow and Spleen for Subsequent Cell Tracking by MRI. *Tomography*, 2(1):26–34, 2016, doi:10.18383/j.tom.2016.00103.

[41] R. Lawaczeck, H. Bauer, T. Frenzel, M. Hasegawa, Y. Ito, K. Kito, N. Miwa, H. Tsutsui, H. Vogler, and H.-J. Weinmann. Magnetic Iron Oxide Particles Coated with Carboxydextran for Parenteral Administration and Liver Contrasting. *Acta Radiologica*, 38(4):584–597, 1997, doi:10.1080/02841859709174391.

[42] G. A. Soares, J. V. C. Faria, L. A. Pinto, A. G. Prospero, G. M. Pereira, E. G. Stoppa, L. P. Buranello, A. F. Bakuzis, O. Baffa, and J. R. A. Miranda. Long-Term Clearance and Biodistribution of Magnetic Nanoparticles Assessed by AC Biosusceptometry. *Materials*, 15(6):2121, 2022, doi:10.3390/ma15062121.

[43] H. Paysen, N. Loewa, K. Weber, O. Kosch, J. Wells, T. Schaeffter, and F. Wiekhorst. Imaging and quantification of magnetic nanoparticles: Comparison of magnetic resonance imaging and magnetic particle imaging. *Journal of Magnetism and Magnetic Materials*, 475:382–388, 2019, doi:10.1016/j.jmmm.2018.10.082.

[44] M. Graeser, P. Ludewig, P. Szwargulski, F. Foerger, T. Liebing, N. D. Forkert, F. Thieben, T. Magnus, and T. Knopp. Design of a head coil for high resolution mouse brain perfusion imaging using magnetic particle imaging. *Physics in Medicine & Biology*, 65(23):235007, 2020, doi:10.1088/1361-6560/abc09e.

[45] H. Suzuka, A. Mimura, Y. Inaoka, and K. Murase. Magnetic Nanoparticles in Macrophages and Cancer Cells Exhibit Different Signal Behavior on Magnetic Particle Imaging. *Journal of Nanoscience and Nanotechnology*, 19(11):6857–6865, 2019, doi:10.1166/jnn.2019.16619.

[46] M. Utkur, Y. Muslu, and E. U. Saritas. Relaxation-based color magnetic particle imaging for viscosity mapping. *Applied Physics Letters*, 115(15):152403, 2019, doi:10.1063/1.5110475.

[47] H. Arami, R. M. Ferguson, A. P. Khandhar, and K. M. Krishnan. Size-dependent ferrohydrodynamic relaxometry of magnetic particle imaging tracers in different environments. *Medical Physics*, 40(7):071904, 2013, doi:10.1118/1.4810962.

[48] E. Teeman, C. Shasha, J. E. Evans, and K. M. Krishnan. Intracellular dynamics of superparamagnetic iron oxide nanoparticles for magnetic particle imaging. *Nanoscale*, 11(16):7771–7780, 2019, doi:10.1039/C9NR01395D.