Proton Density Fat Fraction (PDFF) Measurement of Myelofibrosis in Mouse Tibia

version 20230214

I. Executive Summary

The goal of this Co-Clinical Imaging Research Program (CIRP) pre-clinical imaging procedure entitled, "Proton Density Fat Fraction (PDFF) Measurement of Myelofibrosis in Mouse Tibia", is to provide detailed description of key steps used to achieve a stated level of performance embodied in "Claims", for MRI measurement of PDFF in mouse tibia of myelofibrosis mouse models. This pre-clinical imaging procedure document will be referred to as a "profile" since it has been designed to have some common features with a Quantitative Imaging Biomarkers Alliance (QIBA) Profile targeting standardization of human quantitative imaging procedures (http://qibawiki.rsna.org/index.php/Main_Page).

Myelofibrosis (MF) is a chronic, ultimately fatal myeloproliferative neoplasm caused by genetic mutations in hematopoietic stem cells leading to systemic inflammation and progressive fibrosis disrupting normal architecture and composition of the bone marrow[1, 2]. Bone marrow biopsy, which is painful and subject to sampling error, remains the default method to assess MF disease in humans. The University of Michigan (UM) CIRP U24 CA 237683 project involves a longitudinal study design in human MF patients and mouse MF models to develop noninvasive quantitative bone marrow MRI methods sensitive to alteration of bone marrow composition due to myelofibrosis evolution and response to MF treatments[3]. The UM CIRP project involves measurement of three image-based metrics (apparent diffusion coefficient (ADC), PDFF, and magnetization transfer ratio (MTR)) that have potential to objectively document MF disease status. Profiles corresponding to ADC and MTR measurement of MF mouse tibia are also available on the UM CIRP website (UMU24CIRP (umich.edu)) with links to Protocols.io. Normal bone contains red marrow where red blood cells, platelets and white blood cells are created, and yellow marrow that contains fat [4-6]. Not only is PDFF change a potential reflection of disease evolution or response to treatment, but coexistence of water and fat constituents in MRI signal can have a profound impact on water mobility measurements being employed in the UM CIRP project. This document details procedures for PDFF measurement in MF mouse tibia to achieve stated performance claims. In this profile PDFF values are expressed in "% units" on a 0 to 100% scale, such that bone marrow containing dominant red marrow and dominant yellow marrow have PDFF values of 0-few% and 80-100%, respectively.

II. Pre-Clinical Imaging Claims

Tibia bone marrow composition in MF mouse models has gradation going from proximal to distil ends of the tibia, therefore separate claims are made for volume of interest (VOI) analysis of PDFF maps for each of three distinct sections along the length of the tibia (see Figure 1):

Section 1 (proximal) ≡ VOI (~4-5mm^3) within 9mm of proximal end of tibia
Section 2 (transition) ≡ VOI (~0.4-0.5mm^3) from 10 to 12mm of proximal end of tibia

Section 3 (distil) ≡ VOI (~0.1-0.2mm^3) from 13 to 14mm of proximal end of tibia

Claim 1: A measured change in the mean PDFF in Section 1 VOI of MF mouse model tibia that exceeds ±1.6% indicates a true biological change has occurred in the tibia bone marrow with 95% confidence.

Claim 2: A measured change in the mean PDFF in Section 2 VOI of MF mouse model tibia that exceeds ±15.5% indicates a true biological change has occurred in the tibia bone marrow with 95% confidence

Claim 3: A measured change in the mean PDFF in Section 3 VOI of MF mouse model tibia that exceeds ±25.5% indicates a true biological change has occurred in the tibia bone marrow with 95% confidence

The claims hold when:

- **Scanner hardware, proton density-weighted multi gradient-echo data acquisition method and parameters, image reconstruction, and data-reduction procedures are equivalent (or superior) to those detailed in section III.**
- **Use of the same animal model and interventions to induce myelofibrosis are performed as detailed in section V.**
- **PDFF change is assessed on an individual animal basis where each animal undergoes identical procedures on the same MRI system over longitudinal timepoints.**

III. MR Imaging Process Specifications

1. MRI Scanner Hardware
    i. Bruker BioSpec® MRI Console Paravision 7.0.0 software installed on 64 bit Linux multicore workstation, 16 GB RAM, 1TB hard disk.
    ii. 7Tesla, 30cm bore magnet model “7T/310/AS” System (Agilent) with compact Faraday RF-shielding cabinet is attached to the magnet service end
    iii. System gradient/shim coil set model “B-GA12S HP” with standard 300V/200A gradient amplifier and standard 5A shim amplifier:
        a. Inner diameter 114mm
        b. Gradient strength 440 mT/m
        c. Max slew rate 3,440 T/m/s
        d. 10 shim channels, up to 4th order shim coils
    iv. Large Transmit/Receive RF Volume Coil: Outer/inner diameter 112mm/86mm
    v. Medium Transmit/Receive RF Volume Coil: Outer/inner diameter 75mm/40mm RF RES 300 1H 075/040 QSN TR; model 1P T13161V3.
    vi. Small Receive: CryoProbe™ 4 Element Array RF Coil Kit for Mice cryogenically cooled to 20-30^{o}K with parallel receiver upgrade.

2. Acquisition Technique

i. Multi-echo chemical shift encoding to decompose water and fat signal constituents at 7T requires echo-to-echo spacing <0.4ms which is difficult to achieve in a single echo train while maintaining high spatial resolution to image mouse tibia. Instead, four consecutive series were acquired for retrospective combination, where each series is a 3-echo, gradient-echo sequence (MGE) via "Method=<Bruker:MGE>".
ii. Echo spacing between the 3 echoes is held constant at 2.1739ms, although TE of the first echo was changed over the 4 consecutive series: Ser1_TE1 = 1.4757ms; Ser2_TE1 = 1.793ms; Ser3_TE1 = 2.11ms; Ser4_TE1 = 2.427ms.
iii. Hardware settings, acquisition geometry, shim, transmit gain and receiver gains were held constant over the 4 consecutive series such that the 4 series data may be retrospectively combined and sorted by TE into an effective single 12-echo train. Assuming the spectrometer is stable, this scenario samples the evolution between water and fat signal constituents with 0.3173ms temporal resolution which is adequate for chemical shift signal decomposition at 7T.
iv. Three-D multi-slice in coronal plane with geometry:

Table I: MGE geometry

Matrix	Acquired Voxel Size (μm)	FOV (mm)
256 (freq enc on z-axis)	90 on z	23.04 on z
128 (phase enc on x-axis	75 on x	9.6 on x
64 (phase enc on y-axis)	94 (slice thickness) on y	6.0 on y

v. Contrast Control: TR/[TE]=50/[1.4757; 1.7930; 2.1100; 2.4270; 3.6496; 3.9669; 4.2839; 4.6009; 5.8235; 6.1408; 6.4578; 6.7748ms]; Flip angle = 5°. 2 NSA. Four MGE series total scan time = 13.7min x 4 = 55min.
vi. Conventional 3D sequential cartesian k-space trajectory, 1 phase-encode / TR.
vii. Full k-space acquisition, no acceleration, no multi-band, no turbo spin-echo.
viii. No physiologic synchronization. Anesthetized mouse leg is held in place between 3D-printed, leg-shaped mold on posterior side and CryoProbe™ on anterior side.

3. Image Reconstruction
   i. Time-domain MGE are reconstructed to complex-valued space-domain on the Bruker MRI system using standard 3DFT image reconstruction routines within the Paravision 7.0.0 environment. Three-D images for all 3 echoes of each series are stored "pdata" data tree in "2dseq" datafile, a native Bruker format. Associated Bruker-generated text files "acqp", "AdjStatePerScan", "configscan", "method", "pulseprogram", "specpar", "id", "methreco", "procs", "reco", and "visu_pars" are stored in series-level folders within the exam-level folder. Exam-level folders, coded by group+mouseID+acqdate+acqtime, are transferred to UM-maintained drives for archival and subsequent analysis described below.

4. Biomarker Map Generation

i. Conversion from Bruker MGE complex-valued 2dseq-format to PDFF-related ITK-compatible format images is performed using a custom MATLAB (ver ≥2016b) "ProcessLegTLC.m" script that calls scripts within Bruker-provided "pvmatlab" (ver 2013) MATLAB package for handling ParaVision data.
ii. Initially, ProcessLegTLC surveys all "method" text files for parameter values of keywords to create a catalog of the entire exam. The catalog is stored in a "GroupID_MouseID_DateTime_Catalog.tsv" text file containing key attributes for each series: [Series#; ScanTime; Protocol; Method; TR; TE; NEcho; Nave; PVM_Matrix; Thk; Nslc; MTsatFreq; MTstate; Receiver_Gain; ReconSize; bvalue].
iii. For all series cataloged with keyword "Method" = "MGE", the ProcessLegTLC script calls "FAT_ProcessLegTLC.m" script which loads all MGE 2dseq data into complex-valued 4-dimensional arrays (x, y, z, TE) for sorting by increasing TE-values. Note, at this point the 4th dimension contains 12 echoes.
iv. Water and fat signals are separated using MatLab "ISMRM Fat-Water Toolbox". Specifically the "fw_i2cm1i_3pluspoint_hernando.m" graphcut script is used [7]. Inputs to this script are the 4-dimensional complex-valued image array, list of TEs, and Graphcut Algorithm Parameters (GAP) given in Appendix II.
v. Success of fat-water decomposition depends on the GAP set, B0 shim and data quality. Occasionally, fat and water identities are erroneously reversed on some slices, where an alternative GAP set may perform better. Our PDFF processing pipeline determines when an erroneous water-fat swap occurs by integrating fat (F) and water (W) signal across all pixels in the given slice. If F > W, the alternative GAP set was automatically applied. Which GAP set was used for each slice was recorded in a "mat file".
vi. The following 3D images were stored on disk in ITK-compatible MetaImage format [8] comprised of paired header (eg. "ABCD.mhd") and binary data (eg. "ABCD.raw"): Fat Only (arb units); Water Only (arb units); Mean MGE (average of all echoes in arb units); R2* (in Hz); PDFF (in % units); and B0 Field Map (in Hz). MetaImage format 3D volumes provide compactness and portability to commonly used medical/scientific image viewing platforms, such as ITK-SNAP and 3DSlicer [9]. Geometry content (slice location, extent and angulation) are retained in the mhd header structure.

5. Additional Map Conditioning

Data reported in this profile *DID NOT* undergo additional conditioning beyond that detailed in sections 6 through 9 below. However, spatial registration of image volumes will be performed in UM CIRP U24 CA 237683 study of MF in mouse tibia to longitudinally follow disease evolution and response to treatment. For completeness, the process for spatial registration of a mouse tibia over time is described below:

i. MATLAB scripts with a GUI interface were developed in the UM Center for Molecular Imaging (CMI) to facilitate MATLAB calls to Elastix (ver 4.8) image registration software (elastix: download (lumc.nl)). This tool was designed to input/output MetaImage format 3D volumes as created by ProcessLegTLC script.

ii. In addition to multi-slice PDFF-related images, the MF tibia MRI protocol includes a variety of imaging contrasts including: DWI and magnetization-transfer (MT) RF pulses "OFF" and "ON" for calculation of magnetization transfer ratio (MTR). The 3D MTOFF is used as the reference "FIXED" image volume to which the Mean MGE, here called "MOVER", is transformed to spatially align with the FIXED volume.

iii. The tibia is effectively a rigid body, whereas surrounding muscle may be deformed over serial MRI sessions. Therefore, the tibia is manually segmented on the FIXED image baseline timepoint using the CMI GUI (or 3DSlicer) and saved in MetaImage format. A dilated version of the tibia VOI segmentation serves as a mask to drive the Elastix rigid-body volume registration routine based on tibia anatomic features without regard to muscle features.

iv. The CMI GUI prompts user to identify 4 points on the FIXED tibia along with 4 homologous points on the MOVER tibia to initialize the Euler (rigid body) transformation. For this step, two landmarks on the distil end of the tibia and two mid-tibia are used.

v. Main Elastix parameters are:
   a. FIXED = "*_MT_Off.mhd"
   b. MOVER = "*_MGEmean.mhd"
   c. MASK = "*_MT_Off_VOI.mhd"
   d. Sample: RandomSparseMask
   e. Metric: Mutual Information
   f. Transform: Euler
   g. Final Interpolation: BSpline-3
   h. Iterations: 4000
   i. NSamples: 5000

vi. Complete Elastix parameter file is provided in Appendix I. Since all PDFF-related 3D images have identical acquisition geometry, the final spatial transformation that aligns "MGEmean" (mover) to fixed reference is applied to other PDFF-related 3D volumes. All spatially registered volumes are output to disk with their original filename amended by an "_R". Elastixlog.txt, ElastixParameters.txt, and TransformParamters.txt are also retained in sub-folders for future use.

6. Region / Volume of Interest (ROI / VOI) Segmentation
   i. For longitudinal datasets spatially-registered to common (FIXED) dataset, the original tibia segmentation saved in "*_MT_Off_VOI.mhd" may be applied to any co-registered datasets. This *WAS NOT DONE* nor was it necessary for data, analysis and results reported in this profile.
   ii. For this profile reporting repeatability of PDFF measurement in MF mouse tibia using Test-retest design, the tibia was manually segmented independently for each Test and ReTest dataset. Also note, the CMI GUI or other 3D image analysis platform such as 3DSlicer may be used to manually segment tibia with the resultant output stored in MetaImage format with "*_VOI.mhd" filename.

For this profile, both Test and ReTest datasets from all mice and timepoints were manually segmented using the CMI GUI by one individual (KH).

iii. Gradation in PDFF from proximal to distil ends of the tibia was dealt with by separate analysis of three sections along the length of the tibia, each with its own repeatability claim.

iv. To systematically define these sections, a dedicated MATLAB script was created to read-in the VOI mask and automatically detect a landmark (knee) defined as the Z-slice containing largest VOI cross-section in X-Y plane. Then given known slice thickness, three sections were parsed by Z-dimension from the VOI as displayed in Figure 1.

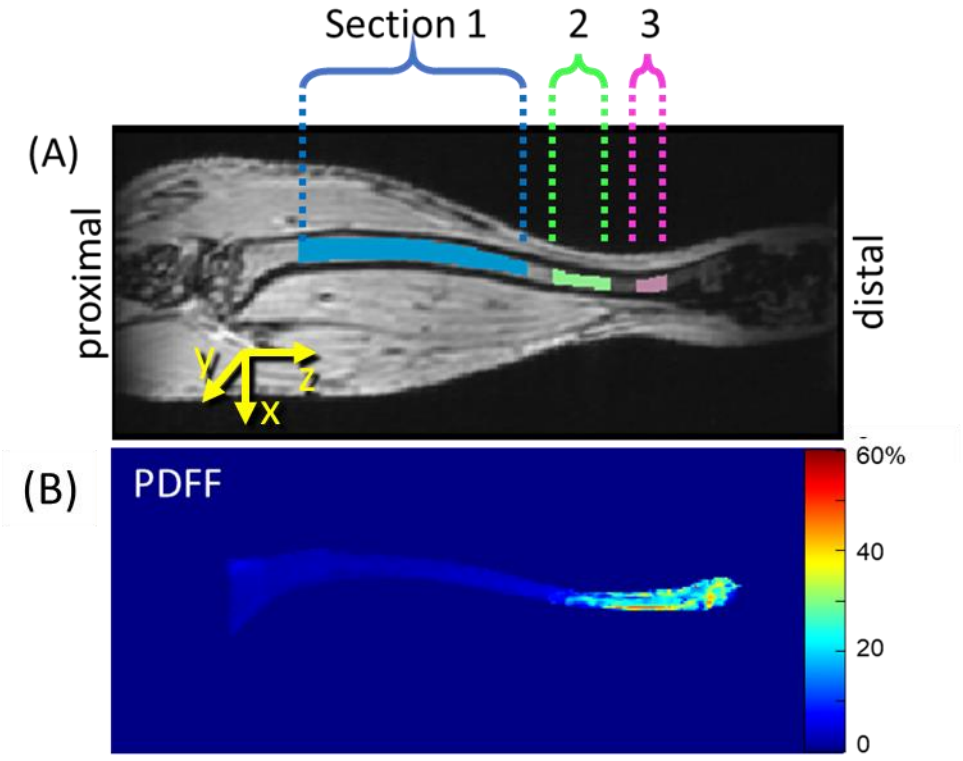


Figure 1: (A) Coronal slice through volumes of interest (VOI) along length of tibia relative to reference point determined by greatest axial tibia area assigned location "z=0mm". Relative to this reference, Section 1 spans z = 1.8 to 9mm, Section 2 z = 9.8 to 11.7mm, and Section 3 z = 12.6 to 13.5mm for all mice. (B) Coronal projection of mean PDFF value within the tibia are shown in % units.

v. PDFF map quality is dependent on shim and SNR. Occasionally, artifactual PDFF maps result. Severely artifactual PDFF maps are apparent by inspection and will be eliminated in future longitudinal studies. Therefore QC vetting was performed for this repeatability study, wherein a PDFF "quality score" (A=Good; B=Acceptable; C=Poor) was assigned to each dataset for each timepoint (Figure 2). A Test-retest pair was eliminated from analysis if there was a Poor quality score on either date. Quality of Sections 1, 2, and 3 were scored independently.

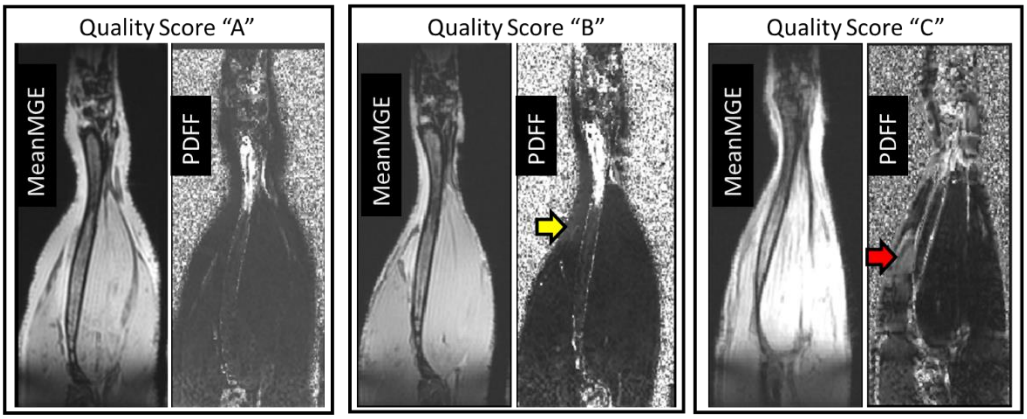
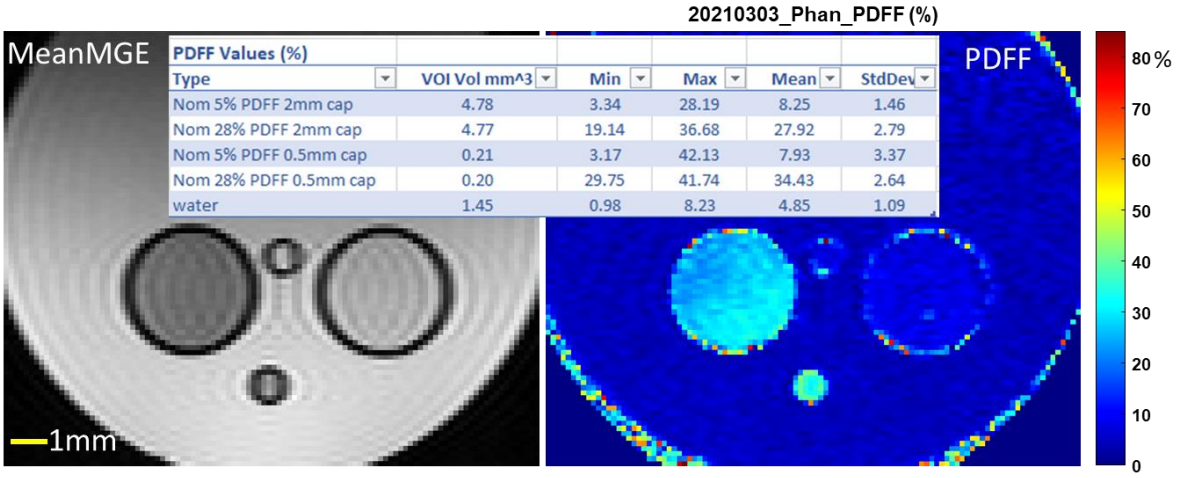


Figure 2: Representative PDFF quality scores A=Good; B=Acceptable; C=Poor. Uniformly low PDFF in muscle adjacent to the bone is an indication of good quality. Note, slightly elevated PDFF marked by yellow arrow and high artifact marked by red arrow. High PDFF in the distil end of tibia is not artifact as confirmed by histology.

7. Biomarker Metric
    i. Three-D masks for Sections 1, 2, and 3 were applied independently to PDFF maps having Quality Score A or B for both Test and ReTest scans. The mean PDFF of voxels in each section is taken as the "biomarker". Repeatability of mean PDFF for each section was analyzed by Test – Retest study design where MF model mice were scanned on two consecutive days where biological change in the tibia over 1 day is considered to be small/insignificant.

8. Imaging System Performance Validation
    i. PDFF phantom materials: Water based emulsifying gel (95% water) was mixed with ionic surfactants and cetearyl alcohols (5% solids) at 79° C. The gel and variable amounts of peanut oil were heated and vortexed to emulsify the mixture to create low PDFF (~5%) and moderate PDFF (~25%) material. The emulsions were added to large sample tubes and capillary tubes for PDFF measurement on a 3T clinical system and the 7T pre-clinical system, respectively. This 3T system was used in a round-robin study of a commercial PDFF phantom and was shown to be accurate in PDFF measurement[10], thus provided reference standard PDFF values for the same materials used in a pre-clinical phantom to validate the 7T pre-clinical methods. Consistent with mixture targets, reference PDFF values measured at 3T were 5% (±0.6%) and 28% (±1%) (data not shown).
    ii. Chemical-shift-encoding requires multiple gradient echoes with short echo-to-echo spacing which difficult to achieve at 7T while maintaining high spatial resolution. To demonstrate accurate PDFF quantification of small targets a phantom was constructed within a 10mm ID test-tube. In it, two large (2mm diam) and two small (0.5mm diam) capillary tubes were filled with nominal 5% and 28% PDFF materials. Note the 0.5mm capillary is comparable in size to mouse tibia (0.8-1.8mm). Ideally, PDFF measured in large and small capillaries would be equivalent and match reference values. As displayed in Figure 3, PDFF

values measured on the 7T pre-clinical systems were reasonably close the reference standards, although 7T PDFF was overestimated by a few %, even for pure water. There was no difference between capillaries for the low PDFF material, though there was a significant overestimation in moderate PDFF material in the small capillary probably due to inclusion of edge-artifact pixels.



PDFF Values (%)

Type	VOI Vol mm^3	Min	Max	Mean	StdDev
Nom 5% PDFF 2mm cap	4.78	3.34	28.19	8.25	1.46
Nom 28% PDFF 2mm cap	4.77	19.14	36.68	27.92	2.79
Nom 5% PDFF 0.5mm cap	0.21	3.17	42.13	7.93	3.37
Nom 28% PDFF 0.5mm cap	0.20	29.75	41.74	34.43	2.64
water	1.45	0.98	8.23	4.85	1.09

Figure 3: Axial slice through PDFF phantom (meanMGE on left, PDFF map on right) with VOI statistics in the table insert. At low PDFF, values on the pre-clinical system are a few percent higher than the reference standards. Moderate PDFF (nominal 28%) is correctly measured in the large (2mm diam) capillary, whereas PDFF is overestimated in the 0.5mm capillary by approximately 5%.

IV. Quantitative Metrics, Statistical Methods and Data Supporting Claim(s)

i. Repeatability of VOI mean PDFF in MF mouse tibia was assessed using a "test-retest" design study [11, 12].

ii. A total of 15 female BALB/c JAK2 mutant MF model mice were included in this study. MF disease model was induced by whole body irradiation to ablate bone marrow, followed by same day bone-marrow transplant (BMT) a 50% - 50% combination of normal and diseased cells per the protocol summarized in Section IV below. DWI was acquired over 4 to 10 weeks following BMT as the disease developed in the marrow space. Each test-retest DWI dataset pair was acquired on two consecutive days for each animal, followed by 9 to 14 days without DWI. Typically, test-retest pairs were spaced 10 to 14 days apart, such that a total of 37 test-retest pairs were collected from these 15 mice. Since the disease develops relatively slowly over the 10 weeks following BMT, the biological status of bone marrow over any given 24 hour period (i.e. a test-retest pair) is assumed to be effectively constant. Moreover, while correlation between pairs from a given animal is possible, the 37 test-retest pairs were treated as independent measurements to simplify test-retest analysis.

iii. To limit inclusion of artifactual data in the analysis, quality of PDFF maps were prospectively assessed and scored as Good, Acceptable, or Poor, to eliminate test-retest pairs scored by "Poor" on either scan date. Sections 1, 2, and 3 were

scored independently, and higher prevalence of Poor scores in distil sections meant fewer usable test-retest pairs in Sections 2 and 3 relative to Section 1.

iv. Explicit steps to assess repeatability followed Bland-Altman and QIBA-recommended procedures summarized as [11-13]:

a. Calculate mean (M) and variance (V) for each test-retest pair

$$M = \frac{(MeanPDFF_{test} + MeanPDFF_{retest})}{2};$$
$$V = \frac{(MeanPDFF_{test} - MeanPDFF_{retest})^2}{2}$$

b. For each of N-pairs, calculate $V/(M^2)$, take the mean over all N-pairs, then take square root to get within-subject coefficient of variation, wCV (in % units):

$$wCV = 100\% \cdot \sqrt{\frac{1}{N}\sum_i^N \frac{V_i}{M_i^2}}$$

c. Note, wCV is a relative (dimensionless) repeatability metric that becomes unreliable as M→0. For non-relative repeatability of PDFF, we use within-subject standard deviation, wSD, and repeatability coefficient, RC given by:

$$wSD = \sqrt{\frac{1}{N}\sum_i^N V_i}$$
$$RC = 2.77 \cdot wSD$$

d. The RC is a measure of precision and useful to infer minimum threshold of observed biomarker change attributable to true biologic change (with 95% confidence), as opposed to measurement error.

e. The corresponding 95% confidence interval (α = 0.05) to RC are given by multiplicative lower-bound (LB) and upper-bound (UB) factors given by ChiSqr function for N-1 degrees of freedom [11, 12]:

$$95\%\ CI\ of\ RC = RC \cdot \left[\frac{1}{\sqrt{(N-1)\cdot\ \chi^2_{0.975}}};\ \frac{1}{\sqrt{(N-1)\cdot\ \chi^2_{0.025}}}\right]$$

f. The same multiplicative factors were used to estimate 95% CI's for wCV and wSD indicated by [lower bound (LB), upper bound (UB)]. Test-retest results are summarized in Table II:

Table II	Section 1	Section 2	Section 3
N Test Retest Pairs	37	27	22
Mean PDFF (%)	3.4	21.9	50.3
Bias (PDFFretest – PDFFtest) (%)	+0.1	-3.5	-6.9
wCV [LB, UB] (%)	17.3 [14.0, 22.4]	24.1 [19.6, 31.3]	24.7 [20.1, 32.1]
wSD [LB, UB] (%)	0.6 [0.5, 0.7]	5.6 [4.5, 7.3]	9.2 [7.5, 11.9]
RC [LB, UB] (%)	1.6 [1.3, 2.0]	15.4 [12.6, 20.1]	25.5 [20.8, 33.2]

g. Bland-Altman x-y plots provide a graphic view of repeatability [13], where x = ($PDFF_{retest}$ + $PDFF_{test}$)/2 and y = ($PDFF_{retest}$ - PDF_{test}). Mean of y is a measure of bias (dotted line), or apparent PDFF change between test and retest conditions. Ideally, bias is close to zero thereby supporting the assumption that bone marrow was biologically constant between test and retest measurements. Dashed lines (bias ± 1.96*std($PDFF_{retest}$ - PDF_{test})) provide a graphical indication of measurement precision. Negative bias in Sections 2 and 3 over the 1-day test-retest interval is consistent loss of fat cells displaced by disease cells as the disease developed over 8wks.

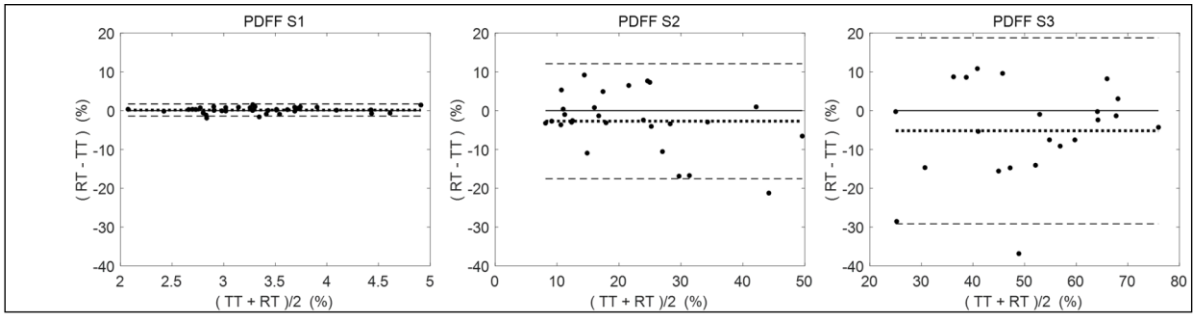


Figure 2: Bland-Altman plots graphically illustrate biomarker repeatability between Test (TT) and Retest (RT) paired measurements for PDFF. Plots for Sections S1, S2, and S3 are in left, center, right panels, respectively. Mean difference (RT–TT) shown as dotted lines, and dashed lines represent confidence range defined as mean difference ± 1.96 * std(RT-TT). Vertical scale is held constant for each given biomarker to illustrate repeatability dependence on section (see Figure 1). Refer to Table 2 for number of TT, RT-pairs in each plot.

V. Animal Model Specifications

1. Species: mouse
2. Strain: C57BL/6 purchased from Charles River Laboratories (Wilmington, MA).
3. Sex: Female
4. Disease induction: The JAK2 V617F (Jak2+/VF) animal model of myelofibrosis (MF) was generated using resultant 8-10 week old female donor offspring from a cross between Jak2+/Fl mice (B6N.129S6(SJL)-Jak2tm1.2Ble/AmlyJ; Jackson Laboratory Stock No. 031658) and Mx-Cre mice (B6.Cg-Tg(Mx1-cre)1Cgn/J; Jackson Laboratory Stock No. 003556), similar to previously described methods [9, 20]. Briefly, whole bone marrow cells were isolated from donor mice, and mixed 1:1 with whole bone marrow cells isolated from age- and gender-matched wild-type C57Bl/6 mice. A total of 1 x 107 mixed bone marrow cells was injected retro-orbitally into lethally irradiated 6 week old female recipient C57Bl/6 mice. Polyinosinic-polycytidylic acid (10 mg/kg) was administered intra-peritoneally 10 days post-bone marrow transplant (post-BMT) for induction of Cre recombinase-mediated replacement of the floxed endogenous exon with the mutated exon of Jak2 for expression of the JAK2 V617F mutant allele.
5. Therapeutic intervention – Not Applicable.
6. Animal age at start of MRI data acquisition: ~12 weeks (approximately 28 days post- BMT).
7. Animal prep and during imaging: 1.5% Isofluorane/air inhalation.

8. Animal monitoring/support during imaging:
    i. Thermoregulated heating bed during imaging
    ii. Respiratiory montoring (SAI monitor)
9. Animal recovery: isolated cage until full recovery, then back to communal cage
10. Imaging schedule: Each test-retest PDFF dataset pair was acquired on two consecutive days for each animal starting ~28 days post-BMT. Test-retest pairs were spaced approximately 10 to 14 days apart, such that a total of 37 test-retest pairs were collected from 15 mice.
11. UM Laboratory Animal Medicine Approval Code: #PRO00010851 and date: 09/26/22.

VI. Outcome Specifications

1. Time to moribund/survival
2. Longitudinal body weight and spleen volume measurements (last MRI) and spleen weight at sacrifice
3. Tibia and spleen tissues were harvested for flow cytometry and histological analysis
    i. Complete Blood Count (CBC); flow cytometry for immune cell populations
    ii. Liver and spleen weight
    iii. Liver, spleen and femur/tibia histology preparation and staining
    iv. Immune cell populations of spleen and bone marrow
    v. Immuno-blotting of spleen tissue

References

1. Garmezy, B., et al., *A provider's guide to primary myelofibrosis: pathophysiology, diagnosis, and management.* Blood Rev, 2021. **45**: p. 100691.
2. Schaefer, J.K., et al., *Primary myelofibrosis evolving to an aplastic appearing marrow.* Clin Case Rep, 2018. **6**(7): p. 1393-1395.
3. Luker, G.D., et al., *A Pilot Study of Quantitative MRI Parametric Response Mapping of Bone Marrow Fat for Treatment Assessment in Myelofibrosis.* Tomography, 2016. **2**(1): p. 67-78.
4. Bani Hassan, E., et al., *Bone Marrow Adipose Tissue Quantification by Imaging.* Curr Osteoporos Rep, 2019. **17**(6): p. 416-428.
5. Li, X. and A.V. Schwartz, *MRI Assessment of Bone Marrow Composition in Osteoporosis.* Curr Osteoporos Rep, 2020. **18**(1): p. 57-66.
6. Rosen, C.J., et al., *Marrow fat and the bone microenvironment: developmental, functional, and pathological implications.* Crit Rev Eukaryot Gene Expr, 2009. **19**(2): p. 109-24.
7. Hernando, D., et al., *Robust water/fat separation in the presence of large field inhomogeneities using a graph cut algorithm.* Magn Reson Med, 2010. **63**(1): p. 79-90.
8. *Metaimage MHD Format*. Available from: https://itk.org/Wiki/ITK/MetaIO/Documentation#:~:text=MetaImage%20is%20the%20text%2Dbased,library%20is%20known%20at%20MetaIO.
9. *3D Slicer*. Available from: https://www.slicer.org/#what-is-3d-slicer.
10. Hu, H.H., et al., *Linearity and Bias of Proton Density Fat Fraction as a Quantitative Imaging Biomarker: A Multicenter, Multiplatform, Multivendor Phantom Study.* Radiology, 2021. **298**(3): p. 640-651.
11. Raunig, D.L., et al., *Quantitative imaging biomarkers: a review of statistical methods for technical performance assessment.* Stat Methods Med Res, 2015. **24**(1): p. 27-67.
12. Winfield, J.M., et al., *Extracranial Soft-Tissue Tumors: Repeatability of Apparent Diffusion Coefficient Estimates from Diffusion-weighted MR Imaging.* Radiology, 2017. **284**(1): p. 88-99.
13. Bland, J.M. and D.G. Altman, *Statistical methods for assessing agreement between two methods of clinical measurement.* Lancet, 1986. **1**(8476): p. 307-10.

Appendix I: Elastix Parameters (automatically retained in “ElastixParameters.txt” file)

```
(FixedInternalImagePixelType "float")
(FixedImageDimension  3)
(MovingInternalImagePixelType "float")
(MovingImagedimension  3)
(UseDirectionCosines "false")
(WriteTransformParametersEachIteration "false")
(WriteTransformParametersEachResolution "true")
(WriteResultImageAfterEachResolution "false")
(WriteResultImage "true")
(Registration "MultiResolutionRegistration")
(Metric "AdvancedMattesMutualInformation")
(UseJacobianPreconditioning "false")
(FiniteDifferenceDerivative "false")
(ShowExactMetricValue "false")
(UseFastAndLowMemoryVersion "false")
(NumberOfHistogramBins  32)
(NumberOfFixedHistogramBins  32)
(NumberOfMovingHistogramBins  32)
(FixedLimitRangeRatio  0)
(MovingLimitRangeRatio  0)
(FixedKernelBSplineOrder  1)
(MovingKernelBSplineOrder  3)
(ImageSampler "RandomSparseMask")
(NumberOfSpatialSamples  5000)
(NewSamplesEveryIteration "true")
(UseRandomSampleRegion "false")
(CheckNumberOfSamples "true")
(Interpolator "LinearInterpolator")
(ResampleInterpolator "FinalBSplineInterpolator")
(FinalBSplineInterpolationOrder  3)
(Resampler "DefaultResampler")
(ResultImageFormat "mhd")
(ResultImagePixelType "float")
(ErodeFixedMask "false")
(ErodeMovingMask "false")
(DefaultPixelValue  0)
(Transform "EulerTransform")
(AutomaticTransformInitialization "false")
(AutomaticScalesEstimation "true")
(HowToCombineTransforms "Compose")
(Optimizer "StandardGradientDescent")
(NumberOfSamplesForSelfHessian  100000)
(NumberOfGradientMeasurements  0)
```

```
(NumberOfJacobianMeasurements 2700)
(NumberOfSamplesForExactGradient 100000)
(MaximumNumberOfIterations 4000)
(MaximumNumberOfSamplingAttempts 0)
(SP_a 2)
(SP_alpha 0.60200000)
(SP_A 50)
(NumberOfResolutions 1)
(FixedImagePyramid "FixedSmoothingImagePyramid")
(FixedImagePyramidSchedule 1 1 1)
(MovingImagePyramid "MovingSmoothingImagePyramid")
(MovingImagePyramidSchedule 1 1 1)
(Scales 10000 10000 10000 1 1 1)
```

Appendix II: Graphcut Parameters: Input to "fw_i2cm1i_3pluspoint_hernando.m" Matlab script within the "ISMRM Fat-Water Toolbox"

```
% ***************** Primary parameter set #1 ******************************
params1 = struct('species',struct('name',{'water','fat'},'frequency',{0,wF},'relAmps',{1,aF}),...
                 'size_clique',1,...
                 'range_r2star',[0 1000],...
                  'NUM_R2STARS',n_r2star,…
                 'range_fm',[-1000 2000],...
                 'NUM_FMS',301,...
                 'NUM_ITERS',40,...
                 'SUBSAMPLE',2,...
                 'DO_OT',1,...
                 'LMAP_POWER',2,...
                 'lambda',0.05,...
                 'LMAP_EXTRA',0.05,...
                 'TRY_PERIODIC_RESIDUAL',0);
% ********************************************************************************

% ***************** Alternative parameter set #2 ******************************
params2 = struct('species',struct('name',{'water','fat'},'frequency',{0,wF},'relAmps',{1,aF}),...
                 'size_clique',1,...
                 'range_r2star',[0 1000],...
                 'NUM_R2STARS',n_r2star,...
                 'range_fm',[-1500 1500],...
                 'NUM_FMS',301,...
                 'NUM_ITERS',40,...
                 'SUBSAMPLE',2,...
                 'DO_OT',1,...
                 'LMAP_POWER',2,...
                 'lambda',0.05,...
                 'LMAP_EXTRA',0.05,...
                 'TRY_PERIODIC_RESIDUAL',0);
% ********************************************************************************
```