Titre: Title:	Bringing Quantitative Magnetic Resonance Imaging Under One Umbrella
Auteur: Author:	Agah Karakuzu
Date:	2022
Type:	Mémoire ou thèse / Dissertation or Thesis
Référence: Citation:	Karakuzu, A. (2022). Bringing Quantitative Magnetic Resonance Imaging Under One Umbrella [Thèse de doctorat, Polytechnique Montréal]. PolyPublie. https://publications.polymtl.ca/10224/

Document en libre accès dans PolyPublie
Open Access document in PolyPublie

URL de PolyPublie: PolyPublie URL:	https://publications.polymtl.ca/10224/
Directeurs de recherche: Advisors:	Nikola Stikov
Programme: Program:	Génie biomédical

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Bringing quantitative magnetic resonance imaging under one umbrella

AGAH KARAKUZU

Institut de génie biomédical

Thèse présentée en vue de l'obtention du diplôme de *Philosophiæ Doctor*

Génie biomédical

Mars 2022

POLYTECHNIQUE MONTRÉAL

affiliée à l'Université de Montréal

Cette thèse intitulée :

Bringing quantitative magnetic resonance imaging under one umbrella

présentée par **Agah KARAKUZU**
en vue de l'obtention du diplôme de *Philosophiæ Doctor*
a été dûment acceptée par le jury d'examen constitué de :

Julien COHEN-ADAD, président
Nikola STIKOV, membre et directeur de recherche
Pierre BELLEC, membre
Peder LARSON, membre externe

DEDICATION

To my beloved family and cats...

ACKNOWLEDGEMENTS

closest collaborator in the years to come.

I owe Drs. Julien Cohen-Adad and Pierre Bellec a debt of gratitude for all the support and guidance they provided throughout my PhD. I have never seen anyone who can master time as efficiently as Julien does, from minutes in real-life as an advisor, to milliseconds in MRI as an engineer. I am confident that Dr. Cohen-Adad will remain as my role model in these regards for foreseeable future. Behind all the cool software projects I found out about is Dr. Bellec. Pierre's energy and motivation for engaging us BrainHack members to learn reproducible software practices has always been second to none. As the co-supervisor of my research projects, Dr. Bellec has made substantial contributions to shaping up my research. I would like to thank Pierre for all his support and patience in making this thesis a reality.

It was undoubtedly a big excitement of a new beginning, but it weighs heavy on my heart whenever I think about the first time I was leaving the most important people in my life thousands of miles behind. Canim annem Çigdem, I am most grateful to you for supporting my decision, despite how inexplicably difficult it was for you. I know I took "yuksek yuksek tepelere ev kurmasinlar" to a whole new level and the lost time is one thing we cannot make up for. I am deeply sorry for whenever you needed me, but I wasn't there for you. From the bottom of my soul, I am sorry for the times I got carried away with work and sounded evasive when you sent me virtual Turkish coffee invitations. In my heart have the biggest spot, you and all the cute things you do to make me smile. Regardless of how it came through a whatsapp message or a video call, you've always lifted my mood in the most difficult times. I am proud to be your son. Longing was a price we both paid in return of what I was trying to achieve here, but I hope that this thesis makes it worth it to some extent. I used green jellybeans to explain to you why we had to live an ocean apart for years in Chapter 2. You are the biggest, yet the hidden sponsor of this thesis. You did not make it to the sponsors list in my articles just because your help is priceless. If it was not for your self-devotion to our family, I could not graduate from high-school, let alone finishing a PhD abroad. I hope that things will go back to normal one day, and we will visit Canada together. You mean the world to me.

Canim gardasim Eda, you would argue that jellybeans were your thing. That's why I wrote the explanation in Chapter 2 around a lawsuit (don't worry, I really worked on MRI stuff in Canada, did not do illegal jellybean trafficking). I am so proud of you becoming a lawyer before I could finish my PhD! But I am mostly proud because how uniquely we are bounded as siblings. From the simplest kaybis nights to fancy night outs, you've always made everything more enjoyable beyond my imagination. You would only fail at being a sister in English when the word is used synonymously with "caregiver", as you almost made me disrupt my sutures

laughing at your "sandevicini yiyim sen de seyret" joke. Kidding aside, I am the luckiest brother I know and I am truly happy that you inherited the smart genes preventing you from grad studies. I am sorry for all the summer nights we could not ride kaybis and eat topkek together, and all the sunny winter days we could not go skiing. You are my favourite person, and I hope that this upcoming chapter of your life will bring you all the joy and happiness the world has to offer. Publishing this thesis will make me a "Dr." for some people to take me more seriously (I guess), you will get married and change your societal status, but I am certain that our gardaslique will remain the same forever. Cunku o benim gardasim.

Canim babam Melih, I could not find a metaphor from the English Literature to explain what quantitative MRI is, but I think you have your contribution to every analogy I come up with, given that I grew up around your unique sense of humor. If it was not for you and for the amazing environment you provided for our family in Erzurum, I may have never gotten interested in programming. Believe it or not, it is thanks to a Visual Basic book I found on your desk and to your interest towards computer technologies that I started learning how to code. I remember the days you were taking me to your office, showing me what an academic environment looked like. I've always found the idea of having one's own room at the faculty so cool, and always wondered what it felt like. I hope with this thesis, I am now one step closer to it! I am sorry for all the time we could not spend together eating mohkem kebab and walking around the campus to have some father-son moments. You've always supported my creative development, be it playing guitar or skiing, and always been on my side to fight bureaucratic nonsense that I could never comprehend. I am endlessly grateful for all the support you provided for me, from my first day in Eras dormitory in Kayseri to my first night in Hotel Viger in Montreal. We have a precious family bond that makes time and space irrelevant, and it is all because HOW LEL WE ARE! Thank you, my dear family, for everything.

Speaking of family, I would like to give special mention to my closest friends in Montreal, Tommy Boshkovski and Atef Badji. Tommy, the living representation of the song "malo viski malo coca cola", dismissing the fact that he had more coke than the whole NeuroPoly combined between 2017-2019. I wish that the song followed as is for him, but it was rather "mnogo analize na moga stola", like it is for any other PhD student. I will never forget the day we shared that pizza in Cotes des Neiges, or our Parc Extension adventures. You spent years (basically :p) weighing the connectomes with R1, but it took no effort for us to connect. I hope that our Balkan-weigthed and Montreal-reinforced connection remains as strong as it is today, against the inevitable plasticity of life. Thank you for making the other end of the world more familiar and for being a great friend.

Atef, your absence had the most noticeable impact on NeuroPoly's atmosphere. Since you've left for Sweden, the lab started feeling like a typical engineering lab or how Montreal looks in winters, monochromatic. The joy and colors you brought to our lab were unmatchable, especially around Christmas. Thanks to you, we had the privilege of working around someone with endless positive energy for a while. I hope that sparkle of yours never fades away, but only grows up with your son! Thank you so much for being there for me in my difficult times, especially when we were regulars at Mamie Clafoutis (nope, I still cannot pronounce it correctly). You and Tommy were 80% of my social life, even before Covid was a thing. With the same nostalgic vibes emanated from the song by Evanthia Reboutskia: we will meet again.

I cannot deny the role of moral support I received from my four-legged feline friends in pushing this thesis (off the table) to the finish line. I am sure that Theo would have enjoyed chewing on the hardcover binding of this thesis and Maple would show his criticism by scratching around invisible litter to fake cover it. Luna would only care to take a look at it if I covered this thesis in catnip, and Leo would follow her at the expense of getting slapped in the face. There is nothing that annoys Benek as much as my presence, so she would enjoy this dissertation as a token of my absence. Last but not the least, Kardelen would not believe her amber eyes seeing me graduating from PhD, as our friendship dated back to my high school days. Unfortunately we lost her during my second year in Canada, but her heartwarming memories are here to stay. Thank you all my furry friends for being the cutest epitome of disregarding things we think are important by playfully pushing them off the table. I still have a lot to learn from you.

The years I spent in Montréal during my PhD gave me a lot of amazing friends and unforgettable memories. Dr. Mathieu Boudreau, one of my closest collaborators who has conributed valuable time and effort to improve my project. Even though Mathieu worked remotely, I was lucky to have his support by my side whenever I needed. Thank you Mathieu! Drs. Elizabeth DuPre and Gregory Kiar, my dear friends from McGill who helped me explore not only the realm of reproducible neuroimaging research, but also a lot of cool and bikable places in Montréal. My special thanks go to my close friend Zeynep Baytaz, who always had my back in tougher times and shared happy memories whenever we found them. It is comforting to know that Maple is receiving the best attention he could possibly get from any hooman being by living together with you, thank you for everything! My life in Montréal would be unimaginably more dull and boring if it was not for Avsar Eymirlioglu (a.k.a., Altan). If I ever wished for an older brother, that would be him. I am sure that we will keep ranting about things for years to come and make retirement plans to live in Bodrum. Speaking of retirement, I hope I can be half as fit and cheerful as Paule Samson in my late 50s. Thanks

to Paule, my scan hours at the MHI was always fun and educative. I can't deny the big role of her and Marie-Elaine Clavet's help in obtaining an IRB approval from MHI. I thought that I've met people who know their way around coding until I met the amazing Christian Perone. I learnt something new from Chris whenever I stopped by at his desk, and I believe that he was the coolest software developer our lab had to chance to work with! After Chris left NeuroPoly, I was lucky to meet Jonathan Armoza to meet my newfound requirement of being acquaintance with a world-class developer. Jonathan's fascinating digital humanities projects and hearing him making quotations from movies and literature during our coffee walks are among the reasons I am excited to spend a few more years in Mtl. Before finishing this section, I would like to thank all the NeuroPoly members with whom our paths crossed: Tanguy Duval, Gabriel Mangeat, Harris Nami, Ryan Topfer, Charley Gros, Dominique Eden, Sara Dupont, Aldo Zaimi, Ariane Saliani, Darya Morozov, Pascale Beliveau, Simon Levy, Nibardo Lopez Dios, Alexandru Foias, Matteo Mancini and to the newest member of our team, Juan-Jose Velazquez-Reyes. Last but not the least, I would like to thank Elif Ciftcioglu for making the last chapter of my PhD more special than it was ever before, despite that we were an ocean apart and I was grumpy all the time trying to muddle through all the projects I took on. Your heart, kindness and loyalty is a rare combination that one would witness once in a lifetime. Come rain or shine, I wish with all my heart that our paths will remain connected.

RÉSUMÉ

Les systèmes conventionnels d'imagerie par résonance magnétique (IRM) ne sont pas des appareils de mesure. L'IRM clinique repose sur l'évaluation visuelle des images représentées sur des niveaux de gris arbitraires. Redéfinir les scanners de RM en tant qu'appareils de mesure est un objectif ambitieux qui a motivé un certain nombre de méthodes d'IRM quantitative (IRMq). Ces méthodes visent à compléter les images obtenues par RM avec des unités significatives qui sont informatives de la microstructure fondamentale de tissu. Cependant, la mise à niveau des mesures dans les scanners RM commerciaux nécessite des efforts coordonnés de développement et de normalisation. Un obstacle majeur est la variabilité de l'IRMq en raison des implémentations de la boîte noire des séquences d'impulsions d'IRM. Un autre défi provient du manque de flux de travail de traitement IRMq interopérables et reproductibles.

Le grand objectif de cette thèse de doctorat était de résoudre les deux problèmes décrits ci-dessus en réunissant l'IRMq sous un même parapluie. Cela a été réalisé grâce au développement d'un flux de travail neutre pour le fournisseur qui peut normaliser les séquences d'impulsions, les données et les méthodes de traitement pour rationaliser de manière transparente les protocoles d'IRM quantitative. Bien que la thèse ait porté sur l'imagerie de la myéline, la solution de bout en bout offerte par ce travail peut profiter à la plupart des applications d'IRMq.

La première partie de la thèse consiste à développer un progiciel open source pour traiter, simuler et analyser les données IRMq: qMRLab (https://qmrlab.org). La base de code de qMRLab a été développée sur la base de principes de conception collaborative et modulaire pour favoriser une intégration facile de toute méthode IRMq. En plus de la documentation complète des utilisateurs et des développeurs, un exemple de jeu de données a été fourni pour chaque méthode. Équipé d'un pipeline complet d'intégration et de déploiement continus, chaque version du logiciel a été assurée de fonctionner de manière cohérente sur des environnements logiciels portables pour des tutoriels et des publications interactifs. À ce jour, qMRLab comprend 24 implémentations ajoutées par plus de 10 contributeurs et est utilisé par des dizaines de laboratoires de recherche à travers le monde.

La deuxième contribution de cette thèse a été de diriger l'élaboration d'une norme mondiale de données d'IRMq. Une collaboration communautaire entre les concepteurs de la méthode de RM et les chercheurs en neuroimagerie a été établie pour étendre la structure de données d'imagerie cérébrale (BIDS) pour les données et les métadonnées de l'IRMq : qMRI-BIDS. Des réunions mensuelles ont eu lieu avec la participation de plus de 30 chercheurs sur une

période de quatre ans. La version finale de la proposition d'extension a normalisé 18 cartes quantitatives ainsi que leurs unités, a ajouté de nouveaux champs de métadonnées pour améliorer l'enregistrement de la provenance, a défini un nouveau principe commun de collectes de fichiers pour organiser les données d'entrée et les métadonnées pour 15 méthodes d'IRMq, et a établi un ensemble de règles pour étendre la spécification pour les futurs cas d'utilisation de l'IRMq. Enfin, un exemple de jeu de données a été créé et l'extension a été fusionnée dans la version principale de BIDS v1.5.0 (https://bids-specification.readthedocs.io/).

Les contributions ci-dessus ont permis le développement de qMRFlow, une collection de pipelines portables et compatibles qMRI-BIDS qui peuvent rationaliser qMRLab avec d'autres logiciels de pré- et post-traitement via des conteneurs. En tant que dernière pièce du puzzle de flux de travail de bout en bout, une séquence neutre pour le fournisseur (VENUS) a été développée sur une plateforme d'IRM en temps réel approuvée par la FDA et rendue publique (https://github.com/qmrlab/venus) pour l'acquisition de cartes sensibles à la myéline de T1, du rapport de transfert d'aimantation et de la saturation (MTR et MTsat). Pour communiquer avec qMRFlow, VENUS a été conçu pour reconstruire et exporter des données au format qMRI-BIDS. Pour activer un contrôle de balayage qMRI convivial, nous avons équipé VENUS d'une interface utilisateur unifiée. Le pipeline de bout en bout développé (qMRFlow + VENUS) a été déployé sur un scanner GE et deux scanners Siemens (3T) pour tester l'hypothèse selon laquelle les séquences neutres pour les fournisseurs réduisent la variabilité inter-fournisseur des cartes T1, MTR et MTsat. Trois participants en bonne santé et un fantôme T1 de référence ont été scannés à l'aide de VENUS et de ses homologues natifs du fournisseur sur les trois scanners. Les résultats in vivo ont montré que la variabilité entre les fournisseurs était considérablement réduite pour toutes les cartes (p = 0,015). Dans le fantôme, non seulement les différences entre les fournisseurs ont été réduites (8 - 19,4% à 0,2 - écart de 5%)), mais aussi une amélioration globale a été obtenue dans la précision T1 (7 - 11% à 0,2 - 4% d'erreur).

Cette thèse démontre une intégration puissante entre les séquences neutres pour les fournisseurs, les normes de données communautaires et les logiciels open source pour améliorer la fiabilité de l'IRM quantitative. Les flux de travail développés réduisent la variabilité à plusieurs niveaux, en commençant par la mise en œuvre de la séquence d'impulsions d'IRM et en continuant avec la reconstruction, le pré- et le post-traitement. En outre, ces flux de travail sont construits sur une plateforme approuvée par la FDA qui peut héberger des méthodes d'IRMq avec une voie rapide vers l'application clinique. L'IRM quantitative doit contourner les boîtes noires du fournisseur pour fournir ce qu'elle a à offrir, et ce travail montre la voie à suivre.

ABSTRACT

Conventional magnetic resonance imaging (MRI) systems are not measurement devices. Clinical MRI relies on visual assessment of images represented on an arbitrary grayscale. Redefining MR scanners as measurement devices is an ambitious goal that has motivated a number of quantitative MRI (qMRI) methods. These methods aim at supplementing MR images with meaningful units that are informative of the underlying tissue microstructure. However, retrofitting measurement into commercial MR scanners requires coordinated development and standardization efforts. A major stumbling-block is the variability of qMRI due to the black-box vendor implementations of MRI pulse sequences. Another challenge originates from the lack of interoperable and reproducible qMRI processing workflows.

The grand objective of this PhD work was to address both problems described above by bringing qMRI under one umbrella. This was achieved through the development of a vendor-neutral workflow that can standardize pulse sequences, data and processing methods to seamlessly streamline quantitative MRI protocols. While the focus of the thesis was on myelin imaging, the end-to-end solution offered by this work can benefit most qMRI applications.

The first part of the thesis consists of developing an open-source software package to process, simulate and analyze qMRI data: qMRLab (https://qmrlab.org). The codebase of qMRLab was developed based on collaborative and modular design principles to foster easy integration of any qMRI method. In addition to extensive user and developer documentation, an example dataset was provided for each method. Equipped with a thorough continuous integration and deployment pipeline, each release of the software was ensured to run consistently on portable software environments for interactive tutorials and publications. As of today, qMRLab includes 24 implementations added by more than 10 contributors and is being used by tens of research labs around the globe.

The second contribution of this thesis was leading the development of a global qMRI data standard. A joint-community collaboration between the MR method developers and neuroimaging researchers was established to extend the brain imaging data structure (BIDS) for qMRI data and metadata: qMRI-BIDS. Monthly meetings were held with the participation of more than 30 researchers over the course of four years. The final version of the extension proposal standardized 18 quantitative maps along with their units, added new metadata fields to enhance provenance recording, defined a new common principle of file collections to organize input data and metadata for 15 qMRI methods, and established a set of rules to extend the specification for future qMRI use cases. Finally, an exam-

ple dataset was created and the extension was merged into the main BIDS release v1.5.0 (https://bids-specification.readthedocs.io).

The above contributions enabled the development of qMRFlow, a collection of portable and qMRI-BIDS compatible pipelines that can streamline qMRLab with other pre- and post processing software via containers. As the final piece of the end-to-end workflow puzzle, a vendor-neutral sequence (VENUS) was developed on an FDA-approved real-time MRI platform and made publicly available at https://github.com/qmrlab/venus for acquiring myelin sensitive maps of T1, magnetization transfer ratio and saturation (MTR and MTsat). To communicate with qMRFlow, VENUS was designed to reconstruct and export data in the qMRI-BIDS format. To enable a user-friendly qMRI scan control, we equipped VENUS with a unified user interface. The developed end-to-end pipeline (qMRFlow + VENUS) was deployed on one GE and two Siemens scanners (3T) to test the hypothesis that vendor-neutral sequences reduce inter-vendor variability of T1, MTR and MTsat maps. Three healthy participants and a reference T1 phantom were scanned using VENUS and its vendor-native counterparts on all three scanners. In vivo results showed that the variability between vendors was significantly reduced for all maps (p=0.015). In the phantom, not only were inter-vendor differences reduced (8 - 19.4% to 0.2 - 5% deviation), but also an overall improvement was achieved in T1 accuracy (7 - 11% to 0.2 - 4% error).

This thesis demonstrates a powerful integration between vendor-neutral sequences, community data standards and open-source software to improve the reliability of quantitative MRI. The developed workflows reduce the variability at multiple levels, starting with the MRI pulse sequence implementations and continuing with reconstruction, pre- and post-processing. In addition, these workflows are built on an FDA-approved platform that can host qMRI methods with a fast track to clinical translation. Quantitative MRI needs to bypass the vendor black boxes to deliver what it has to offer, and this work shows the way forward.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF SYMBOLS AND ACRONYMS

B_0	Main magnetic field
M	Net magnetization
M_0	Equilibrium magnetization
γ	Gyromagnetic ratio
θ	Excitation flip angle
μ	Magnetic moment
T	Tesla
ADC	Analog to digital converter
AVF	Axonal volume fraction
$B1_+$	Transmit radiofrequency field
$B1_-$	Receive radiofrequency field
bSSFP	Balanced Steady-state Free Precession
CSF	Cerebral spinal fluid
CSF	Cerebral spinal fluid
DWI	Diffusion-weighted imaging
DTI	Diffusion tensor imaging
EEG	Electroencephalography
FA	Flip Angle
FID	Free Induction Decay
FOV	Field of view
GE	General Electric
GRE	Gradient Echo
ICC	Intraclass correlation coefficient
ihMT	Inhomogeneous magnetization transfer
IR	Inversion recovery
MET2	Multiexponential T2
MP2RAGE	Magnetization Prepared 2 Rapid Acquisition Gradient Echoes
MRI	Magnetic resonance imaging
MS	Multiple sclerosis
MT	Magnetizaton transfer
MTV	Macromolecular tissue volume
MTR	Magnetization transfer ratio
MTsat	Magnetization transfer saturation

MVF	Myelin volume fraction
MWF	Myelin water fraction
NODDI	Neurite orientation dispersion and density imaging
NMR	Nuclear magnetic resonance
PD	Proton Density
qMRI	Quantitative magnetic resonance imaging
qMT	Quantitative magnetization transfer
QSM	Quantitative susceptibility mapping
R1	Longitudinal relaxation rate
RF	Radiofrequency
ROI	Region of interest
SE	Spin Echo
SIRFSE	Selective Inversion Recovery Fast Spin Echo
SPGR	Spoiled Gradient Echo
SWI	Susceptibility-weighted imaging
T1	Longitudinal relaxation time
T2	Transverse relaxation time
TE	Echo time
TR	Repetition time
VFA	Variable flip angle

LIST OF APPENDICES

CHAPTER 1   INTRODUCTION

With its unparalleled soft-tissue contrast, versatile imaging modalities, and absence of ionizing radiation, magnetic resonance imaging (MRI) is a cut above other non-invasive imaging methods in research and clinics. Although transforming scanners from mere imaging apparatus to tissue characterization devices is the next overarching goal across MRI disciplines, neuroimaging applications have been at the forefront with the developments from relaxometry, magnetization transfer (MT), diffusion and quantitative susceptibility imaging applications. Going beyond conventional diagnostic contrasts, these quantitative MRI (qMRI) methods have in common the ability to provide objective information about tissue microstructure in the living brain. Over the last four decades, quantitative brain imaging has become a world of its own, such that there are more than 30 MRI-based digital biomarkers developed in order to estimate myelin content (Lazari and Lipp, 2021; Mancini et al., 2020). Just to list a few examples, these quantitative myelin biomarkers have driven several clinical trials (Ashton, 2010), applied neuroimaging studies (Boshkovski et al., 2021) and, multicenter studies to determine the range of parameters in health (Bojorquez et al., 2017; Lee et al., 2019) and disease (Goebel et al., 2016; Young, 2004).

As the number of qMRI biomarkers grew remarkably (Mancini et al., 2020), the need for validation and correction strategies have also increased. These strategies include phantom manufacturing, development of imaging methods to characterize instrumental imperfections and multi-center and longitudinal studies. One of the major leaps forward in the qMRI phantom manufacturing was initiated by a collaboration between the International Society of Magnetic Resonance in Medicine (ISMRM) ad-hoc Committee on Standards for Quantitative MRI and the National Standards and Technology (NIST). Being the first traceable qMRI calibration reference artifact, the ISMRM/NIST system phantom (Stupic et al., 2021) has become a driving force to conduct large-scale performance assessment studies for trending quantification methods, such as T1 mapping (Bane et al., 2018; Keenan, Gimbutas, Dienstfrey, Stupic, et al., 2021). On the one hand, this phantom highly benefits method developers to fine-tune acquisition sequences and protocols across sites. On the other hand, it has been shown that the relaxometry measurements are prone to lower error rates in phantoms, as their construction inevitably downplays the in-vivo intricacy of biological structures (Wang et al., 2006). Therefore, phantom validation studies are often considered complementary to in-vivo quality assurance of qMRI techniques.

When a human subject is placed inside the scanner, the biochemical complexity of the tis-

sue coupled with abrupt magnetic susceptibility transitions between air-tissue compartments lead to considerable local variations in the static magnetic field and radiofrequency (RF) transmission profile. As a result, the accuracy of the parameter estimation becomes spatially dependent (Stikov, Boudreau, et al., 2015). Although novel phantom manufacturing techniques aim at reducing this gap between phantoms and in-vivo structure by using more realistic biochemical compounds (Jona et al., 2021) or anatomy mimicking 3D printed scaffolds (Gopalan et al., 2021), field mapping methods remain an indispensable part of in-vivo performance assessment studies. RF field inhomogeneity maps are particularly important for the correction of commonly used spoiled gradient-echo (SPGR) based T1 mapping methods (Stikov, Boudreau, et al., 2015). It has been shown that the impact of field mapping method selection is not trivial on the parameter quantification (Boudreau et al., 2018). Nonetheless, most of these sequences are not commercially available, posing a challenge for unifying spatial correction methods in multi-center qMRI studies (Lee et al., 2019).

Given that establishing the ground-truth values of tissue relaxation time constants is still not perfectly known in the living brain (Bojorquez et al., 2017; Wright et al., 2008) mostly due to the technical challenges summarized above, the reliability of qMRI relaxometry measurements across time points has found widespread acceptance as a performance indicator. Frequently assessed by intraclass correlation coefficient (ICC), the within-site test-retest performance of relaxometry and magnetization transfer qMRI metrics showed substantial reliability according to the benchmarks defined in (Zou, 2012). For example, it has been reported that ICC of T1 measurements ranges from 0.88 to 0.99 at 3T in the human brain, as calculated in white matter (WM) and gray matter (GM) of 20 healthy subjects using a single scanner (Okubo et al., 2016). For the same ROIs, an ICC of 0.92 has been reported at 3T for a wider age interval using a highly under-sampled acquisition method, suggesting that reliable T1 quantification can be achieved within clinically feasible scan durations (Grafe et al., 2021).

Despite such high intra-vendor reliability scores, the literature reveals that vendor-specific implementations cause notable multi-site variations in most qMRI applications (Gracien et al., 2020; Lee et al., 2019; Schmeel et al., 2019; Teixeira et al., 2020). Since vendors often conceal pulse sequence implementation details to protect their commercial interests, several unknowns are introduced in multi-center qMRI studies. For example, parameter settings for RF spoiling schemes of SPGR sequences are typically not available to the users. However, when used for T1 mapping, the configuration of these parameters has a substantial impact on the resultant quantitative maps (Yarnykh, 2007b). Let alone hidden parameters and implementation differences, it may not even be possible to equalize user-accessible sequence settings across vendors, such as the configuration of parallel imaging. This methodological variability is exacerbated by vendor variant availability of field mapping sequences. Although

major MRI vendors provide developer access for researchers to implement their own pulse sequences, identical across vendor implementations are highly demanding, if not impossible, to achieve in practice. To dispense with the need for using vendor-exclusive software libraries, several open-source pulse sequence development platforms have been proposed (Cordes et al., 2020; Jochimsen and Von Mengershausen, 2004; Layton et al., 2017; Magland et al., 2016; Nielsen and Noll, 2018; Ravi, Geethanath, et al., 2019; Stöcker et al., 2010). These platforms have in common the ability to interpret and translate the same sequence logic for multiple vendors, considerably reducing multi-center development efforts and minimizing implementation variability. Earliest qMRI studies performed using these vendor-neutral pulse sequence development platforms have reported an overall improvement in multicenter reproducibility within the confines of a single vendor (Herz et al., 2021; Tong et al., 2021). However, the efficiency of such vendor-neutral solutions in reducing inter-vendor variability remains to be evaluated.

Reducing sequence implementation variabilities is a critical, yet not the sole factor to improve the reproducibility of qMRI. Downstream processes included in a typical qMRI workflow, such as image reconstruction, pre- and post-processing are among other potential sources of variability. Hence, to clear the path to reproducible multi-center qMRI studies, there is a need for a thorough solution capable of bringing transparency and ease of use to the whole process. To that end, this thesis develops and integrates vendor-neutral acquisitions with open-source processing software through a global data standard, establishing transparent qMRI workflows that can extend from scanner to publication.

The literature review provided in Chapter 2 starts by expanding on the fundamental aspects of the problems outlined above, then proceeds to introduce a variety of myelin imaging methods. Chapter 3 establishes the link between the reviewed literature and three main objectives of this PhD project. These objectives are achieved by studies presented in Chapters 4, 5 and 6, bringing qMRI under one umbrella through open-source software development, data standardization and vendor-neutral acquisitions, respectively.

CHAPTER 2   LITERATURE REVIEW

2.1   Basics of MRI and quantitative MRI

This section starts by explaining the distinction between MRI and quantitative MRI (qMRI), which is of essence to the central premise of this thesis. Next, it aims at delivering an intuitive understanding of how MRI works by using cartoons, simulations and example applications, all introduced in the context of overarching concepts from physics and everyday life. For a more theoretical introductory explanation, the reader is referred to (Nishimura, 1996). After covering the basics of MRI, the relationship between data acquisition and parameter estimation will be explained based on two basic qMRI applications: T1 and T2 mapping. Finally, we will look at three aspects of qMRI that need to be improved for clinical translation.

2.1.1   The million dollar question: Pixels have values, then why is MRI not quantitative?

We will start answering this question by looking at the most prominent use cases of MRI from two distinctive fields: diagnostic radiology and food engineering. In the clinics, MRI stands out as one of the most preferred imaging methods, because it can generate detailed images with superb soft tissue contrast, without using ionizing radiation or cutting open the human body. Surprisingly, MRI scanners have also been extensively used in food science to study soft tissue. For example, several studies used MRI to observe how moisture migrates towards the center of jellybeans over time (Troutman et al., 2001; Ziegler et al., 2003).

Be it in diagnostic radiology, or in food science, it is the superior soft tissue contrast that makes MRI appealing. In routine diagnostic readings, the radiologists browse through MR images to capture abnormalities that may be resolved by conventional MRI contrasts, i.e. T1- or T2-weighted images. As a result, the detection of pathological patterns depends on a radiologists' visual assessment, which is then transferred to a written report – a narration of observations such as "*T2 hyperintense appearance in the left parieto-occipital lobe suggests hemorrhagic infarction*" (Figure 2.1f). Here, the word hyperintense implies a relative comparison. Figure 2.1e illustrates that cropping the tumorous region away from the image removes the basis of comparison and makes the hyperintense appearance irrelevant. This is because the pixel brightness of conventional MR images is assigned using an arbitrary scale consisting of shades of gray. Due to the lack of a calibrated measurement scale, conventional MRI is considered to be qualitative.
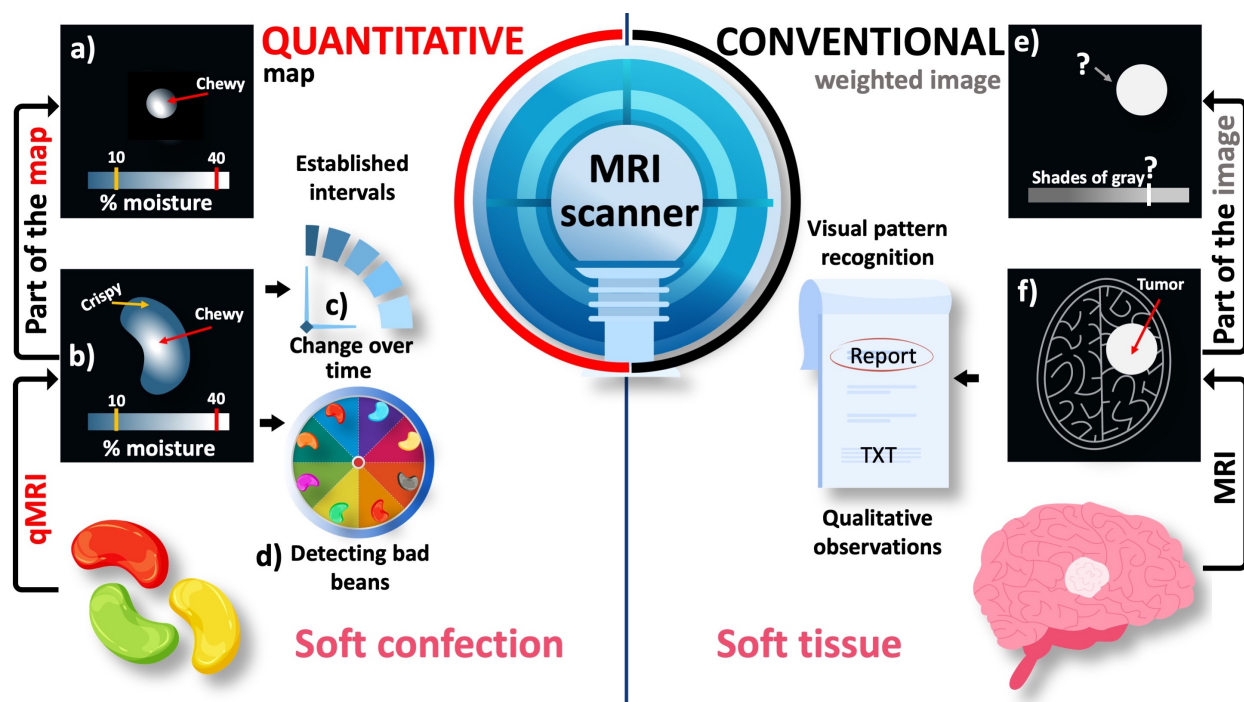
Figure 2.1 An illustrative comparison between the conventional and quantitative MRI (qMRI). The pixel brightness of conventional MR images is defined in an arbitrary grayscale **(e,f)**. As a result, only a qualitative pattern recognition is possible when the suspected region (i.e., the tumor) is assessed against the background of the target anayomy (i.e., the brain). On the other hand, quantitative maps spatially resolve a meaningful metric **(a,b)** to detect changes over time **(c)** and between different samples **(d)**. On a standardized measurement scale of percent moisture, the texture characteristics of a jellybean (e.g., crispy or chewy) can be objectively determined even from a randomly selected part of the image.

Using the same MRI scanner, it is possible to assign meaningful numbers to the images and this approach turns out to be the most common MRI method in food engineering (Mariette et al., 2012; Ziegler et al., 2003). Figure 2.1 illustrates the added value of quantitative MRI (qMRI) when applied to a sample familiar to everyone: a jellybean. The moisturization map indicates that the jellybean has formed a crispy shell while remaining chewy at the center, which is the desired texture (2.1b). Given that the level of chewiness is determined by a threshold on a standardized measurement scale, a randomly selected part of the image can be still characterized by comparing selected pixel values against the established threshold (2.1a). This feature of qMRI offers an objective insight into how the texture of this soft confection changes over time, which would help determine its best before date (2.1c, prognosis).

The ability to reveal what underpins the appearance of visually similar samples is yet another powerful feature of qMRI. In a Bean-Boozled challenge, which is a Russian roulette of jellybean flavors, tasty flavors are mixed with nauseous look-alikes (Gambon, 2015). For

example, a green jellybean may taste like lime (tasty) or lawn clippings (nauseous) in the Bean-Boozled game (Figure 2.2). Therefore, no matter how experienced the player is, the chances of picking up a lime-flavored bean is as good as tossing a coin. Conventional MR images of a handful of green jellybeans do not offer a distinguishing feature, but only reveal their structure. As a result, the chances of making an unfortunate choice remain the same (Figure 2.2).
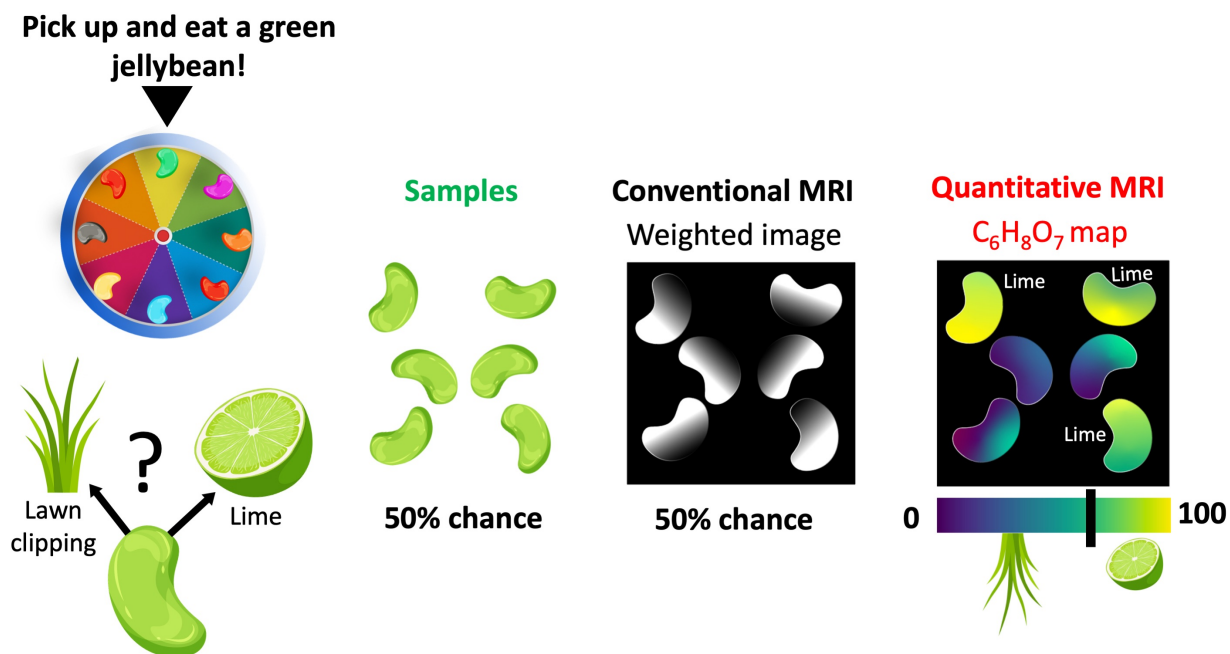
Figure 2.2 The comparison of conventional and quantitative MRI in a Bean-Boozled game, where the task is picking up and eating a green jellybean. Out of 6 green beans, half of them taste like lime (tasty) and the remaining are lawn clipping flavored (nauseous). Conventional MRI is not sensitive to either aroma, as a result the beans show similar contrasts. On the other hand, a quantitative mapping method sensitive to citric acid ($C_6H_8O_7$) can help reveal lime-flavored beans.

On the other hand, spatially resolving a quantitative property that is sensitive to either flavor would step up our game in making the right decision. For example, a qMRI method capable of mapping the distribution of citric acid ($C_6H_8O_7$) – the chemical compound that gives citrus fruits a sour taste – would help distinguish lime-flavored green beans. Even though the grass-flavored beans may contain a slight amount of $C_6H_8O_7$ (used as lawn fertilizer), establishing a threshold can help make an informed decision (Figure 2.2), giving the players a competitive edge in the bean-boozled game (diagnosis).

Note that the pixel brightness of the conventional (weighted) image has contributions from the $C_6H_8O_7$ concentration. However, it is also affected by several other factors, such as water

density and glucose content. Therefore, understanding the relationship between the pixel brightness and the flavour depends on the experience and subjective interpretation of the observer.

Although the distinction between the values and pixel brightness has become clearer in the 50 years since the invention of MRI, their definitions were conflated at the time when nuclear magnetic resonance (NMR) was evolving into MRI. For General Electric (GE), one of the largest MRI manufacturers in the world, the cost of this blurry line between the numbers and pixel brightness was \$129,000,000 (detailed below).

NMR is a spectroscopy method that gives information about the chemical makeup of the analyzed substance. In analogy with the jellybean example, an NMR measurement is similar to quantifying the amount of $C_6H_8O_7$ *from the fragments of* a green jellybean (Figure 2.2). If we used NMR for the bean-boozled challenge, we would be looking at a list of values to pick up a lime-flavoured bean instead of looking at a map. Beginning in the late 1930s, researchers have been using NMR to characterize chemical compounds. The history of MRI begins with the idea of using NMR to tell a cancerous tissue sample from a healthy one.

The evolution of NMR into MRI is a turbulent story (Dreizen, 2004) that starts with the development of the first MRI scanner (1980) and leads to a Nobel Prize (2003). In 1971, Damadian published a study on the use of NMR-based T1 and T2 values for detecting malignant tumors (Damadian, 1971). Based on this work, he issued a patent application titled "an apparatus and method for detecting cancer in tissue" in 1972, which was accepted in 1974 (Damadian, 1974).

However, an actual image was out of the picture until Lauterbur was finally able to publish his work in 1973, showing a crude weighted-image of two liquid filled tubes (Lauterbur, 1989). Lauterbur's publication was delayed because the initial submission was desk-rejected by the editors of Nature, and his university did not find his work valuable enough to submit a patent (Dawson, 2013). Around the same time but an ocean apart, Mansfield was applying NMR to image crystals by borrowing a concept dubbed k-space from 2D crystal structures (Turner, 2017). This approach led Mansfield to develop a fast image generation method, bringing MRI closer to practical reality (Mansfield and Maudsley, 1977).

With undeniable insight from the studies of Lauterbur and Mansfield, Damadian's team built the first human MRI scanner in 1978 and made it commercially available in two years. Around the same time, GE started manufacturing scanners without paying royalty to Damadian as consideration for the patent. In the decade that follows, GE sold nearly 600 scanners, for which Damadian's company Fonar filed a patent infringement lawsuit in the late 1990s and awarded \$128,705,766 as a compensation for pecuniary damages (Cir., 1996).

Returning to the jellybean analogy, Damadian's patent was mainly describing a device to scan whole jellybeans for a complete $C_6H_8O_7$ measurement. The key invention of the patent was to collect multiple measurements at different locations of the jellybean without fragmenting them. Later on, Lauterbur and Mansfield developed the methodology to create weighted images of the jellybeans. This was followed by Damadian marketing the first device that can generate weighted images of whole jellybeans (Figure 2.2) and the U.S. Supreme Court reached the verdict that GE infringed Damadian's patent. The original judgement on the verdict reads:

"*On May 27, 1997 the Honorable Wm. H. Rehnquist, Chief Justice, the United States Supreme Court, enforced the Order of the Federal Circuit Court of Appeals and ordered GE to pay Fonar. GE paid Fonar $128,705,766 for patent infringement. GE was further restrained from any use of Fonar technology.*"

"*The Court found that GE had infringed U.S. Patent 3,789,832, MRI's first patent, which was filed with the U.S. Patent Office in 1972 by Dr. Damadian. The Court concluded that MRI machines rely on the tissue NMR relaxations that were claimed in the patent as a method for detecting cancer, and that MRI machines use these tissue relaxations to control pixel brightness and supply the image contrasts that detect cancer in patients.*"

To paraphrase the reasoning behind this decision using the jellybean analogy:

- GE manufactured and sold 600 scanners capable of generating weighted images of the jellybeans,

- the weighted images are influenced by the $C_6H_8O_7$ concentration,

- lime-flavored jellybeans have higher concentrations of $C_6H_8O_7$,

- thus, GE scanners are designed to identify lime-flavored jellybeans, which infringes on Damadian's patent.

Although the U.S. Supreme Court decision implied that Damadian owns the intellectual property rights for MRI scanners at that time, the 2003 Nobel Prize in Medicine was shared between Lauterbur and Mansfied only. Damadian spent nearly $300,000 for full page ads in popular print media outlets to claim his rights to the 2003 prize, yet the situation has remained unchanged up to this date. The notes on why Damadian was not included in the prize will be available in 2053 (Harris, 2003).

The court's interpretation of the difference between the Fonar's patent and GE MRI scanners perfectly captures the essence of qMRI, which is to enable objective and consistent compar-

isons by tagging each pixel with a precisely defined score that ranks a physical characteristic. Although the physical property estimated by qMRI contributes to the pixel brightness of conventional images, the conventional images are presented in an arbitrary grayscale. This is the reason why individual pixels of a weighted image have numbers, but no (physical) value.

To conclude, there is a critical difference between detecting abnormalities based on pixel brightness (conventional MRI) and tissue characterization using quantitative metrics (qMRI), and the court's interpretation of this difference cost GE $129 million. With this central distinction in mind, the following sections will introduce how we can use the same MRI scanner for both qualitative and quantitative imaging.

2.1.2 A pictorial and historic journey into how MRI works, from quantum to macro-scales

The human body can be seen as a complex compartmentalization of water, fat, protein and minerals at every level of its organization from atoms to organs (Siri, 1956). At the atomic scale, nearly 63% of the atoms in the human body are hydrogen atoms (Osmera and Vanicek, 1940), which consists of only one proton and one electron. It is this simplicity that makes hydrogen the most studied atomic structure in quantum mechanics, which eventually lead to the development of MRI.

Once the mass and charge of the hydrogen particles were known, nobody suspected that there was another hydrogen property to be discovered. The study of hydrogen was revolutionized when two scientists in their mid 20s, Uhlenbeck and Goudsmit, wrote together the first article about the *nuclear* spin in 1925 (Uhlenbeck and Goudsmit, 1925). They were afraid to submit their work, because the celebrity physicist Lorentz deemed their idea "unphysical". At that time, Uhlenbeck and Goudsmit were working under the supervision of Zeeman and Ehrenfest, yet Zeeman and Ehrenfest omitted their names from the article, deciding to use the youth of their protegees as a shield against any backlash (Halpern, 2017). What was feared to be a foolish mistake was then accepted as a built-in feature of any fundamental particle in the universe, just like the mass and charge. This quantum property – spin – is the key to understanding how MRI works.

Getting on the same wavelength with a single hydrogen in the quantum realm

28 years before his student would give an underlying explanation, Zeeman showed that the energy levels in an atom change under the influence of a *magnetic* field, an effect known as Zeeman's splitting (Zeeman, 1897). This is particularly important for hydrogen, because

unlike a magnetic field, an electric field does not lead to an energy difference between its spin configurations (i.e., does not split) (Feynman et al., 2015). Hydrogen is assumed to have four (ground) spin states, created by the combinations of two up and two down spins from its proton and electron. Each combination constitutes a different energy level. The difference between these energy levels is so small that their separation is defined by the term "hyperfine splitting". What encourages the hydrogen to exhibit more than one energy level is the presence of a uniform magnetic field (B_0). This effect opens a communication line to interact with hydrogen, yet it takes some special effort to start a conversation. To explain these requirements, we will use an online dating metaphor.

The hydrogen atom comes online only when standing up under the influence of a magnetic field (Figure 2.3a,b). The chances of getting a response from the hydrogen firstly depends on whether our message kindles just the right amount of excitement for it to switch between those hyper-finely separated energy levels (Figure 2.3c). Although it almost seems like the hydrogen is sidestepping a conversation, all it takes is finding the right wavelength to meet this first requirement. Six years after the introduction of the spin concept, Rabi and Breit finally discovered that to resonate with hydrogen's energy levels, we need to send our messages in the radiofrequency (RF) range of the electromagnetic spectrum (Breit and Rabi, 1931).

However, not all hydrogen atoms behave the same. We need to be familiar with the peculiarities of the hydrogen we are in touch with. There are two key attributes: where is the hydrogen from, and in which energy state it is when our resonating message is delivered (Figure 2.3c). The last nuance to get on the same wavelength is finding the right angle to approach it. If we meet all the requirements, we will see the hydrogen getting excited and responding to us within a certain RF bandwidth.

For the imaging of the human body using MRI, we will be mostly communicating with the hydrogen from the water (Figure 2.4). In general, water hydrogens are more easy-going because their electron spin states are balanced, so we are only concerned with the energy levels emerging from their *nuclei*. This is why we call this pick-up line the nuclear magnetic resonance (NMR)[3].

Figure 2.4 shows that the hydrogen from water has two energy levels under a uniform magnetic field: low energy and high energy. If the resonating message flips the hydrogen's energy

---

[2]Schrödinger's cat is a thought experiment, describing a cat closed in a box with a deadly gas vial. At any time, the cat may break the vial and die. We can only find out about the state of the cat by opening the box (i.e., by measuring the system). Therefore, unless we open the box, the cat is both dead and alive at the same time (i.e., the quantum state)

[3]After this method gained popularity in the clinics, the word "nuclear" was omitted to avoid creating a false impression that it uses radioactive energy for imaging.
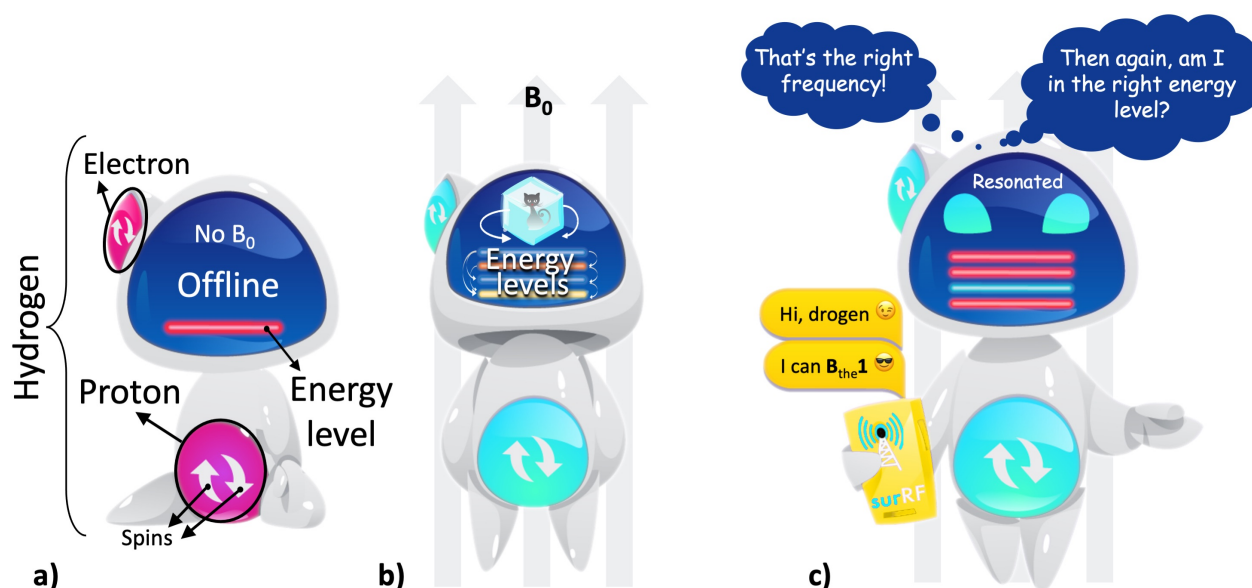
Figure 2.3 **a)** A hydrogen atom has one proton and one electron, with each particle is assumed to have two possible spin states for simplicity (up or down). In the absence of a magnetic field, the atom is at a random orientation (offline) and shows a single energy level (the red line). **b)** Under a uniform magnetic field (B_0), the atom is aligned with the field (online) and shows multiple energy levels (the colored lines). These energy levels are in quantum superposition: all the levels simultaneously exist and we can only determine one state upon observation. This behaviour of the energy levels is represented by the Schrödinger's cat[2]. **c)** Once it is online, the hydrogen atom can be contacted through a communication line that operates within the radiofrequency (RF) range.

state from higher (red) to lower (cyan) state, the result is a "radio silence" (Figure 2.4a), i.e. the hydrogen gives up a small amount of excess energy instead of a response. But if the resonating message elevates the hydrogen's mood from lower (red) to the higher (cyan) state, the hydrogen gets excited (Figure 2.4b). After the message is delivered, we will finally get a response as the excitement quickly fades away. This final process is termed "relaxation", which is of essence to the MRI contrast, because the message carries information about where that hydrogen is from.

The tone in the response shifts slightly if the hydrogen is from non-water molecules, e.g., fat or protein. This slight difference is caused by the amount of negative energy a hydrogen is surrounded by, which interferes with the contribution of hydrogen's electron to its energy state, namely shielding. The more negative energy around the hydrogen, the shorter the response. It appears that the positive effects of being near water on the energy levels transcends scales, from humans themselves (Cracknell et al., 2016) to the hydrogen atoms that make up them (Lawrence and McDonald, 1966).
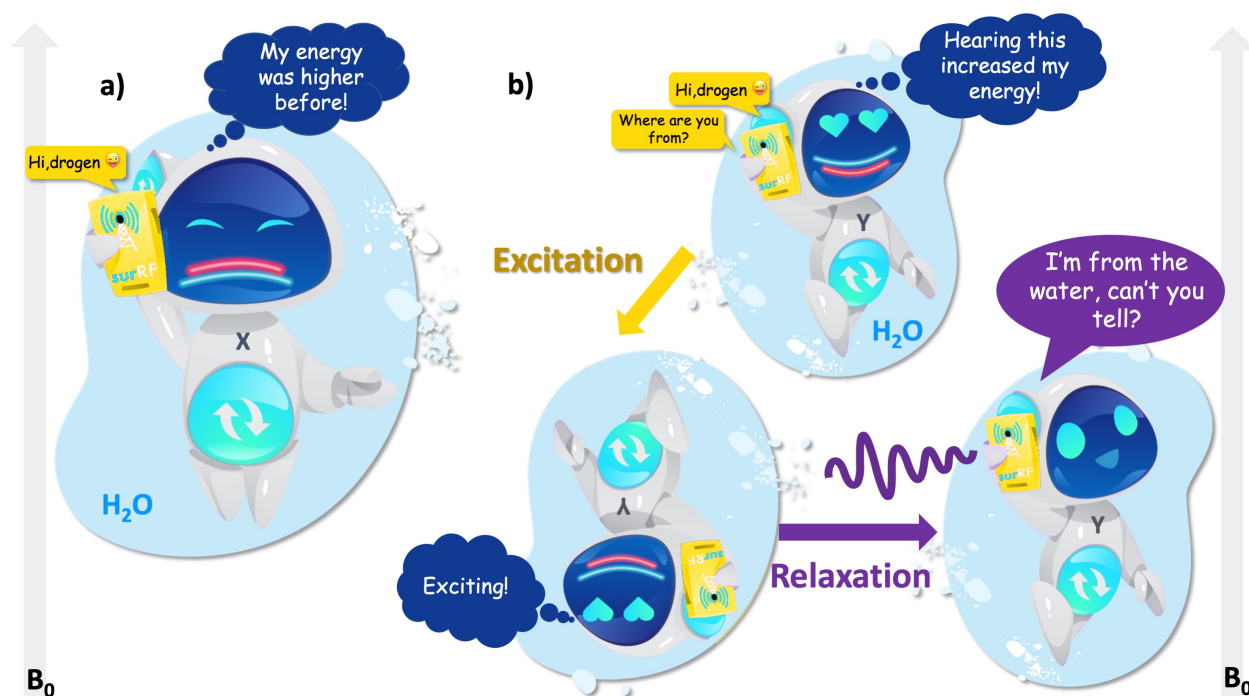
Figure 2.4 **a)** After receiving the call, the hydrogen atom from water switches from the higher (red) to the lower (cyan) energy level and refuses to answer. **b)** When the RF input switches its energy from the lower (red) to the higher level (cyan), the atom gets excited. While returning to its initial state, the atom responds.

So far we looked at the quantum-level interactions between the hydrogen atom and RF energy, and tied it with the NMR phenomenon. However, in reality, we don't have access to observe NMR effects at such a fine-grained level; because no appropriate instrumentation exists, and quantum-jitters make such instrumentation nearly impossible (Erkintalo, 2021). For example, we cannot detect the uncertainty of a single hydrogen atom's energy levels. The best we can do is to visualize the concept using metaphorical illustrations, such as using Schrödinger's cat to imply the quantum state of the hydrogen's energy levels in Figure 2.3b. Until we observe the consequences (whether the hydrogen will respond to our resonant message or not), all the energy levels are assumed to be in quantum superposition, even for only two energy states of proton spins as shown in Figure 2.4. To achieve observational accuracy, we need to move from the uncertainties of the energy levels to a probability of getting a response to our resonant message, which is what we will look at in the following section.

Communicating with millions of hydrogen atoms at the macro level

We can harness the benefits of NMR without having a complete theoretical understanding of the underlying quantum interactions (see proton spin crisis (Siegel, 2017)), because these effects simply smooth over at the macroscale. At this point we will leave the online dating analogy behind, because even at the smallest macroscopic scale, "there are plenty of fish in the sea" (Figure 4a). At the level where an NMR measurement is technically feasible, we will be concerned with a large pool of hydrogen atoms. Here, the individual behaviour of particles becomes useless for characterizing the system in aggregate. This concept of integrating over microscopic details to achieve a compact and useful system description is coarse-graining (Figure 2.5).
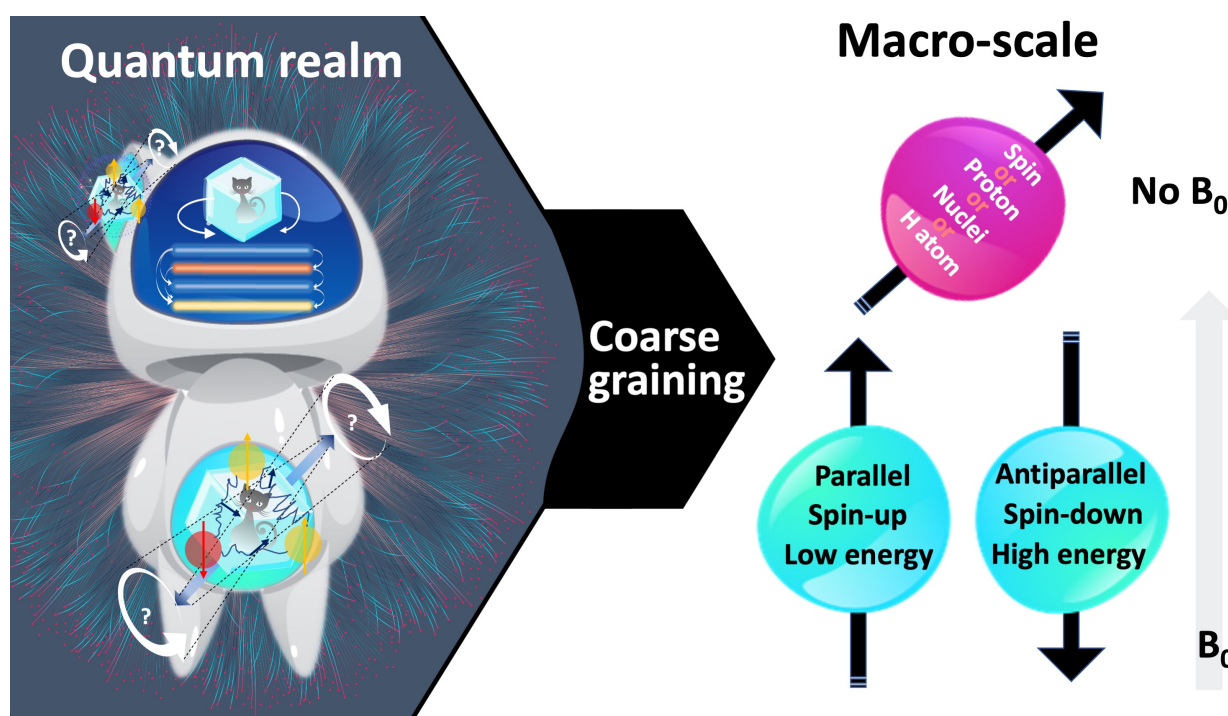
Figure 2.5 At the macro-scale, the quantum mechanical behaviour of individual hydrogen atoms is averaged over, i.e., coarse-grained. Figure 2 illustrates a fine-grained description of the behaviour of a single hydrogen atom in absence and presence of a magnetic field. After coarse graining, the representation simplifies to an arrow passing through the center of the proton, which is aligned parallel (i.e., spin-up, low energy) or antiparallel (i.e., spin-down, high energy) with the magnetic field (blue), or at random if the field is absent (pink). Following coarse-graining, the terms spin, proton, nuclei or hydrogen can be used interchangeably.

Figure 2.5 shows how the representation of a hydrogen atom is changed after coarse-graining microscopic details of spin interactions. From this point onward in this document, hydrogen will be illustrated as shown in Figure 2.5, and used interchangeably with the terms spin,

proton and nuclei. To refer to the energy level associated with a single hydrogen atom, we will describe the magnetic moment (Figure 2.6a): If the proton was merely a rigid body, its rotation about an imaginary axis that passes through its center would create a small angular momentum aligned with that axis. Given that the proton is a charged particle, it also creates a microscopic magnetic moment (μ) as a result of this rotation, which is a vector in the same direction (Figure 2.6a). The ratio between the angular momentum and the magnetic moment yields the gyromagnetic ratio (γ), which is 42.59 MHz for the hydrogen at 1 Tesla (T) magnetic field. This frequency at which the spins rotate is commensurate with the magnetic field strength. The product of the gyromagnetic ratio and the field strength (γB_0) yields the Larmor frequency, at which the RF energy must be delivered to achieve nuclear resonance.
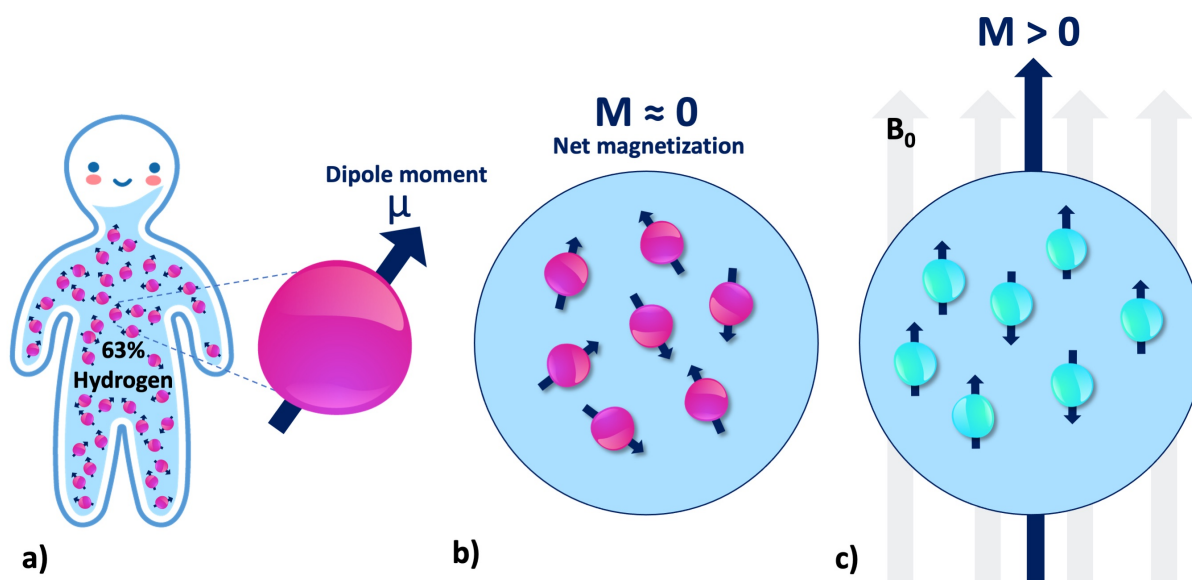
Figure 2.6 **a)** There are millions of protons (i.e., spin or nuclei) even at the smallest macroscopic unit volume relevant to the MR imaging of the human body. Each individual proton exhibits an infinitesimally small magnetic moment (μ). **b)** Without B_0, the protons in a spin pool exhibit random alignment. **c)** In presence of B_0, the spins are aligned with the magnetic field (parallel or antiparallel), giving rise to a net magnetization (M).

At the macroscale, we will be concerned with millions of protons at once (Figure 2.6a), even for a unit volume of 1mm^3. Absent an external magnetic field, magnetic moment vectors are oriented at random (Figure 2.6b). When a magnetic field is applied, they are aligned either parallel (spin-up, low energy) or antiparallel (spin-down, high energy) with the applied field (Figure 2.6c). Given the vast abundance of these spins, the relevant question becomes: which spin configuration is dominant? According to the second law of thermodynamics, if the spin system is at thermal equilibrium, i.e., no energy enters or leaves the system, the entropy of

that spin system increases (Carnot et al., 1899). This omnipresent tendency toward disorder favors low energy (Ferris, 2019). Therefore, the spin system tends to have slightly more low-energy hydrogen atoms (spin-up, parallel). Although the difference is as small as 40 per million protons (Webb, 2016), a net magnetic magnetization (Figure 2.6c) can be observed by a real-world NMR experiment.
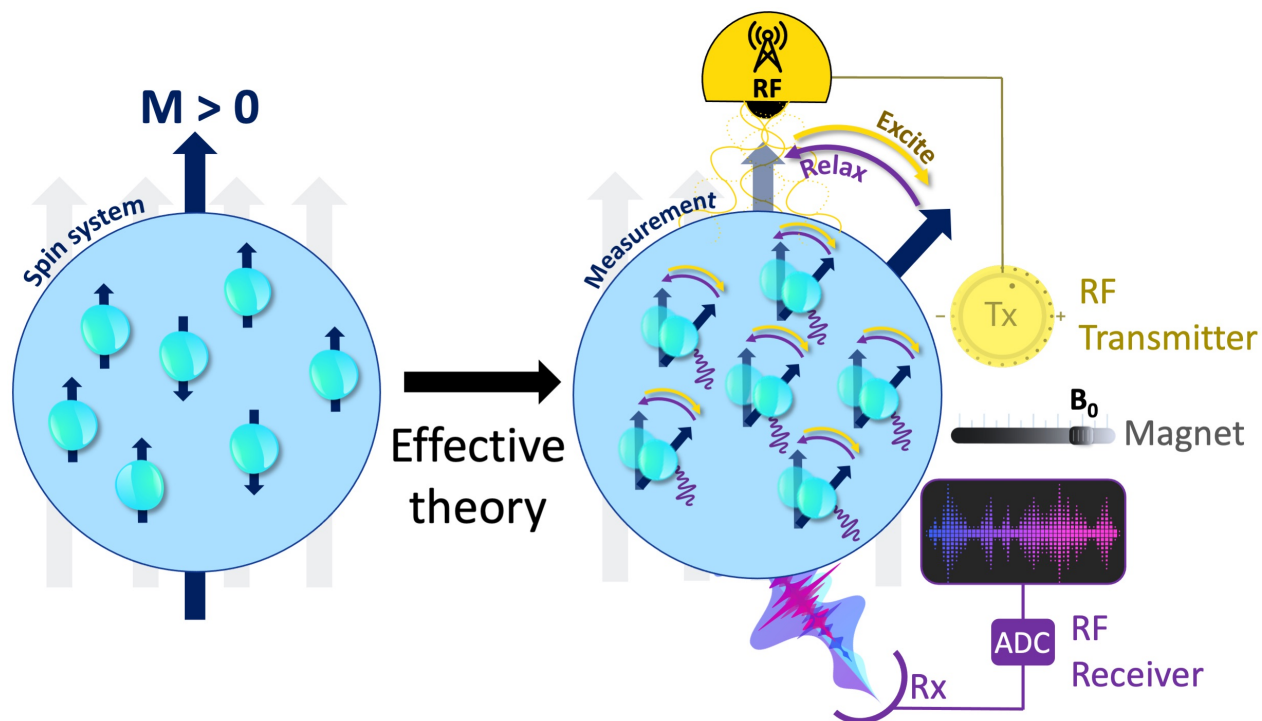
Figure 2.7 The measurement instrumentation is concerned with the relevant degrees of freedom and an effective theory that explains how these coarse-grained variables respond to the perturbations of the measurement system. Some fundamental components of an MRI measurement include a uniform magnetic field generator (i.e., a magnet), an RF transmitter (Tx, yellow), an RF receiver (Rx, purple) and an analog-to-digital converter (ADC).

To perform a real-world measurement, an effective theory is needed to describe how the coarse-grained features of the targeted system (Figure 2.6) changes upon interaction with the measurement instrumentation (Figure 2.7). The effective theory of relaxation for the bulk magnetization M was described by Felix Bloch in 1957, laying one of the cornerstones to bring MRI to reality (Bloch, 1957). A basic instrumentation to study the relaxation of a spin system is depicted in Figure 6c, including a uniform magnetic field (B0) generator, an RF transmission system tuned to the Larmor frequency and an RF receiver coil followed by an analog-to-digital converter (ADC). The following section describes how Bloch equations explain the macroscopic behaviour of the net magnetization and how MRI scanners make use of this effective theory to create images.

Measuring and encoding the MRI signal

Following an excitation of the spin system by an RF energy at the Larmor frequency (γB_0), the macroscopic Bloch equation describes how net magnetization evolves over time in a fixed cartesian reference (xi, yj, zk), under B_0 is given by:

$$\frac{\mathrm{d}M}{\mathrm{d}t} = \gamma M \times B_0 - \frac{M_x\hat{\mathbf{i}} + M_y\hat{\mathbf{j}}}{T2} - \frac{(M_z - M_0)\hat{\mathbf{k}}}{T1} \tag{2.1}$$

where M_0 is the initial magnetization of the spin system and T1 and T2 are the time constants for the longitudinal and relaxational components of relaxation. The first term of the Equation 2.1 is precessional, and the last two terms are the relaxational components of the Bloch equation. Recall that in thermal equilibrium, the net magnetization is aligned with the applied magnetic field (Figure 2.6c), where the longitudinal component is the net magnetization ($M = M_z = M_0$) and the transverse component equals zero ($M_x = M_y = 0$). In this case, the last two terms of the equation vanish, leaving the precessional component of the equation. When the phenomenological Equation 2.1 is solved for the longitudinal (M_z) and the transverse (M_{xy}) components of the macroscopic magnetization, the explicit solutions are given by:

$$M_z(t) = M_z(0)e^{\frac{-t}{T1}} + M_0(1 - e^{\frac{-t}{T1}}) \tag{2.2}$$

$$M_{xy}(t) = M_{xy}(0)e^{\frac{-t}{T2}} \tag{2.3}$$

Note that Equation 2.2 describes an exponential recovery for Mz to return the equilibrium after excitation. On the other hand, Equation 2.3 states that the transverse magnetization follows an exponential decay, quickly converging to zero. The relationship between these components is analogous to how the capacitor of a defibrillator charges and discharges. As soon as the paramedic delivers the shock, the capacitor abruptly empties (Figure 2.8a). At this stage, the paramedic is focused on how the energy (M_{xy}) is dissipated across the patient's body (transverse plane). The time elapses from the beginning of the delivery to when 37% of the energy ($1/e = 0.37$) remains to dissipate corresponds to T2 of the M_{xy}. This is a momentary process as in a click sound. To deliver the next shock, the capacitor should be charged again to a desired level (Figure 2.8b). The time taken for the battery to reach 63% of its total capacity ($1-1/e$) corresponds to T1 of the M_z.

The use of the defibrillator by a paramedic explains the relationship between the T1 and T2 in the context of a repetitive energy deposition. However, it lacks the measurement aspect of an NMR experiment. To complete the analogy, we will design a calibration setup to measure
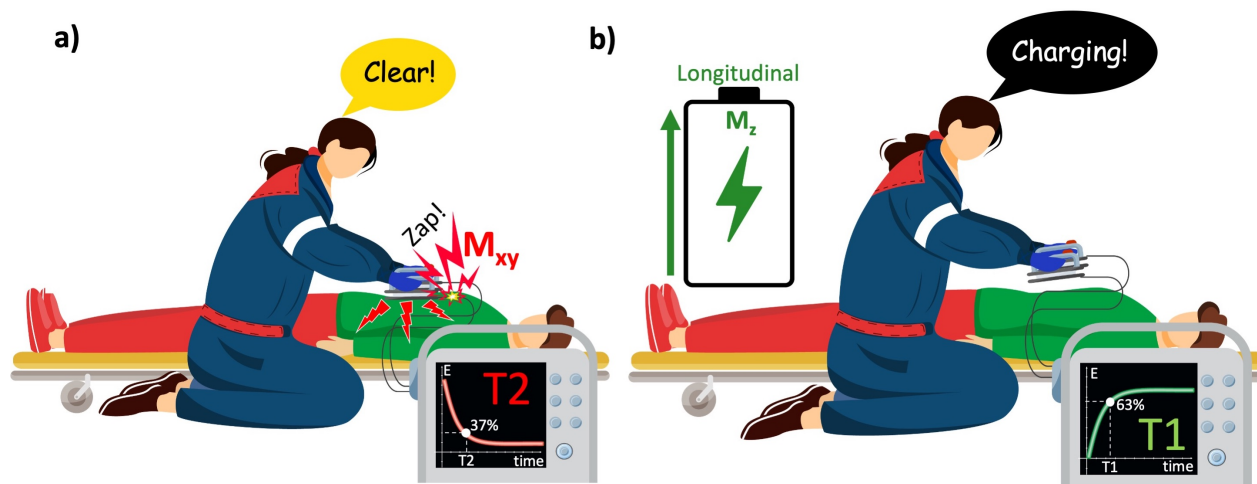
Figure 2.8 Time-dependent behaviour of the transverse and longitudinal magnetization can be compared with how the capacitor of a defibrillator empties and recharges. **a)** When the paramedic activates the shock paddles, the capacitor quickly discharges its energy to the transverse plane (patient's body). **b)** To deliver the shock again, the capacitor must be recharged, which happens quickly, yet relatively much slower than it discharges.

the energy delivered to the patient's torso. To measure this indirectly, a loop will be placed under the stretcher and the current induced in the loop as a result of the delivered shock to the patient's body will be recorded (Figure 2.9). The signal observed at the end of each shock corresponds to the free induction decay (FID) in an NMR experiment.

In a spin system, a hydrogen atom precesses at its Larmor frequency (γB_0). Following an excitation pulse, M_{xy} can be observed, because the on-resonance RF energy nudges all the spins toward rotating synchronously (i.e., in-phase). As they fall out of phase (i.e., dephased), the measured signal fades out. If the biochemical composition of the excited volume varies spatially, the measured FID will be a summation of slightly varying frequency components. For example, a hydrogen atom from the water has a longer response than a hydrogen from the fat ($T2_{fat} < T2_{water}$). When the frequency components of the FID are observed by applying a Fourier transform, the respective peaks will be separated by a certain extent in the NMR spectrum, depending on the field strength (Figure 2.10).

With the ability to precisely reveal molecular signatures, NMR is one of the most popular methods of chemical spectrum analysis, which is still in active use for a broad range of applications. A familiar example would be the benchtop NMR spectrometers at the airports that are used for tracing explosives and narcotics. However, spectral analyses take place in one dimension, in which the application of magnetic resonance was restricted for 25 years.
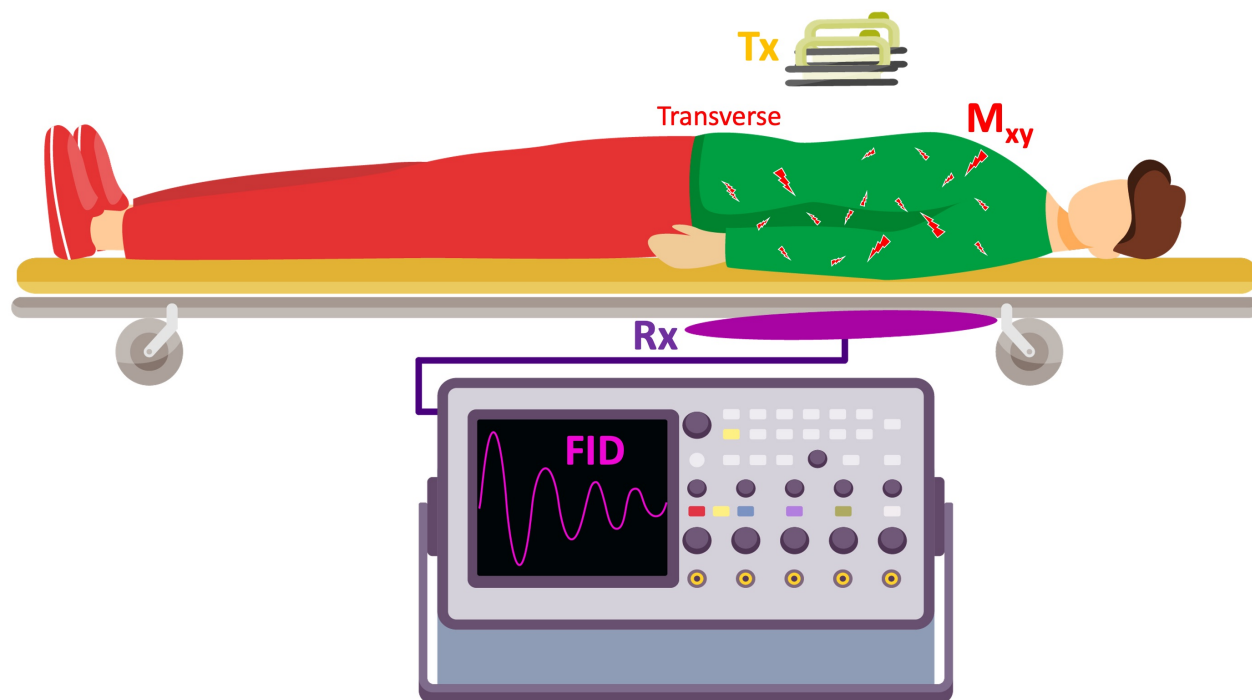
Figure 2.9 A hypothetical calibration setup: To have a measure of the energy delivered to the patient, a conductive loop is placed under the stretcher. After the shock is delivered, the current induced in the loop is captured by a oscilloscope.

Before looking at the new dimensions of NMR, a brief historical perspective will be presented about how they came to light.

A magnetic field gradient refers to a gradual change in B_0 in any desired direction. This is achieved by flowing high-amplitude electric currents through coils installed in three orthogonal directions. For example, Z gradients (head-foot direction) are a pair of circular Helmholtz coils (green rings in Figure 2.11) that can generate a gradually increasing or decreasing magnetic field by running currents in opposite directions. The higher the current, the steeper the magnetic field difference between the opposite ends of the scanner's bore. The magnitude of this gradient field can be adjusted such that the spins precess at the Larmor frequency of B_0 only at a certain region (selected volume). This way, only the spins from the selected volume will absorb the RF energy deposited at the resonance frequency. To achieve this "spatial localization", the RF transmitter is turned on concurrently with the gradient coils, sustained briefly (typically in a micro- to milliseconds scale), then turned off. Operating hardware components in this fashion is termed playing (RF or gradient) pulses. Timing of such events is described by pulse sequence diagrams. Figure 2.12 shows the sequence diagram describing the spatial localization illustrated in Figure 2.11.
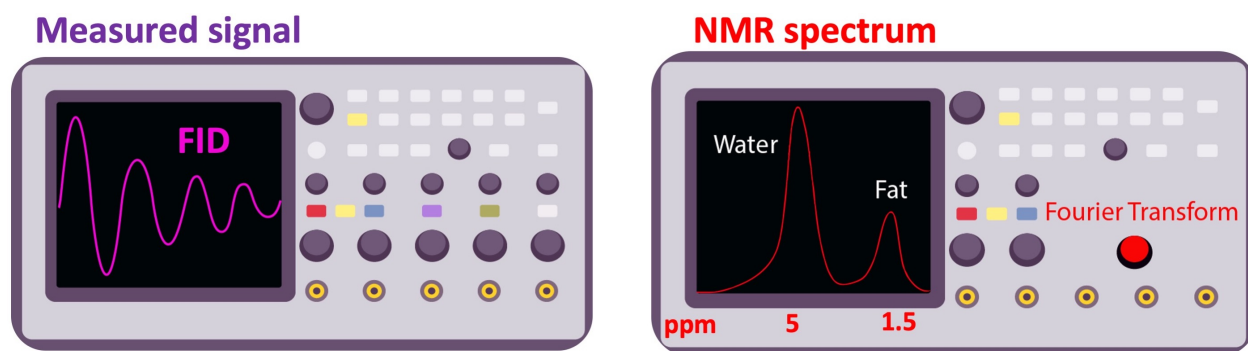
Figure 2.10 The free induction decay (FID) signal (left), and its frequency spectrum (right). As the electron of the hydrogen atom is more shielded in fat, its peak appears on the lower end (right, by convention) of the chemical shift spectrum. The chemical shift of water and fat is separated by 3.5 part-per-million (ppm), which corresponds to 146 Hz frequency difference at 1T ($3.5^{-6}\gamma$).

However, the slice selection procedure does not encode the measured signal with positional information. If an ADC event (i.e. signal measurement) was followed soon after the RF pulse was turned off, the receiver coil (Rx, purple, Figure 2.11) would collect information from the whole excited region at once. To form an image, the received signal must be encoded in-plane, which is along the x (row) and y (column) axes (axial plane) for the selected region.

Figure 2.13b explains how spatial encoding in the row direction is performed by playing a gradient along the x-axis (G_x, teal) while the signal is being measured (ADC, purple). As a result of this, spin locations across a single row (the red box outlined in Figure 2.13c) are uniquely sorted out as a function of their frequency, namely the frequency encoding. The process of acquiring data using frequency encoding is termed a readout (purple box) and the acquired data is referred to as an observation (Figure 2.13d). During a readout, the receiver coil picks up a sinusoidal electromagnetic signal as an observation (purple sinusoid in Figure 2.13d), composed of the frequency components encoded per voxel (blue sine waves in Figure 2.13d). Therefore, the ADC must ensure that the observation is sampled at a high enough rate to resolve all 10 frequency components. In this example, the observation must be sampled at least at 20 locations to satisfy the Nyquist condition (purple squares in Figure 2.13d). The ADC hardware of modern MRI scanners is fast enough to achieve high sampling rates up to 500kHz (Graessner, 2013) and smart enough to perform an IQ sampling, separating the magnitude and phase components of each data point (Kirkhorn, 1999). This offers the convenience to place an observation to its location in a special data plane: the k-space (Figure 2.13e).

By convention, the location in the horizontal axis of the k-space is determined by the spatial
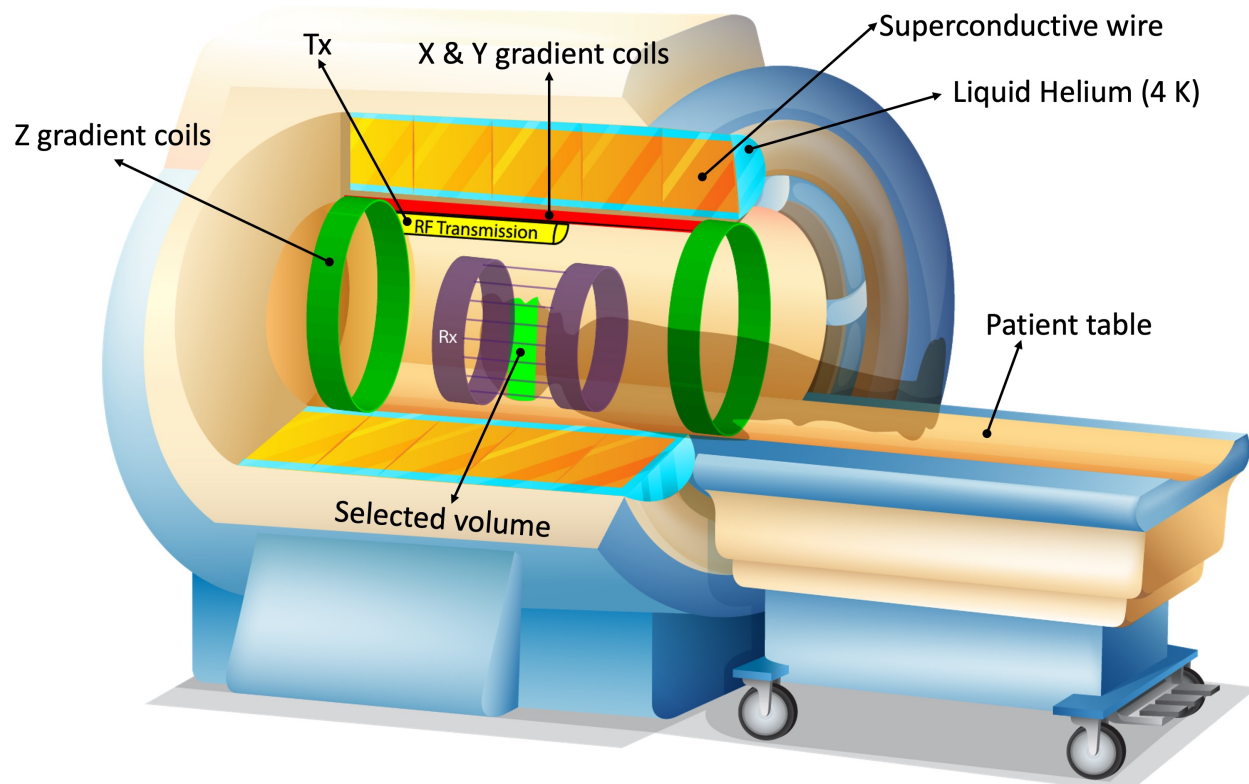
Figure 2.11 Hardware components of a modern MRI scanner using a superconductive magnet. Z-axis gradients (green rings) spatially vary the magnetic field, such that only the spins at a limited region (light green area in the patient's head) precess at the Larmor frequency (γB_0).

frequency and the value of each cell is proportional with the magnitude of the respective signal component. The k-space is arranged such that the higher frequency components are located around the skirts, whereas the lower frequency components are closer to the center. In a sense, placing an observation in its k-space location corresponds to adding in the contribution of that acquired portion to the whole image, but in the frequency domain. Hence, there is not a pointwise correspondence between the k-space and the MR image it represents. Instead, every single cell in the k-space (each hexagon in Figure 2.13e) carries information about the whole image. The concept of k-space involves several layers of abstraction that are beyond the scope of this introduction; therefore, the reader is referred to (Mezrich, 1995) for an intuitive understanding of k-space. For the next step of our MR image generation example, we will be concerned with how to fill out multiple rows of the k-space (along the red axis).

Recall that an excitation RF pulse nudges all the spins toward rotating synchronously, so that $M_x y$ across all the pixels share the same phase. Whenever a gradient is played, an opposite effect comes into play: phase of the spins along the gradient direction experiences a location-dependent shift. Each phase shift moves the location of the observation in k-space
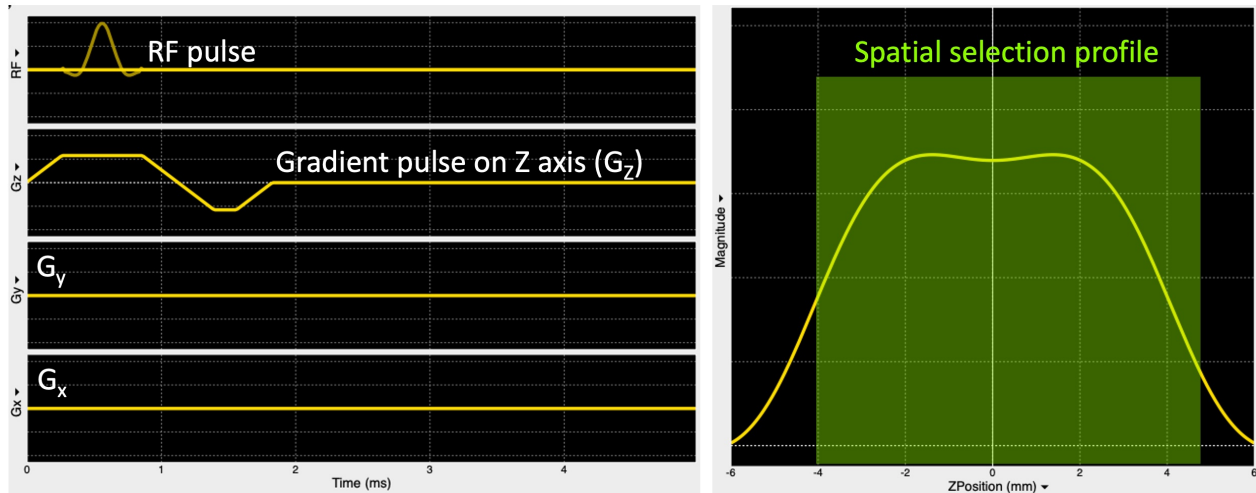
Figure 2.12 A pulse sequence diagram (left) showing an RF pulse to excite the spins only in a plane selected along the z-axis (G_z) and the respective slice profile (right).

along the ky axis (Figure 2.14c), upwards or downwards depending on the polarity of the applied gradient. The amount of dephasing (i.e., the number of Δ_{ky} steps) is proportional with the gradient area, and its effect can be rewinded to restore the transverse magnetization by playing a gradient of the same area with opposite polarity. Figure 2.14b shows a phase-encoding gradient (red) played with negative polarity before the readout in order to shift the observation to the next lower row in the kspace (Figure 2.14e). To rewind this dephasing effect, the readout is followed by another Gy gradient with positive polarity.

By altering the gradient amplitude stepwise, the whole k-space can be sampled (Figure 2.14f). Therefore, the time that it takes to scan the plane shown in Figure 2.14c is a product of the number of rows and the duration of each repetition, namely the repetition time (TR). Finally, the MR image is reconstructed by applying a 2D inverse Fourier transform to the fully sampled k-space, where each sample corresponds to a grid location.

For the sake of simplicity, the examples in Figure 2.13 and 2.14 assumed that the signal observations followed by an excitation pulse can be used to form an image. However in practice, an FID signal is short-lived; therefore, it cannot directly contribute to reconstructing an MR image. This brings us to two milestone discoveries in NMR by Erwin Hahn, which echoed into the beginning of MRI from a quarter-century behind and have become an indispensable part of its everyday use since then: spin- and gradient-echo. For further reading on the history of these methods, the reader is referred to an MRM Highlights interview with Erwin Hahn (Feinberg, 2018). The following section will briefly introduce the spin-echo, then will look at a gradient-echo sequence which is of importance to the third article of this thesis.
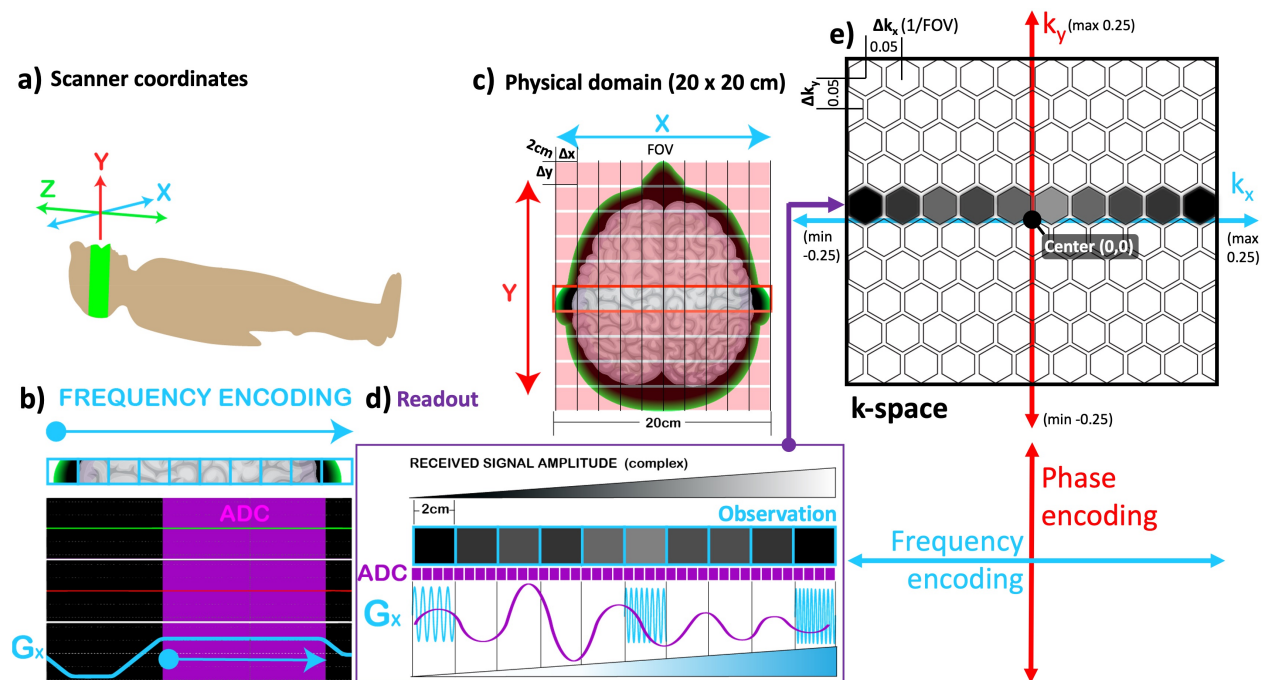
Figure 2.13 The correspondance between the scanner coordinates (**a**) and the selected imaging plane (**c**) is illustrated along with the pulse sequence diagram for frequency-encoding (**b**). The observation obtained by the readout (**d**) is shown in the k-space (**e**).

2.1.3   Coming full circle back to two values with two pulse sequences

"*The Court concluded that MRI machines rely on the tissue NMR relaxations that were claimed in the patent as a method for detecting cancer, and that MRI machines use these [T1 and T2 values] to control pixel brightness...*" (GE vs Fonar 1996, U.S. Fed. Cir.)

Although going from values to brightness and back again is a chicken-and-egg problem, GE could have stood a better chance by defending that the pixel brightness can be used to obtain T1 and T2 values, but T1 and T2 values cannot be directly utilized to obtain pixel brightness by MRI machines. Indeed, T1-weighted and T2-weighted contrasts are primarily determined based on the contribution of T1 and T2 (or neither). Nevertheless, to adjust those contributions, MRI machines use pulse sequences and parameters. This section will introduce two essential sequences, spin-echo and gradient-echo to show how contrasts are determined (conventional MRI) and the values are calculated (qMRI).

In their seminal article titled "atomic memory", Brewer and Hahn introduce the spin-echo (Hahn, 1949) by tapping into an intellectual conflict between the 2nd law of thermodynamics and the time reversal symmetry (Brewer and Hahn, 1984). The discussion is illustrated in Figure 2.15 to portray the paradoxical nature of the discussions on the physical phenomenon
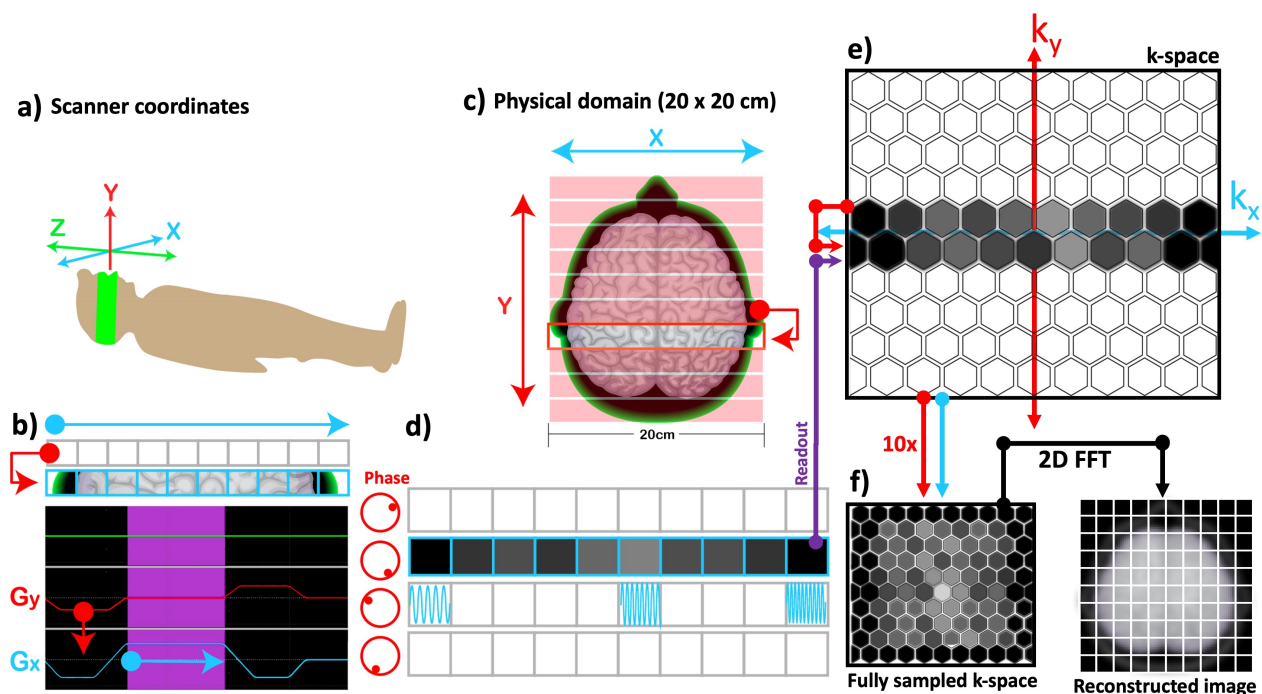
Figure 2.14 The correspondence between the scanner coordinates (**a**) and the selected imaging plane (**c**) is illustrated along with the pulse sequence diagram for phase encoding (**b**). Once the whole k-space is sampled by incrementing the phase-encoding gradient (red) (**b**) multiple times, a 2D inverse Fourier transform is applied to reconstruct the MR image (**f**).

giving rise to a spin-echo. This is yet another effect exploited by the MRI, without necessarily having a complete understanding of the underlying "hidden order" effect emerging from the microscale interactions. For further reading on the conflict between the time symmetry and the entropy, the reader is referred to a recent blog post (Siegel, 2019).

Following an excitation pulse, the spin pool goes towards disorderliness and the transverse magnetization decays (the pancake batter expands, Figure 2.15). In their article, Brewer and Hahn discuss resurfacing of an ordered state out of this increasing entropy by reversing the phase order of the spins (the cooked pancake contracts). By retaining the pancake analogy, Figure 2.16 simulates a spin-echo (SE) pulse sequence and shows the spin evolution at certain time points. At the point (B), a 90°excitation pulse rotates the net magnetization to transfer plane, and after some time, the spins dephase (C, arrows fanning out in the x-y spin scatter plot) as the FID disappears. The "refocusing pulse" rotates the net magnetization by 180°in the transverse plane (D) and reverses the phase order of the spins, analogous to flipping the pancake (Figure 2.16). After a period equal to the duration between the 90°and 180°RF pulses, a spin echo is formed (F). The time elapses between the center of the excitation pulse and the peak of the echo signal is termed the echo time (TE).
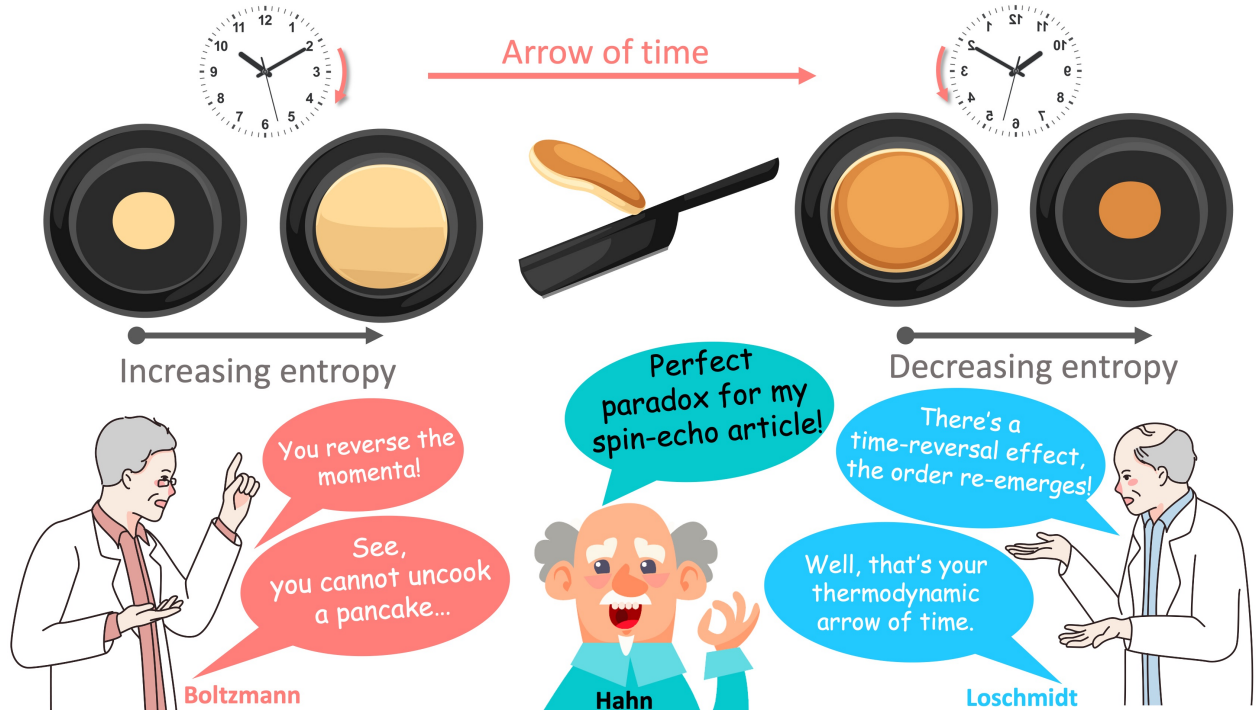
Figure 2.15 The second law of the thermodynamics (Boltzmann) vs the time reversal symmetry (Loschmidt) and the relation of this conflict to spin-echo (Hahn). The pancake analogy is followed in Figure 2.16 for completeness.

$$S = M_0 * (1 - e^{\frac{-TR}{T1}}) * e^{\frac{-TE}{T2}} \frac{sin\theta}{1 - cos\theta * e^{\frac{-TR}{T1}}} \tag{2.4}$$

Equation 2.4 shows the signal representation of a standard SE acquisition. Given that the brightness of the image pixels is determined by the magnitude of the signal S, the relevant contribution of T1 and T2 to the image contrast can be adjusted by changing the TR and TE. For TE $\to 0$ (i.e. short TE), the last term of the equation converges to identity, reducing the T2 contribution. If this is coiled with a long TR (TR $\to$ inf), the exponential in the second term of the equation converges to zero, reducing the T1 contribution. Therefore when the TE is short and the TR is long, the contribution to image contrast comes from the density of the spins, i.e. proton density. On the other hand, to increase the T2-weighting by keeping TR the same (long), the TE must be increased (the importance of the last term increases). Figure xx exemplifies this by showing the same image across 6 echoes, where substances with longer T2 (e.g., eyes and the cerebrospinal fluid (CSF)) appear gradually brighter compared to the other structures in the image as the TE increases.

Since the signal representation is known for the basic SE acquisition, the signal (the pixel
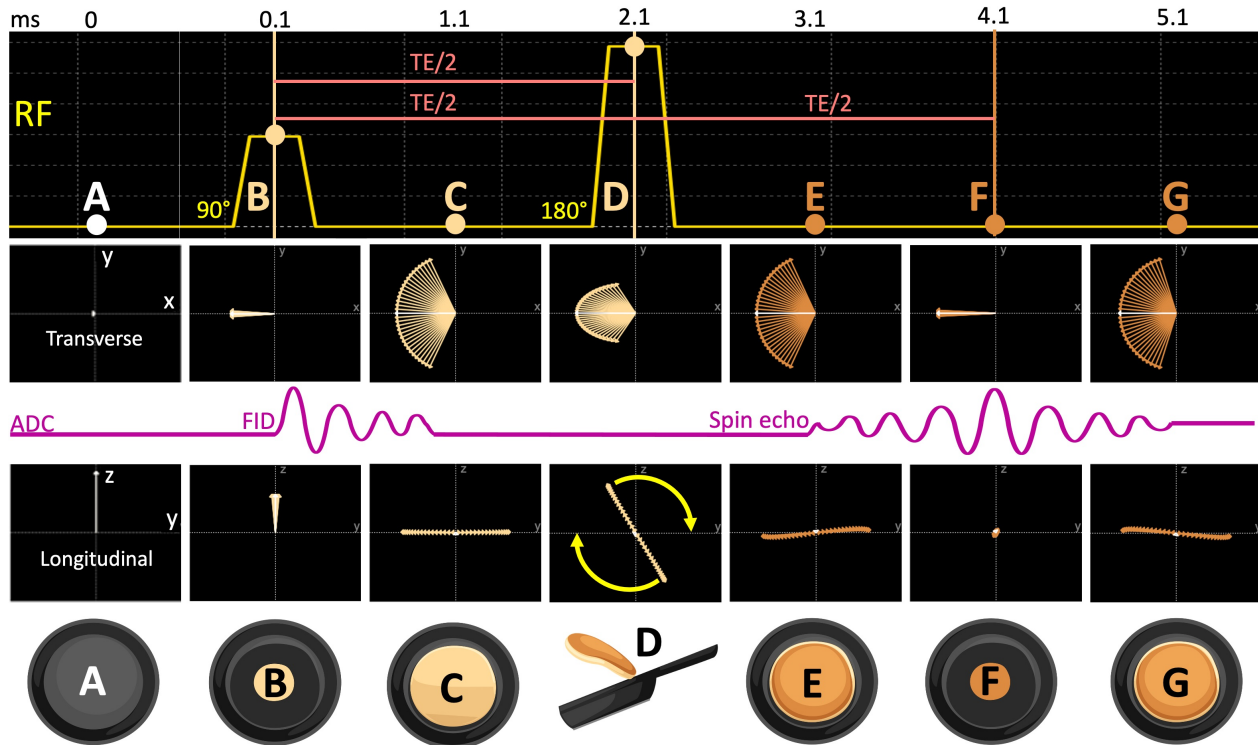
Figure 2.16 The spin evolutoin diagram of a spin-echo sequence is shown **(A)** before the excitation, **(B)** at the peak of the excitation pulse, **(C)** one millisecond after the 90°pulse, **(D)** at the peak of the refocusing pulse, **(E)** one millisecond after the 190°pulse, **(F)** at the echo time (TE) and **(G)** following the echo.

brightness) can be sampled at several TE's and fitted to the exponential decay defined by the last term of the equation to calculate T2, namely T2 mapping. The second term of the equation will not be taken into account, as the TR will be kept constant across the samples. Figure 16 shows a T2 mapping example, where an axial image of the human brain was collected across 32 TE's ranging from 12 to 380ms.

So far we looked at how an echo forms out of a 90-180°pulse pair, and created a T2 map based on its signal representation. Although the images created using SE convey good soft-tissue contrast and are robust against motion artifacts, they often come at the cost of long acquisition time (TR ranges from half to a few seconds) and high RF energy deposition. Fortunately, MRI offers yet another way to generate echoes by taking advantage of the FID following a single RF pulse. The reversal effect needed for echoing the signal is achieved by dephasing and rephasing the spins with the use of a bipolar gradient (Figure 2.18). Therefore, this method is named the gradient echo (GRE).

Unlike SE-based sequences, GRE sequences allow for shorter TE (a few milliseconds) and
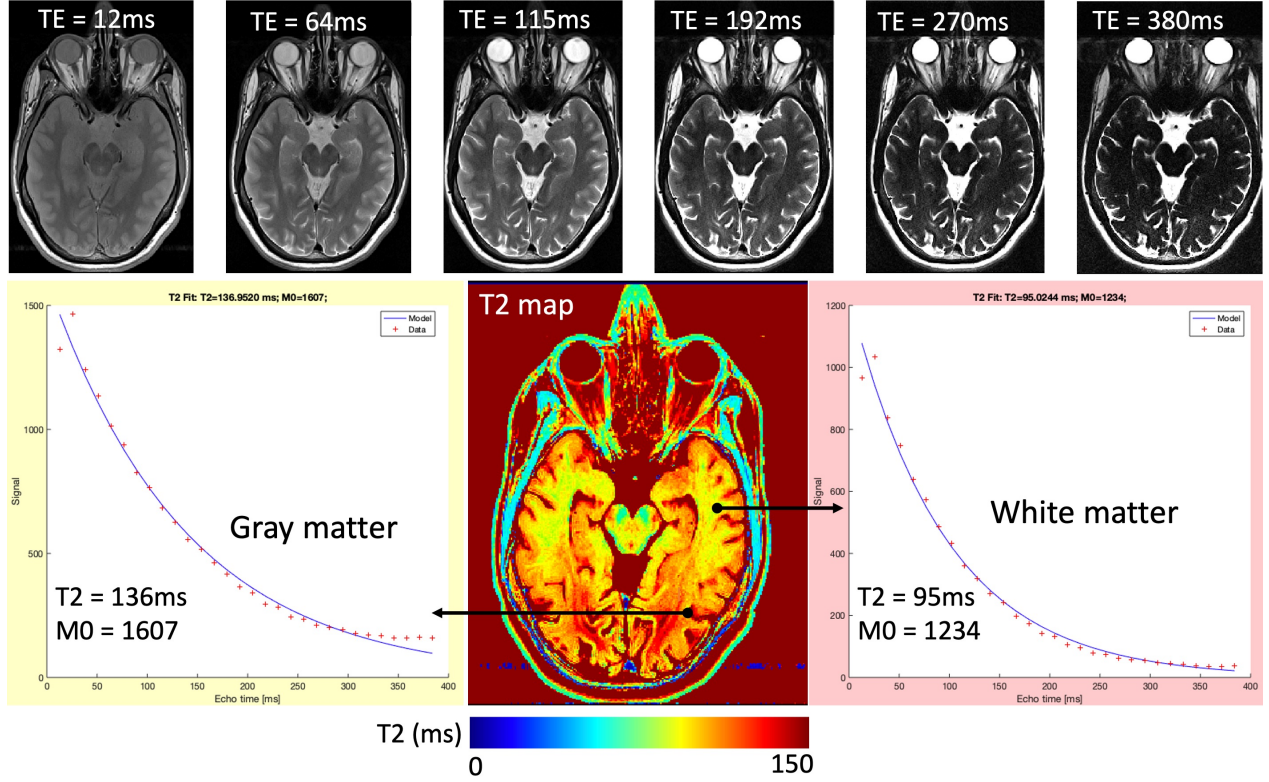
Figure 2.17 An example T2 map, estimated by fitting voxel-wise brightness values (red plus signs) across 32 echo times to the exponential decay (blue line) defined by Equation 2.4. The top row shows how conventional image contrast changes from proton-density to T2-weighted as the TE increases from 12ms to 380ms.

TR (10-50 milliseconds), resulting in faster acquisitions. In conventional imaging, one of the most popular GRE sequences is the spoiled gradient-echo (SPGR) sequence (Haase et al., 1986), as it allows large volumetric coverage within clinically feasible scan durations. SPGR is also widely used as a basis for various qMRI methods as the basis, because it provides a simple signal representation:

$$S = M_0 \frac{sin\theta * (1 - e^{\frac{-TR}{T1}})}{1 - cos\theta * e^{\frac{-TR}{T1}}} e^{\frac{-TE}{T2^*}} \tag{2.5}$$

Unlike the SE signal representation (Equation 2.4), Equation 2.5 does not include a term explaining the decay of the transverse magnetization by T2. Instead, the last term of the SPGR signal representation indicates that the TE governs the signal contribution of T2* – the effective T2. As the GRE is restored from the FID (Figure 2.18), the resulting echo is susceptible to slight variations in the main magnetic field. These variations may originate from the hardware-related imperfections of the B_0, or from the field disruptions induced
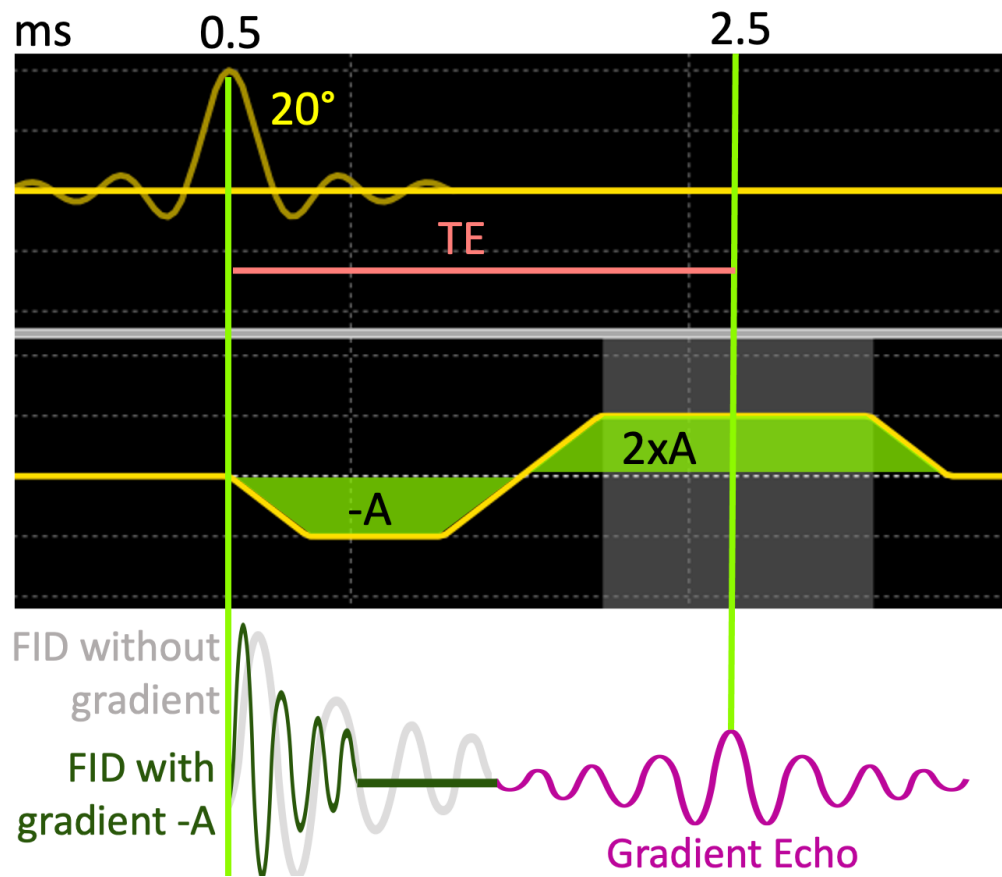
Figure 2.18 The formation of gradient echo by playing a bipolar gradient followed by a 20°RF pulse.

by adjacent substances with distinct magnetization levels, such as the air-tissue interfaces around the nasal cavity. As a result, T2 weighting cannot be achieved with SPGR. Instead, the second term of the Equation 2.5 indicates that the SPGR sequence is primarily T1-weighted, which can be controlled by changing the flip angle (FA) of the excitation pulse (θ) or the TR.

After a number of GRE excitation pulses, the longitudinal magnetization reaches a dynamic equilibrium, i.e. steady-state. When the steady state is reached, the spin system looks static to the observations at a macro-scale, whereas the underlying spin interactions carry on by balancing out each other. This is similar to a person walking up (T1 recovery) an escalator that is going down (the excitation pulse). When this equilibrium is accounted for by the Equation 2.5, the FA that maximizes the signal for a fixed TR is given by:

$$\theta_E = arccos(e^{\frac{-TR}{T1}}) \tag{2.6}$$

where θ_E is known as the Ernst Angle (Ernst and Anderson, 1966). Figure 18 shows this relationship by simulating an SPGR signal across multiple FA at a fixed TR of 30ms for a substance with T1=800ms. It can be seen that the signal is maximized around an FA of 15°, as given by the Equation 2.6 for these settings.
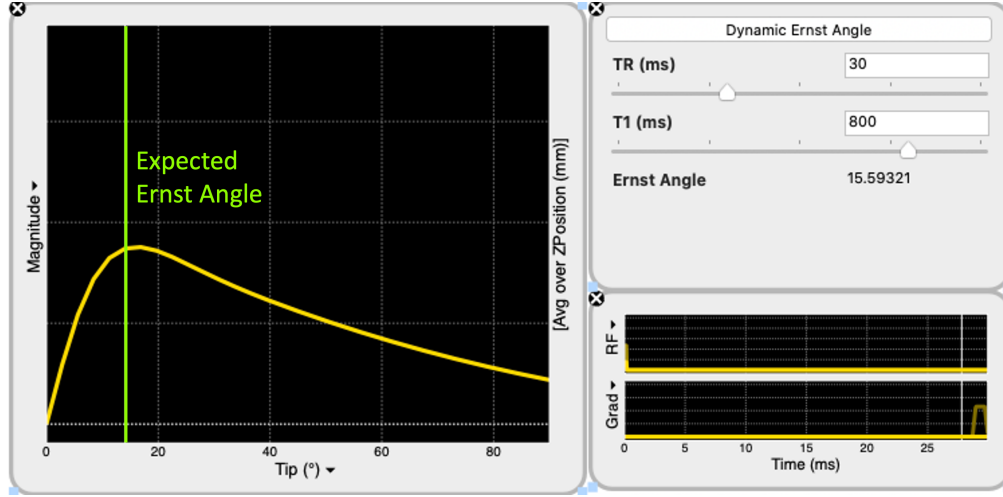
Figure 2.19 Spoiled gradient-echo (SPGR) signal is simulated for excitation flip angle (FA) ranging from 0 to 90°. With the repetition time set to 30ms, the Ernst Angle (θ_E) is calculated at 16° for a spin system with T1 of 800ms (2.6). The simulated signal agrees with the calculation, indicating that the signal is the highest when the FA is around 16° (green line).

However, this relationship is valid under the assumption that the transverse magnetization is zero. Therefore, any residual transverse magnetization during the readout violates the steady-state conditions of the signal representation given by Equation 2.5. To avoid this, the entire spin system is dephased at the end of each TR by using a spoiler gradient, hence the name the "spoiled" GRE. In real-world pulse sequence implementations, the gradient spoiling is coupled with RF spoiling by phase cycling the excitation from TR to TR at predetermined increments (Zur et al., 1991). Figure 2.20 displays Bloch simulation results, indicating the critical role of selecting the spoiler gradient area (a), enabling RF spoiling (b) and the selection of a proper phase increment value in disrupting the residual transverse magnetization. The effect of spoiling efficiency becomes particularly important when multiple SPGR acquisitions are performed at different flip angles to calculate a T1 map, namely variable flip angle (VFA) method (Fram et al., 1987). This is simply because when the observed data deviates from the expected signal representation (Figure 2.20), the fitted parameters become inaccurate. A more theoretical explanation of the VFA T1 mapping method, its accuracy aspects and limitations are available as an interactive qMRLab blog post (Boudreau, 2018).

Figure 20 shows an example T1 mapping application using the SPGR sequence in a phantom
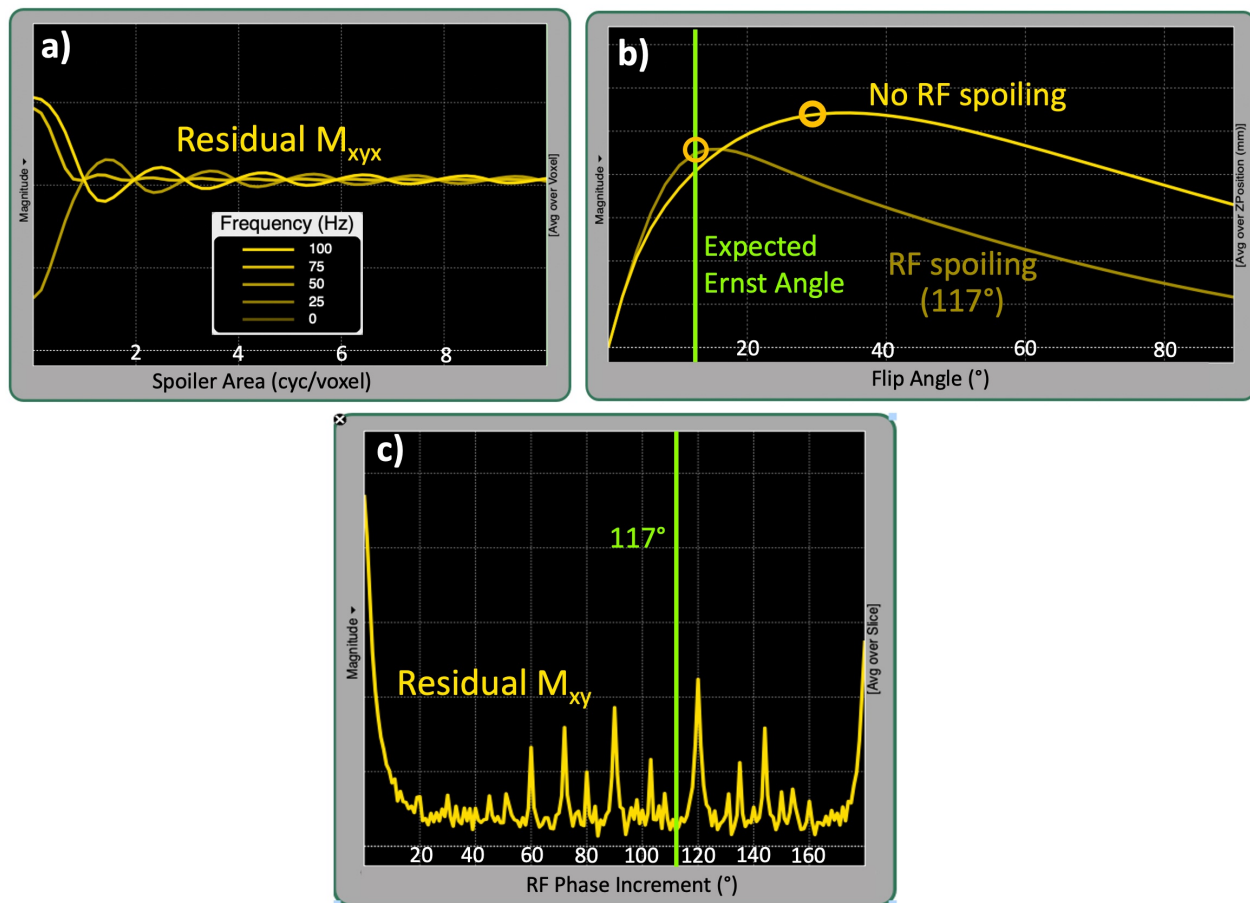
Figure 2.20 The influence of spoiling on SPGR signal is simulated for **(a)** the spoiling gradient area ranging from 0 to 10 (cyc/voxel), **(b)** enabling/disabling RF spoiling at 117°quadratic phase increment and **(c)** different phase increment values ranging from 0 to 180°.

with known values. The acquisitions were performed at two flip angles of 20° (a) and 6° (b), and TR=32ms. Both images show the center of a plate accommodating 14 spheres with T1 values ranging from 0.1 to 1.99 seconds in clockwise ascending order (R1 to R10). In agreement with the Equation 2.4, the image acquired at the higher FA shows superior T1 contrast (a), as the pixel brightness varies inversely with the reference T1 values (d). On the other hand, spheres in the lower FA image show similar brightness (b), in proportion with the spin density of the spheres. From this image pair, a T1 map (c) was estimated using a linear fit as described in (Mathieu), exhibiting good accuracy across the reference values.

In this section, we covered two fundamental pulse sequences: SE and SPGR. Starting from spin-level interactions, we looked at the signal representations of each sequence and how acquisition parameters are associated with the contrast characteristics of the resulting conventional images. By doing so, we analyzed how pixel brightness emerges from two values,
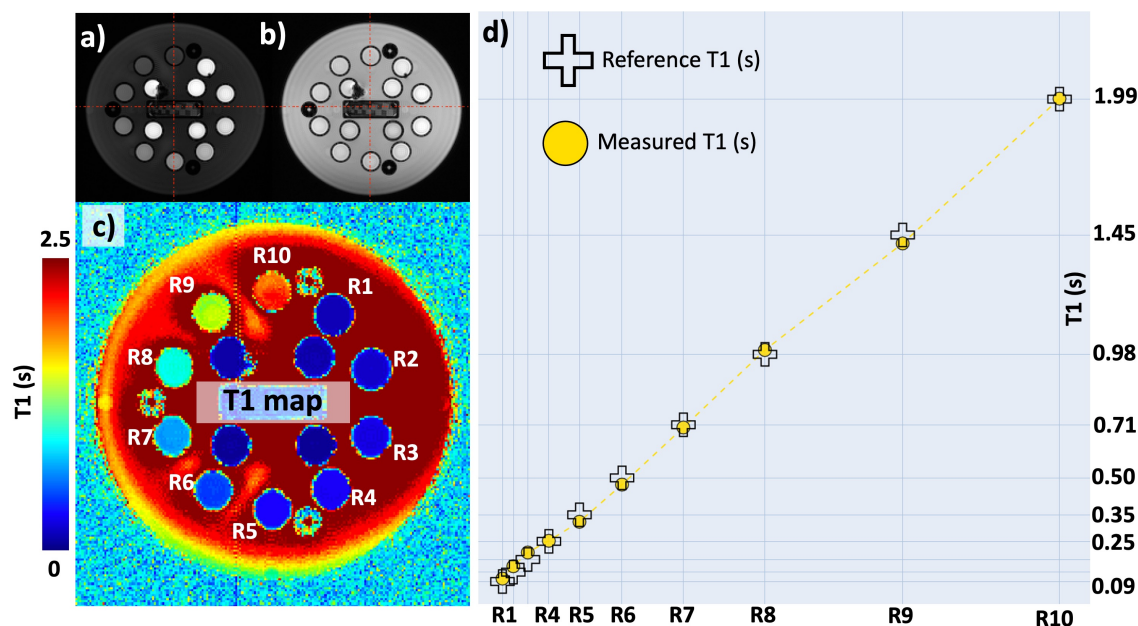
Figure 2.21 Variable flip angle (VFA) using SPGR sequence. **a)** T1-weighted and **b)** PD-weighted images of ISMRM/NIST system phantom acquired at 18 and 32ms, respectively. **c)** T1 map estimated by fitting images (a) and (b) to the relaxational component of the Equation 2.5. **d)** The comparison of the estimated and reference T1 values.

T2 and T1. Then using the same signal representations, we came full circle back to these values from the pixel brightness by applying two fundamental qMRI methods, multi-echo SE for T2 and VFA for T1 mapping. A more entertaining summary of the SE and GRE sequences is unintentionally delivered by two indie rock albums (Figure 2.22).

Figure 2.22 Two album covers by the Artic Monkeys and the Districts summarizing spin-echo (SE) and spoiled gradient-echo (SPGR) sequences.

2.1.4 If qMRI is possible and powerful, why is clinical imaging still conventional?

"*The court found that Fonar failed to establish the existence of standard T1 and T2 values, which are limitations of the asserted claims...*" (GE vs Fonar 1996, U.S. Fed. Cir.)

After decades of research and development since this court decision that saved GE millions of dollars, the standard values for T1 and T2 are still not well known (Bojorquez et al., 2017) and qMRI has yet to find one clinical application. This is partly because of the inherently complex make-up of the human body, where sensitivity alone is not enough to tease out biological variability. Quantifications should also be specific to the targeted microstructure, such as the myelin in the living human brain. For this purpose alone, the literature offers more than 30 methods for quantifying myelin at varying methodological complexity, yet they all appear to be statistically indistinguishable in specifying myelin (Mancini et al., 2020). This indicates that a lack of methodological extensity is not the culprit preventing qMRI from clinical use. Quite the contrary, there is an abundance of solutions, yet we cannot make an informed decision about which method to use. This problem has multiple roots and Figure 2.23 outlines the components of a qMRI study for identifying them.
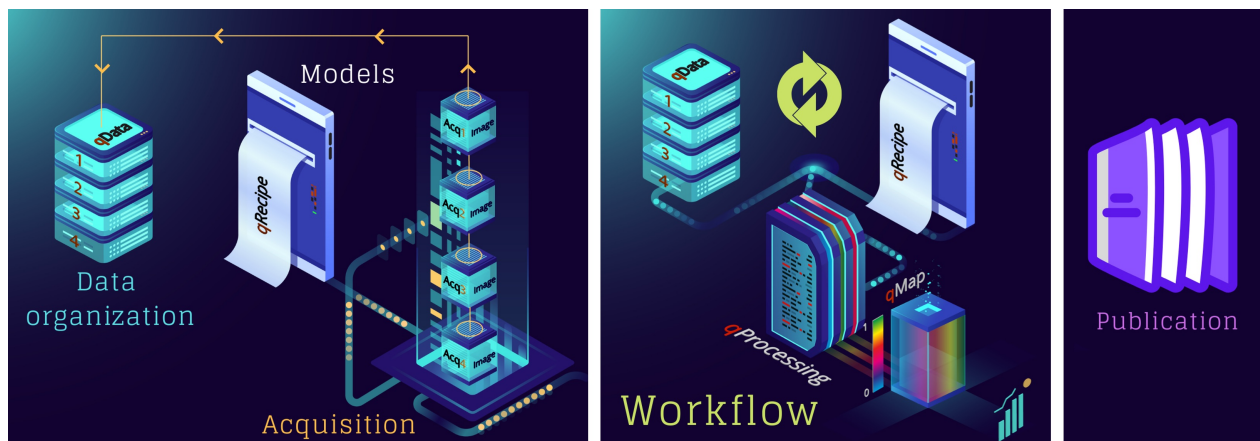
Figure 2.23 An illustration of the components that make up a qMRI study.

Every qMRI study starts with the acquisition of a set of conventional images. This is often achieved by altering protocol parameters according to the signal representation of the respective pulse sequence, i.e. successive runs. As previously shown in the SPGR example for T1 mapping (Figure 2.20), there are various parameters that are vital to the measurement accuracy and precision. In general, strict metrological standards are established for the manufacturing process of any medical device expected to fulfill some accuracy requirements. For example, all the ventilator vendors are obliged to disclose their measurement uncertainty for

inspiratory oxygen concentration (ISO 80601-2-12:2011). However, MRI is exempt from such a class of essential performance assessments on the accuracy and precision, given that the medical diagnoses using conventional MRI depend on qualitative feature recognition (Figure 2.1). In turn, design considerations that matter to the reliability of qMRI measurements fall through the cracks of the device manufacturing and programming processes. Although this is understandable from a vendor's cost-effectiveness standpoint, it bears dire consequences on the quantitative applications.
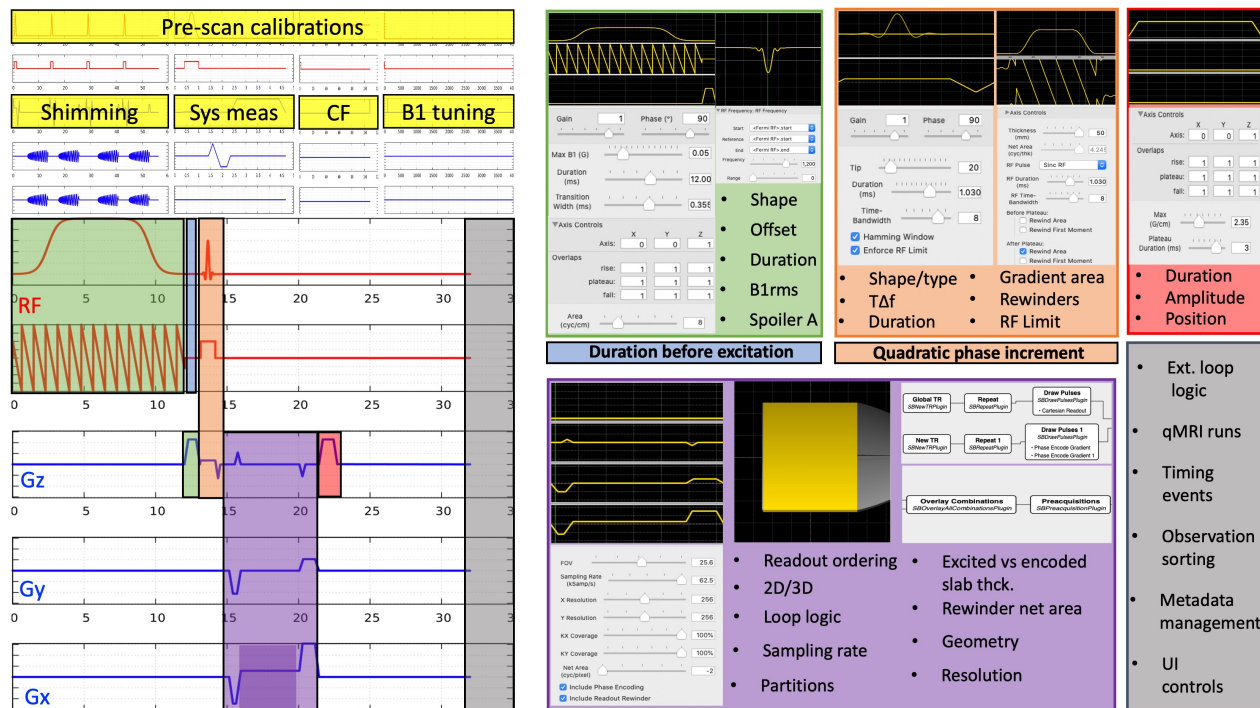
Figure 2.24 Choices involved in the implementation of a magnetization-transfer weighted spoiled gradient echo (SPGR) sequence are shown for all the gradient and RF waveforms involved.

Even for the simplest sequence implementations, there may be several parameters that matter to quantification, but are hidden from the end user. Figure 2.24 shows the implementation-level parameters that are available after the type/shape selections were made for an SPGR sequence with a magnetization-transfer saturation pulse. In addition to the sequence itself, pre-scan calibrations such as shimming, center frequency tuning and transmit gain adjustments are other factors that affect the measurement accuracy. For example, neither of the major vendor implementations enable the selection of spoiling gradient area (Figure 2.20a), RF spoiling regime (Figure 2.20b,c), magnetization transfer (MT) pulse specifications, excitation pulse type or the ordering of the observations (Figure 2.24). The more advanced the sequence, the more implementation choices come to the surface. These restrictions and

unknowns brought by vendor-specific sequence implementations trap tens of qMRI measurements into a maze of variability and prevent them from reaching the clinics (Figure 2.25).

Figure 2.25 The current landscape of quantitative MRI is a maze of variability for amazing methods. A complete recipe is needed to chart out the path towards clinics.

After the raw data is acquired, it has to be reconstructed to generate images, which is yet another process with a wide range of options to choose from. Therefore, properly formatting and making the raw data accessible is a non-trivial interim step for the provenance of the following steps. In addition, the relationship between the reconstructed images and the grouping logic entailed by the qMRI model should be retained along with sufficient metadata. Even though there is an emerging data standard for organizing the raw data (Inati et al., 2017), there is not a community consensus on how to organize qMRI datasets (Figure 2.23). This creates idiosyncrasies, challenging interoperability and decreasing efficiency of processing qMRI data.

There are dozens of publications introducing new qMRI methods, yet a majority of these implementations are kept within their labs of origin, making it challenging to reproduce qMRI studies. This is partly caused by the vendor restrictions. Nonetheless, it is generally possible to share the workflow components of the developed methods. Figure 2.23 shows that all

qMRI methods share a common methodology at their core: a signal representation (qRecipe) that relates a set of parametrically linked MR images (qData) to some microstructural and physical features (qMap) by computation (qProcessing). Although these ingrained attributes exist at a conceptual level in the source code developed by independent developers, there is not a consensus on how to represent them in a programming paradigm. Even though 80% of the source code made available by the MRI developers share the same programming language (MATLAB) (Boudreau, 2019), there is still a need for a common framework for the development of qMRI methods in MATLAB to make implementations easier.

To summarize the problems mentioned above, there are three outstanding issues that hinder the standardization of qMRI:

- Most methods are developed using in-house code that is difficult to distribute, challenging the accessibility and reproducibility of qMRI studies.

- The lack of a qMRI data standard poses an interoperability challenge for open-source solutions aiming at making qMRI methods publicly accessible.

- The unknowns involved in the implementation of commercial pulse sequences constitute a substantial source of (vendor-specific) variability in fitting quantitative parameters using pixel brightness data.

2.2 Myelin Imaging

The myelin sheath is a composition of glial cell membranes rolled like coiled spring pins, blanketing axons to insulate them from their surroundings thanks to its low capacitance and accelerating conduction velocity by interspacing capacitance regions along them (Waxman, 1982). A considerable portion of its mass is comprised of water molecules trapped between the spiral folds of the glial membranes. Non-aqueous parts of this substance is primarily composed of lipids, whereas proteins and nucleic acids make up one third of its dry mass (Norton and Autilio, 1966). Moreover, myelin is colocalized with iron molecules (Fukunaga et al., 2010), given that iron is one of the myelin building components used by oligodendrocytes (Todorich et al., 2009). As a result, myelin creates a complex susceptibility environment where weakly and strongly magnetized compounds are present along with water associated with macromolecules (lipids). Apparent contrast in most conventional brain MR images is dominated by the distribution of iron and myelin (Stüber et al., 2014). On one hand, these properties facilitates the sensitivity of MRI contrasts against changes in myelin state and in the iron content. On the other hand, it hampers the isolation of their relative contribution.

Quantitative susceptibility mapping has been at the forefront of iron concentration mapping (Bilgic et al., 2012). As for distribution of myelin concentration, several approaches have been developed in the realm of proton density (PD) mapping, relaxivity measurement and MT.

2.2.1 If not water, then it contains myelin: Proton Density

Among all myelin mapping techniques, one of the most widely accessible methods is the measurement of macromolecular tissue volume (MTV). Images acquired by standard SE sequences with relatively short TE and TR give PD weighting, indicating the amount of water content in a voxel (Abbas et al., 2015). Ascribing PD with water content implies the probing of the free water that are not associated with macromolecules. Remainder of content in a voxel after subtracting the portion that accounts for water content is attributed to MTV (Mezer et al., 2013). Voxels in the cerebrospinal fluid are assumed as 100% water to use values from those voxels as a calibration constant. Moreover, coil sensitivity maps are exported after scan to correct for B1 field receive inhomogeneity (B1-). It is worth noting that water protons are the signal source of all type of MR images. Therefore a complementary PD metric cannot be qualified as myelin specific marker irrespective of which corrections are applied. Nonetheless, MTV can be regarded as the absolute myelin content in healthy human spinal cord (Duval, Perraud, et al., 2016).

2.2.2 Which water belongs to myelin: Relaxometry

The T1 weighted images in the longitudinal assessment of developmental myelination has been widely used (Paus et al., 2001). However, all developmental changes may not be specific to the myelination. The gradual increase in the myelin content causes a gradual decrease in the native T1 values over time due to macromolecular content in the tissue. However, this goes hand in hand with a decrease in T2 and PD values (Deoni, 2011). Therefore, myelin specificity is not likely to be achieved by T1 mapping. However, some reports contradict this by indicating strong associations between T1 values and myelin content (Geyer et al., 2011; Sigalovsky et al., 2006). At any rate, quantification of native T1 yields an essential input for some other qMR metrics reviewed below. This can be performed by implementations of different sequences including inversion recovery (Barral et al., 2010), Look-Locker (Henderson et al., 1999) and variable flip angle approach with spoiled gradient-echo acquisition (Deoni, Rutt, and Peters, 2003).

Relaxation times of proton spins in water molecules depend highly on the biochemical composition of the environment where water resides. Compartmentalization of water content in

the human body is a long known phenomenon, explaining that more than half of the water is contained intracellularly, whereas one-third of the total water is hosted by extracellular space (Levitt and Gaudino, 1950). Boundaries separating these compartments exhibit permeable characteristics, enabling exchange of the water molecules. As a result, residence lifetime of a molecule at a given compartment becomes a relevant metric for MRI methods aiming at performing compartment specific measurement of the water molecules (Quirk et al., 2003). Therefore, to distinguish signals coming from multiple compartments using relaxometry based approaches, one should consider the difference between relaxation time and water exchange rate. If water exchange happens faster than the relaxation, this effect cannot be captured by the relaxometry measurements. In this case, signal is represented by a mono-exponential curve as exchange is confined to a single compartment, known as fast exchange regime. In the opposite case (slow exchange regime), signal from individual compartments can be resolved by fitting signal to a multiexponential curve (Donahue et al., 1997). This measurement necessitates the use of a water exchange model. MacKay et al. (1994) have introduced a multiexponential T2 (MET2) measurement to probe the proton spins from water molecules related to the myelin (MacKay, Whittall, et al., 1994). Their implementation makes use of a multi-echo spin echo sequence with regularly interspaced refocusing pulses, spanning a wide range of echo times within a long TR to minimize T1 effects. In the CNS, a three pool model can be used to classify compartments with respect to T2 values where short T2 compartment (10-40 ms) stands for water caged between glial cell membranes of myelin sheath, intermediate T2 compartment (80ms) describes intra- and extracellular water and longer T2 values are ascribed to the CSF (MacKay and Laule, 2016). Ratio of the signal from the first compartment to the total signal is described as myelin water fraction (MWF) which constitutes an indirect, yet a reliable marker of myelin (Laule, Kozlowski, et al., 2008). This metric has found widespread utilization cases such as development (Melbourne et al., 2013), normal appearing white matter in multipl sclerosis (MS) (Laule, Vavasour, et al., 2004) and chronic ischemic stroke (Borich et al., 2013). One of the drawbacks of this method comes from the assumption that there is not water exchange before the signal decay completes. Simulations have shown that MWF may be underestimated due to exchange effects (Levesque and Pike, 2009a). The other drawback with multiexponential T2 method is the requirement of lengthy acquisition times for a single slice observation. Deoni et al. (2008) have suggested a more efficient MWF mapping approach, namely as mcDESPOT, that enables whole brain coverage with higher spatial resolution (Deoni, Rutt, and Jones, 2008). Although MWF parameter as estimated by mcDESPOT has been validated on a demyelinated animal model, concerns exist that computational errors due to the complexity of the data fitting may swamp over the detection of meaningful microstructural changes (Lankford and Does, 2013).

2.2.3   Irradiate MRI-invisible macromolecules: Magnetization transfer

Macromolecular T2 values are in the range of microseconds, which renders impossible the direct measurement of signal from these sources (Henkelman, Stanisz, et al., 2001). Protons in close association with the surface of these macromolecules are defined as the restricted pool, which can be saturated by off-resonance excitation pulses of broad bandwidth. On the other hand, conventional MRI images are reconstructed from the signals received from protons that are not bound to macromolecules, constituting the free pool. Restricted pool transmits the RF energy deposited upon itself to the free pool to reach equilibrium state. This physical phenomenon is known as magnetization transfer and leads to an attenuation in the water signal measured by conventional MRI (Wolff and Balaban, 1989b). Given that myelin constitutes a substantial portion of the macromolecular population in brain (Bjarnason et al., 2005), MT can be exploited to provide an indirect measure of myelin content. The most practical application is MT ratio (MTR), which is a semi-quantitative measure calculated by normalizing an image acquired with MT effect to its conventional counterpart (Fralix et al., 1991). Although MTR maps show promising sensitivity to alterations in myelin content under pathological conditions (Giacomini and Arnold, 2008), measurements can be confounded by the changes in tissue relaxivity measures due to modifications in bulk water content caused by edema or inflammation (Gareau et al., 2000). Moreover, sequence and scanner dependencies of the MT pulses can hamper the reproducibility of MTR measurement in multicenter and longitudinal studies. To tackle with this problem, Brown et al. (2011) have suggested a statistical normalization that makes use of a reference scan from a healthy person at a given site (Brown et al., 2011). Although such approaches can provide practical solutions, a systematic reduction in the site dependency of MT imaging can be achieved by quantitative modeling of the MT parameters. Multiple MT acquisitions with varying off-resonance frequencies and RF amplitudes permits fitting resultant quantitative MT (qMT) data to a MT model (Henkelman, Huang, et al., 1993). Along with several other parameters decoupling the magnetization exchange rate and relaxation properties, qMT provides a parameter that is particularly linked to myelin (Tozer et al., 2003), which is pool size ratio (F) that defines the ratio of which MR-visible protons reside in the macromolecules. Whole brain coverage can be obtained by utilizing a qMT acquisition with moderate spatial resolution. Several approaches exist to acquire a full qMT data. One of these approaches make use of off-resonance saturation pulses (qMT-SPGR) with a dependency on field and T1 corrections offers short acquisition times at expense of a complex post processing (Sled and Pike, 2000). Confining image to a single slice and at the cost of a longer acquisition, qMT data can be collected in absence of an MT pulse and supplementary acquisitions. In this approach developed by Ou et al. (2007), restricted pool is saturated at part through the

employment of selectively inverting the free pool (qMT-SIRFSE) (Ou and Gochberg, 2008). Lastly, taking advantage of the inherent MT weightning of the steady-state free precession (bSSFP) sequence, Gloor et al. (2008) have developed a qMT method (Gloor et al., 2008). Although can be collected within clinically feasible acquisition times and at high temporal resolution with high SNR, qMT-bSSFP can further benefit from corrections for transmit field and relaxation properties. Overall, qMT methods provide comprehensive information about the mechanisms involved in MT interactions. It comes with the advantages of insensitivity to water exchange (Levesque and Pike, 2009b) and across-site variabilities. Nevertheless, parameters exhibit changes based on the selection of the model and implementation (Campbell et al., 2018).

It has been shown that F shows a marked non-linear association with MTR, whereas displaying a high linear correlation (r = 0.80) with MT saturation index (MTsat), which reduces the effect of T1 from MTR (Campbell et al., 2018). Although MTsat is also a semi-quantitative measure, it substantially reduces some of the residual contrast attained by MTR, yielding an easy to acquire and a promising parameter for myelin mapping (Helms, Dathe, Kallenberg, et al., 2008).

Quantitative mapping of microstructural parameters has opened a venue to combine multiple biophysical models, bringing additive value useful to a more specific characterization of microstructure. An outstanding example is g-ratio weighted imaging (Stikov, Campbell, et al., 2015) that brings more than a century-old metric to the realm of in-vivo imaging (Donaldson and Hoke, 1905). The g-ratio is described as the ratio of the inner diameter of an axon to its outer diameter (Rushton, 1951), a physiologically highly relevant metric to assess nerve conductivity. The g-ratio model includes two quantitative metrics myelin volume fraction (MVF) and intra-axonal volume fraction (AVF) that can provide voxelwise aggregate measure of g-ratio, ceasing the need for explicit estimation of axon diameter and myelin sheath thickness in a micrometer-scale. Another metric that is useful for g-ratio estimation is the fiber volume fraction (FVF), the sum of AVF and MVF. So far, g-ratio weighted imaging has been studied in developing brain (Dean et al., 2016), healthy spinal cord (Duval, Smith, et al., 2018), MS lesions and ex-vivo corpus callosum (Stikov, Campbell, et al., 2015). Findings from these studies indicated considerable agreement between ex-vivo histology and in-vivo measurements, and promising sensitivity to characterize myelin alterations. However, there exist multiple ways to calculate metrics involved in g-ratio estimation. For example, MVF can be represented by MTsat (Campbell et al., 2018), qMT imaging (Cercignani et al., 2017), MTV (Duval, Smith, et al., 2018) and MWF (MacKay, Whittall, et al., 1994) and FVF can be incorporated into g-ratio formula using different diffusion models (Ellerbrock and Mohammadi, 2018). The existence of multiple options may result in a choice overload,

raising concerns about making an optimum metric and calibration procedure selection by taking accuracy and cost-efficiency into account (Campbell et al., 2018).

In conclusion, last decades have witnessed significant advances in the in-vivo myelin mapping methods using MRI. To keep up with these techniques evolving toward faster acquisitions, higher resolutions and increased sensitivity, establishing elaborate statistical testing frameworks is a critical step for the reliability assessments and between method comparisons.

2.2.4 Finding out the relevant variability in myelin imaging

Establishing tissue specific normative parameters lies at the heart of qMRI applications. However, assessment of the accuracy and precision of a biomarker cannot be undertaken before establishing reliability of the quantitative metrics involved. Localizing acquisition and modeling related variability sources, and investigating their impact on qMRI metrics with sound statistical approaches holds great importance in distilling essential information provided by quantitative maps. Therefore, a considerable amount of effort has also been made towards evaluating the reliability of a wide range of qMRI methods. Nonetheless, variabilities in the qMR measurements are not merely originated from acquisition and modeling. Outputs from data intensive scientific workflows are subjected to a higher level variability based on the diversity of computational methods

Sources of variability

A recent review has formulated essential problems that prevent admission of microstructural mapping to the clinical zone (Novikov et al., 2018). Although diffusion imaging based microstructural mapping was chosen as the central subject of discussion, it is reasonable to generalize concepts presented in the paper by Novikov et al. to nearly all qMRI based methods. From this window, quantitative parameter extraction from any given qMRI dataset is based either on a mathematical approximation to the acquired signal or on a genuine biophysical model that roots from a solid theoretical background. Drawing the line between a model and a representation has critical implications for the validation of these methods, as only models can be validated (Novikov et al., 2018). Another outstanding point that merits distinctive attention is the selection of relevant degrees of freedom in a model. Comparison of MET2 and steady-state multicomponent relaxometry (mcDESPOT) based MWF estimation approaches constitute a good example to this point. Under the non-exchanging water assumption, MET2 fitting gives a reliable estimation of myelin water (Laule, Kozlowski, et al., 2008). Although it permits faster and volumetric acquisition, unreliability in the fitting due to high complexity of mcDESPOT model may shadow the expressive changes in MWF

(Lankford and Does, 2013). This example puts an emphasis on the identification of variability sources. In a qMRI study, variations can conceivably originate from differences in i) inherent measurement errors, ii) modeling, iii) acquisition sequence and parameters, iv) hardware and v) parameter estimation. Although the foremost can be characterized by statistical approaches, there is no straightforward way to analyze the effect of the remaining variability sources. Nonetheless, they can be mitigated to some extent by developing standardization strategies.

Statistical measures of reliability

Measurement of a continuous variability is inevitably subjected to an error due to inherent precision limitations of the measurement device. This type of random errors are easy to detect and their confounding influence can simply be minimized by averaging. Systematic errors, on the other hand, are persistent inaccuracies that cause consistent over- or underestimation of a reference measure. Therefore establishing validity hinges on the presence of a standard measure that agrees well with the measurement under question. Then one can make a decision on which measurement method to use by comparing their cost effectiveness (Bartlett and Frost, 2008). However this is not feasible for myelin imaging due to lack of a standard measure. Although histological metrics can be regarded as a reference measure for comparison, they cannot be qualified as gold standard due to variations between different staining methods and imperfect sample preparation (Shepherd et al., 2009). Therefore, microstructural mapping validations have begun to lean primarily towards electron microscopy scans (Jespersen et al., 2010). This is expected to become a more prevalent choice thanks to the freely accessible state-of-the-art segmentation methods powered by deep learning (Zaimi et al., 2018). Nonetheless, when true values are unknown, which is the most common case under in-vivo conditions, reliability becomes a relevant statistic to provide a basis to assess degree of agreement between methods.

If a method is highly reliable, measurement error of that method is said to be substantially lower than the inherent variability of the measured construct (Bartlett and Frost, 2008). Reliability is defined by the ratio of variability between subjects to the total variability, known as intraclass correlation coefficient (ICC) (McGraw and Wong, 1996). ICC depends on the heterogeneity of the population, therefore it should be reported along with descriptive statistics (Bartlett and Frost, 2008). ICC can be used in the intermediary step of the calculation of minimum detectable change (de Vet et al., 2006), which sets an acceptance threshold for changes in the measured parameter. ICC and its derivative reliability parameters are of critical importance in clinical studies to have a high likelihood of finding a significant pattern in

the data, if such trend is actually present. However, inconsistent use of statistical terminology can hamper examination of quantitative imaging biomarkers (Kessler et al., 2015). For example, reliability is often taken synonymous with repeatability (Kim, 2013) . Ability of a method to provide stable and consistent measurements can be an outcome of high reliability. Repeatability test is a specific type of reliability assessment, which can be performed only if the testing conditions are kept identical over more than one measurements (Taylor and Kuyatt, 1994). In this case, standard deviation of the differences between two test sessions multiplied by z-score for a given confidence level provides a relevant repeatability coefficient. If multiple test-retest measurements are acquired on a series of subjects from a single grouping condition, Bland-Altman plots can be used to characterize repeatability (Bland and Altman, 1986). This provides advantages over a single subject measurement, because proportional errors can be observed. For more than two repeat measurements, extended Bland-Altman plots can be used (Jones et al., 2011).

Levy et al. (2018) have studied the reliability of myelin imaging using statistical methods briefly described in the previous paragraph and noted the importance of using minimum detectable change over coefficient of variation to compare individual measurements (Lévy et al., 2018). Their main finding was that relaxometry and PD based myelin mapping techniques show higher reliability than semi-quantitative MT metrics in human spinal cord WM regions. Besides from high reliability across different sites, acquisition protocols and measurements, these metrics have also shown good agreement with histology in brain tissue (Mezer et al., 2013). Whole brain MWF based myelin mapping performed by a novel sequence with improved B1 uniformity showed good within-site reproducibility (Nguyen et al., 2016). Smith et al. (2014) have shown that test-retest relaibility of MTR is lower than that of F as quantified by qMT data with a single RF offset (Smith, Dortch, et al., 2014). The same group have studied reproducibility of MT in the different tracts of the spinal cord and concluded that tract specific reproducibility of MT is promising for column-specific prognosis.

Recent statistical approaches to assess reliability

Statistical analyses of studies reviewed above had intra-class comparison methods in common rather than inter-class correlation analysis for paired observations. This in good agreement with the suggestions by Bland and Altman (1986) for comparing clinical measurement methods (Bland and Altman, 1986). Nonetheless, reliability and reproducibility analyses of myelin imaging methods can benefit further from more recently developed statistical tools (Wilcox and Rousselet, 2018). For example, relationship between test-retest sessions can be assessed using shift function analysis for pairwise differences in a single subject reliability study (Rous-

selet et al., 2017). This becomes particularly useful when deciles of shift function analysis are mapped back to the image space, enabling a visual inspection to see if they correspond to known structures. Moreover, test-retest agreement of different methods as determined by Pearson and Concordance analyses can be compared using percentile bootstrapping to infer statistical significance. Probability masks can also be used to generate correlation curves to account for the effect of partial voluming on repeatability. This enables setting a threshold level to establish more robust reliability testing frameworks for certain quantitative biomarkers in certain tissue types. The added value of a similar approach has been recently shown by Ellerbrock et al. (2018), indicating that repeatability of g-ratio increases as threshold level is increased (Ellerbrock and Mohammadi, 2018).

CHAPTER 3   METHODOLOGY

Following the introduction of MRI fundamentals, the literature review presented in Chapter 2 identifies three major issues that hamper the reliability and accessibility of qMRI: i) the in-house nature of the developed qMRI methods, ii) the lack of a data standard to establish interoperable workflows with unified data access and iii) the hidden implementation details of vendor pulse sequences. To bring quantitative MRI under one umbrella, this project seeks solutions to each one of these problems with a particular focus on myelin imaging. Therefore, Chapter 2 also includes a review on several myelin imaging methods and introduces statistical analysis tools to capture systematic variabilities in the context of qMRI. The research objectives determined based on the problems (i), (ii) and (iii) are listed in the following section.

3.1   Research objectives

This thesis has two main objectives:

1. O1. **Develop open-source software that can implement, simulate and compare a wide range of qMRI methods with the help of a user-friendly graphical and command line interface:** The codebase adopts an object-oriented architecture to ease adding new qMRI methods. In addition, the development ecosystem is equipped with a continuous integration and delivery pipeline that allows deploying the software on high performance computing clusters and cloud computing infrastructures, ensuring consistent functionality. This will facilitate online and offline reproducibility of qMRI data processing and provide the users with an array of qMRI methods that are easy to reach and implement in a standardized fashion.

2. O2. **Extend the brain imaging data structure (BIDS) to implement findability, accessibility, interoperability and reusability (FAIR) principles for qMRI data and metadata:** The developed data standard provides a unified convention for qMRI analysis software to infer the processing tasks from the file collection descriptions. In addition, a set of requirements will be determined to standardize the units and identification of the quantitative maps, the input file collections, and the name/type/unit definitions of qMRI-relevant metadata for dataset validation.

3. O3. **Develop a vendor-neutral pulse sequence for myelin imaging and deploy it on multiple scanners from different manufacturers to test the hypothesis**

that vendor-neutral acquisitions reduce inter-vendor variability of qMRI:
The pulse sequence is developed as an RTHawk application and exports the data according to the standards established by O2. This will enable streamlining a workflow through seamless integration of the software developed in O1 with the data acquisition platform. The acquisitions will be controlled through a unified user interface and the workflow will be triggered based on the standardized description of the exported data. The hypothesis will be tested by imaging healthy subjects and a reference qMRI phantom.

3.2   Plan of the thesis

Chapter 4 achieves O1, presenting a software package developed for qMRI data analysis and simulation: qMRLab (https://qmrlab.org). With qMRLab, we provide an open-source, MATLAB-based (Octave compatible), object-oriented software package with separate modules for each qMRI technique. It is envisioned as a standard platform with a growing list of contributors, where the qMR community can replicate and cross-validate a wide range of qMR methods. qMRLab includes a user-friendly graphical user interface with a region-of-interest toolbox for summary statistics, online executable user documentations with example datasets and a comprehensive developer documentation. The software can be used to fit and check the quality of qMR data, to optimize protocols, compare fitting models, and simulate the effects of various model assumptions.

Chapter 5 achieves O2, presenting a global data standard developed for quantitative MRI (qMRI-BIDS), which has been a part of the main BIDS specification (https://bids-specification. readthedocs.io/) since the release v1.5.0. The standard covers a wide range of qMRI applications under the anatomical (or structural) imaging and field mapping categories. One of the key contributions of qMRI-BIDS to the main specification is the development of a new common principle named file collections, which defines a protocol-based grouping logic for multiple MR acquisitions through the use of file naming entities. For example, the file collection of variable flip angle (VFA) groups together multiple images acquired at different flip angles; therefore the use of the flip entity is required to distinguish each individual component. In addition to describing a convention for standardizing inputs, qMRI-BIDS establishes a comprehensive set of requirements to ensure that quantitative map units are standardized and includes sufficient metadata for provenance recording. With this study, FAIR principles of scientific data have been enabled for qMRI studies.

Chapter 6 achieves O3, presenting the development of a vendor-neutral sequence (VENUS) for acquiring myelin sensitive maps of T1, magnetization transfer and saturation (MTR and

MTsat) to test the hypothesis that that vendor-neutral sequences reduce inter-vendor variability of T1, MTR and MTsat maps (https://github.com/qMRLab/venus). In this study, we run a multicenter experiment across three different scanner models from two manufacturers at 3 tesla. Vendor-native counterparts of the VENUS protocol were also acquired on each scanner to compare their multicenter reproducibility in three healthy participants and in a qMRI phantom. This work demonstrated that VENUS coupled with fully-transparent qMRLab workflows can significantly reduce qMRI variability across vendors.

CHAPTER 4 ARTICLE 1: QMRLAB: QUANTITATIVE MRI ANALYSIS, UNDER ONE UMBRELLA

Authors: A. Karakuzu[1,4], M. Boudreau[1,4], T. Duval[1], T. Boshkovski[1], I.R. Leppert[2], J.F. Cabana[6], I. Gagnon[1], P. Beliveau[7], G.B. Pike[2,5], J. Cohen-Adad[1,3], N. Stikov[1,4]

Affiliations:

1. NeuroPoly Lab, Institute of Biomedical Engineering, Polytechnique Montreal, Montreal, Canada

2. McConnell Brain Imaging Center, Montreal Neurological Institute, McGill University, Montreal, Canada

3. Functional Neuroimaging Unit, CRIUGM, University of Montreal, Montreal, Canada

4. Montreal Heart Institute, University of Montréal, Montréal, Canada

5. Departments of Radiology and Clinical Neuroscience, Hotchkiss Brain Institute, University of Calgary, Calgary, Canada

6. Chaudière-Appalaches Integrated Health and Social Services Center, Sainte-Marie, Canada

7. Department of radio-oncology, Research Center of the University of Montreal Hospital Center, Montreal, Canada

4.1 Summary

Magnetic resonance imaging (MRI) has revolutionized the way we look at the human body. However, conventional MR scanners are not measurement devices. They produce digital images represented by "shades of grey", and the intensity of the shades depends on the way the images are acquired. This is why it is difficult to compare images acquired at different clinical sites, limiting the diagnostic, prognostic, and scientific potential of the technology.

Quantitative MRI (qMRI) aims to overcome this problem by assigning units to MR images, ensuring that the values represent a measurable quantity that can be reproduced within and across sites. While the vision for quantitative MRI is to overcome site-dependent variations, this is still a challenge due to variability in the hardware and software used by MR vendors to produce quantitative MRI maps.

Although qMRI has yet to enter mainstream clinical use, imaging scientists see great promise in the technique's potential to characterize tissue microstructure. However, most qMRI tools for fundamental research are developed in-house and are difficult to port across sites, which in turn hampers their standardization, reproducibility, and widespread adoption.

4.2 Statement of need

To tackle the problem described above, we developed qMRLab, an open-source software package that provides a wide selection of qMRI methods for data fitting, simulation and protocol optimization (Figure 4.1). It not only brings qMRI under one umbrella, but also facilitates its use through https://qmrlab.readthedocs.io/en/master/ that features online executable notebooks, a user friendly graphical user interface (GUI), interactive tutorials and blog posts.

MATLAB is the native development language of qMRLab, primarily because it is by far the most common choice among MRI methods developers. However, we have made a strong effort to lower licensing and accessibility barriers by supporting Octave compatibility and Docker containerization.

qMRLab started as a spin-off project of qMTLab (Cabana et al., 2015). In the meantime, a few other open-source software packages were developed, addressing the lack of qMRI consistency from different angles. QUIT (Wood, 2018) implemented an array of qMRI methods in C++, which is highly favorable as an on-site solution because of its speed. The hMRI toolbox (Tabelow et al., 2019) was developed as an SPM (Ashburner, 2012) module that expands on the multi-parametric mapping method (Weiskopf and Helms, 2008). Other tools such as

Figure 4.1 qMRLab is an open-source software for quantitative MRI analysis It provides a myriad of methods to characterize microstructural tissue properties, from relaxometry to magnetization transfer.

mrQ (Mezer et al., 2013) and QMAP (Alexander et al., 2011) are also primarily designed for brain imaging. In addition to the arrival of community-developed tools for multi-modal qMRI processing, the field of diffusion MRI (dMRI) has recently witnessed an increase in the development of open-source software, bringing more transparency to the diffusion-based microstructural characterization of the brain. For example, the Dmipy Toolbox implemented an array of multi-compartment diffusion models in Python (Fick et al., 2019), another Python package DIPY brought together dMRI processing methods at multiple levels (Garyfallidis et al., 2014), Camino was developed in Java to provide users with yet another dMRI pipeline (Cook et al., 2006) and TractoFlow (Theaud et al., 2020) introduced container-based reproducible dMRI workflows developed in Nextflow pipeline orchestration tool (Di Tommaso et al., 2017). Yet, brain imaging is not the only qMRI area slowed down by lack of consistency. Recently we published a preprint demonstrating notable disagreements between

cardiac qMRI methods (Hafyane et al., 2018). Open-source software can go a long way in explaining these discrepancies, and the cardiac imaging community was recently introduced to TOMATO (Werys et al., 2020), an open C++ framework for parametric cardiac MRI.

As open-source practices in the realm of qMRI become more popular, the need for effective communication of these tools also increases. This is important not only because we need consistency and transparency in the implementations, but also because non-specialist qMRI users would benefit from better understanding of the methodology. To this end, we envision qMRLab as a powerful tool with which users can easily interact with various techniques, perform simulations, design their experiments and fit their data. We reinforce this vision through our web portal (https://qmrlab.org) that includes interactive tutorials, blog posts and Jupyter Notebooks running on BinderHub, all tailored to a wide range of qMRI methods. The qMRLab portal is open for community contributions.

Currently, qMRLab is used by dozens of research labs around the world, mostly, but not limited to, application in brain and spinal cord imaging. A list of published studies using qMRLab is available on our GitHub repository.

4.3   Conclusion

While closed solutions may be sufficient for qualitative MRI (shades of grey lack standardized units), quantitative MRI will not realize its potential if we cannot peek inside the black box that generates the numbers. With qMRLab we want to open the black boxes developed in-house and reach a critical mass of users across all MR vendor platforms, while also encouraging developers to contribute to a central repository where all features and bugs are in the open. We hope that this concept will level the field for MR quantification and open the door to vendor-neutrality. We've been sitting in our MR cathedrals long enough. It is now time to join the MR bazaar (Raymond, 1999)!

4.4   Code repository, archive and the review

The latest versoin of the codebase of qMRLab is publicly available on GitHub and the release (v2.4.1) associated with this publication is archived at https://zenodo.org/record/4012104. The Journal of Open Source Software (JOSS) reviews are publicly available. The review for this publication can be accessed https://github.com/openjournals/joss-reviews/issues/2343.

4.5   Supplementary materials

A more detailed description of qMRLab modules is provided in Appendix A.

4.6   Acknowledgements

4.7   References

Alexander, A. L., Hurley, S. A., Samsonov, A. A., Adluru, N., Hosseinbor, A. P., Mossahebi, P., Tromp, D. P., Zakszewski, E., & Field, A. S. (2011). Characterization of cerebral white matter properties using quantitative magnetic resonance imaging stains. *Brain connectivity*, *1*(6), 423–446.

Ashburner, J. (2012). Spm: a history. *Neuroimage*, *62*(2), 791–800.

Cabana, J.-F., Gu, Y., Boudreau, M., Levesque, I. R., Atchia, Y., Sled, J. G., Narayanan, S., Arnold, D. L., Pike, G. B., & Cohen-Adad, J. (2015). Quantitative magnetization transfer imaging made easy with q mtl ab: software for data simulation, analysis, and visualization. *Concepts in Magnetic Resonance Part A*, *44*(5), 263–277.

Di Tommaso, P., Chatzou, M., Floden, E. W., Barja, P. P., Palumbo, E., & Notredame, C. (2017). Nextflow enables reproducible computational workflows. *Nature biotechnology*, *35*(4), 316–319.

Fick, R. H., Wassermann, D., & Deriche, R. (2019). The dmipy toolbox: diffusion mri multi-compartment modeling and microstructure recovery made easy. *Frontiers in neuroinformatics*, *13*, 64.

Garyfallidis, E., Brett, M., Amirbekian, B., Rokem, A., Van Der Walt, S., Descoteaux, M., & Nimmo-Smith, I. (2014). Dipy, a library for the analysis of diffusion mri data. *Frontiers in neuroinformatics*, *8*, 8.

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., & Halchenko, Y. O. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific data*, *3*(1), 1–9.

Hafyane, T., Karakuzu, A., Duquette, C., Mongeon, F.-P., Cohen-Adad, J., Jerosch-Herold, M., Friedrich, M. G., & Stikov, N. (2018). Let's talk about cardiac t1 mapping. *bioRxiv*, 343079.

Karakuzu, A., Boudreau, M., Duval, T., Boshkovski, T., Leppert, I., Cabana, J.-F., Gagnon, I., Beliveau, P., Pike, G. B., Cohen-Adad, J., & Stikov, N. (2020). Qmrlab: quantitative mri analysis, under one umbrella. *Journal of Open Source Software*, *5*(53), 2343.

Mezer, A., Yeatman, J. D., Stikov, N., Kay, K. N., Cho, N. J., Dougherty, R. F., Perry, M. L., Parvizi, J., Hua le, H., Butts-Pauly, K., & Wandell, B. A. (2013). Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nat Med*, *19*(12), 1667–72. https://doi.org/10.1038/nm.3390

Raymond, E. (1999). The cathedral and the bazaar. *Knowledge, Technology & Policy*, *12*(3), 23–49.

Tabelow, K., Balteau, E., Ashburner, J., Callaghan, M. F., Draganski, B., Helms, G., Kherif, F., Leutritz, T., Lutti, A., & Phillips, C. (2019). Hmri–a toolbox for quantitative mri in neuroscience and clinical research. *Neuroimage*, *194*, 191–210.

Theaud, G., Houde, J.-C., Boré, A., Rheault, F., Morency, F., & Descoteaux, M. (2020). Tractoflow: a robust, efficient and reproducible diffusion mri pipeline leveraging nextflow & singularity. *NeuroImage*, *218*, 116889.

Weiskopf, N., & Helms, G. (2008). Multi-parameter mapping of the human brain at 1mm resolution in less than 20 minutes. *Proceedings of ISMRM*, *16*, 2241.

Werys, K., Dragonu, I., Zhang, Q., Popescu, I., Hann, E., Puchta, H., Kubik, A., Polat, D., Wu, C., & Moon, N. O. (2020). Total mapping toolbox (tomato): an open source library for cardiac magnetic resonance parametric mapping. *SoftwareX*, *11*, 100369.

Wood, T. (2018). Quit: quantitative imaging tools. *Journal of Open Source Software*, *3*(26), 656.

CHAPTER 5   ARTICLE 2: QMRI-BIDS: AN EXTENSION TO THE BRAIN IMAGING DATA STRUCTURE FOR QUANTITATIVE MAGNETIC RESONANCE IMAGING DATA

Authors: A. Karakuzu[1,2], S. Appelhoff[3], T. Auer[4], M. Boudreau [1,2], F. Feingold[5], A.R. Khan[6], A. Lazari[7], C. Markiewicz[5], M.J. Mulder[8], C. Phillips[9], T. Salo[10], N. Stikov[1,2], K. Whitaker[12], G. Hollander[12,13]

Affiliations:

1. Neuropoly Lab, Polytechnique Montreal, Montreal, Canada

2. Montreal Heart Institute, Montreal, Canada

3. Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

4. NeuroModulation Lab, School of Psychology, University of Surrey, Guildford, UK

5. Stanford University, Stanford, CA, USA

6. Department of Medical Biophysics, Robarts Research Institute, University of Western Ontario, London, Canada.

7. Wellcome Centre for Integrative Neuroimaging, FMRIB, Nuffield Department of Clinical Neurosciences, University of Oxford

8. Department of Experimental Psychology, Utrecht University, Utrecht, the Netherlands

9. GIGA Cyclotron Research Centre in vivo imaging, GIGA Institute, University of Liege, Liege, Belgium

10. Florida International University, Miami, FL, USA

11. The Alan Turing Institute, London, UK

12. Zurich Center for Neuroeconomics (ZNE), Department of Economics, University of Zurich, Zurich, Switzerland

13. Spinoza Centre for Neuroimaging, Amsterdam, The Netherlands

5.1 Abstract

The Brain Imaging Data Structure (BIDS) established community consensus on the organization of data and metadata for several neuroimaging modalities. Traditionally, BIDS had a strong focus on functional magnetic resonance imaging (MRI) datasets and lacked guidance on how to store multimodal structural MRI datasets. Here, we present and describe the BIDS Extension Proposal 001 (BEP001), which adds a range of quantitative MRI (qMRI) applications to the BIDS. In general, the aim of qMRI is to characterize brain microstructure by quantifying the physical MR parameters of the tissue via computational, biophysical models. By proposing this new standard, we envision standardization of qMRI which makes multi-center dissemination of interoperable data possible. As a result, BIDS can act as a catalyst of convergence between qMRI methods development and application-driven neuroimaging studies that can help develop quantitative biomarkers for neural tissue characterization. Finally, our BIDS extension offers a common ground for developers to exchange novel imaging data and tools, reducing the practical barriers to standardization that is currently lacking in the field of neuroimaging.

5.2 Introduction

The brain imaging data standard (BIDS) is an open source initiative from the neuroimaging community that aids in standardizing neuroimaging data sets. BIDS was originally developed with functional MRI (fMRI) applications in mind, describing experimental task blocks in relation to a hierarchical organization of reconstructed MR images (Gorgolewski, Auer, et al., 2016). This convention engaged researchers to share hundreds of open fMRI data on the openneuro platform (Markiewicz et al., 2021; Poldrack et al., 2013) and develop interoperable processing workflows that can seamlessly process these datasets (Gorgolewski, Alfaro-Almagro, et al., 2017). Popular examples include the MRIQC (Esteban, Birman, et al., 2017) and fmriprep (Esteban, Markiewicz, et al., 2019) pipelines, which can be executed even online for any valid BIDS fMRI dataset. Similarly, the development of an MRI k-space data standard, ISMRM-RD (Inati et al., 2017), led open-source MRI reconstruction packages to adapt this convention and now aids potential users in performing advanced reconstruction tasks with minimal effort (Hansen and Sørensen, 2013; Maier, Baete, et al., 2021). These success stories from open science exemplify how data standards can change the landscape of community-driven software for the better, leading to a collective change in researchers' behaviour to adhere with FAIR (findability, accessibility, interoperability and reusability) principles of scientific data (Wilkinson et al., 2016). Here we present our work extending the

BIDS to include multi-contrast MRI acquisitions. BIDS Extension Proposal 001 (BEP001) was merged into the standard (on 23 February 2021) and focuses on quantitative MRI (qMRI) applications.

Quantitative MRI methods map physical magnetic properties of the (brain) tissue. Their application consists of two steps: i) collecting multiple MRI images, where the contributions of effective micrometer-level MRI parameters is systematically manipulated by adapting very specific acquisition parameters, and ii) fitting the resultant voxel intensity variations across the images to a computational (biophysical) model (Novikov et al., 2018). The results are a single or multiple quantitative map of the estimated parameters across the imaged volume. The effective MRI parameters that are typically studied include longitudinal and transverse relaxation time constants (T1 and T2, respectively), proton density (PD), magnetization transfer (MT), and local diffusion coefficient (e.g., fractional anisotropy, FA, or mean diffusivity, MD). The multi-parametric mapping (Weiskopf, Suckling, et al., 2013) (MPM) protocol offers a set of acquisitions that can quantify more than one MR parameter at a time. Another popular technique used in qMRI is field mapping, which characterizes inhomogeneities in MRI radiofrequency (RF) transmit (B1+) and receive (B1-) profiles, as well as static magnetic field (B0) to correct qMRI parameter estimation errors due to these field inhomogeneities.

The earliest qMRI applications date back to the late 70's (Gupta, 1977) and primarily focused on relaxometry, mapping of quantities such as T1 and T2* relaxation time. Since then, the field has witnessed multiple waves of methods development, driven by technological advances and emerging trends in MRI research (Lundervold and Lundervold, 2019; Stikov, Trzasko, et al., 2019). Recently, with the surge of deep learning methods, the gamut of parameter estimation methods have become much larger than ever before (Golkov et al., 2016; Lundervold and Lundervold, 2019; Lyu and Wang, 2018; Wu et al., 2020; Yoo et al., 2018). Interestingly, however, we still do not precisely know the healthy range of relaxation time values in a multicenter setting (Bojorquez et al., 2017) nor do we know how to establish diagnostically reliable tissue typing protocols. This discrepancy highlights that multicenter standardization should be a critical step toward evaluating the clinical potential of decades-long improvements in the acquisition and processing of qMRI data.

Under more controlled research settings, qMRI offers obvious advantages over conventional MRI contrasts (e.g. T1 weighted images) in structural feature extraction. Given that MRI is not a direct measurement of in vivo anatomical structures, voxel-wise morphometry analyses are subjected to various biochemical and physiological confounders affecting the voxel intensity (Weinberger and Radulescu, 2015).Hence, the capacity of disentangling MRI signal

components lands qMRI as a more reliable approach to study structural variations (Lorio et al., 2016). This makes qMRI particularly useful for comparisons of the brain anatomy of different (clinical) groups (Draganski et al., 2011; Lommers et al., 2021; Weiskopf, Mohammadi, et al., 2015) and for more consistent, unbiased automated anatomical segmentation (Dinse et al., 2015; Haast et al., 2016; Lutti et al., 2014; Weiskopf, Callaghan, et al., 2014). The same principle can be exploited to make qMRI sensitive to tissue microstructure, such as iron concentration or myelination. Recent meta analyses revealed that a majority of qMRI methods are comparably sensitive to the myelin content (Lazari and Lipp, 2021; Mancini et al., 2020), although certain parameters such as myelin water fraction (MWF, relaxometry-based) and macromolecular pool fraction (MPF, MT-based) appear to be more specific.

Given the advantages offered by parametric maps in providing structural information and the current landscape of myelin imaging methods, it seems likely that many more myelin imaging methods leveraging the potential of qMRI will be developed in the future. This leads to one of our four main motivations behind covering qMRI methods in BIDS: to bring FAIR principles to a variety of qMRI data that are finding widespread use in neuroimaging research. Other motivations include i) driving open-source qMRI tools to adapt a consolidated input/output convention, ii) creating standardized databases that can help simplify the use of qMRI in clinical and translational research, and iii) stimulating an open provision of qMRI data that can be collected by imaging equipment that is available to a small group of researchers.

Drawing upon the principles outlined in BIDS, we introduce the first consensus data and metadata organization standard for qMRI. This work is a culmination of years of effort (the earliest drafts of the BEP-001 date back to 2017) and discussion between neuroimaging researchers and MRI methods developers around the globe. Our extension will not only aid in organizing qMRI data, but will also facilitate multi-center collaborative work, encourage neuroscientists to adapt advanced MR techniques and go a long way toward the standardization of qMRI methods.

5.3   Results

5.3.1   A new BIDS common principle: entity-linked file collections

The majority of qMRI methods necessitate the grouping of a set of similar images where specific acquisition parameters are carefully varied. Furthermore, the images that are collected for qMRI application do not usually have a clear "weighting" description (e.g., T1w, T2w), like conventional structural images. The novel concept of file collections decouples the semantics of logical group identification from contrast weighting labels or acquisition

sequence names that are not originally developed for qMRI (e.g., FLASH). Instead, suffixes for such logical units may indicate a generic MRI readout type (e.g., multi-echo gradient echo: MEGRE), a qMRI sequence name (e.g, magnetization prepared two rapid gradient echoes, MP2RAGE) or a qMRI data collection framework (e.g., variable flip angle, VFA). Table-1 lists file collection suffixes for various qMRI and fieldmap data, and the quantitative parameters they can derive. These suffixes span a wide range of qMRI applications including relaxometry, MT imaging, multiparametric mapping, and RF field mapping. Application scope can be extended without necessarily adding more suffixes. The BIDS qMRI appendix presents a set of rules and suggestions to add new qMRI suffixes to the specification (https://bids-specification.readthedocs.io/en/stable/99-appendices/11-qmri.html).

Note that the use of file collections is not exclusive to qMRI, anatomy imaging data, or even MRI. Any imaging modality calling for a file grouping logic to define a quantitative or qualitative application can benefit from this principle by specifying a descriptive suffix and filename entity. Such changes would require additional BIDS extensions to create a valid file collection.

Table 5.1 File collections of anatomy imaging data to derive parametric maps of longitudinal, transverse and observed-transverse relaxation times (T1, T2 and T2*, respectively), proton density (PD), magnetization transfer ratio and saturation index (MTR and MTsat) and myelin water fraction (MWF). Relaxation rates (e.g., $T1_{-1}$ and $T2_{-1}$) and residual terms (e.g., M0) are excluded from the table for brevity.

qMRI application	Suffix	Derived maps	BIDS folder	Reference
Magnetization prepared two rapid gradient echoes (MP2RAGE)	MP2RAGE	T1	anat	(Marques et al., 2010)
Multiparametric mapping (MPM)	MPM	T1, T2*, PD, MT	anat	(Weiskopf, Suckling, et al., 2013)
Variable flip angle (VFA)	VFA	T1, T2	anat	(Gupta, 1977)

Inversion recovery for T1 mapping (IRT1)	IRT1	T1	anat	(Barral et al., 2010)
Multi-echo spin-echo (MESE)	MESE	T2, MWF	anat	(Carr and Purcell, 1954), (MacKay, Whittall, et al., 1994)
Multi-echo gradient-echo (MEGRE)	MEGRE	T2*	anat	(Ma and Wehrli, 1996)
Magnetization transfer ratio (MTR)	MTR	MT%	anat	(Wolff and Balaban, 1989a)
Magnetization transfer saturation index (MTS)	MTS	MTsat	anat	(Helms, Dathe, Kallenberg, et al., 2008)
Double angle B1+ mapping	TB1DAM	B1+	fmap	(Insko and Bolinger, 1993)
B1+ mapping with 3D echo-planar imaging (EPI)	TB1EPI	B1+	fmap	(Jiru and Klose, 2006)
Actual flip angle imaging (AFI)	TB1AFI	B1+	fmap	(Yarnykh, 2007b)
Rapid B1+ mapping with TurboFLASH readout	TB1TFL	B1+	fmap	(Chung et al., 2010)

Saturation-prepared with 2 rapid gradient echoes (SA2RAGE)	TB1SRGE	B1+	fmap	(Eggenschwiler et al., 2012)
Inter-scan motion correction using receive field modulation	RB1COR	B1-	fmap	(Papp et al., 2016)

To distinguish individual files of a file collection, we introduced filename entities that are associated with commonly altered acquisition parameters (e.g., flip angle) or with inherent components of the same data (e.g., phase information), hence the name "entity-linked file collection" (Table 5.2).

Table 5.2 Filename entities representing an MRI acquisition parameter or designating an inherent part of the reconstructed image (e.g., magnitude or phase).

Entity forma	Entity values	Associated acquisition parameter	Associated qMRI file collections
echo-<index>	01,02,03,...,n	EchoTime	MEGRE, MESE, MPM
flip-<index>	01,02,03,...,n	FlipAngle	VFA, MTS, MPM
inv-<index>	01,02,03,...,n	InversionTime	IRT1, MP2RAGE
mt-<label>	on/off	MTState	MTR, MTS, MPM
part-<label>	mag/phase	N/A	MP2RAGE

It is important to highlight that these entities cannot store acquisition parameter values in the filename, but can only index or categorize them. Respective parameter values are stored in so-called "sidecar JSON"-files. Requirement level of these entities in relation to file collections are presented in the BIDS entity table appendix (https://bids-specification. readthedocs.io/en/stable/99-appendices/04-entity-table.html).

5.3.2 Data organization for qMRI file collections and quantitative parametric maps

By combining entities in the filename that represent different acquisition parameters (Table 5.2) with entity-linked file collection suffixes (Table 5.1), BEP001 provides an intuitive way to organize filenames of most existing qMRI data. For example, raw data from MP2RAGE acquisitions comprises both magnitude and phase reconstructed images, acquired at two successive inversion times (Figure 5.1a). The respective file collection for MP2RAGE (Figure 5.1c) clearly defines these components via part and inv-components, which are required for the MP2RAGE file collection. Note how the BIDS inheritance-rules do allow for using a single JSON-file to describe both phase- and magnitude-images, since these have identical acquisition parameters. In addition, the same collection suffix can be extended to specify its multi-echo variant (Caan et al., 2019) using the echo entity, which is made optional to MP2RAGE. For clarity, these specific use cases are defined in the BIDS qMRI appendix.

The same logic applies to the raw images of double-angle B1+ mapping, identified by the TB1DAM suffix (Figure 5.1c). In this case, the maximum value of the flip entity indicates that the data is collected over two flip angles. We recognize that an alternative approach to organize such data is stacking images at each flip angle into the 4th dimension of a Nifti-file, and storing the corresponding metadata in vector form using a single JSONfile. This approach offers a less crowded file list for this particular example. However, indexing acquisition parameter dependent variations across additional dimensions is less favorable for comprehensive qMRI methods. For example, MPM collects raw data at different echo times, flip angles and MT preparations with the option of phase reconstruction. After extended debates that took more than a year, the qMRI-BIDS extension group ultimately concluded that this approach is less favourable for human-readability of qMRI datasets, especially for multiparametric acquisition methods where the number of images per protocol can go into the dozens.
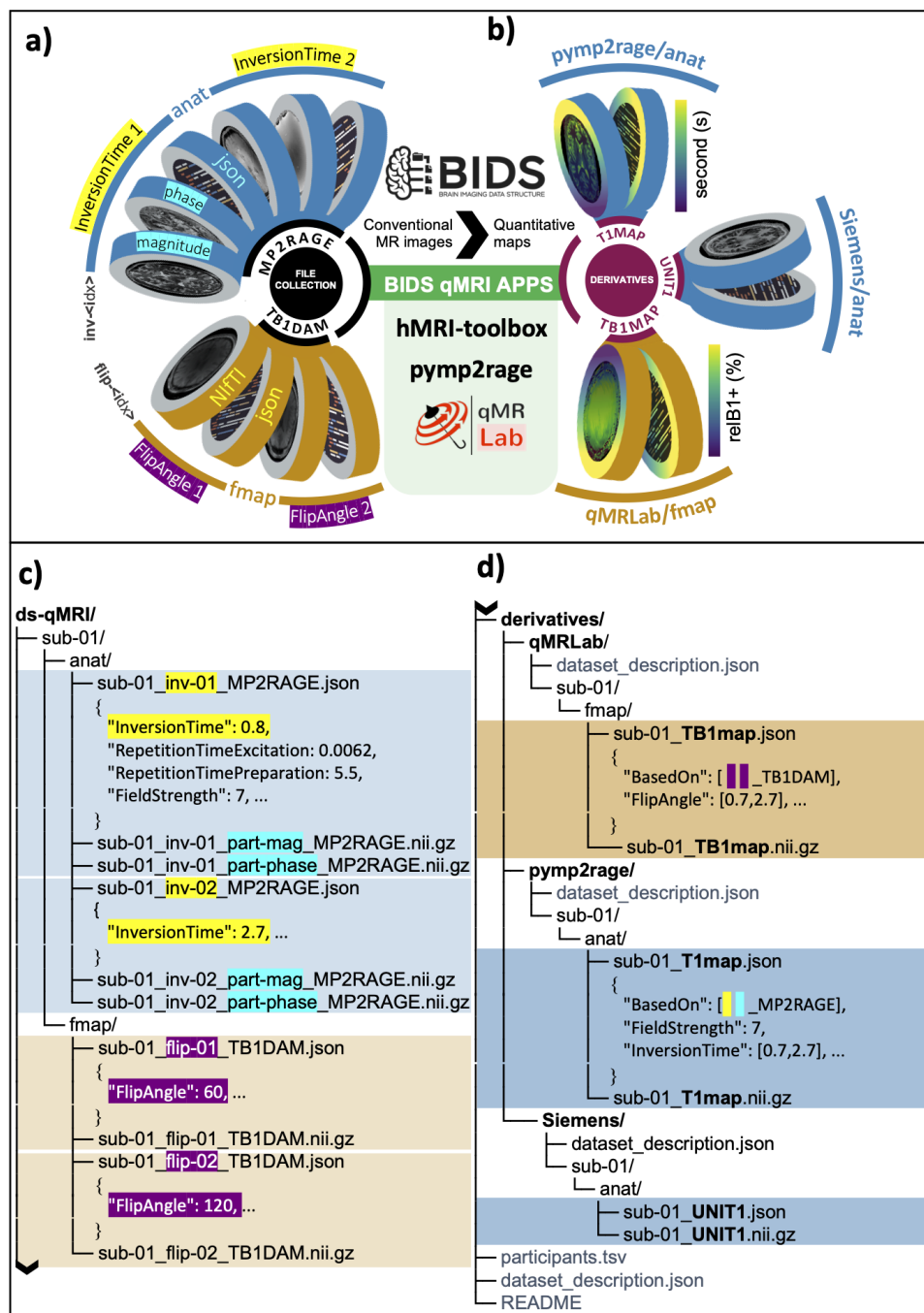
Figure 5.1 **a)** Schematic of BIDS formatted raw quantitative MRI (qMRI) data representing MP2RAGE (anat) and TB1DAM (fmap) file collections, for which entity-linked metadata fields are highlighted for the InversionTime (yellow), the FlipAngle (purple) and for the reconstructed image type (cyan). **b)** Derivatives of MP2RAGE and TB1DAM file collections generated by using pymp2rage and qMRLab to calculate T1 and B1+ maps, respectively, including a vendor-native derivative of UNIT1 images. **c)** File organization of raw qMRI data for MP2RAGE and TB1DAM file collections, where respective linking entities are highlighted for the inv entity (yellow, InversionTime), the flip entity (purple, FlipAngle) and the part entity (cyan, magnitude/phase). **d)** File organization of qMRI derivatives indicating how sidecar JSON files of quantitative maps generated by open-source software keeps a log of the input files (the BasedOn field) and associated acquisition parameters (FlipAngle in TB1map and InversionTime in B1map).

5.3.3 Metadata requirements for file collections and quantitative parametric maps

For the file collections, linking entities (Table 5.2) indicate a requirement for the respective acquisition parameters that are subject to change from image-to-image. Therefore, the entity table appendix lists such parameters as required in relation to the corresponding file collection suffix based on the descriptions made in the BIDS schema. Note that not all the parameters that change across file collection images are captured by a linking entity, but may still be required for data fitting. For example, the value of the *FlipAngle* parameter might (but does not necessarily) covary with that of *InversionTime* between MP2RAGE file pairs; however, the filenames are distinguished solely by the inv entity (since that is the crucial parameter that is swept over, whereas the flip angle could in principle remain the same). In addition, certain parameters that are constant across file collection images may be required as well. For example, *RepetitionTimeExcitation* and *RepetitionTimePreparation* are required metadata for an MP2RAGE acquisition. Such parameters are required when they are strictly necessary to calculate the qMRI-maps that a specific acquisition scheme was designed to obtain; e.g., a T1-map in case of MP2RAGE. BEP001 added an array of new metadata fields that may be required for certain file collections (e.g. *MTState*, specifying whether an MT preparation is enabled in an MPM acquisition, associated with the mt linking entity) or provide supporting information (e.g., *SpoilingRFPhaseIncrement*, specifying the amount of incrementation applied to the phase of an excitation pulse). The complete list of metadata fields and their requirement levels for all the qMRI file-collections are included in the BIDS release v1.5.0 and later. Currently, metadata conversions for some of these required fields have been implemented in dcm2niix (Li et al., 2016), a commonly used DICOM to NIfTI converter to create BIDS-compatible datasets.

Certain quantitative parameters cannot be interpreted in absence of fundamental scanner specifications. For example, to interpret relaxometry maps (e.g., T1map), the magnetic field strength must be known. The BEP001 ensures that such requirements are met (again, see the qMRI Appendix in BIDS release v1.5.0 and later). Moreover, sidecar JSON files of quantitative maps contain all the metadata values involved in the fitting by representing varying parameters in vector form and inheriting the constant ones from the raw images. To supplement the provenance recording of parameter estimation process with software-relevant details, the derived dataset and pipeline rules are respected as outlined in the modality agnostic files section of the main specification.

Finally, the units and range of the fitted parameters have been standardized by BEP001 to define interchangeable qMRI maps. For relaxometry-based parameters (e.g., T1map or

T2map), the time is described in seconds and the rate in reciprocal seconds or Hz. Wherever applicable, unitless ratio maps are described in percentage (e.g., MTRmap or MWFmap). For quantitative susceptibility maps (i.e., Chimap) the local magnetic susceptibility is represented in parts per million. The RF transmit maps (i.e., TB1map) are specified in relative percentage units, where 100% denotes the ideal case (i.e., measured flip angle equals the nominal value). Any deviations from 100% convey proportional deviations from the intended field strength. Please note that certain quantitative parameters are described in arbitrary units, where the acceptable range of values vary based on the target anatomy (e.g., MTsat).

5.3.4 Community software for qMRI-BIDS data acquisition, conversion and processing

As of release v1.5.0, the BIDS validator can perform on BEP001-compatible qMRI data at the directory and filename level rules, based on the entity requirement levels specified per file collection suffix. However, metadata-level validation rules have not been implemented yet. This is mainly because multi-vendor extraction of qMRI related metadata fields (e.g., *MTState* or *RepetitionTimePreparation*) is not supported by commonly used converters. Recently, we started working with dcm2niix (Li et al., 2016) and BIDSme (https://github.com/CyclotronResearchCentre/bidsme) developers to identify and map vendor-specific header information to BEP001-compatible metadata.

Nevertheless, some metadata entities that are of profound importance to the accuracy of quantitative maps cannot be typically found in the vendor-native DICOM headers. For example, the BIDS fields of *RFSpoilingPhaseIncrement* and *SpoilingGradientMoment* are two major determinants of T1 and B1+ estimation accuracy using spoiled gradient echo based applications (Yarnykh, 2010). Although this information is not provided by vendors, open-source pulse sequence development frameworks such as Pulseq (Layton et al., 2017), PyPulseq (Ravi, Geethanath, et al., 2019), Gammastar (Cordes et al., 2020), TOPPE (Nielsen and Noll, 2018), SequenceTree (Magland et al., 2016), ODIN (Jochimsen and Von Mengershausen, 2004) and RTHawk (Santos et al., 2004) can make a qMRI-tailored metadata annotation possible. An example implementation is qMRPullseq, a collection of publicly available vendor-neutral pulse sequences that are designed to export images in accordance with BEP001 format without hidden acquisition parameters (Karakuzu, Boudreau, Cohen-Adad, et al., 2020b). We highly encourage open-source MRI pulse sequence developers to use and contribute to the qMRI metadata annotations. This simple consensus can remove proprietary roadblocks from disseminating qMRI datasets that incorporate key information on the reproducibility of data acquisition.

Most qMRI methods can benefit from a plethora of BIDS applications (Gorgolewski, Alfaro-Almagro, et al., 2017) to prepare data for parameter estimation and downstream statistical analyses. There are several open-source tools emerging to perform qMRI fitting at multiple levels, like the hMRI-toolbox (Tabelow et al., 2019), qMRLab (Karakuzu, Boudreau, Duval, et al., 2020), QUIT (Wood, 2018), PyQMRI (Maier, Spann, et al., 2020), QMRTools (Froeling, 2019), mrQ (Mezer et al., 2013), Madym (Berks et al., 2021), MITK-ModelFit (Debus et al., 2019), ROCKETSHIP (Barnes et al., 2015), DCEMRI.jl (Smith, Li, et al., 2015) and DCE@urLAB (Ortuño et al., 2013). Giving these tools the ability to operate on BIDS formatted data is an important step towards establishing interoperable qMRI processing pipelines.

5.3.5   The role of BIDS in wider adoption and accessibility of quantitative MRI

Quantitative MRI offers a rapidly developing set of techniques that can inform us about brain (micro)structure beyond what conventional MRI techniques have to offer (Weiskopf, Edwards, et al., 2021). We believe that, in coming years, qMRI will become increasingly important to both clinical and fundamental brain science. Therefore, a concrete standard for organizing and thereby also disseminating open qMRI data sets is much warranted. BEP001 extends the framework of the existing and very successful BIDS standard, to develop a standard for qMRI in the form of a "BIDS extension proposal". To aid actual user adoption of this standard, it includes very precise descriptions of how to use it in many real-life qMRI use-cases, as well as many example data sets.

Currently, obtaining qMRI data is still expensive and needs considerable expertise, which is not readily available at many MRI facilities. Therefore, we also hope that BEP001 will aid researchers that do not have easy access to such facilities to get familiar with qMRI data and potentially can even use open qMRI data sets for their particular research questions.

Finally, the popularity of BIDS is likely in large part also due to some software packages that are designed around this standard and therefore extremely easy-to-use, when one's data adheres to the BIDS standard (Yarkoni et al., 2019). We hope that the success of BIDS in the domain of functional MRI will also inspire and encourage MRI software developers to work on similar "BIDS apps" to make it easier to work with qMRI data, as well as make processing pipelines more open and transparent.

5.4 Data records

During and since the development of BIDS extension proposal 001, multiple data sets have been adapted to the new qMRI standard, in part also to stress-test the developing file naming schemes. The example data repository we created by converting publicly available qMRI data into the developed BIDS format is freely available (https://doi.org/10.17605/OSF.IO/K4BS5) under CC-BY 4.0 License and the datasets are described below.

5.4.1 Common conventions in the shared qMI-BIDS datasets

Each dataset is a folder named using the prefix ds- and a suffix indicating the name of the respective qMRI method and contains only one subject. Therefore, raw qMRI file collections are found in the sub-01 folder. In this folder, file collections intended for structural imaging and field mapping are stored in the anat and fmap directories, respectively. Quantitative maps are calculated from these raw file collections using open-source software tools and placed at a subdirectory under the derivatives folder, indicating the name of the software carried out the processing. In compliance with the BIDS format (Gorgolewski, Auer, et al., 2016), each NIfTI file (*.nii or *.nii.gz if compressed) stores the image data and co-exist with a sidecar JSON file of identical name, reporting image metadata. Located at the root directory of all dataset folders, dataset_description is a JSON file reporting the name, compatible BIDS version and the author(s) who converted the datasets from the original sources listed in the ReferencesAndLinks field.

5.4.2 The ds-mp2rage dataset

The raw data is an example MP2RAGE (Marques et al., 2010) file collection (4 images) denoted by the MP2RAGE suffix. Two entities describe the raw data: inv and part indicating that the file collection is comprised of images acquired at multiple InversionTime settings (two in total) and reconstructed in two parts (magnitude or phase), respectively. Quantitative maps were calculated using the pymp2rage toolbox and include a longitudinal relaxation time (T1) map and a unified T1 contrast image, denoted by T1map and UNIT1 suffixes, respectively.

5.4.3 The ds-mp2rageme dataset

The raw data is a multi-echo variant (Caan et al., 2019) of the ds-mp2rage dataset, yielding a file collection of 10 images. In addition to the inv and part entities, the echo entity remarks

that the readout associated with the second InversionTime (inv-2) has multiple images exported at 4 EchoTime durations. Quantitative maps were calculated using the pymp2rage toolbox. In addition to the ds-mp2rage derivatives, outputs include an R2starmap and a T2starmap, reporting observed transverse relaxation rate (R2*) and time (T2*), respectively. This dataset exemplifies how an existing file collection can be extended for different use cases without introducing a new suffix as described in the qMRI appendix.

5.4.4 The ds-mpm dataset

The raw data is a multiparametric mapping (Callaghan et al., 2019) (MPM) file collection, comprising 20 structural images of varying contrast acquired at multiple FlipAngle, MT-State and EchoTime settings. Therefore, individual files are distinguished by flip, mt and echo naming entities, respectively. Note that different flip and mt combinations yield 3 unique contrast groups, commonly denoted as PDw (6 echoes), MTw (6 echoes) and T1w (8 echoes). To reflect this association, filenames are supplemented by the freeform acq entity. In addition, the fmap folder contains file collections for calculating transmit (TB1EPI) and receive (RB1COR) radiofrequency field inhomogeneity maps. Quantitative maps were calculated using hMRI-Toolbox. The structural maps (hmri/anat folder) include R1map, R2starmap, PDmap and MTsat, and the field maps (hmri/fmap folder) include TB1map and RB1map. Units of the quantities reported by these maps are in reciprocal seconds (or Hertz) for R1map and R2starmap, in arbitrary units for PDmap and MTsat, and in relative percentage units for TB1map and RB1map. Sidecar JSON files of the quantitative maps indicate the raw images from which they are calculated.

5.4.5 The ds-mts dataset

The raw data is a magnetization transfer saturation index (MTS) file collection (3 images), which can be seen as a subset of the ds-mpm dataset, with each iteration of multiple FlipAngle and MTState parameters are acquired at a single EchoTime. Therefore, individual files are distinguished by fllip and mt entities, respectively. Quantitative maps were calculated using qMRLab and includes T1map and MTsat, in seconds and arbitrary units, respectively.

5.4.6 The ds-qsm dataset

For this dataset, only the derivatives folder is provided to exemplify how quantitative susceptibility maps (chiMap) are represented by qMRI-BIDS. The susceptibility map is in the unit of parts per million (ppm) and was calculated using qMRLab from a high-resolution

anatomical scan. There are numerous methods to acquire raw data to calculate a chiMap and qMRI-BIDS does not provide specialized file collection suffixes for this purpose. However, many of these methods can benefit from the existing suffixes (e.g., MEGRE if a multi-echo framework is used) and naming entities (e.g., the part entity to share wrapped phase images in radians).

5.4.7 The ds-vfa dataset

The raw data is a variable flip angle (VFA) file collection, comprising two images acquired at two FlipAngle settings, using the flip entity to distinguish individual files. Quantitative mapping was performed using qMRLab to generate a T1map (in seconds).

To create a lightweight version of this data repository, we contributed to the BIDS-examples (https://github.com/bids-standard/bids-examples), a set of placeholder files for example qMRI dataset that are punctually organised according to BEP001 and therefore a good reference for developers and users.

5.4.8 Other (third-party) qMRI-BIDS datasets

Since the release of BIDS v1.5.0, some projects have already adapted qMRI-BIDS to organize quantitative imaging data:

- The spine generic project (Cohen-Adad, Alonso-Ortiz, Abramovic, Arneitz, Atcheson, Barlow, Barry, Barth, Battiston, and Büchel, 2021) shared multi-center MTS data for standardized quantitative imaging of human spinal cord, publicly available at https://doi.org/10.5281/zenodo.4299140.

- The Neuromod (Boyle et al., n.d.) project collects longitudinal MTS and MP2RAGE data from 6 individuals for training artificial neural networks on human brain and behaviour. The dataset is accessible upon request (https://www.cneuromod.ca).

- Using the ISMRM-NIST system phantom, multi-center data was collected for comparing the accuracy and reproducibility of MTS and TB1AFI acquisitions between vendor-native and vendor-neutral sequence implementations (Karakuzu, Boudreau, Cohen-Adad, et al., 2020b). The dataset is publicly available at https://doi.org/10.17605/OSF.IO/5N3CU.

- The hMRI-Toolbox has made an MPM dataset publicly available for in-vivo histology using MRI at https://doi.org/10.1016/j.dib.2019.104132.

5.4.9 Code availability

The main codebase of qMRI-BIDS (https://github.com/bids-standard/bep001) was a GitHub fork of the main BIDS specification (https://github.com/bids-standard/bids-specification), where all the development history of BEP001 is publicly available through GitHub Issues. Following the completion of the development, upstream contribution to the main BIDS specification was made in 5 Pull Requests (ID 668, 671, 672, 680 and 690), where conversations with the project maintainers and codebase changes can be accessed. Source code to generate quantitative maps in the example dataset are provided by qMRLab (https://github.com/qMRLab/qmrlab), hMRI-Toolbox (https://hmri-group.github.io/hMRI-toolbox/) and pymp2rage (https://github.com/Gilles86/pymp2rage). Each software provides extensive user documentation, which were followed to create derivative datasets.

5.5 Methods

5.5.1 Community-driven development of BEP001

The development history of BEP001 spanned nearly 5 years. This extension was initiated following mailing list discussions about standardizing MP2RAGE(Marques et al., 2010) datasets and including multi-echo MRI acquisitions in late 2016 (https://bit.ly/bids_mailing). These discussions revealed that BIDS was still lacking a generic convention for specifying structural acquisitions yielding multiple contrasts. In the summer of 2018, meeting were held to hear concerns and questions from interested participants and to set an action plan for the development during the annual INCF NeuroInformatics conference in Montréal/Canada (http://www.neuroinformatics2018.org/) and the OHBM meeting in Singapore (https://www.humanbrainmapping.org/i4a/pages/index.cfm?pageID=3821). As a first action, a joint-community meeting was organized between MRI and neuroimaging scientists on 4 October 2018 (https://www.ismrm.org/virtual-meetings/virtual-meetings-archive/) , where a consensus decision was made on extending the specification for a variety of qMRI methods. After this meeting, a standard operational procedure was established and followed to advance the proposal, focusing on both transparency and accessibility to other researchers (Figure 5.2).
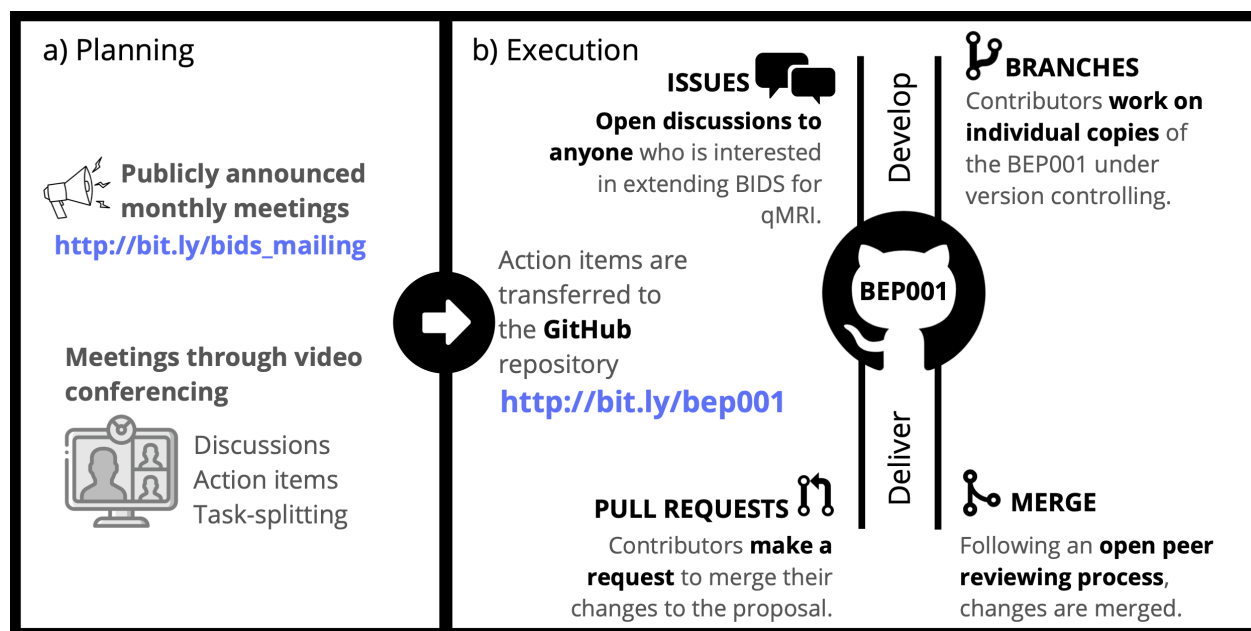
Figure 5.2 Summary of the standard operational procedure for improving BEP001. Outcomes from the monthly meetings **(a)** are transferred to a central GitHub repository, opened for more elaborate public discussions via issues and merged into the proposal through peer-reviewed pull requests **(b)**. BEP001 is inclusive to all communities who would like to contribute to the proposal or keep themselves up-to-date with the latest developments.

Interim outcomes from the development were presented in the 2020 annual conferences of OHBM (Hollander et al., 2020) and ISMRM (Karakuzu, Hollander, et al., 2020) to reach out more neuroimaging and MRI physics researchers, respectively. Following another year of development on the specification, example datasets and applications, BIDS incorporated and released BEP001 as part of their version 1.5.0. The main problems identified and resolved during the development are outlined in the following section, laying out the methodology of how qMRI can be incorporated into BIDS.

5.5.2   Extending an existing standard for new use cases

BIDS traditionally focused on conventional anatomical images that are collected in functional MRI experiments and whose contrast characteristics are well-defined (i.e., mostly T1-weighted images). This posed a challenge for the naming scheme of collections of multimodal images used in qMRI. Unlike conventional structural data, qMRI inputs are usually formed by collections of images where specific acquisition parameters are systematically manipulated. Moreover, the line separating contrast characteristics between these images is blurred. A concrete example: in a multi-echo GRE acquisition with a long TRs, early echoes will

be mostly PD- and B1+/B1- signal-weighted, whereas later echoes will be increasingly T2*-weighted. Most echoes will show a contrast that is the result of a mixture of underlying physical properties. This ambiguity disqualifies MRI weightings (e.g., T1w or T2starw) as suffix labels to specify interchangeable qMRI datasets. The use of often proprietary acquisition sequence names like "FLASH" (fast low angle shot) or "GRE" (gradient-recalled echo) as a suffix turned out to also be undesirable, because different MRI vendors use different naming conventions and, moreover, one type of sequence can often be used for numerous qMRI applications. To address this problem, BEP001 introduced a new common principle: file collections.

A second challenge that BEP001 addressed pertains to standardizing the data organisation of quantitative parametric maps. One central challenge of such maps is that the calculations on which they are based can be made both by proprietary vendor software run on the scanner system, or offline using open-source workflows. The resultant map can be described as derivative data in either case, yet the former lacks provenance of the whole calculation process and may not export the raw inputs to the calculation.

5.6 Acknowledgements

5.7 Author contributions

A.K., G.H. and K.W. prepared the original manuscript; A.K., G.H. and K.W. developed the initial draft of the standard and managed community contributions. A.K. merged the extension proposal to the main BIDS specification. G.H. and K.W. supervised the project. A.K., S.A, T.A., M.B., F.F., A.K., A.L., C.M., M.M., C.P., T.S., N.S., K.W. and G.H. contributed to meetings and drafts outlining the extension proposal. A.K., S.A, T.A., M.B., F.F., A.K., A.L., C.M., M.M., C.P., T.S., N.S., K.W. and G.H. revised the original manuscript.

5.8 Competing interests

The authors declare no competing financial interests.

5.9 References

Barnes, S. R., Ng, T. S. C., Santa-Maria, N., Montagne, A., Zlokovic, B. V., & Jacobs, R. E. (2015). Rocketship: a flexible and modular software tool for the planning, processing and analysis of dynamic mri studies. *BMC medical imaging*, *15*(1), 1–20.

Barral, J. K., Gudmundson, E., Stikov, N., Etezadi-Amoli, M., Stoica, P., & Nishimura, D. G. (2010). A robust methodology for in vivo t1 mapping. *Magn Reson Med*, *64*(4), 1057–67. https://doi.org/10.1002/mrm.22497

Berks, M., m Parker, G. J., Little, R., & Cheung, S. (2021). Madym: a c++ toolkit for quantitative dce-mri analysis. *Journal of Open Source Software*, *6*(66), 3523.

Bojorquez, J. Z., Bricq, S., Acquitter, C., Brunotte, F., Walker, P. M., & Lalande, A. (2017). What are normal relaxation times of tissues at 3 t? *Magn Reson Imaging*, *35*, 69–80. https://doi.org/10.1016/j.mri.2016.08.021

Boyle, J., Pinsard, B., Boukhdir, A., Belleville, S., Brambatti, S., Chen, J., Cohen-Adad, J., Cyr, A., Fuente, A., Rainville, P., & Bellec, P. (n.d.). The courtois project on neuronal modelling - first data release. *26th Annual Meeting of the Organization for Human Brain Mapping.*

Caan, M. W. A., Bazin, P.-L., Marques, J. P., de Hollander, G., Dumoulin, S. O., & van der Zwaag, W. (2019). Mp2rageme: t1, t2*, and qsm mapping in one sequence at 7 tesla. *Human brain mapping*, *40*(6), 1786–1798.

Callaghan, M. F., Lutti, A., Ashburner, J., Balteau, E., Corbin, N., Draganski, B., Helms, G., Kherif, F., Leutritz, T., & Mohammadi, S. (2019). Example dataset for the hmri toolbox. *Data in brief*, *25*, 104132.

Carr, H. Y., & Purcell, E. M. (1954). Effects of diffusion on free precession in nuclear magnetic resonance experiments. *Physical review*, *94*(3), 630.

Chung, S., Kim, D., Breton, E., & Axel, L. (2010). Rapid b1+ mapping using a preconditioning rf pulse with turboflash readout. *Magnetic resonance in medicine*, *64*(2), 439–446.

Cohen-Adad, J., Alonso-Ortiz, E., Abramovic, M., Arneitz, C., Atcheson, N., Barlow, L., Barry, R. L., Barth, M., Battiston, M., & Büchel, C. (2021). Generic acquisition protocol for quantitative mri of the spinal cord. *Nature protocols*, *16*(10), 4611–4632.

Cordes, C., Konstandin, S., Porter, D., & Günther, M. (2020). Portable and platform-independent mr pulse sequence programs. *Magnetic resonance in medicine*, *83*(4), 1277–1290.

Debus, C., Floca, R., Ingrisch, M., Kompan, I., Maier-Hein, K., Abdollahi, A., & Nolden, M. (2019). Mitk-modelfit: a generic open-source framework for model fits and their exploration in medical imaging–design, implementation and application on the example of dce-mri. *BMC bioinformatics*, *20*(1), 1–18.

Dinse, J., Härtwich, N., Waehnert, M. D., Tardif, C. L., Schäfer, A., Geyer, S., Preim, B., Turner, R., & Bazin, P. L. (2015). A cytoarchitecture-driven myelin model reveals area-specific signatures in human primary and secondary areas using ultra-high resolution in-vivo brain mri. *Neuroimage*, *114*, 71–87.

Draganski, B., Ashburner, J., Hutton, C., Kherif, F., Frackowiak, R. S. J., Helms, G., & Weiskopf, N. (2011). Regional specificity of mri contrast parameter changes in normal ageing revealed by voxel-based quantification (vbq). *Neuroimage*, *55*(4), 1423–1434.

Eggenschwiler, F., Kober, T., Magill, A. W., Gruetter, R., & Marques, J. P. (2012). Sa2rage: a new sequence for fast b1+-mapping. *Magnetic resonance in medicine*, *67*(6), 1609–1619.

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). Mriqc: advancing the automatic prediction of image quality in mri from unseen sites. *PloS one*, *12*(9), e0184661.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., & Snyder, M. (2019). Fmriprep: a robust preprocessing pipeline for functional mri. *Nature methods*, *16*(1), 111–116.

Froeling, M. (2019). Qmrtools: a mathematica toolbox for quantitative mri analysis. *Journal of Open Source Software*, *4*(38), 1204.

Golkov, V., Dosovitskiy, A., Sperl, J. I., Menzel, M. I., Czisch, M., Sämann, P., Brox, T., & Cremers, D. (2016). Q-space deep learning: twelve-fold shorter and model-free diffusion mri scans. *IEEE transactions on medical imaging*, *35*(5), 1344–1351.

Gorgolewski, K. J., Alfaro-Almagro, F., Auer, T., Bellec, P., Capotă, M., Chakravarty, M. M., Churchill, N. W., Cohen, A. L., Craddock, R. C., & Devenyi, G. A. (2017). Bids apps: improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS computational biology*, *13*(3), e1005209.

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., & Halchenko, Y. O. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific data*, *3*(1), 1–9.

Gupta, R. K. (1977). A new look at the method of variable nutation angle for the measurement of spin-lattice relaxation times using fourier transform nmr. *Journal of Magnetic Resonance (1969)*, *25*(1), 231–235. https://doi.org/https://doi.org/10.1016/0022-2364(77)90138-X

Haast, R. A. M., Ivanov, D., Formisano, E., & Uludağ, K. (2016). Reproducibility and reliability of quantitative and weighted t1 and t2* mapping for myelin-based cortical parcellation at 7 tesla. *Frontiers in neuroanatomy*, *10*, 112.

Hansen, M. S., & Sørensen, T. S. (2013). Gadgetron: an open source framework for medical image reconstruction. *Magnetic resonance in medicine*, *69*(6), 1768–1776.

Helms, G., Dathe, H., Kallenberg, K., & Dechent, P. (2008). High-resolution maps of magnetization transfer with inherent correction for rf inhomogeneity and t1 relaxation obtained from 3d flash mri. *Magn Reson Med*, *60*(6), 1396–407. https://doi.org/10.1002/mrm.21732

Hollander, G., Karakuzu, A., Appelhof, S., Auer, T., Boudreau, M., Feingold, F., Khan, A., Lazari, A., Philips, C., Stikov, N., & Whitaker, K. (2020). A standard for the organization of quantitative mri data: bids extension proposal 001. *26th Annual Meeting of the Organization for Human Brain Mapping*.

Inati, S. J., Naegele, J. D., Zwart, N. R., Roopchansingh, V., Lizak, M. J., Hansen, D. C., Liu, C.-Y., Atkinson, D., Kellman, P., & Kozerke, S. (2017). Ismrm raw data format: a proposed standard for mri raw datasets. *Magnetic resonance in medicine*, *77*(1), 411–421.

Insko, E. K., & Bolinger, L. (1993). Mapping of the radiofrequency field. *Journal of Magnetic Resonance, Series A*, *103*(1), 82–85.

Jiru, F., & Klose, U. (2006). Fast 3d radiofrequency field mapping using echo-planar imaging. *Magnetic Resonance in Medicine*, *56*(6), 1375–1379. https://doi.org/https://doi.org/10.1002/mrm.21083

Jochimsen, T. H., & Von Mengershausen, M. (2004). Odin—object-oriented development interface for nmr. *Journal of Magnetic Resonance*, *170*(1), 67–78.

Karakuzu, A., Boudreau, M., Cohen-Adad, J., & Stikov, N. (2020b). Thinking outside the blackbox: a fully transparent t1 mapping pipeline. *Proceedings of ISMRM 28th Annual Meeting.*

Karakuzu, A., Boudreau, M., Duval, T., Boshkovski, T., Leppert, I., Cabana, J.-F., Gagnon, I., Beliveau, P., Pike, G. B., Cohen-Adad, J., & Stikov, N. (2020). Qmrlab: quantitative mri analysis, under one umbrella. *Journal of Open Source Software*, *5*(53), 2343.

Karakuzu, A., Hollander, G., Appelhof, S., Auer, T., Boudreau, M., Feingold, F., Khan, A., Lazari, A., Philips, C., Stikov, N., & Whitaker, K. (2020). A joint-community effort to standardize quantitative mri data: updates from the bids extension proposal. *Proceedings of ISMRM 28th Annual Meeting.*

Layton, K. J., Kroboth, S., Jia, F., Littin, S., Yu, H., Leupold, J., Nielsen, J.-F., Stöcker, T., & Zaitsev, M. (2017). Pulseq: a rapid and hardware-independent pulse sequence prototyping framework. *Magnetic resonance in medicine*, *77*(4), 1544–1552.

Lazari, A., & Lipp, I. (2021). Can mri measure myelin? systematic review, qualitative assessment, and meta-analysis of studies validating microstructural imaging with myelin histology. *Neuroimage*, 117744.

Li, X., Morgan, P. S., Ashburner, J., Smith, J., & Rorden, C. (2016). The first step for neuroimaging data analysis: dicom to nifti conversion. *Journal of neuroscience methods*, *264*, 47–56.

Lommers, E., Guillemin, C., Reuter, G., Fouarge, E., Delrue, G., Collette, F., Degueldre, C., Balteau, E., Maquet, P., & Phillips, C. (2021). Voxel-based quantitative mri reveals spatial patterns of grey matter alteration in multiple sclerosis. *Human Brain Mapping*, *42*(4), 1003–1012. https://doi.org/https://doi.org/10.1002/hbm.25274

Lundervold, A. S., & Lundervold, A. (2019). An overview of deep learning in medical imaging focusing on mri. *Zeitschrift für Medizinische Physik*, *29*(2), 102–127.

Lutti, A., Dick, F., Sereno, M. I., & Weiskopf, N. (2014). Using high-resolution quantitative mapping of r1 as an index of cortical myelination. *Neuroimage*, *93*, 176–188.

Lyu, Q., & Wang, G. (2018). Quantitative mri: absolute t1, t2 and proton density parameters from deep learning. *arXiv preprint arXiv:1806.07453.*

Ma, J., & Wehrli, F. W. (1996). Method for image-based measurement of the reversible and irreversible contribution to the transverse-relaxation rate. *Journal of Magnetic Resonance, Series B*, *111*(1), 61–69.

MacKay, A., Whittall, K., Adler, J., Li, D., Paty, D., & Graeb, D. (1994). In vivo visualization of myelin water in brain by magnetic resonance. *Magn Reson Med*, *31*(6), 673–677.

Magland, J. F., Li, C., Langham, M. C., & Wehrli, F. W. (2016). Pulse sequence programming in a dynamic visual environment: sequencetree. *Magnetic resonance in medicine*, *75*(1), 257–265.

Maier, O., Baete, S. H., Fyrdahl, A., Hammernik, K., Harrevelt, S., Kasper, L., Karakuzu, A., Loecher, M., Patzig, F., & Tian, Y. (2021). Cg-sense revisited: results from the first ismrm reproducibility challenge. *Magnetic resonance in medicine*, *85*(4), 1821–1839.

Maier, O., Spann, S. M., Bödenler, M., & Stollberger, R. (2020). Pyqmri: an accelerated python based quantitative mri toolbox. *Journal of Open Source Software*, *5*(56), 2727.

Mancini, M., Karakuzu, A., Cohen-Adad, J., Cercignani, M., Nichols, T. E., & Stikov, N. (2020). An interactive meta-analysis of mri biomarkers of myelin. *eLife*, *9*. https://doi.org/10.7554/elife.61523

Markiewicz, C. J., Gorgolewski, K. J., Feingold, F., Blair, R., Halchenko, Y. O., Miller, E., Hardcastle, N., Wexler, J., Esteban, O., Goncalves, M., Jwa, A., & Poldrack, R. A. (2021). Openneuro: an open resource for sharing of neuroimaging data. *bioRxiv*, 2021.06.28.450168. https://doi.org/10.1101/2021.06.28.450168

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P. F., & Gruetter, R. (2010). Mp2rage, a self bias-field corrected sequence for improved segmentation and t1-mapping at high field. *Neuroimage*, *49*(2), 1271–81. https://doi.org/10.1016/j.neuroimage.2009.10.002

Mezer, A., Yeatman, J. D., Stikov, N., Kay, K. N., Cho, N. J., Dougherty, R. F., Perry, M. L., Parvizi, J., Hua le, H., Butts-Pauly, K., & Wandell, B. A. (2013). Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nat Med*, *19*(12), 1667–72. https://doi.org/10.1038/nm.3390

Nielsen, J.-F., & Noll, D. C. (2018). Toppe: a framework for rapid prototyping of mr pulse sequences. *Magnetic resonance in medicine*, *79*(6), 3128–3134.

Novikov, D. S., Kiselev, V. G., & Jespersen, S. N. (2018). On modeling. *Magn Reson Med*, *79*(6), 3172–3193. https://doi.org/10.1002/mrm.27101

Ortuño, J. E., Ledesma-Carbayo, M. J., Simões, R. V., Candiota, A. P., Arús, C., & Santos, A. (2013). Dce@ urlab: a dynamic contrast-enhanced mri pharmacokinetic analysis tool for preclinical data. *BMC bioinformatics*, *14*(1), 1–17.

Papp, D., Callaghan, M. F., Meyer, H., Buckley, C., & Weiskopf, N. (2016). Correction of inter-scan motion artifacts in quantitative r1 mapping by accounting for receive coil sensitivity effects. *Magnetic resonance in medicine*, *76*(5), 1478–1485.

Poldrack, R., Barch, D., Mitchell, J., Wager, T., Wagner, A., Devlin, J., Cumba, C., Koyejo, O., & Milham, M. (2013). Toward open sharing of task-based fmri data: the openfmri

project. *Frontiers in Neuroinformatics*, *7*, 12. https://www.frontiersin.org/article/10.3389/fninf.2013.00012

Ravi, K. S., Geethanath, S., & Vaughan, J. T. (2019). Pypulseq: a python package for mri pulse sequence design. *Journal of Open Source Software*, *4*(42), 1725.

Santos, J. M., Wright, G. A., & Pauly, J. M. (2004). Flexible real-time magnetic resonance imaging framework. *1*, 1048–1051.

Smith, D. S., Li, X., Arlinghaus, L. R., Yankeelov, T. E., & Welch, E. B. (2015). Dcemri. jl: a fast, validated, open source toolkit for dynamic contrast enhanced mri analysis. *PeerJ*, *3*, e909.

Stikov, N., Trzasko, J. D., & Bernstein, M. A. (2019). Reproducibility and the future of mri research. *Magn Reson Med*, *82*(6), 1981–1983. https://doi.org/10.1002/mrm.27939

Tabelow, K., Balteau, E., Ashburner, J., Callaghan, M. F., Draganski, B., Helms, G., Kherif, F., Leutritz, T., Lutti, A., & Phillips, C. (2019). Hmri–a toolbox for quantitative mri in neuroscience and clinical research. *Neuroimage*, *194*, 191–210.

Weiskopf, N., Suckling, J., Williams, G., Correia, M. M., Inkster, B., Tait, R., Ooi, C., Bullmore, E. T., & Lutti, A. (2013). Quantitative multi-parameter mapping of r1, pd(*), mt, and r2(*) at 3t: a multi-center validation. *Front Neurosci*, *7*, 95. https://doi.org/10.3389/fnins.2013.00095

Weiskopf, N., Callaghan, M. F., Josephs, O., Lutti, A., & Mohammadi, S. (2014). Estimating the apparent transverse relaxation time (r2*) from images with different contrasts (estatics) reduces motion artifacts. *Frontiers in neuroscience*, *8*, 278.

Weiskopf, N., Edwards, L. J., Helms, G., Mohammadi, S., & Kirilina, E. (2021). Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. *Nature Reviews Physics*, *3*(8), 570–588. https://doi.org/10.1038/s42254-021-00326-1

Weiskopf, N., Mohammadi, S., Lutti, A., & Callaghan, M. F. (2015). Advances in mri-based computational neuroanatomy: from morphometry to in-vivo histology. *Current opinion in neurology*, *28*(4), 313–322.

Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., & Bourne, P. E. (2016). The fair guiding principles for scientific data management and stewardship. *Scientific data*, *3*(1), 1–9.

Wolff, S., & Balaban, R. (1989a). Magnetization transfer contrast (mtc) and tissue water proton relaxation in vivo. *Magn. Reson. Med.*, *10*(1), 135–144. https://doi.org/citeulike-article-id:10096204doi:10.1002/mrm.1910100113

Wood, T. (2018). Quit: quantitative imaging tools. *Journal of Open Source Software*, *3*(26), 656.

Wu, Y., Ma, Y., Du, J., & Xing, L. (2020). Accelerating quantitative mr imaging with the incorporation of b1 compensation using deep learning. *Magnetic Resonance Imaging*, *72*, 78–86.

Yarkoni, T., Markiewicz, C. J., de la Vega, A., Gorgolewski, K. J., Salo, T., Halchenko, Y. O., McNamara, Q., DeStasio, K., Poline, J.-B., & Petrov, D. (2019). Pybids: python tools for bids datasets. *Journal of open source software*, *4*(40).

Yarnykh, V. L. (2007b). Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magn Reson Med*, *57*(1), 192–200.

Yarnykh, V. L. (2010). Optimal radiofrequency and gradient spoiling for improved accuracy of t1 and b1 measurements using fast steady-state techniques. *Magn Reson Med*, *63*(6), 1610–1626.

Yoo, Y., Tang, L. Y. W., Brosch, T., Li, D. K. B., Kolind, S., Vavasour, I., Rauscher, A., MacKay, A. L., Traboulsee, A., & Tam, R. C. (2018). Deep learning of joint myelin and t1w mri features in normal-appearing brain tissue to distinguish between multiple sclerosis patients and healthy controls. *NeuroImage: Clinical*, *17*, 169–178.

CHAPTER 6   ARTICLE 3: VENDOR-NEUTRAL SEQUENCES AND FULLY TRANSPARENT WORKFLOWS IMPROVE MULTICENTER REPRODUCIBILITY OF QUANTITATIVE MRI

Authors: A. Karakuzu[1,2], J. Cohen-Adad[1,3,4], N. Stikov[1,2]

Affiliations:

1. NeuroPoly Lab, Institute of Biomedical Engineering, Polytechnique Montréal, Montréal, QC, Canada

2. Montréal Heart Institute, Montréal, QC, Canada

3. Functional Neuroimaging Unit, CRIUGM, Université de Montréal, Montréal, QC, Canada

4. Mila - Quebec AI Institute, Montreal, QC, Canada

6.1   Abstract

Purpose: We developed a transparent end-to-end qMRI workflow that starts with a vendor-neutral acquisition and tested the hypothesis that vendor-neutral sequences (VENUS) decrease inter-vendor variability of T1, MTR and MTsat measurements.

Methods: We developed and deployed a vendor-neutral 3D spoiled gradient-echo (SPGR) sequence on three clinical scanners by two MRI vendors and acquired T1 maps on the NIST phantom, as well as T1, MTR and MTsat maps in three healthy participants. We performed hierarchical shift function analysis in vivo to characterize the differences between scanners when VENUS is used instead of commercial vendor implementations. Inter-vendor deviations were compared for statistical significance to test the hypothesis.

Results: In the NIST phantom, VENUS reduced inter-vendor differences from 8 - 19.4% to 0.2 - 5% with an overall accuracy improvement, reducing ground truth T1 deviation from 7 - 11% to 0.2 - 4%. In vivo we found that the variability between vendors is significantly reduced ($p = 0.015$) for all maps (T1, MTR and MTsat) using VENUS.

Conclusion: We conclude that vendor-neutral workflows are feasible and compatible with clinical MRI scanners. The significant reduction of inter-vendor variability using VENUS has important implications for qMRI research and for the reliability of multicenter clinical trials.

6.2 Introduction

As the invention of MRI approaches its 50th anniversary (Lauterbur, 1989), the notion of image acquisition has almost become synonymous with data collection. A major driving force in the transformation of MR images from mere pictures into mineable data (Gillies et al., 2016) is attributing physiologically relevant physical parameters to the voxels, namely quantitative MRI (qMRI). MRI is not a quantitative measurement device by design. Nonetheless, systematic manipulation of effective micrometer-level MRI parameters via specialized acquisition methods, followed by fitting the resulting data to a signal representation or a biophysical model (Novikov et al., 2018) can yield parametric maps, turning scanners into quantitative diagnostic tools. Despite being as old as MRI itself, most of the qMRI methods have not succeeded to find their widespread use in the clinic, at least in part due to a major multicenter reproducibility challenge.

The introduction is organized around two problems hampering multicenter reproducibility of qMRI, which this study seeks to address:

1. Lack of transparency in vendor implementations and multicenter consistency of pulse sequences that are commonly used in qMRI

2. Technical roadblocks in the way of deploying a standardized pulse sequence along with a unified user interface to multiple imaging sites

T1 relaxometry is a clear example of how availability, transparency and multicenter consistency of pulse sequences influence multicenter reproducibility. Several methods such as inversion-recovery spin-echo (IR-SE) (Hahn, 1949) variable flip angle (VFA-T1) (Gupta, 1977) Look-Locker IR (Look and Locker, 1970) and magnetization-prepared two rapid-echoes (Marques et al., 2010) have gained popularity in MRI research. Although measured T1 values can exhibit up to 30% inter-sequence variability in the same scan session for the same participant (Stikov, Boudreau, et al., 2015), a selected T1 relaxometry method is much more reliable within-site (Grafe et al., 2021; Okubo et al., 2016). As for the multicenter stability, MP2RAGE appears to be a promising T1 mapping method at 7T with a single vendor considered across 8 imaging sites (Voelker et al., 2021). On the other hand, substantial multicenter variability is reported for another popular whole-brain imaging method VFA-T1, both in-vivo and in the system phantom (Bane et al., 2018; Keenan, Gimbutas, Dienstfrey, Stupic, et al., 2021). Several factors contribute to the variability of the VFA-T1 measurement, including B1 field inhomogeneity (Stikov, Boudreau, et al., 2015), incomplete spoiling (Yarnykh, 2010), sequence parameters and bore temperature (**Keenan:2021**), and

uncontrolled magnetization transfer (MT) effects (Teixeira et al., 2020). Because of all these diverse confounders of T1 stability, the healthy range of in-vivo T1 values at 3T remains elusive (Bojorquez et al., 2017; Cohen-Adad, Alonso-Ortiz, Abramovic, Arneitz, Atcheson, Barlow, Barry, Barth, Battiston, Buchel, et al., 2021). This constitutes a critical problem for the potential use of T1 relaxometry in clinics.

Considerable amount of research has focused on measurement bias stemming from the imperfections related to image acquisition. However, reproducibility of the developed techniques is often hindered by problem 1. For example, a simple yet powerful B1 correction framework for VFA-T1 has been established (Liberman et al., 2014), but B1 mapping sequences are typically not available in commercial systems or available sequences are different across vendors. This not only imposes a practical challenge in evaluating the reliability of VFA-T1 measurements across vendors (Keenan, Gimbutas, Dienstfrey, Stupic, et al., 2021; Leutritz et al., 2020) but the differences between vendor-native B1 mapping methods can aggravate the instability (Boudreau et al., 2018). Another example is the spoiling gradient area and RF spoiling phase increment in the commercial implementations of 3D spoiled gradient-echo (3D-SPGR) sequences. Both parameters determine the accuracy of VFA-T1 mapping (Yarnykh, 2010) However, vendors are known to set different defaults for these parameters, rendering stock 3D-SPGR sequences of two major MRI manufacturers unfit for this application (Yarnykh, 2007a). Note that neither parameter can be controlled or displayed at the scanner console. Similarly, fundamental properties of the excitation pulse (e.g., pulse shape, time-bandwidth product, duration) are not disclosed and it is not known how these properties are adjusted under different SAR requirements. To achieve a standardized 3D-SPGR acquisition for T1 mapping, such parameter configurations should be disclosed, made accessible and equalized across scanners for eliminating systematic biases. Recently, Gracien et al. showed a successful example of how this solution can reduce systematic biases in relaxometry mapping between two different scanner models from the same vendor (Gracien et al., 2020).

Addressing inadequacies of model assumptions constitutes another solution toward improving the reliability of qMRI methods (Novikov et al., 2018). For example, balancing the total amount of RF power deposited by each run of a VFA acquisition using multiband non-selective pulses, Teixeira et al. enforced two-pool MT systems to behave like a single pool MT system (Teixeira et al., 2020). They showed that making single pool assumption of relaxometry methods valid through controlled saturation of MT increases the measurement reliability. Although this technique holds important implications for multicenter reproducibility of qMRI, deploying it to multiple sites is not a straightforward process. Moreover, proprietary programming libraries of different scanner manufacturers may not be flexible to allow identical implementations, exemplifying the constraints imposed by problem 2. An-

other model-related improvement has been recently introduced to replace the fixed empirical correction factor of MT saturation (MTsat) B1 correction (Helms, Dathe, and Dechent, 2008) with a correction factor map to reduce B1 dependency of MTsat maps (Rowley et al., 2021). The proposed methodology calls for the details on saturation and excitation pulses (e.g., pulse shape, frequency offset, pulse duration, etc.) as the correction framework is based on simulations. From the standpoint of problem 1, such information is not easily accessible in the stock sequence, so the correction cannot be applied. From the perspective of problem 2, deploying sequences to multiple centers with known saturation and excitation pulse parameters may not be realistic within the confines of vendor development environments. Even though both studies made their code publicly available to facilitate the reproducibility of their work (Stikov, Trzasko, et al., 2019), black-box vendor strategies thwart these valuable efforts.

Fortunately, there are several open-source pulse sequence development platforms to contend with problem 2 (Cordes et al., 2020; Jochimsen and Von Mengershausen, 2004; Layton et al., 2017; Magland et al., 2016; Nielsen and Noll, 2018; Ravi, Geethanath, et al., 2019; Stöcker et al., 2010). These platforms have in common the ability to interpret and translate the same sequence logic for multiple vendors, considerably reducing multi-center development efforts and minimizing implementation variability. Another advantage of these tools is to attract community-driven development. For example, Pulseq has received considerable community attention to motivate the development of sequences in Python (Ravi, Geethanath, et al., 2019), or even going beyond code to graphically assemble (Zwart and Pipe, 2015) Pulseq descriptions using Pulseq-GPI (Ravi, Potdar, et al., 2018). Currently, Pulseq can be operated on two major clinical scanners (Siemens and GE) and three pre-clinical scanner platforms. There is recent literature showing the feasibility of Pulseq for performing multicenter qMRI studies. For example, a standardized chemical exchange saturation (CEST) protocol has been developed and deployed on three Siemens scanners, where two of the systems had different vendor software versions (Herz et al., 2021). Results by Herz et al. showed multicenter consistency for an advanced CEST method, which has been made publicly available for both Python and Matlab users. Another recent Pulseq study performed inversion-recovery T1 mapping and multi-echo spin-echo T2 mapping on two Siemens scanners at 1.5 and 3T using the system phantom (Tong et al., 2021). In that study the reference T1 mapping method (Barral et al., 2010) accurately estimated T1 values within an 8% error band, whereas the T2 accuracy was slightly reduced. Taken together, these studies reveal the vital role of vendor-neutral pulse sequences in standardizing qMRI across centers. However, whether a vendor-neutral approach can improve quantitative agreement between scanners from different vendors has remained an open question.

The focus of earlier open-source pulse sequence platforms was providing a rapid and unified prototyping framework for facilitating interoperability, so some of the most adjusted scan parameters (e.g., field of view) had to be fixed once the sequence was downloaded to the scanner. More recent solutions such as GammaStar (Cordes et al., 2020) can remove this limitation by enabling user interaction trough the vendor's user interface to modify fundamental protocol settings during the imaging session. Offering a more complete solution to problem 2 through on-the-fly sequence updates, GammaStar eases the collaborative sequence development process by providing a web-based interface that can be accessed online, executed in a Docker container or locally. Although such technical improvements reduced the entrance barrier for free sequence development, exchange and standardization, validation for safety and multi-center use has remained elusive. Recently, Tong et al. proposed a framework for testing, documenting and sharing open-source pulse sequences (Tong et al., 2021), which adds an important missing piece to the community-driven MRI development puzzle.

RTHawk (Santos et al., 2004) is another vendor-neutral solution, which is a proprietary platform for MRI software development. As it is utilizing the same infrastructure as an FDA approved (510(k), No: K1833274) cardiac imaging platform, it ensures operation within MRI hardware and safety limits. Unlike the above-mentioned solutions, RTHawk provides a remote procedure call (RPC) server that replaces the vendor's pulse sequence controller to orchestrate vendor-specific low-level hardware instructions. The RPC pulse sequence server receives control commands and relevant sequence components (i.e., RF and gradient waveforms, ADC and timing events designed in SpinBench, as shown in Figure 6.1e directly from a standalone Ubuntu workstation connected to the scanner network (Figure 6.1c). This gives the flexibility to issue synchronous or asynchronous updates to a sequence in real-time, such as scaling/replacing waveforms between TRs or changing the volume prescription. As the sequence control manager is decoupled from the vendor's workstation, RTHawk makes it possible to develop a vendor-neutral unified user interface (UUI) per application (Figure 6.1d). In addition, collected raw data is streamed over to the standalone Ubuntu workstation through a real-time transport protocol (RTP). The RTP data manager enables adding or changing the metadata associated with each observation, which enables exporting raw and reconstructed images in community data standards (Figure 6.1f). Together, these components offer a complete solution to problem 2 by enabling the design of a unified user console that can automate qMRI protocols (e.g., MTsat), dispensing with the manual effort of keeping prescan calibrations constant across multiple runs and removing the need for vendor-specific data handling. Although the application programming interface (API) to develop and the runtime to execute RTHawk applications are proprietary components, the GUI-based sequence design and simulation engine (SpinBench) is freely available for non-commercial use

(Figure 6.1e). In addition, researchers are allowed to share source-code that uses RTHawk libraries for the acquisition control and image reconstruction (Overall and Pauly, 2007). As a result, these sequences can be made open source.

Aside from vendor-neutral experiments, researchers looked at improving qMRI stability by customizing vendor-native implementations and equalizing parameters to the utmost extent possible (Gracien et al., 2020; Lee et al., 2019; Leutritz et al., 2020). However, downstream data harmonization methods were still needed to correct for certain inter-vendor differences (Leutritz et al., 2020), or some of the bias could not be removed altogether, as not all qMRI-relevant settings can be unified using a vendor-native approach (Lee et al., 2019). This approach uses intricate vendor programming interfaces and becomes less feasible in a multicenter setting. In addition, there is no guarantee of inter-vendor generalization even for scanners with comparable hardware specs, as the selection of sequence design elements is unique to each vendor. Even within the confines of a single vendor, recent studies have demonstrated the promise of a vendor-neutral approach to qMRI in standardizing methods between centers (Herz et al., 2021; Tong et al., 2021).

In this study, we test the hypothesis that our vendor-neutral sequence (VENUS) reduces inter-vendor variability of T1, MTR and MTsat measurements. To test this hypothesis, we compared vendor-native T1, MTR and MTsat maps (Helms, Dathe, Kallenberg, et al., 2008) with those obtained by developing and using VENUS in three healthy participants, across three different scanners models from two manufacturers at 3T.

6.3 Methods

6.3.1 Vendor-neutral pulse sequence development

We deployed vendor-neutral pulse sequences developed in RTHawk v3.0.0 (rc4-28-ge3540dda19) (HeartVista Inc., CA, USA) on three 3T systems: i) GE Discovery 750 software version DV25 (R02_1549.b) (GE Healthcare, Milwaukee, MI, USA), ii) Siemens Prisma software version VE11C (N4_LATEST_20160120) (Siemens Healthineers, Erlangen, Germany) and iii) Siemens Skyra with the same software version as (ii). Throughout the rest of this article, these scanners will be referred to as G1, S1 and S2 (Figure 6.1a). In all the experiments, RTHawk was connected to the vendor's host workstation through the local area network using the SSH protocol to drive the scanner hardware and to collect the raw data (Figure 6.1b).
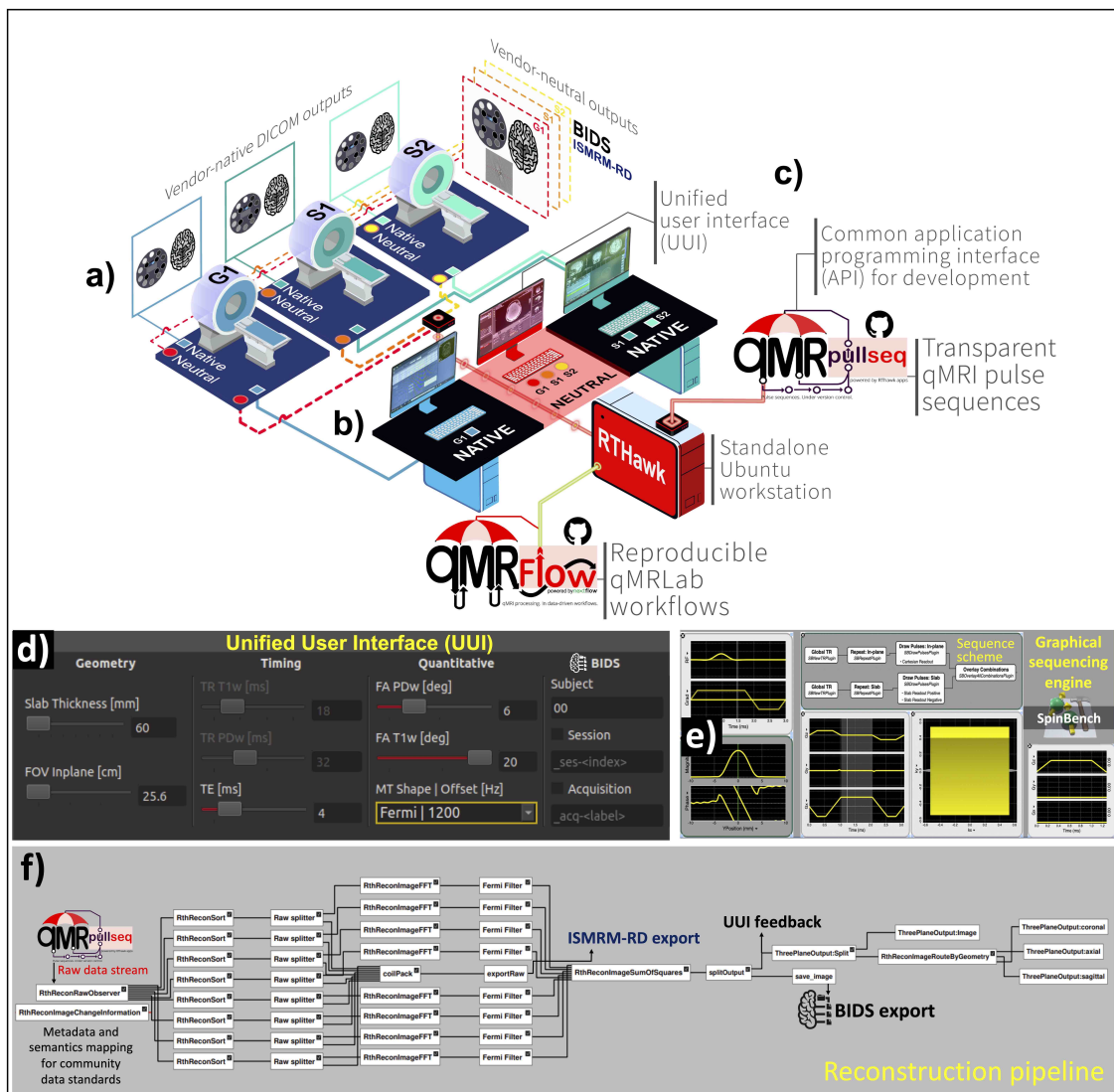
Figure 6.1 Schematic illustration of the experimental design for multicenter data collection using vendor-native and vendor-neutral pulse sequences and pulse sequence development components: **a)** 3 MRI systems are located at 2 different sites and are labeled G1 (GE 750w), S1 (Siemens Prisma) and S2 (Siemens Skyra). Vendor "Native" systems export data in the DICOM format. The proposed vendor-agnostic "Neutral" system can export a complete set of reconstructed images in BIDS and the k-space data in ISMRM-RD format, synchronized across MRI systems. **b)** Connecting to the MRI system(s) over the local network, RTHawk (red workstation) can play open-source qMRI pulse sequences under version control (qMRPullseq). All the sequences are publicly available at https://github.com/qmrlab/pulse_sequences. **c)** Fully containerized qMRFlow data-driven pipelines can connect to the scanner data stream for post-processing on the RTHawk workstation (red workstation). The same pipelines can be reproduced on a local computer, supercomputing clusters or on the cloud. **d)** The acquisitions are controlled using a unified user interface (UUI), providing a consistent user experience across vendors. **e)** RF and gradient waveform stub blocks together with the readout logic is developed using SpinBench. **f)** RTHawk reconstruction pipeline nodes are illustrated for an 8-channel receiver, also indicating how raw and reconstructed data are exported and forwarded to the display tools for on-site visualization.

6.3.2 General design considerations

All vendor-neutral protocols are based on a 3D SPGR pulse sequence (Haase et al., 1986), with the RF, gradient waveforms, and the readout scheme developed from scratch as independent sequence blocks in SpinBench v2.5.2, a graphical user interface (GUI) based pulse sequence design and Bloch simulation software (Figure 6.1e). Sequence stubs that correspond to a combination of these blocks were ordered in a configuration file (application.apd) and linked to a scan control manager script (control.js). Specifically, this script was responsible for shifting the timing blocks and scaling or enabling/disabling the gradient waveforms in the stub to play the sequence following the parameters requested by the user. To enable such interactions, a UUI (control.ui) was developed for quantitative imaging, allowing the user to iterate through relevant acquisition parameters (e.g., FA, TR, and MT pulse for MTsat) from one simple panel that is vendor-neutral (Fig.1d). The RTHawk real-time acquisition controller was configured to synchronously update these parameters at the end of each individual acquisition in the quantitative protocol (e.g., a flip angle image in the VFA-T1 mapping protocol). Identical scan geometry and pre-acquisition settings were transferred between each individual acquisition. To avoid signal clipping, the highest SNR acquisitions (e.g., T1w acquisition of the MTsat protocol) were defined as the first iteration. A simple sum-of-squares multi-coil reconstruction (recon.js) for fully sampled k-space was developed with a Fermi filter (transition width = 0.01, radius = 0.48, both expressed as a proportion of the FOV) (Figure 6.1f).

All the metadata annotations, accumulation logic of the collected data and naming of the exported images were designed in full compliance with the community data standards: the Brain Imaging Data Structure (BIDS) for the reconstructed images (Gorgolewski, Auer, et al., 2016; Karakuzu, Appelhoff, et al., 2021) and ISMRM-RD (Inati et al., 2017) for the k-space data (Figure 6.1a). Moreover, a custom panel in the UUI was provided for users to define freeform BIDS naming entities such as participant ID, session, and acquisition labels (Figure 6.1d). Other naming entities relevant to the quantitative acquisitions (e.g., flip and mt for FA and MT pulse, respectively) were automatically governed.

6.3.3 The vendor-neutral protocol

A slab-selective (thickness = 50 mm, gradient net area = 4.24 cyc/thickness) SINC excitation pulse (time-bandwidth product (TΔf) = 8, duration = 1.03 ms, Hanning windowed) was implemented with a quadratic phase increment of 117°for RF spoiling. This was followed by a fully-sampled 3D cartesian readout with outside-in ordering (i.e., sampling the center last), with a padding factor of 1.2. The default geometry properties were 256x256 acquisition

matrix, 25.6 cm FOV and 20 partitions in the slab-selection direction, yielding 1x1x3 mm resolution. The readout gradient had a rewinder lobe with 2 cyc/pixel net area and was followed by a spoiling gradient with an area of 40 mT·ms/m.

For the magnetization transfer (MT) saturation, a Fermi pulse (duration = 12ms, B1rms = 3.64 μT, frequency offset = 1.2kHz, transition width = 0.35, max B1 = μT, pulse angle= 490°) was designed as an optional block that can be enabled/disabled by the user. Finally, a loop command was defined for the sequence to iterate through three sets of parameters, defined by the user in the UUI for a complete MTsat protocol.

From this protocol we acquired three images: (i) PD-weighted SPGR with no MT, FA = 6°and TR = 32 ms (ii) MT-weighted SPGR with MT, FA = 6°and TR = 32 ms (iii) T1-weighted SPGR without MT, FA = 20°and TR = 18 ms. From images (i) and (iii) we computed a T1 map, from images (i) and (ii) we computed an MTR map, and from images (i), (ii) and (iii) we computed an MTsat map.

6.3.4   Data acquisition

Experimental procedures agreed with the guidelines and regulations concerning human welfare and experimentation set by Montreal Heart Institute Ethics Committee, Montreal, Canada and Sunnybrook Health Center, Toronto, Canada. Three healthy male participants (age = 37 ± 6, height = 178 ± 6 cm and body mass = 81 ± 4 kg) volunteered for multi-center data collection. Following a detailed explanation of the purpose and the methodology of the experiment, the participants gave their written informed consent.

The participants and the ISMRM-NIST system phantom (manufactured by HPD Inc., SN = 42) were scanned on three imaging systems at two imaging sites. In S1 and S2, the phantom was scanned using a 20-channel head coil due to space constraints, whereas a 32-channel coil was used in G1. For in-vivo imaging, 32-channel head coils were used in G1 and S1, whereas a 64-channel coil was used in S2. S1 was equipped with an XR-K2309_2250V_951A (Siemens Healthineers, Erlangen,Germany) gradient system (80 mT/m maximum amplitude and 200 T/m/s slew rate per axis, 48 cm maximum FOV), S2 with an XQ-K2309_2250V_793A (Siemens Healthineers, Erlangen,Germany) gradient system (45 mT/m maximum amplitude and 200 T/m/s slew rate per axis, 48cm maximum FOV) and G1 with a 8920-XGD (GE Healthcare, Milwaukee, USA) gradient system (50 mT/m maximum amplitude and 200 T/m/s slew rate per axis, 48cm maximum FOV). The nominal field strengths on G1 and S1-2 were 3T and 2.89T, respectively. Before the scan, the system phantom was kept in the imaging site for at least a day, and in the scanner room for at least 3 hours. The measured bore temperature in G1, S1 and S2 was 20.1°C, 20.2°C and 20.8°C, respectively.

The acquisition parameters were set according to a generic protocol established for MTsat imaging of neural tissue (Cohen-Adad, Alonso-Ortiz, Abramovic, Arneitz, Atcheson, Barlow, Barry, Barth, Battiston, and Büchel, 2021). Vendor-neutral acquisition parameters were identical on all systems. However, it was not possible to equalize all the parameters between the vendor-native protocols. Comparison of vendor-native and vendor-neutral protocols are presented in Table 6.1 and Table 6.1. To scan the system phantom, prescan measurements were performed as described by Keenan et al. (2021) and the vendor-neutral acquisitions were configured to start the acquisitions with these calibrations (Keenan, Gimbutas, Dienstfrey, Stupic, et al., 2021). For all acquisitions, the prescan settings of the initial T1w acquisition were used for the subsequent PDw and MTw acquisitions on all vendor systems. Necessary care was taken not to perform additional calibration measurements between acquisitions. For the VENUS acquisitions, B0 shimming gradients were set using a spiral multi-echo (four echoes) gradient-echo sequence by fitting the field map obtained after water/fat separation to the first order in 3D. Gradient non-linearity correction was performed as part of the on-site reconstruction pipeline (Figure 6.1f) and the relevant warping coefficients are reported in the image metadata for offline reconstruction. For the systems S1-2 the identical protocol was used by exporting the vendor-native protocol files from S2, whereas the protocols for G1 were set on-site.

Table 6.1 Parameters that are common between VENUS and vendor-native acquisition protocols.

	Common acquisition parameters
FA (°) PDw/MTw/T1w	6/6/20
MT PDw/MTw/T1w	off/on/off
Voxel size (mm)	1x1x3
TR (ms) PDw/MTw/T1w	32/32/18
TE (ms)	4
FOV (cm)	25.6
Readout sampling rate (kSamp/s)	62.5
MT frequency offset (Hz)*	1200

Table 6.2 Parameters that are equal among VENUS, but vary between vendor-native acquisitions.

Scanner ID and sequence type	G1$_{\text{NATIVE}}$	S1$_{\text{NATIVE}}$	S2$_{\text{NATIVE}}$	VENUS
Sequence name	3D SPGR	3D FLASH	3D FLASH	mt_sat (v1.1.0)
MT pulse shape*	Fermi	Gaussian	Gaussian	Fermi
MT pulse duration (ms)*	8	10	10	12
RF phase increment (°)*	115.4	50	50	117

6.3.5 Data processing

Vendor-native DICOM images were converted to BIDS v1.6.0, whereas vendor-neutral images were already exported in the BIDS format. All the processing was performed using data-driven and container mediated qMRFlow pipelines comprised of two docker images (qmrlab/antsfsl:v2.5.0b and qmrlab/minimal:v2.5.0b) that are published by qMRLab using a continuous deployment pipeline. Quantitative fitting was performed in qMRLab (Karakuzu, Boudreau, Duval, et al., 2020) v2.5.0b. Pre-processing steps were performed using ANTs (Avants et al., 2011) for registration and FSL (Jenkinson et al., 2012) for automatic gray-matter (GM) and white-matter (WM) segmentation.
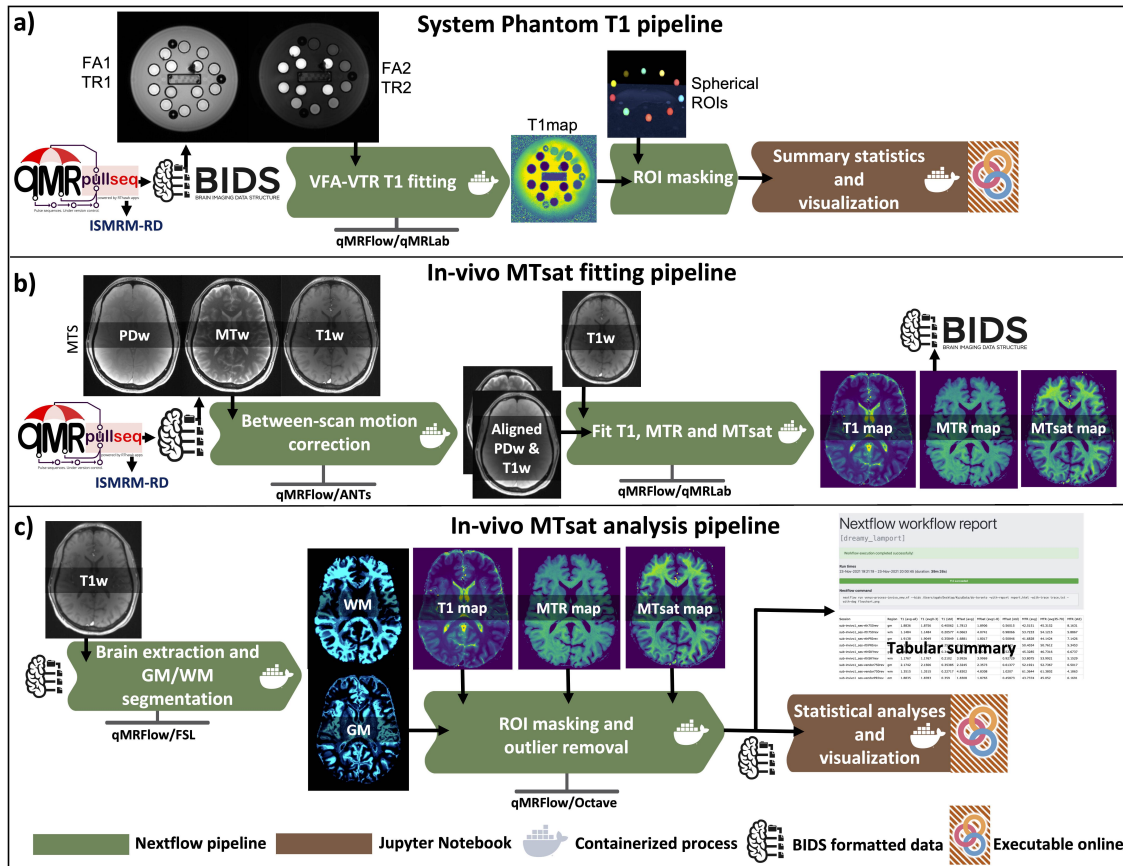
Figure 6.2 Overview of the analysis workflow for phantom scans (a) and in vivo scans (b, c). File collection (MTS) and output map names (T1map, MTsat, MTRmap) follow the BIDS standard v1.6.0. **a)** Vendor-neutral and vendor-native ISMRM system phantom images were acquired at two flip angles and two repetition times. The output data are then subjected to T1 fitting using qMRLab (Docker container image: qmrlab/minimal:v2.5.0b). The resulting T1 maps are masked using manually prescribed 10 spherical ROIs (reference T1 ranging from 0.9 to 1.9s). **b)** PDw and MTw images are aligned to the T1w image to correct for between-scan motion. The aligned dataset is then subjected to MTsat and MTR fitting in qMRLab to generate T1map, MTRmap and MTsat. **c)** Brain extraction and tissue type segmentation is performed on the T1w images using FSL. Following region masking and outlier removal for each map, vector outputs are saved for statistical analysis and visualization in an online-executable Jupyter Notebook (R-Studio and Python) environment. The tabular summary and the Nextflow pipeline execution report are exported. The pipeline execution report is available at https://qmrlab.org/VENUS/qmrflow-exec-report.html.

The system phantom T1 pipeline consisted of linearized T1 mapping in the VFA framework by accounting for varying TRs (Helms, Dathe, and Dechent, 2008). Resultant phantom T1 maps were then masked using spherical ROIs as described in (Keenan, Gimbutas, Dienstfrey, and Stupic, 2019). The center of each sphere was manually prescribed per dataset (3D Slicer

v4.11) at the 10 highest T1 reference spheres of the system phantom (Figure 6.4a). For the in-vivo data, between-scan motion correction was performed by aligning PDw and MTw images onto the T1w, filled by MTsat fitting (Figure 6.2b). Brain region segmentations were performed on the T1w images and ROI masking was performed to prepare data for statistical analyses (Figure 6.2c).

6.3.6 Statistical analyses

All the descriptive statistics were reported by the processing pipeline in tabular format for phantom and in-vivo maps (available at https://osf.io/5n3c). Vendor-neutral and vendor-native phantom measurement performances were compared against the reference (Figure 6.4b,c) and percent deviations from the ground truth were reported (Figure 6.4d).

Kernel density estimates of the T1, MTR and MTsat distributions in WM and GM were visualized as ridgeline plots for one participant (Fig, 5d-i). Before the statistical comparisons in WM, these distributions were filtered to a relevant range discerned from the ridgeline plots (Figure 6.5d-i) to remove spurious values captured by masking. The range was 0 to 3s for T1, 35 to 70% for MTR and 1 to 8 for MTsat. Filtered distributions were then subjected to random sampling (the random number generator was seeded for reproducibility) to obtain an equal number of WM voxels (N = 37,000), enabling a balanced comparison.

Percentile bootstrap based shift function analysis (Rousselet et al., 2017) was performed to compare dependent measurements of T1, MTR and MTsat in WM (N=37,000) between different systems (G1 vs S1, G1 vs S2 and S1 vs S2) for VENUS and for the vendor-native implementations. Deciles of the distributions were computed using Harrell-Davis quantile estimator (Harrell and Davis, 1982) and the decile differences were calculated using 250 bootstrap samples. Note that unlike traditional significance tests that operate on a single point estimate of the central tendency, shift function analysis can characterize differences at any location of the distributions (Figure 6.6a). Even for a single participant, this provides a powerful graphical tool to assess how vendor-neutral and vendor-native measurements differ between different systems throughout the relevant range of the distributions. To supplement the shift function plots, we annotated the 5th decile (median) differences using respective percent differences (Figure 6.6b-d). The shift function was then extended to a hierarchical design to characterize the overall difference between scanners for multiple participants, for VENUS and for the vendor-native WM distributions of all studied metrics. A hierarchical shift function chart shows individual between-scanner differences for each participant, along with averaged group differences plotted against a categorical axis at each decile (Figure 6.7a). Similarly, percent deviations between scanners at the median deciles were annotated

per subject and the average percent deviations were reported (Figure 6.7b-d). Please note that the hierarchical shift function by default calculates high density intervals to test the significance of averaged decile differences against the null value of zero using a t-test at trimmed means. Given the high number of samples and the fact that the hypothesis aims at comparing between-scanner differences in a VENUS vs vendor-native setting, these intervals are excluded from the visualization for simplicity. The reader is welcome to reproduce these figures online, where necessary changes can be made to visualize the intervals at https://github.com/qMRLab/VENUS.

Finally, quantitative measurement discrepancies of vendor-native and VENUS implementations between different vendors (G1-S1 and G1-S2) were compared using Wilcoxon signed rank test. The comparison was performed on the G1-S1 and G2-S2 percent absolute differences of T1, MTR and MTsat in white-matter between vendor-native and vendor-neutral implementations. The level of significance was set at 0.05.

6.4 Results

Figure 6.3 compares basic image quality metrics across phantom images. The contrast transitions across 10 spheres included in the analysis are discernible and qualitatively comparable in all T1w images (Figure 6.3h-m). In addition, VENUS and vendor-native peak SNR values are on a par with each other, ranging from 25.4-27.8 dB (PDw) and 28.5-31.6 dB (T1w, Figure 6.3a). According to the American College of Radiology (ACR) guidelines, the number of visible bright spots within a high-contrast resolution insert determines the resolution performance (of Radiology, 2019). In the system phantom, 4x4 1 mm coarse resolution arrays are drilled into an 8 mm thick polyphenylene sulfide plate (plate 4), which falls into the FOV when the T1 array (plate 5) is centered in a 60mm slab (Stupic et al., 2021). Overall, the resolution markers are conspicuous in vendor-neutral images (Figure 6.3b-d) with a slightly less horizontal resolution compared to the S1-2$_{\text{NATIVE}}$ (Figure 6.3f,g). Although the number of spots is discernible (Figure 6.3e), the insert pattern in the G1$_{\text{NATIVE}}$image appears smoother.
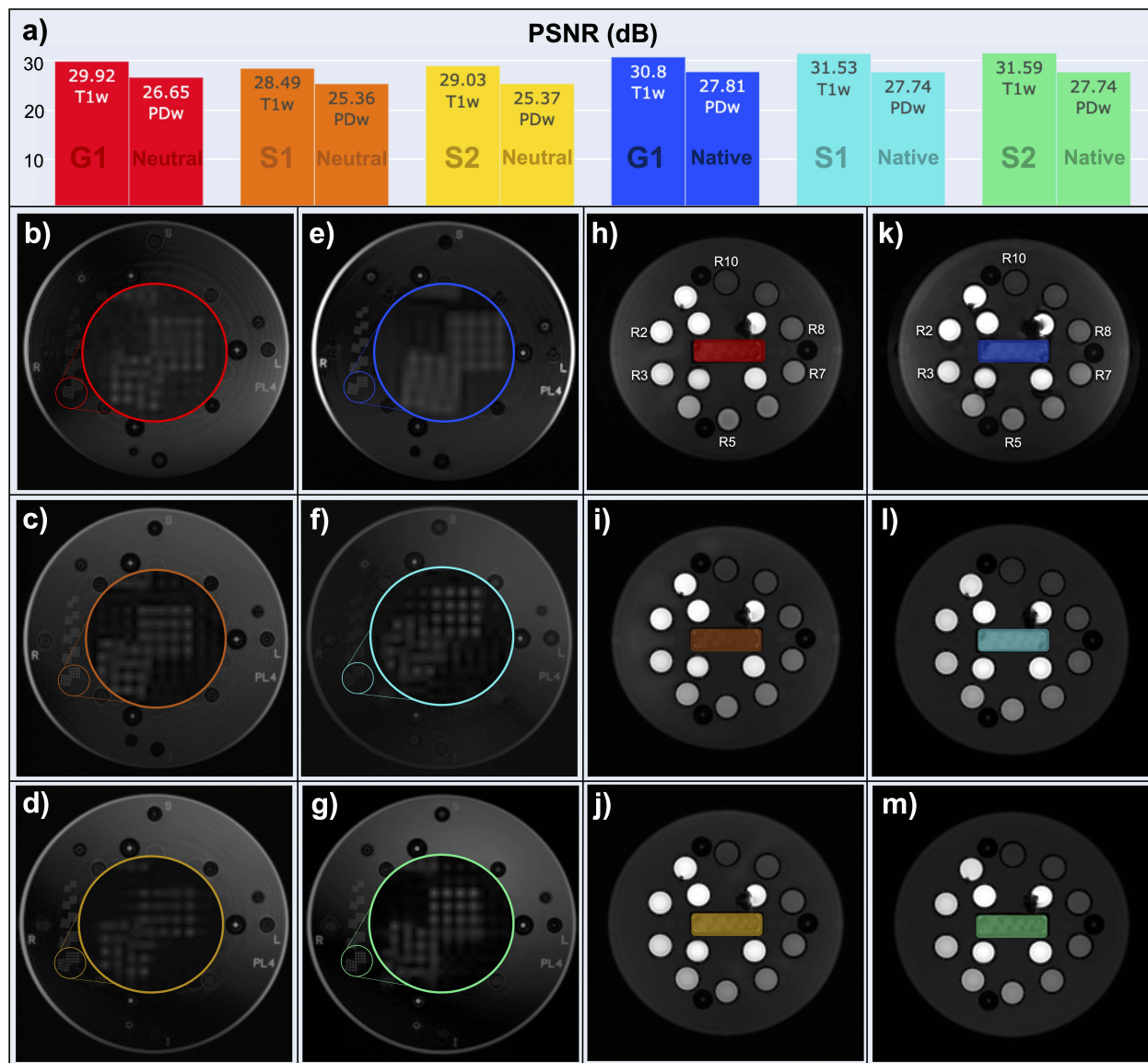
Figure 6.3 Image quality assessment using the system phantom: **a)** Peak SNR values (PSNR) from T1w and PDw phantom images are displayed for vendor-neutral (red, orange, and yellow) and vendor-native (blue, cyan, and teal) G1, S1 and S2 scans, respectively. The same color coding is used in the following panels. **b-g)** Coronal PDw phantom images, with an inset zoom on two 4x4 grids with 1mm spacing. The brightness of the zoomed-in insets is increased by 30% for display purposes. **h-m)** Coronal T1w phantom images showing the center of the reference T1 arrays. The fine resolution (<0.6mm) inserts located at the center of the T1 array (rectangular area) are not relevant for the present resolution level. These inserts are colored following the same convention described in a) for convenience.

Figure 6.4 shows the T1 mapping results from the NIST phantom. Shades of blue (light blue, cyan, teal) stand for the vendor-native T1 measurements, and shades of red (red, orange,

yellow) stand for the VENUS T1 measurements. The ground truth NIST phantom values are marked with a cross.
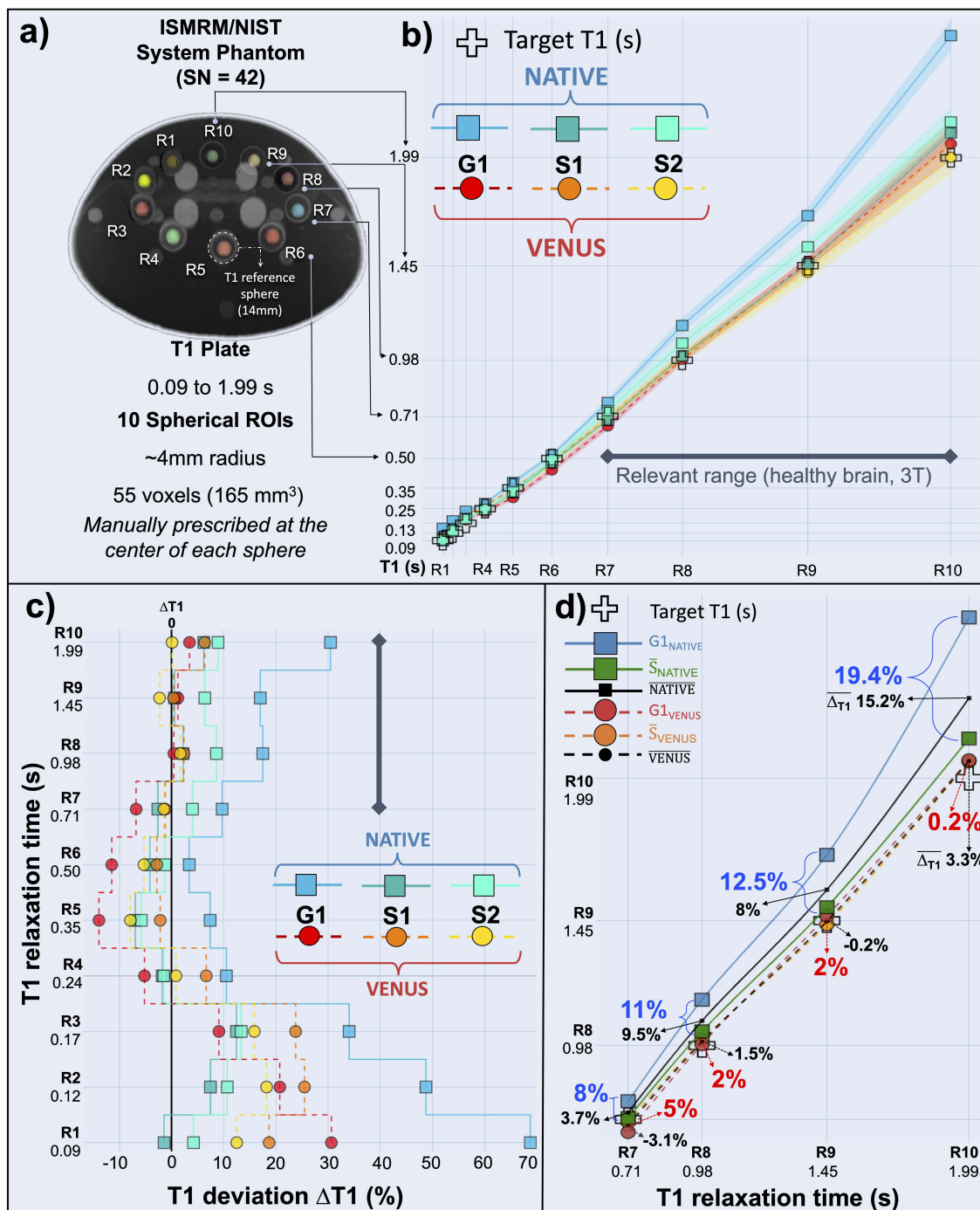
Figure 6.4 Comparison of vendor-native and vendor-neutral T1 measurements in the studied range of ISMRM-NIST system phantom reference values, from 0.09 to 1.99s **(a)**. T1 values from the vendor-native acquisitions are represented by solid lines and square markers in cold colors, and those from VENUS attain dashed lines and circle markers in hot colors. **b)** Vendor-native measurements, especially G1$_{\text{NATIVE}}$ and S2$_{\text{NATIVE}}$, overestimate T1. G1$_{\text{VENUS}}$ and S1-2$_{\text{VENUS}}$ remain closer to the reference. **c)** For VENUS, ΔT1 remains low for R7 to R10, whereas deviations reach up to 30.4% for vendor-native measurements. **d)** T1 values are averaged over S1-2 ($\overline{S}_{\text{NATIVE}}$ and $\overline{S}_{\text{VENUS}}$, green square and orange circle) and according to the acquisition type ($\overline{NATIVE}$ and $\overline{VENUS}$, black square and black circle). Inter-vendor percent differences are annotated in blue (native) and red (VENUS). Averaged percent measurement errors ($\overline{\Delta T1}$) are annotated on the plot (black arrows).

The vendor-native implementations result in overestimation of the ground truth T1 values at different rates, particularly for the physiologically relevant T1 values in the interval from 0.7 to 1.9s (Figure 6.4b). Ground truth T1 deviations (ΔT1) calculated by percent error indicate that $G1_{NATIVE}$ and $S2_{NATIVE}$ exhibit a persistent overestimation trend, whereas $S1_{NATIVE}$ shows better accuracy approaching that of VENUS T1 estimations in the relevant range (Figure 6.4c). Within the same interval, the highest deviation is observed for $G1_{NATIVE}$, ranging from 9.7 to 30.4% for ΔT1. For R4-6, $G1_{NATIVE}$ and $G1_{VENUS}$ T1 measurements straddle the reference, where $G1_{VENUS}$ shows 5.1-13.8% underestimation and the $G1_{NATIVE}$ overestimation remains within the 3.4-10.5% interval (Figure 6.4c). For lower T1 reference values (T1 < 170ms), all measurements indicate higher deviations, with $S1\text{-}2_{NATIVE}$ performing better than $S1\text{-}2_{VENUS}$. When the measured T1 values are averaged over S1-2 ($\overline{S}$), the differences between $G1_{NATIVE}$ and $\overline{S}_{NATIVE}$ are 8, 11, 12.5 and 19.4%, whereas the differences between $G1_{VENUS}$ and $\overline{S}_{VENUS}$ are 5, 2, 2 and 0.2% for R7-10, respectively (Figure 6.4d). This reduction in between-vendor differences brought by VENUS is coupled with an improvement in accuracy. When averaged according to the implementation type, average VENUS deviation ($\overline{\Delta T1}_{VENUS}$) falls within the 0.2 - 4% range and $\overline{\Delta T1}_{NATIVE}$ ranges from 7 to 11%. Even though $G1_{NATIVE}$ has the dominant contribution to the higher $\overline{\Delta T1}_{NATIVE}$ values, Figure 6.4d shows that $\overline{S}_{VENUS}$ is closer to the reference than $\overline{S}_{NATIVE}$ for the majority of the R7-10 (ΔT1 of 7.6, 3.5, 5.4, 0.7% and 3.2, 0.9, 2, 1.3% for $\overline{S}_{NATIVE}$ and $\overline{S}_{VENUS}$, respectively). As a result, VENUS reduces between-vendor differences with an overall accuracy improvement.
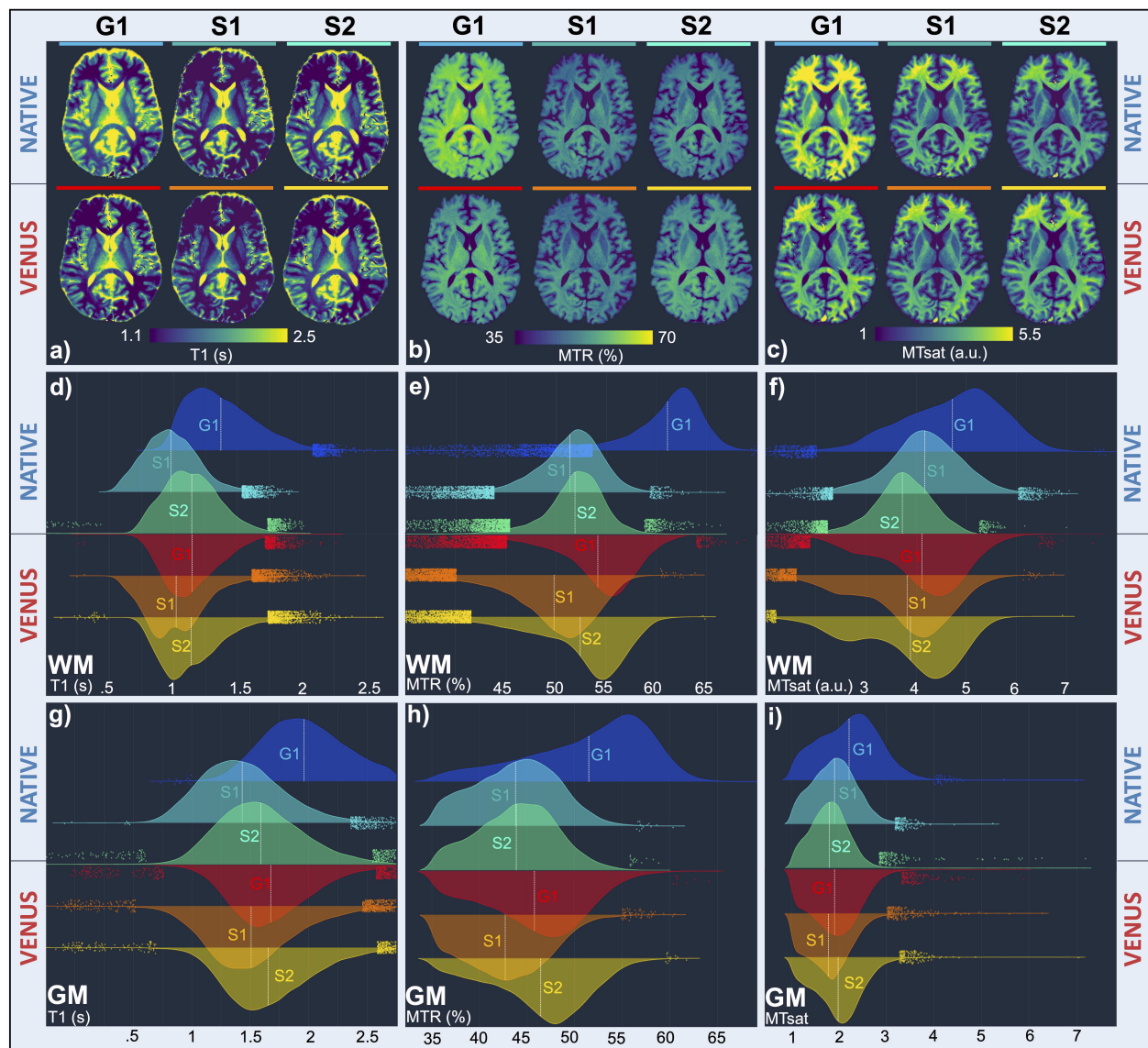
Figure 6.5 Vendor-native and VENUS quantitative maps from one participant are shown in one axial slice (a-c). Distributions of quantified parameters in white matter (d-f) and gray matter (g-i) are shown using ridgeline plots of kernel density estimations. **a-c)** Inter-vendor images (G1 vs S1 and G1 vs S2) are appear more similar in VENUS (lower row) than in native (upper row). **d-f)** Distribution shapes and locations agree with visual inspection from (a), indicating closer agreement among VENUS sendistributions. **g-i)** Superior between-scanner agreement of VENUS persists in GM as well. Compared to WM, GM distributions are in the expected range (higher T1, lower MTR and MTsat values).

Figure 6.5 shows in vivo T1, MTR and MTsat maps from a single participant (P3). While most of the improvements are evident from the maps (5a - 5c), the ridgeline plots (5d – 5i) make it easier to appreciate the distribution range of values and VENUS vs vendor-native

between-scanner agreement in the GM and WM per metric. Conforming to the higher myelin content in WM, T1 values are lower (around 1.1 ± 0.2s, Figure 6.5d), whereas MTR and MTsat values are higher (around $50 \pm 8\%$ and 3.8 ± 0.9 a.u., Figure 6.5e,f) in comparison to those in GM (1.9 ± 0.4s, $40 \pm 2\%$ and 1.8 ± 0.5, for T1, MTR and MTsat, respectively, Figure 6.5g-i). The general trend observed in the images is captured by ridgeline plots, showing better agreement between VENUS distributions of G1, S1 and S2. This is further supported by between-scanner coefficient of variation (CoV) per metric (Table 6.3), showing that VENUS reduces the CoV from 16.5, 10.1 and 12.5% to 6.1, 4.1 and 4.1% for T1, MTR and MTsat, respectively. This indicates a sizable decrease in between-scanner variability using VENUS compared with vendor-native measurements and the trend is consistent across participants. However, CoV is a relative statistic of dispersion and does not characterize the nature and absolute magnitude of changes between the measurements. Such characterization is important to evaluate the between-scanner performance of a quantitative measure in distinguishing biological variability. To enable such comparisons, Figure 6.6b-d shows pairwise difference of each metric between scanners using shift functions, which are constructed by absolute differences at each decile of the compared distributions (Figure 6.6a). VENUS and vendor-native between-scanner differences for all metrics are annotated by percent differences at the median (fifth) decile for simplicity.

Table 6.3 Coefficient of variation (%) of vendor-neutral (VNS) and vendor-native (NTV) quantitative measurements between the scanners for each and across participants.

Participants	P1		P2		P3		Across	
Protocol	NTV	VNS	NTV	VNS	NTV	VNS	NTV	VNS
T1	9.8	1.3	11.7	4.0	16.5	6.1	11.36	4.3
MTR	8.5	3.4	8.5	3.4	10.1	4.1	7.9	3.2
MTsat	13.6	5.9	11.9	3.4	12.1	4.1	10.7	4.2

Going from vendor-native (top rows, blue panels) to VENUS (bottom rows, red panels), Figure 6.6b-d indicate a decrease in T1, MTR and MTsat WM differences between scanners from different vendors (G1 vs S1 and G1 vs S2) for P3, without exception and throughout the deciles. As for within-vendor comparisons (S1 vs S2), VENUS reduces difference scores for T1 and MTsat by 5.8 and 7.8% while increasing that for MTR by 5.3%. One can also appreciate the changes in shift function shapes. For example, shift function for $G1_{NATIVE}$ vs $S2_{NATIVE}$ MTsat comparison in Figure 6.6d shows a positive linear trend, indicating that WM voxels with higher MTsat values tend to show a higher between-vendor difference. On

the other hand, G1$_{\text{VENUS}}$ vs S2$_{\text{VENUS}}$ MTsat shift function appears flatter, describing a more uniform (and reduced) bias throughout the WM distribution.
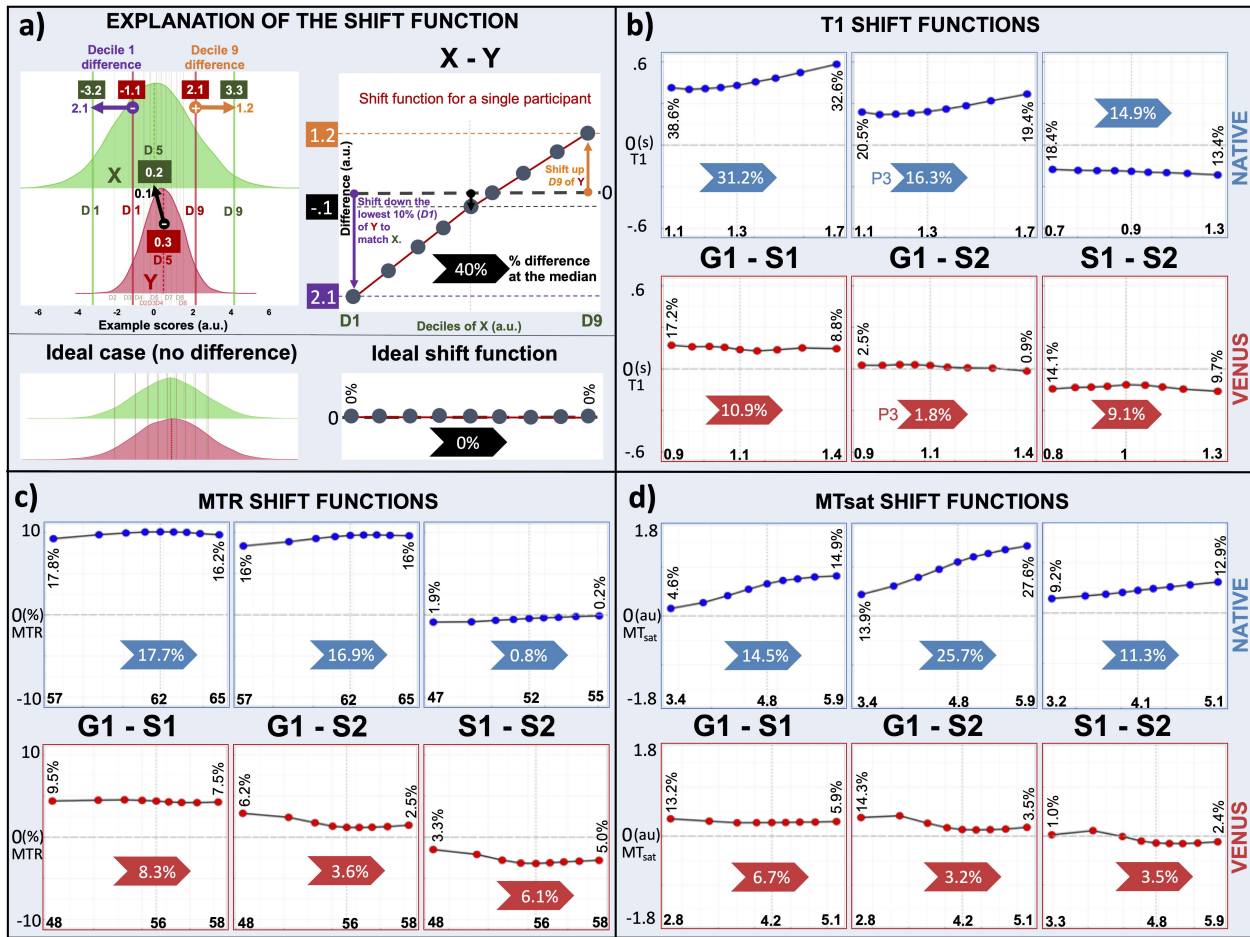
Figure 6.6 Shift function analysis of T1, MTR and MTsat results from a single participant in white-matter (WM). **a)** Shift function analysis is a graphical tool for analyzing differences between two (dependent in this case) measurements at any location of the distributions. It shows 9 markers dividing the distribution into 10 equal chunks; hence the markers represent deciles. The shape of the curve (shift function) obtained by plotting decile differences against the first decile characterizes how distributions differ from each other. **b-d)** Here, shift function plots compare the agreement between different scanners for VENUS (bottom row) and vendor-native (top row) implementations in quantifying T1, MTR and MTsat. Across all the comparisons, the apparent trend is that the VENUS inter-vendor variability is lower than for the vendor-native implementations.

Figure 6.7 expands on Figure 6.6 for multiple participants by overlaying individual shift functions (shades of pink) and illustrating the across-participants trend using group shift functions that are red for VENUS and blue for vendor-native differences (Figure 6.7a). A general observation is that individual shift function shapes are consistent across participants,

indicating that the nature of between-scanner VENUS and vendor-native differences are not affected by anatomical differences. However, the magnitude of the difference is participant-specific. For example, P3 shows the highest $G1_{NATIVE}$ vs $S1_{NATIVE}$ T1 difference of 31.2%, which is followed by 20.9 and 14.2% for P2 and P1, respectively (Figure 6.7b). The same participant order is preserved for $G1_{VENUS}$ vs $S1_{VENUS}$; however, with consistently lower decile differences, which is what stands out in Figure 6.7: participant-averaged between-vendor T1, MTR and MTsat WM differences are smaller for VENUS than they are for vendor-native implementations, without exception and throughout the deciles. Overall, VENUS G1 – S1 and G1 – S2 differences are on the order of 2.3 to 7.9%, which starts from 13.8% and extends up to 25.6% for vendor-native implementations, averaged across participants. The reduction in between-vendor differences achieved by VENUS is significant after correction for multiple comparisons for all maps (p=0.015). Within-vendor difference trends observed for P3 (Figure 6.6b-d) hold true across participants, indicating a 3.7% increase in S1 – S2 MTR difference achieved by VENUS, coupled with a 3.2% and 2% decrease in S1 – S2 T1 and MTsat differences, respectively (Figure 6.7b-d).
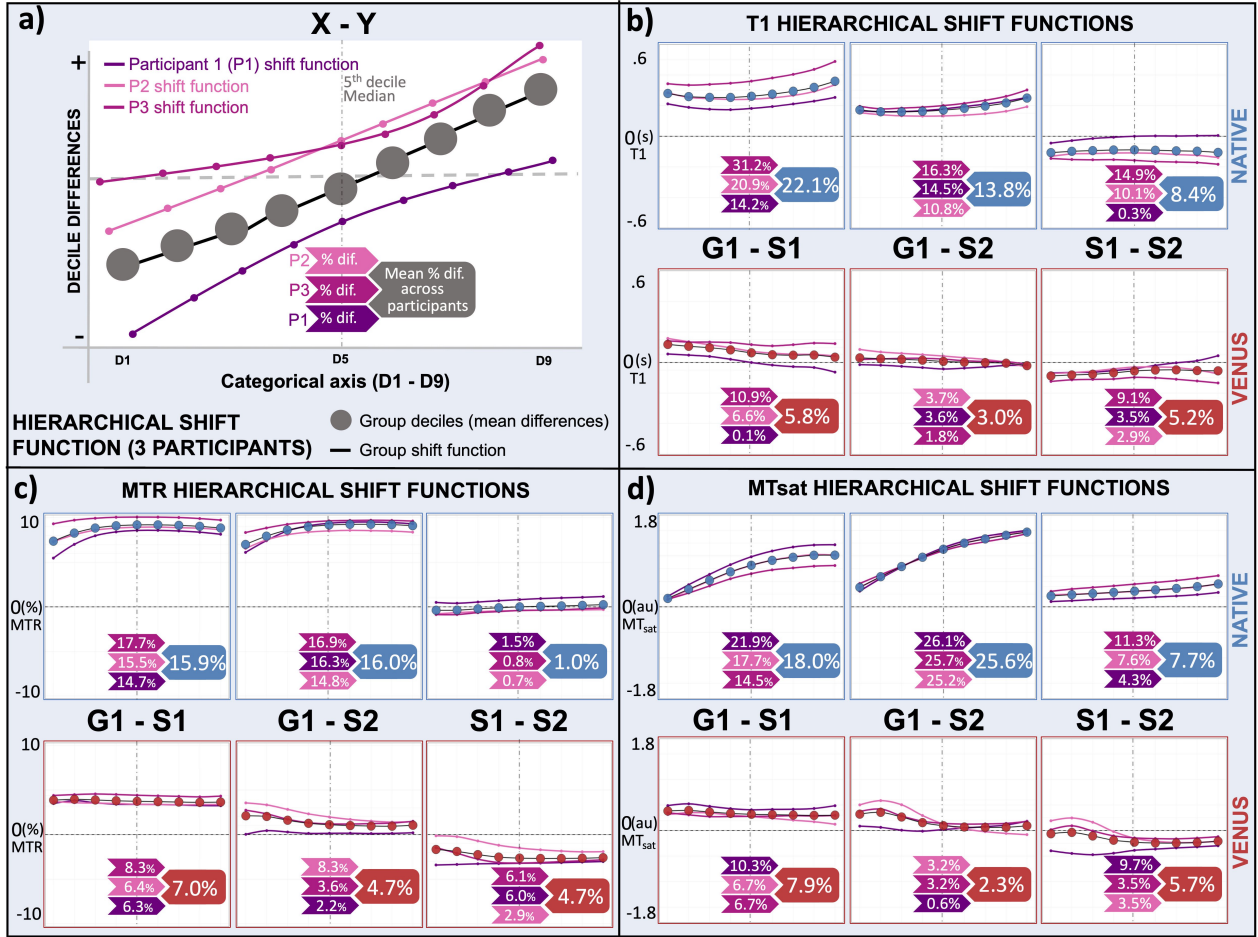
Figure 6.7 Hierarchical shift function analysis of T1, MTR and MTsat results from three participants in the white-matter (WM). **a)** Hierarchical shift function repeats Figure 6.6 for all participants (shades of pink). Group deciles (red and blue markers for VENUS and vendor-native, respectively) show the average trend of inter-scanner differences across participants. **b-d)** G1-S1 and G1-S2 (inter-vendor) agree in VENUS better than they do in vendor-native for all quantitative maps of T1, MTR and MTsat.

6.5   Discussion

In this study, we developed and deployed a vendor-neutral qMRI protocol (VENUS) for T1, MTR and MTsat mapping on three 3T commercial scanners by two vendors. Three healthy participants were scanned to calculate all the metrics, as well as the ISMRM/NIST phantom for T1 mapping and fundamental quality assessments. Our findings confirm the hypothesis that vendor-neutral sequences decrease inter-vendor variability of T1, MTR and MTsat measurements. This key improvement addresses problem 1, as stated in the Introduction, with transparent and publicly available sequence descriptions. The developed sequence can

be run on most GE and Siemens scanners through a UUI that allows users to prescribe customized file naming entities for exporting reconstructed images in the brain imaging data structure (BIDS) (Gorgolewski, Auer, et al., 2016; Karakuzu, Appelhoff, et al., 2021) and k-space data in the ISMRM-RD format (Inati et al., 2017). Conforming with community data standards, providing a user-friendly end-to-end solution with a simplified inter-vendor deployment, this work offers a complete solution for problem 2 and shows a way forward for the standardization of qMRI.

6.5.1 Developing an end-to-end qMRI workflow

First, we created a vendor-native qMRI protocol that is unified across vendors to the greatest extent possible, by keeping contrast, timing, and acquisition geometry identical (Table 6.1). However, other vendor-native implementation details such as RF spoiling, MT and excitation pulse characteristics were different (Table 6.2), as it is commonly the case in multicenter studies (Lee et al., 2019; Yarnykh, 2010). Trying to address these issues is difficult with a vendor-native sequence given that the implementations of commercial stock sequences commonly used for qMRI are not open (problem 1). One candidate solution for this problem is modifying sequences on the vendor's proprietary development environment to equalize implementations as much as possible, which has been shown to improve reproducibility to some extent (Gracien et al., 2020). However, this requires familiarity with multiple sequence development environments and still may fall short in unifying all the aspects of on-site acquisitions (e.g., different parallel imaging implementations). Not only is this approach impractical for the developers, but it is also not a user-friendly solution for clinical use. As we mention in the context of problem 2, reproducibility solutions unifying inter-vendor implementations become more favorable if they are designed with clinicians' needs in mind. To that end, we aimed at providing a unified and smooth user experience by developing VENUS as an RTHawk application, which allows implementation details to be shared publicly starting at the pulse sequence level.

Second, we built from scratch a vendor-neutral sequence that was developed and tested on a single site and then ported to two more scanners from different vendors. In doing so, we adapted a system that is primarily geared toward real-time imaging (RTHawk) to perform quantitative MRI measurements. For example, absolute gradient limits have been allowed to achieve higher spoiling gradient moments and string-valued customized metadata injection has been made possible to comply with community data standards (Gorgolewski, Auer, et al., 2016; Inati et al., 2017; Karakuzu, Appelhoff, et al., 2021). Nonetheless, there is still room for improvement to make the most out of the hardware specifications in using RTHawk for qMRI.

For example, the gradient duty cycle specifications of all the scanners in the present work were suitable for playing high spoiling gradients concurrently long enough to achieve spoiling levels suggested for an accurate AFI B1 mapping (Yarnykh, 2007b). However, it is not possible to implement gradient durations up to 80ms on RTHawk v3.0.0 (rc4-28-ge3540dda19) due to conservative hardware safety checks on the minimum TR, as the acquisition controller assumes that the designed TRs will repeat for an indefinite period. This limitation will be removed in a future release to regulate hardware and patient safety limits by taking a finite sequence run into account.

Third, we created a fully transparent, container-mediated and data-driven workflow (Karakuzu, Boudreau, Cohen-Adad, et al., 2020a; Karakuzu, Boudreau, Cohen-Adad, et al., 2020b) that automates the processing and reduces variability introduced by the operators. By design, the workflow operates according to the BIDS qMRI standard (Karakuzu, Appelhoff, et al., 2021) for picking up all the necessary data and metadata, and generates outputs following a consistent derivative hierarchy. Moreover, the raw data is exported in the ISMRM-RD format by our vendor-neutral sequence, allowing the use of community developed reconstruction tools by simply adding another container at the beginning of our modular workflow. We envision that using open-source reconstruction tools would be highly favourable for vendor-neutral sequences employing under-sampled k-space with complex trajectories to guarantee reproducibility (Assländer et al., 2018; Hansen and Sørensen, 2013; Knopp and Grosser, 2021; Maier, Baete, et al., 2021). In the present work, we analyzed the images reconstructed using the RTHawk pipeline (https://github.com/qMRLab/VENUS), given that we only performed a simple multi-coil reconstruction.

6.5.2 Reducing inter-vendor variability

Stock sequences are optimized for reliable clinical imaging. These optimizations do not necessarily serve for accuracy when the sequences are used for qMRI experiments. For example, phase increment value of S1-2$_{\text{NATIVE}}$ sequences (Table 6.2) are hardcoded to maximize in-vivo signal stability (Preibisch and Deichmann, 2009), not T1 accuracy in phantoms (Heule et al., 2016). On the other hand, that of G1 has been shown to be unsuitable for T1 mapping, exhibiting severe overestimations(Yarnykh 2007b). In this study, we set this value for T1 accuracy (Yarnykh, 2010) while unifying all other aspects of the vendor-neutral acquisition between scanners. Results from phantom analysis clearly demonstrate that VENUS achieves higher accuracy and a notable reduction in inter-vendor variability compared to its native counterparts (Figure 6.4).

Our in vivo measurements are more difficult to evaluate because there is no ground truth mea-

surement. In the absence of an inversion recovery T1 map, we only looked at the agreement between the three implementations and explored whether VENUS brought the T1 values closer across vendors when compared to the vendor-native sequences. Visually (Figure 6.5a), the reduction in T1 variability can be appreciated for VENUS within the dynamic range of T1 adjusted for WM/GM. As supported by the ridgeline plots (Figure 6.5d,g), the $G1_{NATIVE}$ T1 distribution is globally shifted towards higher values compared to $S1-2_{NATIVE}$, and their central tendency differs. As observed in phantom, $G1_{VENUS}US$ alleviates this discrepancy, shifting the T1 distribution closer to those of $S1-2_{VENUS}$. Interestingly, WM T1 distributions show a more unimodal attribute on G1 than they do on S1-2 both for VENUS and vendor-native, with a more pronounced bimodal appearance for $S1-2_{VENUS}$. A plausible explanation for that is vendor and implementation specific differences resulting in discrepancies between B1+ field inhomogeneity. Nevertheless, the VENUS shift functions for G1 vs S1 and G1 vs S2 comparisons are flatter than the vendor-native shift functions (Figure 6.6b), indicating that the inter-vendor WM T1 statistical distribution characteristics are more similar using VENUS.

Table 6.3 indicates that reduction in inter-vendor variability is not limited to T1 but persists for all the metrics across all participants. The inter-vendor variability in MTR and MTsat is relatively easier to appreciate visually (Figure 6.5b,c). The three MTR and MTsat maps from VENUS are in better agreement, and this is most likely because our unified implementation compensated for MT saturation pulse differences (Table 6.2).

S1-2 run the exact same vendor-native software versions, so we do not expect a big intra-vendor difference for any of the parameters. However, T1 and MTsat agreements slightly improved using VENUS (Figure 6.7b,d). On the other hand, within-vendor agreement was better using native sequences for MTR. Even in that case, transforming a vendor-native bias to a vendor-neutral one is more advantageous to establish a standardized measurement. Knowing more about the vendor-native implementations (e.g., why a particular MT pulse was selected as the default stock option for that vendor) would help pinpoint the underlying causes of variability. This appears particularly important for multiparametric mapping methods as the combined effect of different variability sources from individual maps leads to a higher between-vendor disagreement in combined quantitative maps (Figure 6.7d).

As reducing variability matters, so does how we assess it. Shift functions (Rousselet et al., 2017) take the comparison beyond differences in point estimates of centrality and relative spread (CoV) to a robust characterization of differences on the absolute scale of the measurement. Figure 6 makes use of 80% of the data collapsed into a single relative statistic by Table 2 to explain how distributions differ for P3. For example, Table 2 shows that

VENUS reduces CoV from 12.1 to 4.1% for MTsat. Figure 6d explains that majority of that reduction is achieved by decreasing absolute G1 vs S2 MTsat difference from 1.1 to 0.1 (a.u.), corresponding to a reduction from 25.7% to 3.2% inter-vendor difference. In addition, Figure 6.6d indicates that higher deciles benefit from the G1 vs S2 variability reduction more compared to the lower deciles, yielding a flatter shift function for VENUS. This suggests that VENUS not only brings averaged MTsat values closer, but also matches their distribution shape (Figure 6.5f). Implications of vendor-neutrality and the importance of transparency

The most important contribution of this article is the vendor-neutral solution it provides for multi-center reliability of qMRI by significantly reducing inter-vendor variability. This has been a plaguing issue for standardization of qMRI methods in validation (Lazari and Lipp, 2021; Mancini et al., 2020), multi-center clinical trials (Ashton, 2010), establishing protocols (Cohen-Adad, Alonso-Ortiz, Abramovic, Arneitz, Atcheson, Barlow, Barry, Barth, Battiston, and Büchel, 2021), applied neuroimaging studies (Boshkovski et al., 2021), determining the range of parameters in disease typing (Goebel et al., 2016; Young, 2004) and health (Bojorquez et al., 2017; Lee et al., 2019), between scanner upgrades (Keenan, Gimbutas, Dienstfrey, and Stupic, 2019) and even in phantom studies (Bane et al., 2018; Keenan, Gimbutas, Dienstfrey, Stupic, et al., 2021). By reducing such variabilities, VENUS approach can bring qMRI closer to teasing out the true biological variability in quantifying in-vivo tissue microstructure (Weiskopf, Edwards, et al., 2021).

We recognize that part of the RTHawk workflow is proprietary. Nevertheless, we emphasize the importance of the transparency to inter-vendor reproducibility at the level of sequence definitions. RTHawk allows sharing open-source sequences (https://github.com/qMRLab/mt_sat). Indeed, moving from open-source to open-execution (Salah et al., 2018) of pulse sequences is a long distance to cover, but for that we also need open-hardware (Winter et al., 2016). Neither RTHawk nor open-source solutions can access under the hood of vendor-specific drivers to directly interface with the hardware. Although RTHawk's pulse sequence and data management servers give more flexibility to the scanner operation at multiple levels of the workflow (e.g., UUI, customized raw data stream, asynchronous real-time updates to sequences, standalone workstation etc.), conversion of the open-source sequence descriptions to vendor-specific hardware instructions is not transparent. We argue that this is a reasonable trade-off to peel another layer from a vendor-specific ecosystem, as it does not sacrifice the transparency of sources relevant to a pulse sequence description. The accuracy and reliability of the parameter estimation methods depend on these descriptions; therefore, for qMRI to work we need to be able to access, modify, and share the methods (Karakuzu, Boudreau, Cohen-Adad, et al., 2020b). Fortunately, VENUS approach to qMRI is not framework exclusive and satisfies this key requirement.

Namely, using community developed tools such as Pulseq, GammaStar, SequenceTree, ODIN or TOPPE, interoperable qMRI applications can be developed. A critical step to achieve this is an effective communication between method developers to foster compatibility between frameworks. This is nicely exemplified by GammaStar and JEMRIS as both applications can export Pulseq descriptions. Enabling a similar feature by developing a SpinBench plugin is among our future goals. To facilitate discussions on this topic with vendor-neutral framework developers, we created a forum page on the code repository of this article (https://github.com/qMRLab/VENUS).

6.5.3 Limitations and future directions

RF transmission systems were different between the all the scanners used for data collection. This is indeed a likely cause of differences between the agreement and accuracy of T1 and MTsat maps. Therefore, another obvious limitation of this study is the lack of B1+ mapping. Unfortunately, a vendor-native B1+ mapping sequence was not available on G1, and it is also well-known that discrepancies between vendor-native B1+ mapping contribute to between-scanner bias in T1 mapping (Lee et al., 2019). Due to the RTHawk gradient limitations described above, developing an accurate AFI sequence was not possible using the current version of RTHawk for a robust characterization of the B1+ effect on T1 and MTsat (note that MTR does not require B1 correction). Therefore, further investigation is needed to compare vendor-neutral B1+ maps across vendors for isolating the specific contribution of transmit field inhomogeneity.

In addition to the non-uniformity of the B1+ field, another critical factor affecting the accuracy is the calculation of a global RF scaling factor. Vendor-native systems set the transmit gain using their own prescan routine, which may lead to a systematic bias in quantitative mapping. In this work, we implemented prescan for G1 and S1-2 as described by (Keenan, Gimbutas, Dienstfrey, and Stupic, 2019) and configured RTHawk to use the same calibration measurements. Nevertheless, it is possible to make this step vendor-neutral as well. For future work, we plan to develop a double-angle VENUS prescan using the same excitation pulses as the qMRI sequences that follow, to determine a global RF scaling factor. Coupled with the use of anatomy-mimicking quantitative MRI phantoms (Gopalan et al., 2021), this would offer qMRI-ready adaptive prescan routines and help investigate the effect of standardizing calibration measurements on multicenter accuracy and agreement.

Finally, given that the data was acquired using full k-space sampling, we used vendor-native and RTHawk resources to perform image reconstruction. The details of the RTHawk reconstruction pipeline are available (Figure 6.1f), which incorporates gradient non-linearity

correction and exports relevant warping coefficients. However, raw data from vendor-native acquisitions were not available. Open-source reconstruction tools (Assländer et al., 2018; Hansen and Sørensen, 2013; Knopp and Grosser, 2021; Maier, Baete, et al., 2021) are an important asset to investigate the potential effect of reconstruction pipeline differences on image characteristics, such as observed in Figure 6.3b-g. Therefore, future work will enable raw data export from vendor-native systems and add a containerized reconstruction node to the qMRFLow (Karakuzu, Boudreau, Cohen-Adad, et al., 2020a) pipeline for investigating potential sources of reconstruction variability.

6.6 Conclusion

In this article we have demonstrated that vendor-neutral sequences (VENUS) and transparent workflows improve reproducibility in quantitative MRI. Additionally, these workflows can be deployed on an FDA-approved device, which demonstrates the potential for wide clinical adoption. Quantitative MRI needs to bypass the vendor black boxes to make an impact in the clinic, and this work shows the way forward.

6.7 Acknowledgements

6.8 Data availability statement

All the vendor-neutral pulse sequences are publicly available as git submodules at https://github.com/qmrlab/pulse_sequences and can be run on RTHawk systems v3.0.0 and later. The RF and gradient waveforms (spv files) can be inspected and simulated using SpinBench (https://www.heartvista.ai/spinbench). As per the general design principles of fully reproducible qMRFlow pipelines, we adhered to a one-process one-container mapping for the processing of this dataset. Docker images, BIDS and ISMRM-RD compliant dataset from the current study are freely available at https://doi.org/10.17605/osf.io/5n3cu. Finally, the whole analysis and interactive version of all the figures in this article will be available and executable online at https://github.com/qmrlab/venus.

6.9 References

Ashton, E. (2010). Quantitative mr in multi-center clinical trials. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *31*(2), 279–288.

Assländer, J., Cloos, M. A., Knoll, F., Sodickson, D. K., Hennig, J., & Lattanzi, R. (2018). Low rank alternating direction method of multipliers reconstruction for mr fingerprinting. *Magnetic resonance in medicine*, *79*(1), 83–96.

Avants, B. B., Tustison, N. J., Song, G., Cook, P. A., Klein, A., & Gee, J. C. (2011). A reproducible evaluation of ants similarity metric performance in brain image registration. *Neuroimage*, *54*(3), 2033–2044.

Bane, O., Hectors, S. J., Wagner, M., Arlinghaus, L. L., Aryal, M. P., Cao, Y., Chenevert, T. L., Fennessy, F., Huang, W., Hylton, N. M., Kalpathy-Cramer, J., Keenan, K. E., Malyarenko, D. I., Mulkern, R. V., Newitt, D. C., Russek, S. E., Stupic, K. F., Tudorica, A., Wilmes, L. J., . . . Taouli, B. (2018). Accuracy, repeatability, and inter-platform reproducibility of t 1 quantification methods used for dce-mri: results from a multicenter phantom study. *Magnetic Resonance in Medicine*, *79*(5), 2564–2575. https://doi.org/10.1002/mrm.26903

Barral, J. K., Gudmundson, E., Stikov, N., Etezadi-Amoli, M., Stoica, P., & Nishimura, D. G. (2010). A robust methodology for in vivo t1 mapping. *Magn Reson Med*, *64*(4), 1057–67. https://doi.org/10.1002/mrm.22497

Bojorquez, J. Z., Bricq, S., Acquitter, C., Brunotte, F., Walker, P. M., & Lalande, A. (2017). What are normal relaxation times of tissues at 3 t? *Magn Reson Imaging*, *35*, 69–80. https://doi.org/10.1016/j.mri.2016.08.021

Boshkovski, T., Kocarev, L., Cohen-Adad, J., Mišić, B., Lehéricy, S., Stikov, N., & Mancini, M. (2021). The r1-weighted connectome: complementing brain networks with a myelin-sensitive measure. *Network Neuroscience*, *5*(2), 358–372.

Boudreau, M., Stikov, N., & Pike, G. B. (2018). B1 -sensitivity analysis of quantitative magnetization transfer imaging. *Magn Reson Med*, *79*(1), 276–285. https://doi.org/10.1002/mrm.26673

Cohen-Adad, J., Alonso-Ortiz, E., Abramovic, M., Arneitz, C., Atcheson, N., Barlow, L., Barry, R. L., Barth, M., Battiston, M., Buchel, C., Budde, M., Callot, V., Combes, A. J. E., De Leener, B., Descoteaux, M., de Sousa, P. L., Dostal, M., Doyon, J., Dvorak, A., . . . Xu, J. (2021). Open-access quantitative mri data of the spinal cord and reproducibility across participants, sites and manufacturers. *Sci Data*, *8*(1), 219. https://doi.org/10.1038/s41597-021-00941-8

Cohen-Adad, J., Alonso-Ortiz, E., Abramovic, M., Arneitz, C., Atcheson, N., Barlow, L., Barry, R. L., Barth, M., Battiston, M., & Büchel, C. (2021). Generic acquisition protocol for quantitative mri of the spinal cord. *Nature protocols*, *16*(10), 4611–4632.

Cordes, C., Konstandin, S., Porter, D., & Günther, M. (2020). Portable and platform-independent mr pulse sequence programs. *Magnetic resonance in medicine*, *83*(4), 1277–1290.

Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: images are more than pictures, they are data. *Radiology*, *278*(2), 563–77. https://doi.org/10.1148/radiol.2015151169

Goebel, J., Seifert, I., Nensa, F., Schemuth, H. P., Maderwald, S., Quick, H. H., Schlosser, T., Jensen, C., Bruder, O., & Nassenstein, K. (2016). Can native t1 mapping differentiate between healthy and diffuse diseased myocardium in clinical routine cardiac mr imaging? *PLoS One*, *11*(5), e0155591.

Gopalan, K., Tamir, J. I., Arias, A. C., & Lustig, M. (2021). Quantitative anatomy mimicking slice phantoms. *Magnetic Resonance in Medicine*, *86*(2), 1159–1166.

Gracien, R.-M., Maiworm, M., Brüche, N., Shrestha, M., Nöth, U., Hattingen, E., Wagner, M., & Deichmann, R. (2020). How stable is quantitative mri? – assessment of intra- and inter-scanner-model reproducibility using identical acquisition sequences and data analysis programs. *NeuroImage*, *207*, 116364. https://doi.org/10.1016/j.neuroimage.2019.116364

Grafe, D., Frahm, J., Merkenschlager, A., Voit, D., & Hirsch, F. W. (2021). Quantitative t1 mapping of the normal brain from early infancy to adulthood. *Pediatr Radiol*, *51*(3), 450–456. https://doi.org/10.1007/s00247-020-04842-7

Gupta, R. K. (1977). A new look at the method of variable nutation angle for the measurement of spin-lattice relaxation times using fourier transform nmr. *Journal of Magnetic Resonance (1969)*, *25*(1), 231–235. https://doi.org/https://doi.org/10.1016/0022-2364(77)90138-X

Haase, A., Frahm, J., Matthaei, D., Hanicke, W., & Merboldt, K.-D. (1986). Flash imaging. rapid nmr imaging using low flip-angle pulses. *Journal of Magnetic Resonance (1969)*, *67*(2), 258–266.

Hahn, E. L. (1949). An accurate nuclear magnetic resonance method for measuring spin-lattice relaxation times. *Phys. Rev.*, *76*(1), 145–146.

Hansen, M. S., & Sørensen, T. S. (2013). Gadgetron: an open source framework for medical image reconstruction. *Magnetic resonance in medicine*, *69*(6), 1768–1776.

Harrell, F. E., & Davis, C. E. (1982). A new distribution-free quantile estimator. *Biometrika*, *69*(3), 635–640.

Helms, G., Dathe, H., & Dechent, P. (2008). Quantitative flash mri at 3t using a rational approximation of the ernst equation. *Magn Reson Med*, *59*(3), 667–72. https://doi.org/10.1002/mrm.21542

Helms, G., Dathe, H., Kallenberg, K., & Dechent, P. (2008). High-resolution maps of magnetization transfer with inherent correction for rf inhomogeneity and t1 relaxation obtained from 3d flash mri. *Magn Reson Med*, *60*(6), 1396–407. https://doi.org/10.1002/mrm.21732

Herz, K., Mueller, S., Perlman, O., Zaitsev, M., Knutsson, L., Sun, P. Z., Zhou, J., van Zijl, P., Heinecke, K., & Schuenke, P. (2021). Pulseq-cest: towards multi-site multi-vendor compatibility and reproducibility of cest experiments using an open-source sequence standard. *Magnetic resonance in medicine.*

Heule, R., Ganter, C., & Bieri, O. (2016). Variable flip angle t1 mapping in the human brain with reduced t2 sensitivity using fast radiofrequency-spoiled gradient echo imaging. *Magnetic Resonance in Medicine*, *75*(4), 1413–1422. https://doi.org/10.1002/mrm.25668

Inati, S. J., Naegele, J. D., Zwart, N. R., Roopchansingh, V., Lizak, M. J., Hansen, D. C., Liu, C.-Y., Atkinson, D., Kellman, P., & Kozerke, S. (2017). Ismrm raw data format: a proposed standard for mri raw datasets. *Magnetic resonance in medicine*, *77*(1), 411–421.

Jenkinson, M., Beckmann, C. F., Behrens, T. E. J., Woolrich, M. W., & Smith, S. M. (2012). Fsl. *Neuroimage*, *62*(2), 782–790.

Jochimsen, T. H., & Von Mengershausen, M. (2004). Odin—object-oriented development interface for nmr. *Journal of Magnetic Resonance*, *170*(1), 67–78.

Karakuzu, A., Appelhoff, S., Auer, T., Boudreau, M., Feingold, F., Khan, A. R., Lazari, A., Markiewicz, C. J., Mulder, M. J., Phillips, C., Salo, T., Stikov, N., Whitaker, K., & de Hollander, G. (2021). Qmri-bids: an extension to the brain imaging data structure for quantitative magnetic resonance imaging data. *medRxiv*, 2021.10.22.21265382. https://doi.org/10.1101/2021.10.22.21265382

Karakuzu, A., Boudreau, M., Cohen-Adad, J., & Stikov, N. (2020b). Thinking outside the blackbox: a fully transparent t1 mapping pipeline. *Proceedings of ISMRM 28th Annual Meeting.*

Karakuzu, A., Boudreau, M., Duval, T., Boshkovski, T., Leppert, I., Cabana, J.-F., Gagnon, I., Beliveau, P., Pike, G. B., Cohen-Adad, J., & Stikov, N. (2020). Qmrlab: quantitative mri analysis, under one umbrella. *Journal of Open Source Software*, *5*(53), 2343.

Keenan, K. E., Gimbutas, Z., Dienstfrey, A., Stupic, K. F., Boss, M. A., Russek, S. E., Chenevert, T. L., Prasad, P. V., Guo, J., Reddick, W. E., Cecil, K. M., Shukla-

Dave, A., Aramburu Nunez, D., Shridhar Konar, A., Liu, M. Z., Jambawalikar, S. R., Schwartz, L. H., Zheng, J., Hu, P., & Jackson, E. F. (2021). Multi-site, multi-platform comparison of mri t1 measurement using the system phantom. *PLOS ONE*, *16*(6), e0252966. https://doi.org/10.1371/journal.pone.0252966

Knopp, T., & Grosser, M. (2021). Mrireco. jl: an mri reconstruction framework written in julia. *Magnetic Resonance in Medicine.*

Lauterbur, P. C. (1989). Image formation by induced local interactions. examples employing nuclear magnetic resonance. 1973. *Clin Orthop Relat Res*, (244), 3–6. https://www.ncbi.nlm.nih.gov/pubmed/2663289

Layton, K. J., Kroboth, S., Jia, F., Littin, S., Yu, H., Leupold, J., Nielsen, J.-F., Stöcker, T., & Zaitsev, M. (2017). Pulseq: a rapid and hardware-independent pulse sequence prototyping framework. *Magnetic resonance in medicine*, *77*(4), 1544–1552.

Lazari, A., & Lipp, I. (2021). Can mri measure myelin? systematic review, qualitative assessment, and meta-analysis of studies validating microstructural imaging with myelin histology. *Neuroimage*, 117744.

Lee, Y., Callaghan, M. F., Acosta-Cabronero, J., Lutti, A., & Nagy, Z. (2019). Establishing intra- and inter-vendor reproducibility of t1 relaxation time measurements with 3t mri. *Magn Reson Med*, *81*(1), 454–465. https://doi.org/10.1002/mrm.27421

Leutritz, T., Seif, M., Helms, G., Samson, R. S., Curt, A., Freund, P., & Weiskopf, N. (2020). Multiparameter mapping of relaxation (r1, r2*), proton density and magnetization transfer saturation at 3 t: a multicenter dual-vendor reproducibility and repeatability study. *Hum Brain Mapp*, *41*(15), 4232–4247. https://doi.org/10.1002/hbm.25122

Liberman, G., Louzoun, Y., & Ben Bashat, D. (2014). T(1) mapping using variable flip angle spgr data with flip angle correction. *J Magn Reson Imaging*, *40*(1), 171–80. https://doi.org/10.1002/jmri.24373

Look, D., & Locker, D. (1970). Time saving in measurement of nmr and epr relaxation times. *Review of Scientific Instruments*, *41*(2), 250–251. https://doi.org/http://link.aip.org/link/doi/10.1063/1.1684482

Magland, J. F., Li, C., Langham, M. C., & Wehrli, F. W. (2016). Pulse sequence programming in a dynamic visual environment: sequencetree. *Magnetic resonance in medicine*, *75*(1), 257–265.

Maier, O., Baete, S. H., Fyrdahl, A., Hammernik, K., Harrevelt, S., Kasper, L., Karakuzu, A., Loecher, M., Patzig, F., & Tian, Y. (2021). Cg-sense revisited: results from the first ismrm reproducibility challenge. *Magnetic resonance in medicine*, *85*(4), 1821–1839.

Mancini, M., Karakuzu, A., Cohen-Adad, J., Cercignani, M., Nichols, T. E., & Stikov, N. (2020). An interactive meta-analysis of mri biomarkers of myelin. *eLife*, *9*. https://doi.org/10.7554/elife.61523

Marques, J. P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P. F., & Gruetter, R. (2010). Mp2rage, a self bias-field corrected sequence for improved segmentation and t1-mapping at high field. *Neuroimage*, *49*(2), 1271–81. https://doi.org/10.1016/j.neuroimage.2009.10.002

Nielsen, J.-F., & Noll, D. C. (2018). Toppe: a framework for rapid prototyping of mr pulse sequences. *Magnetic resonance in medicine*, *79*(6), 3128–3134.

Novikov, D. S., Kiselev, V. G., & Jespersen, S. N. (2018). On modeling. *Magn Reson Med*, *79*(6), 3172–3193. https://doi.org/10.1002/mrm.27101

of Radiology, A. C. (2019). Phantom test guidance for use of the large mri phantom for the acr mri accreditation program.

Okubo, G., Okada, T., Yamamoto, A., Kanagaki, M., Fushimi, Y., Okada, T., Murata, K., & Togashi, K. (2016). Mp2rage for deep gray matter measurement of the brain: a comparative study with mprage. *J Magn Reson Imaging*, *43*(1), 55–62. https://doi.org/10.1002/jmri.24960

Overall, W., & Pauly, J. (2007). An extensible, graphical environment for pulse sequence design and simulation. *Int. Soc. Magn. Reson. Med., Berlin, Germany.*

Preibisch, C., & Deichmann, R. (2009). Influence of rf spoiling on the stability and accuracy of t1 mapping based on spoiled flash with varying flip angles. *Magn Reson Med*, *61*(1), 125–135.

Ravi, K. S., Geethanath, S., & Vaughan, J. T. (2019). Pypulseq: a python package for mri pulse sequence design. *Journal of Open Source Software*, *4*(42), 1725.

Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2017). Beyond differences in means: robust graphical methods to compare two groups in neuroscience. *European Journal of Neuroscience*, *46*(2), 1738–1748.

Rowley, C. D., Campbell, J. S. W., Wu, Z., Leppert, I. R., Rudko, D. A., Pike, G. B., & Tardif, C. L. (2021). A model-based framework for correcting b 1 + inhomogeneity effects in magnetization transfer saturation and inhomogeneous magnetization transfer saturation maps. *Magn Reson Med*, *86*(4), 2192–2207. https://doi.org/10.1002/mrm.28831

Salah, K., Damiani, E., Al-Fuqaha, A., Martin, T., Taha, K., & Khan, M. K. (2018). Open execution—the blockchain model. *IEEE Blockchain Technical Briefs (December 2018)*.

Santos, J. M., Wright, G. A., & Pauly, J. M. (2004). Flexible real-time magnetic resonance imaging framework. *1*, 1048–1051.

Stikov, N., Boudreau, M., Levesque, I. R., Tardif, C. L., Barral, J. K., & Pike, G. B. (2015). On the accuracy of t-1 mapping: searching for common ground. *Magnetic Resonance in Medicine*, *73*(2), 514–522. https://doi.org/10.1002/mrm.25135

Stikov, N., Trzasko, J. D., & Bernstein, M. A. (2019). Reproducibility and the future of mri research. *Magn Reson Med*, *82*(6), 1981–1983. https://doi.org/10.1002/mrm.27939

Stöcker, T., Vahedipour, K., Pflugfelder, D., & Shah, N. J. (2010). High-performance computing mri simulations. *Magnetic resonance in medicine*, *64*(1), 186–193.

Stupic, K. F., Ainslie, M., Boss, M. A., Charles, C., Dienstfrey, A. M., Evelhoch, J. L., Finn, P., Gimbutas, Z., Gunter, J. L., Hill, D. L. G., Jack, C. R., Jackson, E. F., Karaulanov, T., Keenan, K. E., Liu, G., Martin, M. N., Prasad, P. V., Rentz, N. S., Yuan, C., & Russek, S. E. (2021). A standard system phantom for magnetic resonance imaging. *Magnetic Resonance in Medicine*, *86*(3), 1194–1211. https://doi.org/10.1002/mrm.28779

Teixeira, R. P., Neji, R., Wood, T. C., Baburamani, A. A., Malik, S. J., & Hajnal, J. V. (2020). Controlled saturation magnetization transfer for reproducible multivendor variable flip angle t1 and t2 mapping. *Magnetic resonance in medicine*, *84*(1), 221–236.

Tong, G., Gaspar, A. S., Qian, E., Ravi, K. S., Vaughan Jr, J. T., Nunes, R. G., & Geethanath, S. (2021). A framework for validating open-source pulse sequences. *Magnetic resonance imaging.*

Voelker, M. N., Kraff, O., Goerke, S., Laun, F. B., Hanspach, J., Pine, K. J., Ehses, P., Zaiss, M., Liebert, A., Straub, S., Eckstein, K., Robinson, S., Nagel, A. N., Stefanescu, M. R., Wollrab, A., Klix, S., Felder, J., Hock, M., Bosch, D., . . . Quick, H. H. (2021). The traveling heads 2.0: multicenter reproducibility of quantitative imaging methods at 7 tesla. *Neuroimage*, *232*, 117910. https://doi.org/10.1016/j.neuroimage.2021.117910

Weiskopf, N., Edwards, L. J., Helms, G., Mohammadi, S., & Kirilina, E. (2021). Quantitative magnetic resonance imaging of brain anatomy and in vivo histology. *Nature Reviews Physics*, *3*(8), 570–588. https://doi.org/10.1038/s42254-021-00326-1

Winter, L., Haopeng, H., Barghoorn, A., Hoffmann, W., Hetzer, S., & Winkler, S. (2016). Open source imaging initiative. *Proceedings of the International Society for Magnetic Resonance in Medicine*, *3638.*

Yarnykh, V. L. (2007b). Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magn Reson Med*, *57*(1), 192–200.

Yarnykh, V. L. (2010). Optimal radiofrequency and gradient spoiling for improved accuracy of t1 and b1 measurements using fast steady-state techniques. *Magn Reson Med*, *63*(6), 1610–1626.

Young, I. R. (2004). Significant events in the development of mri. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *20*(2), 183–186.

Zwart, N. R., & Pipe, J. G. (2015). Graphical programming interface: a development environment for mri methods. *Magnetic resonance in medicine*, *74*(5), 1449–1460.

CHAPTER 7   GENERAL DISCUSSION

7.1   The pivotal role of transparency and data standards for reproducible qMRI processing

In 2016, a detailed survey investigated whether researchers from a variety of disciplines were able to independently replicate other scientists' work (Baker, 2016). The results were striking, revealing that nearly 70% of the attempts could not achieve independent replication. More interestingly, the success rate of researchers' ability to replicate their own work was not much higher. The survey concludes that the factors contributing to this reproducibility crisis transcends disciplines and runs deep in the research practices that have been established as the norm to advance scientific literature.

It is plausible that tracing the reproducibility crisis in computational fields is easier (Srinivasan, 2020). The reasoning is that testing a software-driven analysis for reproducibility is relatively undemanding compared to testing a biological experiment. Nevertheless, this assumption is feasible as long as the computational resources including code, data and runtime environment are made accessible.

The field of MRI has shown a rapid trend towards making the computational resources of code, data and runtime publicly accessible (Stikov, Trzasko, et al., 2019). The efforts include the development of a data standard for raw MRI data (Inati et al., 2017), a considerable number of open-source software for data acquisition, reconstruction and processing (Karakuzu, Biswas, et al., 2021). The sub-field of qMRI caught up with some of these developments, especially in the lane of open-source software development. As of early 2022, there are at least 13 open-source software packages for fitting quantitative maps for a variety of applications (Karakuzu, Appelhoff, et al., 2021). These decentralized developments foster the publishing of integrated research objects by independent research groups (DuPre et al., 2022), which is a vital factor for the meta-analysis of these methods (Mancini et al., 2020). However, the landscape of open qMRI research lacked a unifying component to facilitate collaboration between MRI method developers, and to adapt qMRI for clinical and neuroimaging research.

The development of brain imaging data structure (BIDS) (Gorgolewski, Auer, et al., 2016) prompted a large collection of open-source neuroimaging software to evolve into BIDS-apps that can talk the same language for consuming and creating data (Gorgolewski, Alfaro-Almagro, et al., 2017). This enabled chaining an array of BIDS-apps into standardized data processing workflows (Esteban, Birman, et al., 2017; Esteban, Markiewicz, et al., 2019) that

can be deployed on web-based platforms for widespread adoption (Gorgolewski, Esteban, et al., 2017; Sherif et al., 2014). From the standpoint of unifying independent development of open-source solutions, BIDS is a paragon of efficiency. It constitutes evidence that developing a data standard goes a long way towards integrating reproducibility standards into research practices across a research discipline. This was an important motivation for the development of qMRI-BIDS in bringing the open-source solutions closer to interoperable use for wider adoption of qMRI in clinics and research (Karakuzu, Appelhoff, et al., 2021). The developed standard serves as a bridge not only to horizontally integrate qMRI software, but also to vertically streamline transparent solutions across multiple levels of an end-to-end workflow. The following section discusses how qMRI-BIDS and other developments in this thesis enables such a powerful integration.

7.2 Data harmonization only goes so far: Vendor-neutral sequences can standardize qMRI

A traditional qMRI workflow can be seen as a hurdle race, where a researcher needs to overcome several proprietary roadblocks to run a study and to publish the outcomes from the experiment (Figure 7.1a). Availability of the acquisition methods depends on a list of equipment-specific constraints such as the MRI scanner's vendor, model and software version. If it is possible to clear this hurdle with the available equipment, the following steps are data curation and analysis. Based on the implementation, these processes may be a part of the vendor's proprietary pipeline, not allowing the researcher to take a look at the models generating the quantitative maps. The research outcomes from this workflow can be published as a traditional PDF, given that the individual components of the workflow are not publicly available as computational resources. This is a serious impediment for independent or self replication of the research outcomes.

With transparent workflows that begin with a vendor-neutral acquisition (Karakuzu, Biswas, et al., 2021) and extend all the way to the publication of integrated research objects (DuPre et al., 2022), this thesis removes the hurdles from the path of reproducible qMRI research (Figure 7.1b). To reimagine a qMRI experiment as a sprint race provides more than a practical convenience for researchers to reach the finish line faster. Every single step involved in producing the quantitative maps can be traced back to the acquisition, including the details of the pulse sequence implementation. To offer a modular provenance recording, every process is mapped to an isolated and portable runtime in relation with the input-output definitions conforming with community data standards.

The pulse sequences can be cloned from GitHub (qMR*Pull*seq) and run on most General

Electric and Siemens scanners equipped with RTHawk (HeartVista Inc, CA, USA) real-time imaging platform (https://github.com/qMRLab/pulse_sequences). The source code of these sequences is based on RTHawk's application programming interface (API) libraries and are publicly accessible along with all the gradient and RF waveforms played on the scanner. These waveforms are designed and simulated using SpinBench, which is also publicly available for non-commercial use. Using qMRPullseq, no additional step is needed to organize raw or reconstructed data into ISMRM-RD or qMRI-BIDS format. Based on the details provided by the researcher in the unified user interface (e.g., the participant or session ID for the *sub-* or
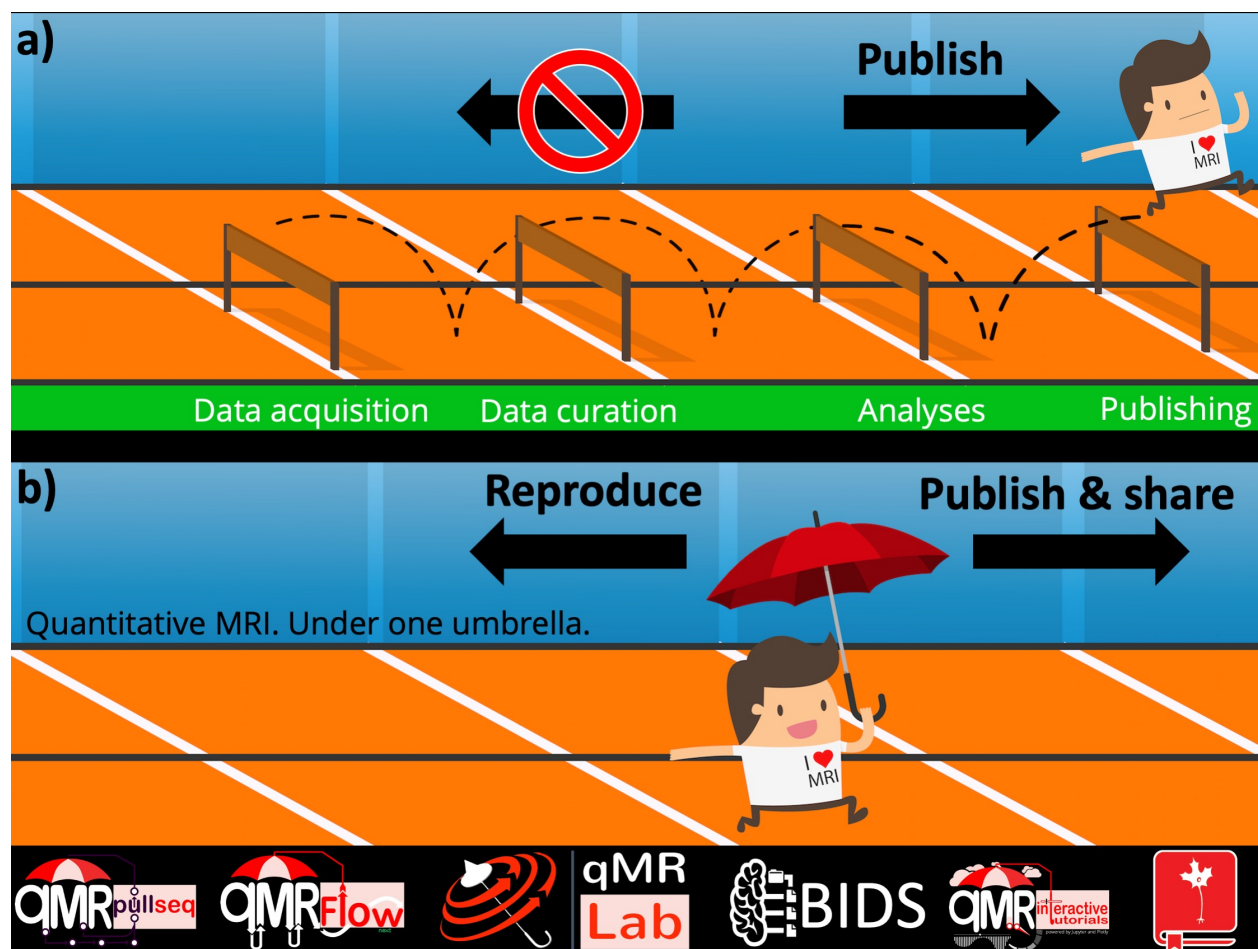
Figure 7.1 **a)** A traditional qMRI workflow is a hurdle race. There are several proprietary barriers that challenge qMRI researchers to reach the finish line and prevent them from following each step back to the start. **b)** Using the solution proposed by this thesis, qMRI workflows become a sprint race. Using vendor-neutral acquisition protocols and transparent processing workflows give the researchers the headstart to spend more time on making their research reproducible. In addition, they can trace every step all the way back to the pulse sequence implementations.

ses-, and freeform descriptions for the *acq-* BIDS entities) and according to the qMRI-BIDS descriptions of the respective sequence (e.g., iterations of *flip-* and *mt-* acquisitions for an MTS file collection, see Table 5.1), the outputs are automatically converted to BIDS-valid NIfTI and JSON file pairs.

The directory where scanner outputs are written (a unix pipe) is asynchronously observed by qMRFlow, a collection of data-driven and container mediated qMRLab workflows (https://github.com/qmrlab/qmrflow). Whenever a set of files match a known qMRI-BIDS file pattern, qMRFlow issues a command to start the respective workflow execution. This way, the data standard also describes an inter-process communication schema, dispensing with the needs for programming sockets and networking protocols. The qMRFlow pipelines are powered by Nextflow, which comes with powerful features such as execution abstraction. Thanks to this feature, the workflow's functional logic is abstracted from the details of the underlying execution system, whether it is a supercomputing cluster, cloud infrastructure or a local machine.

The execution abstraction guarantees that any workflow process executed at the scanner site can be independently reproduced on any desired execution platform provided that:

- i) a standardized data input/output convention is available,

- ii) the software and version dependencies are met,

- iii) the data is available.

To satisfy the condition (i), all the qMRFlow data channels are described by qMRI-BIDS raw/derivative data conventions. The condition (ii) is satisfied by a modular one-process one-container mapping strategy. For example, the preprocessing steps of an MTsat processing workflow are executed in a *qmrlab:antsfsl:v2.5.0* container, then the derivatives are forwarded to a *qmrlab:minimal:v2.5.0* container[1] for fitting quantitative maps. Given that the Nextflow tasks are not handled synchronously as it would happen in imperative pipeline scripts (e.g., shell scripts), each process can be associated with an independent quality control procedure without blocking parallel executions. Be it generating automated reports, or asking for manual user intervention. Finally, all the low level events such as buffer allocation, data queueing, container lifetimes and logging are orchestrated by Nextflow's core engine and combined in a detailed execution report (for an example, see https://qmrlab.org/VENUS/qmrflow-exec-report.html).

---

[1] These containers are published under version control as a part of qMRLab's continuous delivery pipeline (https://hub.docker.com/u/qmrlab).

Meeting the final condition (iii) also benefits from the modular design of qMRFlow. This is because the independently reproducible portion of the workflow is determined by the ethics and policies around sharing sensitive data. For example, if the qMRI study is performed using a phantom, all the processes can be reproduced in their respective containers as there are no ethical concerns (see https://osf.io/5n3cu/). On the other hand, if it is only possible to share higher order derivatives, the modular design (i.e., isolated and portable software environments) allows executing only certain processes from the workflow.

In conclusion; qMRPullseq, qMRLab, qMRFlow and qMRI-BIDS together bring all the qMRI needs under one umbrella, from scanner to the publication of integrated research objects (Figure 7.1b). This fosters easy multi-center deployment of and reproducibility of qMRI applications both in research and clinics.

7.3 Quality meets quantity through workflows that can extend from scanner to publication

As introduced in Chapter 2, the ultimate goal of qMRI is to provide reliable and objective measurements about underlying tissue characteristics across time points and imaging centers. If a qMRI metric can offer satisfactory sensitivity and specificity for a certain pathology, its reliability across time points and imaging centers determines its prognostic and diagnostic performance, respectively. Absent a standardized measurement protocol, achieving this goal becomes highly challenging. In such a scenario, multiple correction models may be needed to bring measurements into a comparable scale not because of the hardware imperfections (e.g., B0 and B1 inhomogeneities), but because of the differences in implementations (e.g., different MT pulses or RF spoiling schemes). Achieving a reduction in accounting variations in this manner is known as data harmonization. Although harmonization can be still useful for exploratory studies when the scanner does not permit access to all functions of the system, this approach is not well aligned with the key promise of qMRI. To bring out the true strength of qMRI, the solutions should focus on moving towards the elimination of any variation, i.e., standardization. Removing the sources of variability with this approach has important implications in neuroimaging research and clinical trials.

In a typical neuroimaging study, high-resolution anatomical images are acquired to extract brain morphometric features, such as the gray matter thickness. However, MRI is not a direct measurement of structure. For example, the perceived boundary between gray- and white-matter on a T1-weighted image is mainly determined by the relaxation characteristics of each compartment. There are several confounders of T1 and T2 (e.g., smoking, antidepressant treatment, alcohol and drug use etc.), which may lead to elusive findings of structural dif-

ferences (Weinberger and Radulescu, 2016). Quantitative MRI provides an array of tools to elucidate the potential origins of such differences. However, to supplement these studies with a more objective comparison basis, qMRI itself needs to be free from systematic variations.

In clinical trials, reducing measurement variability across imaging centers caters to a more robust identification of the true biological variability, if there is any (Ashton, 2010). This in turn can help reduce the number of participants in achieving the same effect size, increasing the efficiency of clinical trials on quantitative biomarkers. Another problem encountered in clinical research is the vendor upgrades of the imaging systems (Keenan, Gimbutas, Dienstfrey, and Stupic, 2019). Given that the release notes are not available for all the aspects of the system upgrades, quantitative estimations are subjected to shift due to unknown changes. Therefore, having a vendor-native access to the scanner hardware gives the user a more confident control over the stability of the measurements.

The final work of this thesis presented in Chapter 6 shows compelling evidence that vendor-neutral implementations can significantly reduce inter-vendor variations by standardizing all the aspects of the acquisition (Karakuzu, Biswas, et al., 2021). This, in return, brings qMRI closer to its promises: In clinics, it can provide reliable in-vivo histological information that can revolutionize the neuroradiological diagnosis and prognosis. In research, it gives a unique window into the understanding of brain structure and function by informing existing methods with unprecedented microstructural information.

7.4   Financial value of the vendor-neutral qMRI approach

Majority of the clinical MRI systems are marketed by three global vendors: Siemens Healthineers, General Electric Healthcare and Philips Healthcare. Even though most of the patents on the conventional MRI have expired (Moritz et al., 2019), the proprietary software products marketed by these key companies account for a considerable portion of a billion dollar global market. Given the current status of the patents, the large market size of the clinical MRI systems coming from nearly 36,000 scanners around the world (Ogbole et al., 2018) and the fact that the most MRI innovations originate from publicly funded research organizations (Moritz et al., 2019), enabling researchers to commercialize their research through a vendor-neutral "app store" can offer a new revenue stream. It is worth noting that this decentralized solution cannot be delivered by vendored development platforms with their current variability management strategies in place (Jaring et al., 2004).

This thesis demonstrates that the vendor-neutral approach to qMRI not only has technical advantages over the vendor-native implementations by reducing variability, but also creates

value for MR researchers, clinicians and software developers by reducing the barrier to the commercialization of MRI research outcomes. Although the proposed framework is based on a proprietary software development platform (RTHawk), the app developers are entitled to select a suitable license to distribute and monetize their applications in an open-source business model (Watson et al., 2008).

CHAPTER 8   CONCLUSION

This thesis brings together all the needs for qMRI methods to find the way out of the maze of variability (Figure 2.25), and to reach the clinics. It demonstrates a powerful integration between vendor-neutral sequences, community data standards and open-source software to improve the reliability of quantitative MRI. The developed workflows reduce the variability at multiple levels, starting with the MRI pulse sequence implementations and continuing with reconstruction, pre- and post-processing. Quantitative MRI needs to bypass the vendor black boxes to deliver what it has to offer, and this work shows the way forward.

REFERENCES

Abbas, Z., Gras, V., Möllenhoff, K., Oros-Peusquens, A.-M., & Shah, N. J. (2015). Quantitative water content mapping at clinically relevant field strengths: a comparative study at 1.5 t and 3 t. *NeuroImage*, *106*, 404–413.

Ashton, E. (2010). Quantitative mr in multi-center clinical trials. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *31*(2), 279–288.

Baker, M. (2016). 1,500 scientists lift the lid on reproducibility. *Nature News*, *533*(7604), 452.

Bane, O., Hectors, S. J., Wagner, M., Arlinghaus, L. L., Aryal, M. P., Cao, Y., Chenevert, T. L., Fennessy, F., Huang, W., Hylton, N. M., Kalpathy-Cramer, J., Keenan, K. E., Malyarenko, D. I., Mulkern, R. V., Newitt, D. C., Russek, S. E., Stupic, K. F., Tudorica, A., Wilmes, L. J., . . . Taouli, B. (2018). Accuracy, repeatability, and inter-platform reproducibility of t 1 quantification methods used for dce-mri: results from a multicenter phantom study. *Magnetic Resonance in Medicine*, *79*(5), 2564–2575. https://doi.org/10.1002/mrm.26903

Barral, J. K., Gudmundson, E., Stikov, N., Etezadi-Amoli, M., Stoica, P., & Nishimura, D. G. (2010). A robust methodology for in vivo t1 mapping. *Magn Reson Med*, *64*(4), 1057–67. https://doi.org/10.1002/mrm.22497

Bartlett, J., & Frost, C. (2008). Reliability, repeatability and reproducibility: analysis of measurement errors in continuous variables. *Ultrasound in Obstetrics and Gynecology: The Official Journal of the International Society of Ultrasound in Obstetrics and Gynecology*, *31*(4), 466–475.

Bilgic, B., Pfefferbaum, A., Rohlfing, T., Sullivan, E. V., & Adalsteinsson, E. (2012). Mri estimates of brain iron concentration in normal aging using quantitative susceptibility mapping. *Neuroimage*, *59*(3), 2625–2635.

Bjarnason, T., Vavasour, I., Chia, C., & MacKay, A. (2005). Characterization of the nmr behavior of white matter in bovine brain. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *54*(5), 1072–1081.

Bland, J. M., & Altman, D. (1986). Statistical methods for assessing agreement between two methods of clinical measurement. *The lancet*, *327*(8476), 307–310.

Bloch, F. (1957). Generalized theory of relaxation. *Physical Review*, *105*(4), 1206.

Bojorquez, J. Z., Bricq, S., Acquitter, C., Brunotte, F., Walker, P. M., & Lalande, A. (2017). What are normal relaxation times of tissues at 3 t? *Magn Reson Imaging*, *35*, 69–80. https://doi.org/10.1016/j.mri.2016.08.021

Borich, M., Mackay, A., Vavasour, I., Rauscher, A., & Boyd, L. (2013). Evaluation of white matter myelin water fraction in chronic stroke. *NeuroImage: Clinical*, *2*, 569–580.

Boshkovski, T., Kocarev, L., Cohen-Adad, J., Mišić, B., Lehéricy, S., Stikov, N., & Mancini, M. (2021). The r1-weighted connectome: complementing brain networks with a myelin-sensitive measure. *Network Neuroscience*, *5*(2), 358–372.

Boudreau, M., Stikov, N., & Pike, G. B. (2018). B1 -sensitivity analysis of quantitative magnetization transfer imaging. *Magn Reson Med*, *79*(1), 276–285. https://doi.org/10.1002/mrm.26673

Boudreau, M. (2018). *Relaxometry series: variable flip angle.* https://qmrlab.org/jekyll/2018/12/11/T1-mapping-variable-flip-angle.html

Boudreau, M. (2019). *On the open source landscape of mrm.* https://qmrlab.org/2019/08/16/mrm-open-source.html

Breit, G., & Rabi, I. (1931). Measurement of nuclear spin. *Physical Review*, *38*(11), 2082.

Brewer, R. G., & Hahn, E. L. (1984). Atomic memory. *Scientific American*, *251*(6), 50–57.

Brown, R., Narayanan, S., Atkins, H., Marder, S., & Arnold, D. (2011). Normalization of magnetization transfer ratio mri for multicentre clinical trials. *Proc Int Soc Magn Reson Med*, *19*, 4082.

Campbell, J. S., Leppert, I. R., Narayanan, S., Boudreau, M., Duval, T., Cohen-Adad, J., Pike, G. B., & Stikov, N. (2018). Promise and pitfalls of g-ratio estimation with mri. *Neuroimage*, *182*, 80–96.

Carnot, S., Clausius, R., & Kelvin, W. T. (1899). *The second law of thermodynamics: memoirs by carnot, clausius, and thomson* (Vol. 6). Harper & brothers.

Cercignani, M., Giulietti, G., Dowell, N. G., Gabel, M., Broad, R., Leigh, P. N., Harrison, N. A., & Bozzali, M. (2017). Characterizing axonal myelination within the healthy population: a tract-by-tract mapping of effects of age and gender on the fiber g-ratio. *Neurobiology of aging*, *49*, 109–118.

Cir., U. F. (1996). *Fonar corporation v. general electric company.* https://caselaw.findlaw.com/us-federal-circuit/1229938.html

Cordes, C., Konstandin, S., Porter, D., & Günther, M. (2020). Portable and platform-independent mr pulse sequence programs. *Magnetic resonance in medicine*, *83*(4), 1277–1290.

Cracknell, D., White, M. P., Pahl, S., Nichols, W. J., & Depledge, M. H. (2016). Marine biota and psychological well-being: a preliminary examination of dose–response effects in an aquarium setting. *Environment and Behavior*, *48*(10), 1242–1269.

Damadian, R. (1971). Tumor detection by nuclear magnetic resonance. *Science*, *171*(3976), 1151–1153.

Damadian, R. (1974). *Apparatus and method for detecting cancer in tissue (us3789832a)*. https://patents.google.com/patent/US3789832?oq=damadian+1971

Dawson, M. J. (2013). *Paul lauterbur and the invention of mri*. MIT Press.

Dean, D. C., O'Muircheartaigh, J., Dirks, H., Travers, B. G., Adluru, N., Alexander, A. L., & Deoni, S. C. (2016). Mapping an index of the myelin g-ratio in infants using magnetic resonance imaging. *Neuroimage*, *132*, 225–237.

Deoni, S. C. (2011). Magnetic resonance relaxation and quantitative measurement in the brain. *Magnetic Resonance Neuroimaging*, 65–108.

Deoni, S. C., Rutt, B. K., & Peters, T. M. (2003). Rapid combined t1 and t2 mapping using gradient recalled acquisition in the steady state. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *49*(3), 515–526.

de Vet, H. C., Terwee, C. B., Ostelo, R. W., Beckerman, H., Knol, D. L., & Bouter, L. M. (2006). Minimal changes in health status questionnaires: distinction between minimally detectable change and minimally important change. *Health and quality of life outcomes*, *4*(1), 1–5.

Donahue, K. M., Weisskoff, R. M., & Burstein, D. (1997). Water diffusion and exchange as they influence contrast enhancement. *Journal of Magnetic Resonance Imaging*, *7*(1), 102–110.

Donaldson, H. H., & Hoke, G. (1905). On the areas of the axis cylinder and medullary sheath as seen in cross sections of the spinal nerves of vertebrates. *Journal of Comparative Neurology and Psychology*, *15*(1), 1–16.

Dreizen, P. (2004). The nobel prize for mri: a wonderful discovery and a sad controversy. *The Lancet*, *363*(9402), 78.

DuPre, E., Holdgraf, C., Karakuzu, A., Tetrel, L., Bellec, P., Stikov, N., & Poline, J.-B. (2022). Beyond advertising: new infrastructures for publishing integrated research objects. *PLOS Computational Biology*, *18*(1), e1009651.

Duval, T., Perraud, B., Vuong, M.-T., Lopez Rios, N., Stikov, N., & Cohen-Adad, J. (2016). Validation of quantitative mri metrics using full slice histology with automatic axon segmentation. *Proceedings of the 24th Annual Meeting of ISMRM (Singapore)*.

Duval, T., Smith, V., Stikov, N., Klawiter, E. C., & Cohen-Adad, J. (2018). Scan–rescan of axcaliber, macromolecular tissue volume, and g-ratio in the spinal cord. *Magnetic resonance in medicine*, *79*(5), 2759–2765.

Ellerbrock, I., & Mohammadi, S. (2018). *Four in vivo g-ratio-weighted imaging methods: comparability and repeatability at the group level* (tech. rep.). Wiley Online Library.

Erkintalo, M. (2021). Got the quantum jitters. *Nature Physics*, *17*(4), 432–434.

Ernst, R. R., & Anderson, W. A. (1966). Application of fourier transform spectroscopy to magnetic resonance. *Review of Scientific Instruments*, *37*(1), 93–102.

Esteban, O., Birman, D., Schaer, M., Koyejo, O. O., Poldrack, R. A., & Gorgolewski, K. J. (2017). Mriqc: advancing the automatic prediction of image quality in mri from unseen sites. *PloS one*, *12*(9), e0184661.

Esteban, O., Markiewicz, C. J., Blair, R. W., Moodie, C. A., Isik, A. I., Erramuzpe, A., Kent, J. D., Goncalves, M., DuPre, E., & Snyder, M. (2019). Fmriprep: a robust preprocessing pipeline for functional mri. *Nature methods*, *16*(1), 111–116.

Feinberg, D. A. (2018). *The transformative genious of erwin hahn.* https://blog.ismrm.org/2018/09/04/qa-with-erwin-hahn/

Ferris, T. (2019). *Entropy: the hidden force that complicates life.* https://fs.blog/entropy/

Feynman, R. P., Leighton, R. B., & Sands, M. (2015). *The feynman lectures on physics, vol. iii: the new millennium edition: quantum mechanics.* Hachette UK.

Fralix, T. A., Ceckler, T. L., Wolff, S. D., Simon, S. A., & Balaban, R. S. (1991). Lipid bilayer and water proton magnetization transfer: effect of cholesterol. *Magnetic resonance in medicine*, *18*(1), 214–223.

Fram, E. K., Herfkens, R. J., Johnson, G. A., Glover, G. H., Karis, J. P., Shimakawa, A., Perkins, T. G., & Pelc, N. J. (1987). Rapid calculation of t1 using variable flip angle gradient refocused imaging. *Magnetic resonance imaging*, *5*(3), 201–208.

Fukunaga, M., Li, T.-Q., van Gelderen, P., de Zwart, J. A., Shmueli, K., Yao, B., Lee, J., Maric, D., Aronova, M. A., & Zhang, G. (2010). Layer-specific variation of iron content in cerebral cortex as a source of mri contrast. *Proceedings of the National Academy of Sciences*, *107*(8), 3834–3839.

Gambon, D. (2015). The crazy world of candy. *Tandartspraktijk*, *36*(7), 17–23.

Gareau, P. J., Rutt, B. K., Karlik, S. J., & Mitchell, J. R. (2000). Magnetization transfer and multicomponent t2 relaxation measurements with histopathologic correlation in an experimental model of ms. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *11*(6), 586–595.

Geyer, S., Weiss, M., Reimann, K., Lohmann, G., & Turner, R. (2011). Microstructural parcellation of the human cerebral cortex–from brodmann's post-mortem map to in vivo mapping with high-field magnetic resonance imaging. *Frontiers in human neuroscience*, *5*, 19.

Giacomini, P. S., & Arnold, D. L. (2008). Non-conventional mri techniques for measuring neuroprotection, repair and plasticity in multiple sclerosis. *Current opinion in neurology*, *21*(3), 272–277.

Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: images are more than pictures, they are data. *Radiology*, *278*(2), 563–77. https://doi.org/10.1148/radiol.2015151169

Gloor, M., Scheffler, K., & Bieri, O. (2008). Quantitative magnetization transfer imaging using balanced ssfp. *Magnetic resonance in medicine*, *60*(3), 691–700.

Goebel, J., Seifert, I., Nensa, F., Schemuth, H. P., Maderwald, S., Quick, H. H., Schlosser, T., Jensen, C., Bruder, O., & Nassenstein, K. (2016). Can native t1 mapping differentiate between healthy and diffuse diseased myocardium in clinical routine cardiac mr imaging? *PLoS One*, *11*(5), e0155591.

Gopalan, K., Tamir, J. I., Arias, A. C., & Lustig, M. (2021). Quantitative anatomy mimicking slice phantoms. *Magnetic Resonance in Medicine*, *86*(2), 1159–1166.

Gorgolewski, K., Esteban, O., Schaefer, G., Wandell, B., & Poldrack, R. (2017). Openneuro— a free online platform for sharing and analysis of neuroimaging data. *Organization for human brain mapping. Vancouver, Canada*, *1677*(2).

Gorgolewski, K. J., Alfaro-Almagro, F., Auer, T., Bellec, P., Capotă, M., Chakravarty, M. M., Churchill, N. W., Cohen, A. L., Craddock, R. C., & Devenyi, G. A. (2017). Bids apps: improving ease of use, accessibility, and reproducibility of neuroimaging data analysis methods. *PLoS computational biology*, *13*(3), e1005209.

Gorgolewski, K. J., Auer, T., Calhoun, V. D., Craddock, R. C., Das, S., Duff, E. P., Flandin, G., Ghosh, S. S., Glatard, T., & Halchenko, Y. O. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Scientific data*, *3*(1), 1–9.

Gracien, R.-M., Maiworm, M., Brüche, N., Shrestha, M., Nöth, U., Hattingen, E., Wagner, M., & Deichmann, R. (2020). How stable is quantitative mri? – assessment of intra- and inter-scanner-model reproducibility using identical acquisition sequences and data analysis programs. *NeuroImage*, *207*, 116364. https://doi.org/10.1016/j.neuroimage.2019.116364

Graessner, J. (2013). Bandwidth in mri. *Magnetom Flash*, *2*, 3–8.

Grafe, D., Frahm, J., Merkenschlager, A., Voit, D., & Hirsch, F. W. (2021). Quantitative t1 mapping of the normal brain from early infancy to adulthood. *Pediatr Radiol*, *51*(3), 450–456. https://doi.org/10.1007/s00247-020-04842-7

Haase, A., Frahm, J., Matthaei, D., Hanicke, W., & Merboldt, K.-D. (1986). Flash imaging. rapid nmr imaging using low flip-angle pulses. *Journal of Magnetic Resonance (1969)*, *67*(2), 258–266.

Hahn, E. L. (1949). An accurate nuclear magnetic resonance method for measuring spin-lattice relaxation times. *Phys. Rev.*, *76*(1), 145–146.

Halpern, P. (2017). *Spin: the quantum property that should have been impossible*. Retrieved October 25, 2017, from https://www.forbes.com/sites/startswithabang/2017/11/21/spin-the-quantum-property-that-nature-shouldnt-possess/?sh=1f2b64116349

Harris, R. F. (2003). Nobel grudge. *Current Biology*, *13*(22), R857.

Helms, G., Dathe, H., Kallenberg, K., & Dechent, P. (2008). High-resolution maps of magnetization transfer with inherent correction for rf inhomogeneity and t1 relaxation obtained from 3d flash mri. *Magn Reson Med*, *60*(6), 1396–407. https://doi.org/10.1002/mrm.21732

Henderson, E., McKinnon, G., Lee, T.-Y., & Rutt, B. K. (1999). A fast 3d look-locker method for volumetric t1 mapping. *Magnetic resonance imaging*, *17*(8), 1163–1171.

Henkelman, R. M., Huang, X., Xiang, Q.-S., Stanisz, G., Swanson, S. D., & Bronskill, M. J. (1993). Quantitative interpretation of magnetization transfer. *Magnetic resonance in medicine*, *29*(6), 759–766.

Henkelman, R., Stanisz, G., & Graham, S. (2001). Magnetization transfer in mri: a review. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo*, *14*(2), 57–64.

Herz, K., Mueller, S., Perlman, O., Zaitsev, M., Knutsson, L., Sun, P. Z., Zhou, J., van Zijl, P., Heinecke, K., & Schuenke, P. (2021). Pulseq-cest: towards multi-site multi-vendor compatibility and reproducibility of cest experiments using an open-source sequence standard. *Magnetic resonance in medicine*.

Jaring, M., Krikhaar, R., & Bosch, J. (2004). Representing variability in a family of mri scanners. *Software: Practice and Experience*, *34*(1), 69–100.

Jespersen, S. N., Bjarkam, C. R., Nyengaard, J. R., Chakravarty, M. M., Hansen, B., Vosegaard, T., Østergaard, L., Yablonskiy, D., Nielsen, N. C., & Vestergaard-Poulsen, P. (2010). Neurite density from magnetic resonance diffusion measurements at ultrahigh field: comparison with light microscopy and electron microscopy. *Neuroimage*, *49*(1), 205–216.

Jochimsen, T. H., & Von Mengershausen, M. (2004). Odin—object-oriented development interface for nmr. *Journal of Magnetic Resonance*, *170*(1), 67–78.

Jona, G., Furman-Haran, E., & Schmidt, R. (2021). Realistic head-shaped phantom with brain-mimicking metabolites for 7 t spectroscopy and spectroscopic imaging. *NMR in Biomedicine*, *34*(1), e4421.

Jones, M., Dobson, A., & O'Brian, S. (2011). A graphical method for assessing agreement with the mean between multiple observers using continuous measures. *International journal of epidemiology*, *40*(5), 1308–1313.

Karakuzu, A., Appelhoff, S., Auer, T., Boudreau, M., Feingold, F., Khan, A. R., Lazari, A., Markiewicz, C. J., Mulder, M. J., Phillips, C., Salo, T., Stikov, N., Whitaker, K., & de Hollander, G. (2021). Qmri-bids: an extension to the brain imaging data structure for quantitative magnetic resonance imaging data. *medRxiv*, 2021.10.22.21265382. https://doi.org/10.1101/2021.10.22.21265382

Karakuzu, A., Biswas, L., Cohen-Adad, J., & Stikov, N. (2021). Vendor-neutral sequences (venus) and fully transparent workflows improve inter-vendor reproducibility of quantitative mri. *bioRxiv*.

Keenan, K. E., Gimbutas, Z., Dienstfrey, A., & Stupic, K. F. (2019). Assessing effects of scanner upgrades for clinical studies. *Journal of Magnetic Resonance Imaging*, *50*(6), 1948–1954.

Keenan, K. E., Gimbutas, Z., Dienstfrey, A., Stupic, K. F., Boss, M. A., Russek, S. E., Chenevert, T. L., Prasad, P. V., Guo, J., Reddick, W. E., Cecil, K. M., Shukla-Dave, A., Aramburu Nunez, D., Shridhar Konar, A., Liu, M. Z., Jambawalikar, S. R., Schwartz, L. H., Zheng, J., Hu, P., & Jackson, E. F. (2021). Multi-site, multi-platform comparison of mri t1 measurement using the system phantom. *PLOS ONE*, *16*(6), e0252966. https://doi.org/10.1371/journal.pone.0252966

Kessler, L. G., Barnhart, H. X., Buckler, A. J., Choudhury, K. R., Kondratovich, M. V., Toledano, A., Guimaraes, A. R., Filice, R., Zhang, Z., & Sullivan, D. C. (2015). The emerging science of quantitative imaging biomarkers terminology and definitions for scientific studies and regulatory submissions. *Statistical methods in medical research*, *24*(1), 9–26.

Kim, H.-Y. (2013). Statistical notes for clinical researchers: evaluation of measurement error 1: using intraclass correlation coefficients. *Restorative dentistry & endodontics*, *38*(2), 98–102.

Kirkhorn, J. (1999). Introduction to iq-demodulation of rf-data. *IFBT, NTNU*, *15*.

Lankford, C. L., & Does, M. D. (2013). On the inherent precision of mcdespot. *Magnetic resonance in medicine*, *69*(1), 127–136.

Laule, C., Kozlowski, P., Leung, E., Li, D. K., MacKay, A. L., & Moore, G. W. (2008). Myelin water imaging of multiple sclerosis at 7 t: correlations with histopathology. *Neuroimage*, *40*(4), 1575–1580.

Laule, C., Vavasour, I., Moore, G., Oger, J., Li, D. K., Paty, D., & MacKay, A. (2004). Water content and myelin water fraction in multiple sclerosis. *Journal of neurology*, *251*(3), 284–293.

Lauterbur, P. C. (1989). Image formation by induced local interactions. examples employing nuclear magnetic resonance. 1973. *Clin Orthop Relat Res*, (244), 3–6. https://www.ncbi.nlm.nih.gov/pubmed/2663289

Lawrence, A., & McDonald, M. (1966). The investigation of lipid-water systems (part i) by classical and nmr methods. *Molecular Crystals and Liquid Crystals*, *1*(2), 205–223.

Layton, K. J., Kroboth, S., Jia, F., Littin, S., Yu, H., Leupold, J., Nielsen, J.-F., Stöcker, T., & Zaitsev, M. (2017). Pulseq: a rapid and hardware-independent pulse sequence prototyping framework. *Magnetic resonance in medicine*, *77*(4), 1544–1552.

Lazari, A., & Lipp, I. (2021). Can mri measure myelin? systematic review, qualitative assessment, and meta-analysis of studies validating microstructural imaging with myelin histology. *Neuroimage*, 117744.

Lee, Y., Callaghan, M. F., Acosta-Cabronero, J., Lutti, A., & Nagy, Z. (2019). Establishing intra- and inter-vendor reproducibility of t1 relaxation time measurements with 3t mri. *Magn Reson Med*, *81*(1), 454–465. https://doi.org/10.1002/mrm.27421

Levesque, I. R., & Pike, G. B. (2009b). Characterizing healthy and diseased white matter using quantitative magnetization transfer and multicomponent t2 relaxometry: a unified view via a four-pool model. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *62*(6), 1487–1496.

Levitt, M. F., & Gaudino, M. (1950). Measurement of body water compartments. *The American journal of medicine*, *9*(2), 208–215.

Lévy, S., Guertin, M.-C., Khatibi, A., Mezer, A., Martinu, K., Chen, J.-I., Stikov, N., Rainville, P., & Cohen-Adad, J. (2018). Test-retest reliability of myelin imaging in the human spinal cord: measurement errors versus region-and aging-induced variations. *PloS one*, *13*(1), e0189944.

MacKay, A., Whittall, K., Adler, J., Li, D., Paty, D., & Graeb, D. (1994). In vivo visualization of myelin water in brain by magnetic resonance. *Magn Reson Med*, *31*(6), 673–677.

MacKay, A. L., & Laule, C. (2016). Magnetic resonance of myelin water: an in vivo marker for myelin. *Brain plasticity*, *2*(1), 71–91.

Magland, J. F., Li, C., Langham, M. C., & Wehrli, F. W. (2016). Pulse sequence programming in a dynamic visual environment: sequencetree. *Magnetic resonance in medicine*, *75*(1), 257–265.

Mancini, M., Karakuzu, A., Cohen-Adad, J., Cercignani, M., Nichols, T. E., & Stikov, N. (2020). An interactive meta-analysis of mri biomarkers of myelin. *eLife*, *9*. https://doi.org/10.7554/elife.61523

Mansfield, P., & Maudsley, A. A. (1977). Planar spin imaging by nmr. *Journal of Magnetic Resonance (1969)*, *27*(1), 101–119. https://doi.org/https://doi.org/10.1016/0022-2364(77)90197-4

Mariette, F., Collewet, G., Davenel, A., Lucas, T., & Musse, M. (2012). Quantitative mri in food science & food engineering. *Encyclopedia of Magnetic Resonance by John Wiley.*

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological methods*, *1*(1), 30.

Melbourne, A., Eaton-Rosen, Z., Bainbridge, A., Kendall, G. S., Cardoso, M. J., Robertson, N. J., Marlow, N., & Ourselin, S. (2013). Measurement of myelin in the preterm brain: multi-compartment diffusion imaging and multi-component t 2 relaxometry. *International Conference on Medical Image Computing and Computer-Assisted Intervention*, 336–344.

Mezer, A., Yeatman, J. D., Stikov, N., Kay, K. N., Cho, N. J., Dougherty, R. F., Perry, M. L., Parvizi, J., Hua le, H., Butts-Pauly, K., & Wandell, B. A. (2013). Quantifying the local tissue volume and composition in individual brains with magnetic resonance imaging. *Nat Med*, *19*(12), 1667–72. https://doi.org/10.1038/nm.3390

Mezrich, R. (1995). A perspective on k-space. *Radiology*, *195*(2), 297–315.

Moritz, M., Redlich, T., Günyar, S., Winter, L., & Wulfsberg, J. P. (2019). On the economic value of open source hardware–case study of an open source magnetic resonance imaging scanner. *Journal of Open Hardware*, *3*(1).

Nguyen, T. D., Deh, K., Monohan, E., Pandya, S., Spincemaille, P., Raj, A., Wang, Y., & Gauthier, S. A. (2016). Feasibility and reproducibility of whole brain myelin water mapping in 4 minutes using fast acquisition with spiral trajectory and adiabatic t2prep (fast-t2) at 3t. *Magnetic resonance in medicine*, *76*(2), 456–465.

Nielsen, J.-F., & Noll, D. C. (2018). Toppe: a framework for rapid prototyping of mr pulse sequences. *Magnetic resonance in medicine*, *79*(6), 3128–3134.

Nishimura, D. G. (1996). *Principles of magnetic resonance imaging.* Stanford University.

Norton, W., & Autilio, L. (1966). The lipid composition of purified bovine brain myelin. *Journal of neurochemistry*, *13*(4), 213–222.

Novikov, D. S., Kiselev, V. G., & Jespersen, S. N. (2018). On modeling. *Magn Reson Med*, *79*(6), 3172–3193. https://doi.org/10.1002/mrm.27101

Ogbole, G. I., Adeyomoye, A. O., Badu-Peprah, A., Mensah, Y., & Nzeh, D. A. (2018). Survey of magnetic resonance imaging availability in west africa. *The Pan African Medical Journal*, *30*.

Okubo, G., Okada, T., Yamamoto, A., Kanagaki, M., Fushimi, Y., Okada, T., Murata, K., & Togashi, K. (2016). Mp2rage for deep gray matter measurement of the brain: a comparative study with mprage. *J Magn Reson Imaging*, *43*(1), 55–62. https://doi.org/10.1002/jmri.24960

Osmera, P., & Vanicek, J. (1940). Gyroscopic behavior of toroidal fractal protons in nmr. *Energy*, *19*(1), 0.

Ou, X., & Gochberg, D. F. (2008). Mt effects and t1 quantification in single-slice spoiled gradient echo imaging. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *59*(4), 835–845.

Paus, T., Collins, D., Evans, A., Leonard, G., Pike, B., & Zijdenbos, A. (2001). Maturation of white matter in the human brain: a review of magnetic resonance studies. *Brain research bulletin*, *54*(3), 255–266.

Quirk, J. D., Bretthorst, G. L., Duong, T. Q., Snyder, A. Z., Springer Jr, C. S., Ackerman, J. J., & Neil, J. J. (2003). Equilibrium water exchange between the intra-and extracellular spaces of mammalian brain. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *50*(3), 493–499.

Ravi, K. S., Geethanath, S., & Vaughan, J. T. (2019). Pypulseq: a python package for mri pulse sequence design. *Journal of Open Source Software*, *4*(42), 1725.

Rousselet, G. A., Pernet, C. R., & Wilcox, R. R. (2017). Beyond differences in means: robust graphical methods to compare two groups in neuroscience. *European Journal of Neuroscience*, *46*(2), 1738–1748.

Rushton, W. (1951). A theory of the effects of fibre size in medullated nerve. *The Journal of physiology*, *115*(1), 101–122.

Schmeel, F. C., Vomweg, T., Träber, F., Gerhards, A., Enkirch, S. J., Faron, A., Sprinkart, A. M., Schmeel, L. C., Luetkens, J. A., & Thomas, D. (2019). Proton density fat fraction mri of vertebral bone marrow: accuracy, repeatability, and reproducibility among readers, field strengths, and imaging platforms. *Journal of Magnetic Resonance Imaging*, *50*(6), 1762–1772.

Shepherd, T. M., Thelwall, P. E., Stanisz, G. J., & Blackband, S. J. (2009). Aldehyde fixative solutions alter the water relaxation and diffusion properties of nervous tissue. *Magnetic*

Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine, *62*(1), 26–34.

Sherif, T., Rioux, P., Rousseau, M.-E., Kassis, N., Beck, N., Adalat, R., Das, S., Glatard, T., & Evans, A. C. (2014). Cbrain: a web-based, distributed computing platform for collaborative neuroimaging research. *Frontiers in neuroinformatics*, *8*, 54.

Siegel, E. (2017). *Why does the proton spin? physics holds a surprising answer.* https://www.forbes.com/sites/startswithabang/2017/04/19/why-does-the-proton-spin-physics-holds-a-surprising-answer/?sh=73c3c4282c3a

Siegel, E. (2019). *No, thermodynamics does not explain our perceived arrow of time.* https://www.forbes.com/sites/startswithabang/2019/11/22/no-thermodynamics-does-not-explain-our-perceived-arrow-of-time/?sh=5229e7883109

Sigalovsky, I. S., Fischl, B., & Melcher, J. R. (2006). Mapping an intrinsic mr property of gray matter in auditory cortex of living humans: a possible marker for primary cortex and hemispheric differences. *Neuroimage*, *32*(4), 1524–1537.

Siri, W. E. (1956). The gross composition of the body. *Advances in biological and medical physics* (pp. 239–280). Elsevier.

Sled, J. G., & Pike, G. B. (2000). Correction for b1 and b0 variations in quantitative t2 measurements using mri. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *43*(4), 589–593.

Smith, A. K., Dortch, R. D., Dethrage, L. M., & Smith, S. A. (2014). Rapid, high-resolution quantitative magnetization transfer mri of the human spinal cord. *NeuroImage*, *95*, 106–116.

Srinivasan, B. (2020). *A thread on independent replication.* https://twitter.com/balajis/status/1337598439266250752?s=20

Stikov, N., Boudreau, M., Levesque, I. R., Tardif, C. L., Barral, J. K., & Pike, G. B. (2015). On the accuracy of t-1 mapping: searching for common ground. *Magnetic Resonance in Medicine*, *73*(2), 514–522. https://doi.org/10.1002/mrm.25135

Stikov, N., Campbell, J., Stroh, T., Lavelée, M., Frey, S., Novek, J., Nuara, S., Ho, M., Bedell, B., Dougherty, R., Leppert, I., Boudreau, M., Narayanan, S., Picard, P., Duva, l. T., Cohen-Adad, J., Gasecka, A., Côté, D., & G.B., P. (2015). In vivo histology of the myelin g-ratio with magnetic resonance imaging. *Neuroimage*, *In press.*

Stikov, N., Trzasko, J. D., & Bernstein, M. A. (2019). Reproducibility and the future of mri research. *Magn Reson Med*, *82*(6), 1981–1983. https://doi.org/10.1002/mrm.27939

Stöcker, T., Vahedipour, K., Pflugfelder, D., & Shah, N. J. (2010). High-performance computing mri simulations. *Magnetic resonance in medicine*, *64*(1), 186–193.

Stüber, C., Morawski, M., Schäfer, A., Labadie, C., Wähnert, M., Leuze, C., Streicher, M., Barapatre, N., Reimann, K., & Geyer, S. (2014). Myelin and iron concentration in the human brain: a quantitative study of mri contrast. *Neuroimage*, *93*, 95–106.

Stupic, K. F., Ainslie, M., Boss, M. A., Charles, C., Dienstfrey, A. M., Evelhoch, J. L., Finn, P., Gimbutas, Z., Gunter, J. L., Hill, D. L. G., Jack, C. R., Jackson, E. F., Karaulanov, T., Keenan, K. E., Liu, G., Martin, M. N., Prasad, P. V., Rentz, N. S., Yuan, C., & Russek, S. E. (2021). A standard system phantom for magnetic resonance imaging. *Magnetic Resonance in Medicine*, *86*(3), 1194–1211. https://doi.org/10.1002/mrm.28779

Taylor, B. N., & Kuyatt, C. E. (1994). Guidelines for evaluating and expressing the uncertainty of nist measurement results.

Teixeira, R. P., Neji, R., Wood, T. C., Baburamani, A. A., Malik, S. J., & Hajnal, J. V. (2020). Controlled saturation magnetization transfer for reproducible multivendor variable flip angle t1 and t2 mapping. *Magnetic resonance in medicine*, *84*(1), 221–236.

Todorich, B., Pasquini, J. M., Garcia, C. I., Paez, P. M., & Connor, J. R. (2009). Oligodendrocytes and myelination: the role of iron. *Glia*, *57*(5), 467–478.

Tong, G., Gaspar, A. S., Qian, E., Ravi, K. S., Vaughan Jr, J. T., Nunes, R. G., & Geethanath, S. (2021). A framework for validating open-source pulse sequences. *Magnetic resonance imaging*.

Tozer, D., Ramani, A., Barker, G., Davies, G., Miller, D., & Tofts, P. (2003). Quantitative magnetization transfer mapping of bound protons in multiple sclerosis. *Magnetic Resonance in Medicine: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *50*(1), 83–91.

Troutman, M. Y., Mastikhin, I. V., Balcom, B. J., Eads, T. M., & Ziegler, G. R. (2001). Moisture migration in soft-panned confections during engrossing and aging as observed by magnetic resonance imaging. *Journal of Food Engineering*, *48*(3), 257–267.

Turner, R. (2017). Peter mansfield (1933–2017). *Nature*, *543*(7644), 180–180. https://doi.org/10.1038/543180a

Uhlenbeck, G. E., & Goudsmit, S. (1925). Ersetzung der hypothese vom unmechanischen zwang durch eine forderung bezüglich des inneren verhaltens jedes einzelnen elektrons. *Die Naturwissenschaften*, *13*(47), 953–954.

Wang, J., Qiu, M., Kim, H., & Constable, T. (2006). T1 measurements incorporating flip angle calibration and correction in vivo. *Journal of Magnetic Resonance*, *182*(2), 283–292. https://doi.org/citeulike-article-id:1430243doi:10.1016/j.jmr.2006.07.005

Watson, R. T., Boudreau, M.-C., York, P. T., Greiner, M. E., & Wynn Jr, D. (2008). The business of open source. *Communications of the ACM*, *51*(4), 41–46.

Waxman, S. G. (1982). Membranes, myelin, and the pathophysiology of multiple sclerosis. *New England Journal of Medicine*, *306*(25), 1529–1533.

Webb, A. (2016). The principles of magnetic resonance, and associated hardware.

Weinberger, D. R., & Radulescu, E. (2016). Finding the elusive psychiatric "lesion" with 21st-century neuroanatomy: a note of caution. *American Journal of Psychiatry*, *173*(1), 27–33.

Wilcox, R. R., & Rousselet, G. A. (2018). A guide to robust statistical methods in neuroscience. *Current protocols in neuroscience*, *82*(1), 8–42.

Wolff, S. D., & Balaban, R. S. (1989b). Magnetization transfer contrast (mtc) and tissue water proton relaxation in vivo. *Magnetic resonance in medicine*, *10*(1), 135–144.

Wright, P., Mougin, O., Totman, J., Peters, A., Brookes, M., Coxon, R., Morris, P., Clemence, M., Francis, S., & Bowtell, R. (2008). Water proton t 1 measurements in brain tissue at 7, 3, and 1.5 t using ir-epi, ir-tse, and mprage: results and optimization. *Magnetic Resonance Materials in Physics, Biology and Medicine*, *21*(1), 121–130.

Yarnykh, V. L. (2007b). Actual flip-angle imaging in the pulsed steady state: a method for rapid three-dimensional mapping of the transmitted radiofrequency field. *Magn Reson Med*, *57*(1), 192–200.

Young, I. R. (2004). Significant events in the development of mri. *Journal of Magnetic Resonance Imaging: An Official Journal of the International Society for Magnetic Resonance in Medicine*, *20*(2), 183–186.

Zaimi, A., Wabartha, M., Herman, V., Antonsanti, P.-L., Perone, C. S., & Cohen-Adad, J. (2018). Axondeepseg: automatic axon and myelin segmentation from microscopy data using convolutional neural networks. *Scientific reports*, *8*(1), 1–11.

Zeeman, P. (1897). The effect of magnetisation on the nature of light emitted by a substance. *nature*, *55*(1424), 347.

Ziegler, G. R., MacMillan, B., & Balcom, B. J. (2003). Moisture migration in starch molding operations as observed by magnetic resonance imaging. *Food Research International*, *36*(4), 331–340.

Zou, G. (2012). Sample size formulas for estimating intraclass correlation coefficients with precision and assurance. *Statistics in medicine*, *31*(29), 3972–3981.

Zur, Y., Wood, M., & Neuringer, L. (1991). Spoiling of transverse magnetization in steady-state sequences. *Magnetic resonance in medicine*, *21*(2), 251–263.

APPENDIX A   SUPPLEMENTARY MATERIALS FOR ARTICLE 1

A.1   qMRLab Architecture

Each quantitative MRI method is modularized as a Matlab class that contains: the protocol parameters, the fitting options, and model options.
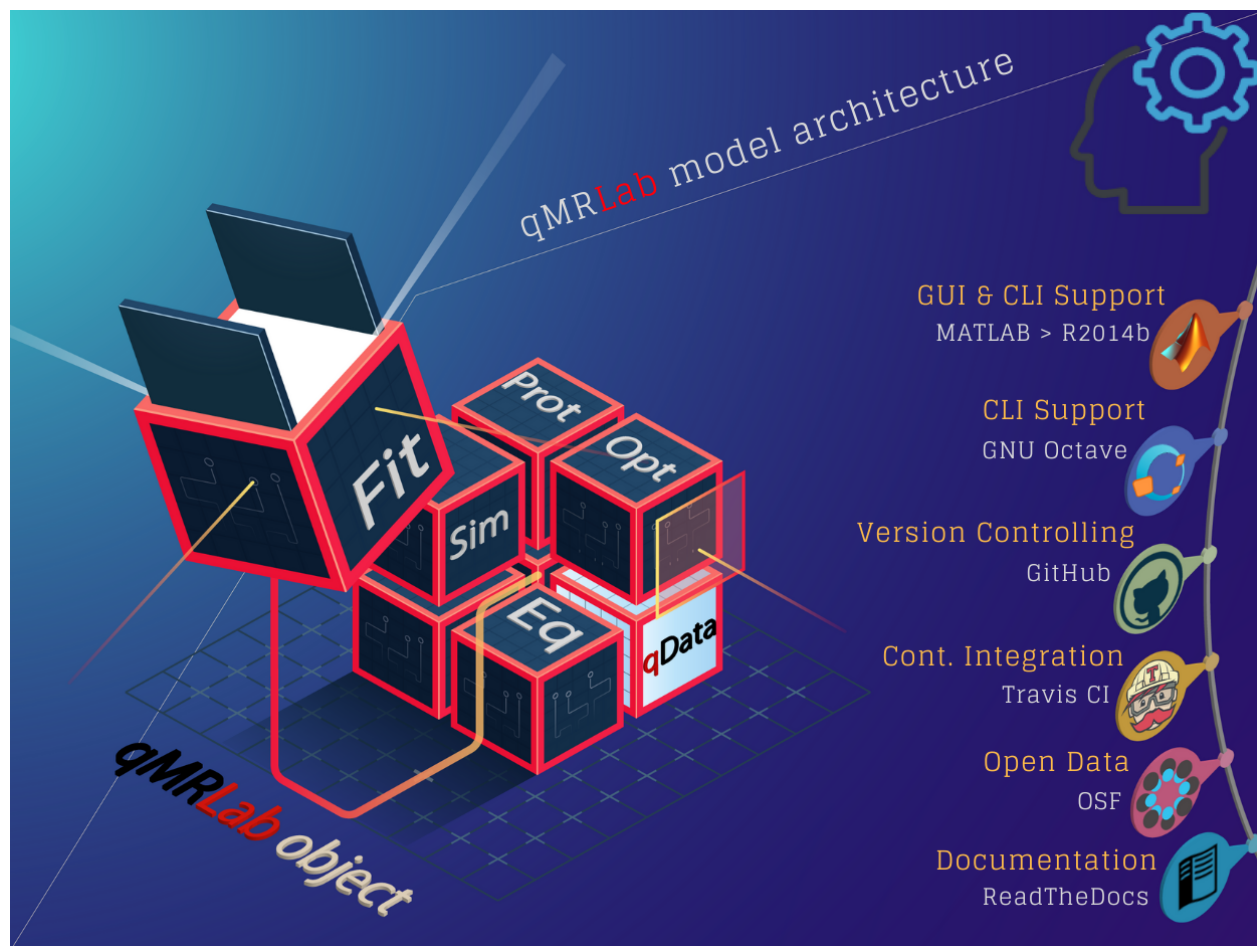
Figure A.1 The qMRLab codebase is based on an object oriented design to facilitate the implementation of qMRI methods.

The class contains the following methods: 1) **equation**, to analytically generate qMR data; 2) **fit**, to fit the experimental data using an analytical equation; 3) **plotModel**, to plot the fitting curves; 4) **plotProt**, to display the protocol (e.g. b-space in diffusion); and 5) **Sim**, to perform various simulations (see next section). The function **qMRusage** provides usage examples for each method and can help guide users who would like to implement scripts or pipelines. A class object can easily be loaded within the GUI for reproducibility and for sharing across centers.

The developer documentation is available at https://github.com/qMRLab/qMRLab/wiki for further details about qMRLab's application programming interface.

A.1.1 Simulations

Several simulation modules are included in qMRLab (folder **Addons**): 1) "Single Voxel Curve", to generate noisy (Rician noise) synthetic MR signals and fit the simulated data, 2) "Sensitivity Analysis", to estimate the precision of fitted parameters for a particular SNR, and for different value of one parameter (all other parameters fixed), 3) "Multi-voxel Distribution", to estimate the precision of fitted parameters for a wide range of parameter combinations, 4) "Protocol Optimization", to produce an optimized protocol using the Cramér-Rao Lower Bound (CRLB) as an objective function, and SOMA all-to-one as an optimizer. Note that modules 2 and 3 above depend on module 1.

A.1.2 Modalities

Currently, qMRLab includes 24 implementations from 8 different categories for qMRI analysis:

1. T1 relaxometry

    (a) Variable flip angle (VFA) T1 mapping

    (b) Inversion-recovery (IR) T1 mapping

    (c) Magnetization prepared two rapid gradient-ehoes (MP2RAGE) T1 mapping

    (d) Macromolecular tissue volume (MTV) quantification in the brain

2. T2 and $T2^*$ relaxometry

    (a) Myelin water fraction (MWF) mapping from multi-exponential T2w data

    (b) Mono-exponential T2 mapping from multi-echo spin-echo data

    (c) $T2^*$ mapping with from multi-echo gradient-echo data

3. Magnetization transfer

   (a) quantitative magnetization transfer (qMT) using balanced steady-state freee precession (bSSFP) acquisition

   (b) qMT using inversion-recovery fast spin-echo (IRFSE) acquisition

   (c) qMT using spoiled gradient-echo (SPGR) acquisition

   (d) Magnetization transfer saturation index (MTsat) mapping

   (e) Magnetization transfer ratio (MTR) mapping

   (f) Inhomogeneous magnetization transfer (ihMT) imaging

4. Diffusion imaging

   (a) Diffusion tensor imaging (DTI)

   (b) Neurite orientation dispersion and density imaging (NODDI)

   (c) Accelerated NODDI

   (d) Composite hindered and restricted model for diffusion (CHARMED)

5. Field mapping

   (a) Actual flip angle imaging (AFI) for B1+ mapping

   (b) Double-angle method for B1+ mapping

   (c) Dual-echo method for B0 mapping

6. Quantitative susceptibility mapping

   (a) Fast quantitative susceptibility mapping using Split-Bregman regularization

7. Denoising

   (a) Noise level histogram fitting within a noise mask

   (b) 4D image denoising and noise map estimation for diffusion-weighted data

8. Processing

   (a) Spatial filtering module for 2D or 3D field maps

All models are tested for consistent results across versions and accurate fitting (using the single voxel curve simulation) and released with examples datasets that are publicly available at https://osf.io/tmdfu/.

A.2　User interfaces
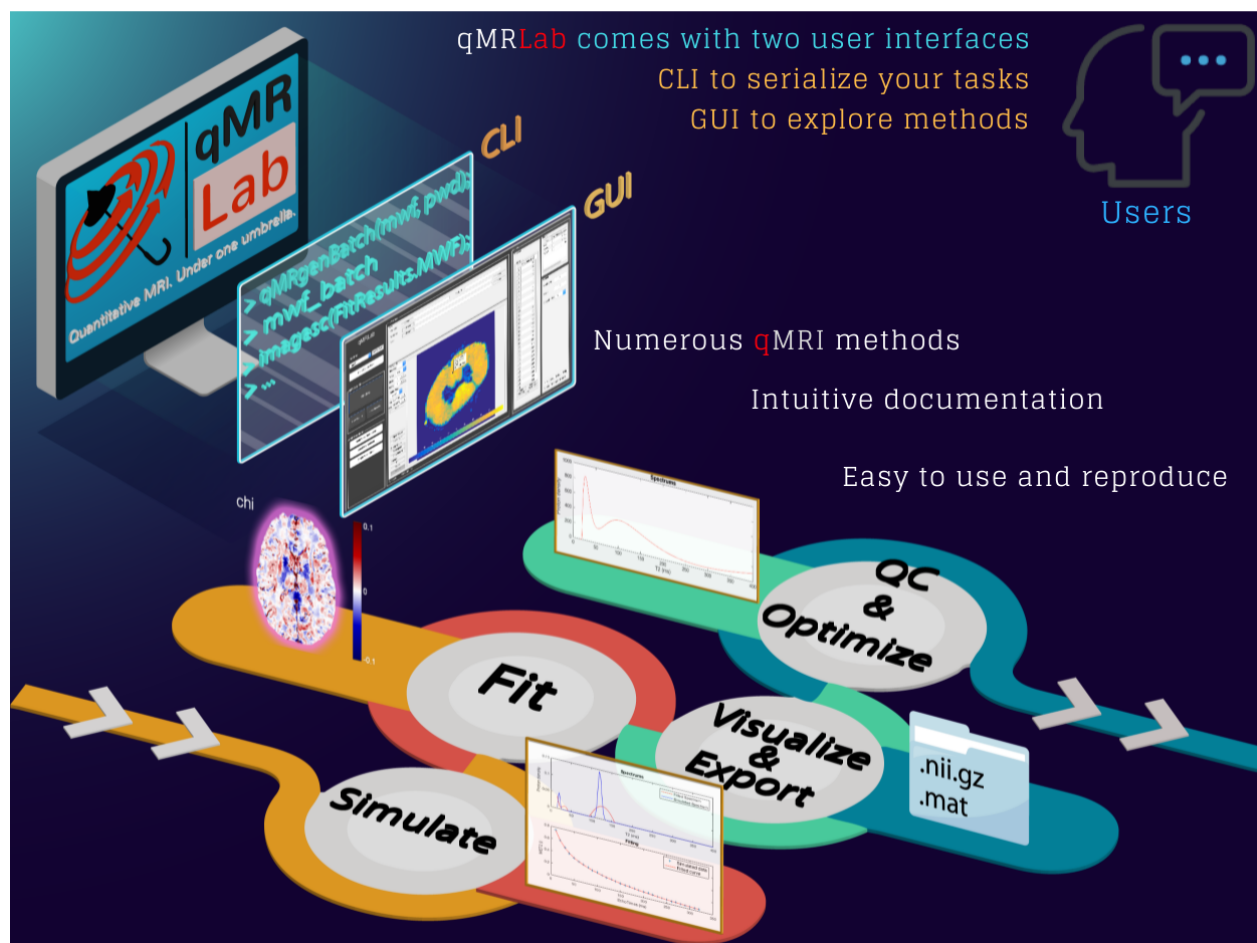
Figure A.2 qMRLab includes a user-friendly graphical user interface (GUI) and a command-line interface (CLI) for serializing tasks.

Automatically generated documentation (based on class definition headers, and batch examples), as well as a video introduction to the basics of qMRLab are available online http://qmrlab.readthedocs.io/.

A.3 Continious integration and deployment

The integrity of the codebase is tested against changes both on MATLAB and Octave-GNU to ensure cross-platform compatibility. The build pipelines are available at https://dev.azure.com/neuropoly/qMRLab and https://github.com/qMRLab/qMRLab/actions for MATLAB and Octave, respectively. A joint code coverage report is published at the end of each build (https://app.codecov.io/gh/qMRLab/qMRLab). On each release, different flavors of Octave-based qMRLab Docker images are published to provide version-controlled portable execution environments (e.g., qmrlab/minimal and qmrlab/octjn for workflows and notebooks). All the Docker images are publicly availabe at https://hub.docker.com/u/qmrlab.

An Azure self-hosted MATLAB release pipeline has been set for qMRLab to compile a standalone MATLAB executable with GUI and to publish the built application in a Docker image (qmrlab/mcrgui). In addition, user documentation is generated along with one Jupyter notebook per module and a BinderHub image is built with these notebooks to publish documentation pages incorporating online-executable content.

APPENDIX B SUPPLEMENTARY MATERIALS FOR ARTICLE 2

B.1 qMRI-BIDS main specification

Beginning from the release v1.5.0, the BIDS main specification presents a clear distinction between the conventional and quantitative MR images. In addition, several metadata fields have been introduced with defined units to enable an array of qMRI applications in BIDS. These additions and definitions are listed in this section.

B.1.1 Metadata fields

Table B.1 Metadata fields added to the main specification for qMRI.

Key name	Requirement level	Data type	Description
NonlinearGradient Correction	RECOMMENDED	[boolean]	Boolean stating if the image saved has been corrected for gradient nonlinearities by the scanner sequence.
MRAcquisitionType	RECOMMENDED, but REQUIRED for Arterial Spin Labeling	[string]	Possible values: 2D or 3D. Type of sequence readout. Corresponds to DICOM Tag 0018,0023 MR Acquisition Type.

Table B.1 continued: Metadata fields added to the main specification for qMRI.

MTState	RECOMMENDED	[boolean]	Boolean stating whether the magnetization transfer pulse is applied. Corresponds to DICOM tag (0018, 9020) Magnetization Transfer.
MTOffsetFrequency	RECOMMENDED if the MTstate is True.	[number]	The frequency offset of the magnetization transfer pulse with respect to the central H1 Larmor frequency in Hertz (Hz).
MTPulseBandwidth	RECOMMENDED if the MTstate is True.	[number]	The excitation bandwidth of the magnetization transfer pulse in Hertz (Hz).
MTNumberOfPulses	RECOMMENDED if the MTstate is True.	[number]	The number of magnetization transfer RF pulses applied before the readout.

Table B.1 continued: Metadata fields added to the main specification for qMRI.

MTPulseShape	RECOMMENDED if the MTstate is True.	[string]	Shape of the magnetization transfer RF pulse waveform. Accepted values: HARD, GAUSSIAN, GAUSSHANN (gaussian pulse with Hanning window), SINC, SINCHANN (sinc pulse with Hanning window), SINCGAUSS (sinc pulse with Gaussian window), FERMI.
MTPulseDuration	RECOMMENDED if the MTstate is True.	[number]	Duration of the magnetization transfer RF pulse in seconds.
SpoilingState	RECOMMENDED	[boolean]	Boolean stating whether the pulse sequence uses any type of spoiling strategy to suppress residual transverse magnetization.
SpoilingType	RECOMMENDED if the SpoilingState is True.	[string]	Specifies which spoiling method(s) are used by a spoiled sequence. Accepted values: RF, GRADIENT or COMBINED.

Table B.1 continued: Metadata fields added to the main specification for qMRI.

SpoilingRFPhaseIncrement	RECOMMENDED if the SpoilingType is RF or COMBINED.	[number]	The amount of incrementation described in degrees, which is applied to the phase of the excitation pulse at each TR period for achieving RF spoiling.
SpoilingGradientMoment	RECOMMENDED if the SpoilingType is GRADIENT or COMBINED.	[number]	Zeroth moment of the spoiler gradient lobe in millitesla times second per meter (mT.s/m).
SpoilingGradientDuration	RECOMMENDED if the SpoilingType is GRADIENT or COMBINED.	[number]	The duration of the spoiler gradient lobe in seconds. The duration of a trapezoidal lobe is defined as the summation of ramp-up and plateau times.

B.1.2 Conventional MR image definitions

Table B.3 Conventional MR image definitions updated by BEP001.

Name	suffix	Description

Table B.3 continued: Conventional MR image definitions updated by BEP001.

| T1 weighted images | T1w | In arbitrary units (arbitrary). The contrast of these images is mainly determined by spatial variations in the longitudinal relaxation time of the imaged specimen. In spin-echo sequences this contrast is achieved at relatively short repetition and echo times. To achieve this weighting in gradient-echo images, again, short repetition and echo times are selected; however, at relatively large flip angles. Another common approach to increase T1 weighting in gradient-echo images is to add an inversion preparation block to the beginning of the imaging sequence (for example, TurboFLASH or MP-RAGE). |

Table B.3 continued: Conventional MR image definitions updated by BEP001.

T2 weighted images	T2w	In arbitrary units (arbitrary). The contrast of these images is mainly determined by spatial variations in the (true) transverse relaxation time of the imaged specimen. In spin-echo sequences this contrast is achieved at relatively long repetition and echo times. Generally, gradient echo sequences are not the most suitable option for achieving T2 weighting, as their contrast natively depends on T2-star rather than on T2.
Proton density (PD) weighted images	PDw	In arbitrary units (arbitrary). The contrast of these images is mainly determined by spatial variations in the spin density (1H) of the imaged specimen. In spin-echo sequences this contrast is achieved at short repetition and long echo times. In a gradient-echo acquisition, PD weighting dominates the contrast at long repetition and short echo times, and at small flip angles.

Table B.3 continued: Conventional MR image definitions updated by BEP001.

T2star weighted images	T2starw	In arbitrary units (arbitrary). The contrast of these images is mainly determined by spatial variations in the (observed) transverse relaxation time of the imaged specimen. In spin-echo sequences, this effect is negated as the excitation is followed by an inversion pulse. The contrast of gradient-echo images natively depends on T2-star effects. However, for T2-star variation to dominate the image contrast, gradient-echo acquisitions are carried out at long repetition and echo times, and at small flip angles.
Fluid attenuated inversion recovery images	FLAIR	In arbitrary units (arbitrary). Structural images with predominant T2 contribution (a.k.a T2-FLAIR), in which signal from fluids (for example, CSF) is nulled out by adjusting inversion time, coupled with notably long repetition and echo times.
Inplane T1	inplaneT1	In arbitrary units (arbitrary). T1 weighted structural image matched to a functional (task) image.

Table B.3 continued: Conventional MR image definitions updated by BEP001.

Inplane T2	inplaneT2	In arbitrary units (arbitrary). T2 weighted structural image matched to a functional (task) image.
PD and T2 weighted images	PDT2	In arbitrary units (arbitrary). PDw and T2w images acquired using a dual echo FSE sequence through view sharing process (https://pubmed.ncbi.nlm.nih.gov/8010268/).
Homogeneous (flat) T1-weighted MP2RAGE image	UNIT1	In arbitrary units (arbitrary). UNIT1 images are REQUIRED to use this suffix regardless of the method used to generate them. Note that although this image is T1-weighted, regions without MR signal will contain white salt-and-pepper noise that most segmentation algorithms will fail on. Therefore, it is important to dissociate it from from T1w. Plase see MP2RAGE specific notes in the qMRI appendix for further information.

B.1.3 Quantitative map definitions and units

Structural MR images whose intensity is represented in a non-arbitrary scale constitute parametric maps.

Table B.5 Quantitative map definitions and units added to the main specification by BEP001.

Name	suffix	Description
Longitudinal relaxation time map	T1map	In seconds (s). T1 maps are REQUIRED to use this suffix regardless of the method used to generate them. See https://qmrlab.org/t1_book/intro for further reading on T1-mapping.
Longitudinal relaxation rate map	R1map	In seconds$_{-1}$> (1/s). R1 maps (R1 = 1/T1) are REQUIRED to use this suffix regardless of the method used to generate them.
True transverse relaxation time map	T2map	In seconds (s). T2 maps are REQUIRED to use this suffix regardless of the method used to generate them.
True transverse relaxation rate map	R2map	In seconds$_{-1}$> (1/s). R2 maps (R2 = 1/T2) are REQUIRED to use this suffix regardless of the method used to generate them.
Observed transverse relaxation time map	T2starmap	In seconds (s). T2-star maps are REQUIRED to use this suffix regardless of the method used to generate them.
Observed transverse relaxation rate map	R2starmap	In seconds$_{-1}$> (1/s). R2-star maps (R2star = 1/T2star) are REQUIRED to use this suffix regardless of the method used to generate them.

Table B.5 continued: Quantitative map definitions and units added to the main specification by BEP001.

Proton density map	PDmap	In arbitrary units (arbitrary). PD maps are REQUIRED to use this suffix regardless of the method used to generate them.
Magnetization transfer ratio map	MTRmap	In arbitrary units (arbitrary). MTR maps are REQUIRED to use this suffix regardless of the method used to generate them. MTRmap intensity values are RECOMMENDED to be represented in percentage in the range of 0-100%.
Magnetization transfer saturation map	MTsat	In arbitrary units (arbitrary). MTsat maps are REQUIRED to use this suffix regardless of the method used to generate them.
T1 in rotating frame (T1 rho) map	T1rho	In seconds (s). T1-rho maps are REQUIRED to use this suffix regardless of the method used to generate them.
Myelin water fraction map	MWFmap	In arbitrary units (arbitrary). MWF maps are REQUIRED to use this suffix regardless of the method used to generate them. MWF intensity values are RECOMMENDED to be represented in percentage in the range of 0-100%.

Table B.5 continued: Quantitative map definitions and units added to the main specification by BEP001.

Macromolecular tissue volume (MTV) map	MTVmap	In arbitrary units (arbitrary). MTV maps are REQUIRED to use this suffix regardless of the method used to generate them.
Combined PD/T2 map	PDT2map	In arbitrary units (arbitrary). Combined PD/T2 maps are REQUIRED to use this suffix regardless of the method used to generate them.
Quantitative susceptibility map (QSM)	Chimap	In parts per million (ppm). QSM allows for determining the underlying magnetic susceptibility of tissue (Chi) (https://onlinelibrary.wiley.com/doi/10.1002/mrm.25358). Chi maps are REQUIRED to use this suffix regardless of the method used to generate them.
RF transmit field map	TB1map	In arbitrary units (arbitrary). Radio frequency (RF) transmit (B1+) field maps are REQUIRED to use this suffix regardless of the method used to generate them. TB1map intensity values are RECOMMENDED to be represented as percent multiplicative factors.

Table B.5 continued: Quantitative map definitions and units added to the main specification by BEP001.

RF receive sensitivity map	RB1map	In arbitrary units (arbitrary). Radio frequency (RF) receive (B1-) sensitivity maps are REQUIRED to use this suffix regardless of the method used to generate them. RB1map intensity values are RECOMMENDED to be represented as percent multiplicative factors.
Observed signal amplitude (S0) map	S0map	In arbitrary units (arbitrary). For a multi-echo (typically fMRI) sequence, S0 maps index the baseline signal before exponential (T2-star) signal decay. In other words: the exponential of the intercept for a linear decay model across log-transformed echos. For more information, please see, for example, the tedana documentation at https://tedana.readthedocs.io/en/latest/. S0 maps are RECOMMENDED to use this suffix if derived from an ME-FMRI dataset.

Table B.5 continued: Quantitative map definitions and units added to the main specification by BEP001.

Equilibrium magnetization (M0) map	M0map	In arbitrary units (arbitrary). A common quantitative MRI (qMRI) fitting variable that represents the amount of magnetization at thermal equilibrium. M0 maps are RECOMMENDED to use this suffix if generated by qMRI applications (for example, variable flip angle T1 mapping).

B.2 qMRI appendix

Quantitative MRI (qMRI) is a collection of methods aiming at generating parametric maps that can characterize underlying tissue properties. Unlike those of conventional MR images (for example, T1w or T2w), intensity values of quantitative maps are not represented in an arbitrary range. Instead, these maps are represented either in absolute physical units (for example, seconds for T1map), or within an application dependent range of arbitrary units (for example, myelin water fraction MWFmap in brain).

Organization of qMRI data in BIDS

Unlike conventional MR images, quantitative maps are not immediate products of the image reconstruction step (from k-space data to structural images). Intensity values of qMRI maps are calculated by fitting a collection of parametrically linked images to a biophysical model or to an MRI signal representation. This processing is typically carried out in the image domain. There are two main ways to obtain a quantitative map:

1. Pre-generated qMRI maps: The qMRI maps are generated right after the reconstruction of required input images and made available to the user at the scanner console. The acquisition scenarios may include (a) vendor pipelines or (b) open-source pipelines deployed at the scanner site.

2. Post-generated qMRI maps: The qMRI maps are generated from a collection of input data after they are exported from the scanner site. This type of processing is commonly carried out using an open-source software such as:

- hMRI toolbox (https://github.com/hMRI-group/hMRI-toolbox)

- mrQ (https://github.com/mezera/mrQ)

- PyQMRI (https://github.com/IMTtugraz/PyQMRI)

- qmap (https://www.medphysics.wisc.edu/ samsonov/qmap/doc/qmap.html)

- qMRLab (https://github.com/qmrlab/qmrlab)

- QUIT (https://github.com/spinicist/QUIT)

Inputs are file collections

The common concept of entity-linked file collections enables the description of a qMRI application by creating logical groups of input files through suffix and certain entities representing acquisition parameters (echo, flip, inv, mt) or file parts (part).

If a qMRI file collection is intended for creating structural quantitative maps (for example, T1map), files belonging to that collection are stored in the anat subfolder. Below is an example file collection for MP2RAGE:

```
sub-01/
└── anat/
    ├── sub-01_inv-1_part-mag_MP2RAGE.nii.gz
    ├── sub-01_inv-1_part-phase_MP2RAGE.nii.gz
    ├── sub-01_inv-1_MP2RAGE.json
    ├── sub-01_inv-2_part-mag_MP2RAGE.nii.gz
    ├── sub-01_inv-2_part-phase_MP2RAGE.nii.gz
    └── sub-01_inv-2_MP2RAGE.json
```

Commonly, RF fieldmaps (B1+ and B1- maps) are used for the correction of structural quantitative maps. As these images do not convey substantial structural information, respective file collections of RF fieldmaps are stored in the fmap subfolder. Below is an example file collection for RF transmit field map TB1EPI:

```
sub-01/
└── fmap/
    ├── sub-01_echo-1_flip-1_TB1EPI.nii.gz
    ├── sub-01_echo-1_flip-1_TB1EPI.json
    ├── sub-01_echo-2_flip-1_TB1EPI.nii.gz
    ├── sub-01_echo-2_flip-1_TB1EPI.json
    └── sub-01_echo-1_flip-2_TB1EPI.nii.gz
```

```
        │
        ├── sub-01_echo-1_flip-2_TB1EPI.json
        ├── sub-01_echo-2_flip-2_TB1EPI.nii.gz
        └── sub-01_echo-2_flip-2_TB1EPI.json
```

Please visit the file collections appendix to see the list of currently supported qMRI applications.

Quantitative maps are derivatives

Regardless of how they are obtained (pre- or post-generated), qMRI maps are stored in the derivatives folder. For example a T1map can be generated from an MP2RAGE file collection using either options.

If the map is post-generated:
```
ds-example/
└── derivatives/
    └── qMRI-software-name/
        └── sub-01/
            └── anat/
                ├── sub-01_T1map.nii.gz
                ├── sub-01_T1map.json
                ├── sub-01_UNIT1.nii.gz
                └── sub-01_UNIT1.json
```

If the map is pre-generated, for example, by a Siemens scanner:
```
ds-example/
└── derivatives/
    └── Siemens/
        └── sub-01/
            └── anat/
                ├── sub-01_T1map.nii.gz
                ├── sub-01_T1map.json
                ├── sub-01_UNIT1.nii.gz
                └── sub-01_UNIT1.json
```

Note: Even though the process from which pre-generated qMRI maps are obtained (vendor pipelines) is not known, vendors generally allow exporting of the corresponding input data. It is recommended to share them along with the vendor outputs, whenever possible for a qMRI method supported by BIDS.

Example datasets

You can find example file collections and qMRI maps organized according to BIDS at https://osf.io/k4bs5/.

Metadata requirements for qMRI data

The table of required entities for qMRI file collections are provided in the entity table. However, viability of a qMRI file collection is determined not only by the naming and organization of the input files, but also by which metadata fields are provided in accompanying json files.

Method-specific priority levels for qMRI file collections

Table B.7 Metadata priority levels for anatomy imaging data.

File collection	Required metadata	Optional metadata
VFA	FlipAngle, PulseSequenceType, RepetitionTimeExcitation	SpoilingRFPhaseIncrement
IRT1	InversionTime	
MP2RAGE[1]	FlipAngle, InversionTime, RepetitionTimeExcitation, RepetitionTimePreperation, NumberShots, MagneticFieldStrength	EchoTime
MESE	EchoTime	
MEGRE	EchoTime	
MTR	MTState	
MTS	FlipAngle, MTState, RepetitionTimeExcitation	
MPM	FlipAngle, MTState, RepetitionTimeExcitation	EchoTime

Explanation of the Table B.7:

- The metadata fields listed in the REQUIRED column are needed to perform a minimum viable qMRI processing for the corresponding file collection.

- Note that some of the metadata fields may be constant across different files in a file collection, yet still required as an input (for example, NumberShots in MP2RAGE). Such metadata fields MUST be provided in the accompanying JSON files.

---

[1]Please see MP2RAGE-specific notes for the calculation of NumberShots and regarding the organization of UNIT1 image.

- The metadata fields listed in the OPTIONAL column can be used to form different flavors of an existing file collection suffix, dispensing with the need for introducing a new suffix. See deriving the intended qMRI application from an ambiguous file collection section for details.

Table B.8 Metadata priority levels for anatomy imaging data.

File collection	Required metadata
TB1DAM	FlipAngle
TB1EPI	EchoTime, FlipAngle, TotalReadoutTime, MixingTime
TB1AFI	RepetitionTime
TB1TFL	
TB1RFM	
TB1SRGE[2]	FlipAngle, InversionTime, RepetitionTime-Excitation, RepetitionTimePreperation, NumberShots
RB1COR	

Metadata requirements for qMRI maps

As qMRI maps are stored as derivatives, they are subjected to the metadata requirements of derived datasets.

An example **dataset_description.json** for a qMRI map derivatives folder:
```
ds-example/
└─ derivatives/
   └─ qMRLab/
   └─ dataset_description.json/
      └─ sub-01/
         └─ anat/
            ├─ sub-01_T1map.nii.gz
            ├─ sub-01_T1map.json
            ├─ sub-01_M0map.nii.gz
            └─ sub-01_M0map.json
```

The corresponding **dataset_description.json** file:

---

[2]Please see TB1SRGE-specific notes for the calculation of NumberShots.

```
{
  "Name": "qMRLab Outputs",
  "BIDSVersion": "1.5.0",
  "DatasetType": "derivative",
  "GeneratedBy": [
  {
    "Name": "qMRLab",
    "Version": "2.4.1",
    "Container": {
    "Type": "docker",
    "Tag": "qmrlab/minimal:2.4.1"
     }
  },{
    "Name": "Manual",
    "Description": "Generated example T1map outputs"
        }],
    "SourceDatasets": [
  {
    "DOI": "DOI 10.17605/OSF.IO/K4BS5",
    "URL": "https://osf.io/k4bs5/",
    "Version": "1"
}]}
```

In addition to the metadata fields provided in the dataset_description.json, qMRI maps are RECOMMENDED to be accompanied by sidecar JSON files that contain further information about the quantified maps. Although this may not be the generic case for common derivative outputs, a proper interpretation of qMRI maps may critically depend on some metadata fields. For example, without the information of MagneticFieldStrength, white-matter T1 values in a T1map become elusive.

- All the acquisition parameters that are constant across the files in a file collection are RECOMMENDED to be added to the sidecar json of the qMRI maps.

- Relevant acquisition parameters that vary across files in a qMRI file collection are RECOMMENDED to be added to the sidecar json of the qMRI map in array form.

- The JSON file accompanying a qMRI map which is obtained by using open-source software is RECOMMENDED to include additional metadata fields listed in Table B.9:

Table B.9 Software provenance metadata tags.

Field name	Definition
BasedOn	List of files in a file collection to generate the map. Fieldmaps are also listed, if involved in the processing.
EstimationReference	Reference to the study/studies on which the implementation is based.
EstimationAlgorithm	Type of algoritm used to perform fitting (for example, linear, non-linear, LM and such)
Units	Units of the maps, in accordance with the BIDS specification.

Example qMRI processing software JSON provenance for a T1map:

```json
{
"BasedOn":["anat/sub-01_flip-1_VFA.nii.gz",
"anat/sub-01_flip-2_VFA.nii.gz",
"anat/sub-01_flip-3_VFA.nii.gz",
"anat/sub-01_flip-4_VFA.nii.gz",
"fmap/sub-01_TB1map.nii.gz"],
"EstimationPaper":"Deoni et. al.MRM, 2015",
"EstimationAlgorithm":"Linear",
"Units": "second",
"MagneticFieldStrength": "3",
"Manufacturer": "Siemens",
"ManufacturerModelName": "TrioTim",
"InstitutionName": "xxx",
"PulseSequenceType": "SPGR",
"PulseSequenceDetails": "Information beyond the sequence type that
    identifies
specific pulse sequence used (VB version, if not standard, Siemens
    WIP
sequence compiled on mm/dd/yyyy/)",
"RepetitionTimeExcitation": "35",
"EchoTime": "2.86",
"SliceThickness": "5",
"FlipAngle": ["5","10","15","20"]
```

```
}
```

Deriving the intended qMRI application from an ambigious file collection

Certain file collection suffixes may refer to a generic data collection regime such as variable flip angle (VFA), rather than a more specific acquisition, for example, magnetization prepared two gradient echoes (MP2RAGE). Such generic acquisitions can serve as a basis to derive various qMRI applications by changes to the acquisition sequence (for example, readout) type or by varying additional scan parameters.

If such an inheritance relationship is applicable between an already existing file collection and a new qMRI application to be included in the specification, the inheritor qMRI method is listed in the table below instead of introducing a new file collection suffix. This approach aims at:

- preventing the list of available suffixes from over-proliferation,

- providing qMRI-focused BIDS applications with a set of meta-data driven rules to infer possible fitting options,

- keeping an inheritance track of the qMRI methods described within the specification.

Table B.10 Conditions for deriving a new qMRI application from an existing suffix.

File-collection suffix	If REQUIRED metadata == Value	OPTIONAL metadata (entity / fixed)	Derived application name (NOT a suffix)
VFA	PulseSequenceType == SPGR		DESPOT1
VFA	PulseSequenceType == SSFP	SpoilingRFPhase (fixed)	DESPOT2
MP2RAGE		EchoTime (echo)	MP2RAGE-ME
MPM		EchoTime (echo)	MPM-ME

Table B.10 denotes whether the OPTIONAL metadata that forms a new flavor of qMRI application for the respective suffix varies across files of a file collection (which calls for using a linking entity) or fixed. If former is the case, the entity is to be added to the files in that

file collection. Note that this addition MUST be allowed by the priority levels given for that suffix in the entity table. If latter (fixed) is the case, filenames will remain the same; however, the optional metadata (third column) may define the flavor of the application (fourth column) along with the conditional value of a required metadata field (second column).

A derived qMRI application becomes available if all the optional metadata fields listed for the respective file collection suffix are provided for the data. In addition, conditional rules based on the value of a given required metada field can be set for the description of a derived qMRI application. Note that the value of this required metadata is fixed across constituent images of a file collection and defined in Method-specific priority levels for qMRI file collections.

For example, if the optional metadata field of PulseSequenceType is SPGR for a collection of anatomical images listed by the VFA suffix, the data qualifies for DESPOT1 T1 fitting. For the same suffix, if the PulseSequenceType metadata field has the value of SSFP, and the SpoilingRFPhaseIncrement is provided as a metadata field, then the dataset becomes eligible for DESPOT2 T2 fitting application.

Please note that optional metadata fields listed in the deriving the intended qMRI application from an ambiguous file collection table are included in the optional (third) column of the priority levels table for the consistency of this appendix.

Introducing a new qMRI file collection

If a qMRI application cannot be interpreted as a subtype of an already existing suffix of a qMRI-related file collection, we RECOMMEND adhering to the following principles to introduce a new suffix:

- All qMRI-relevant file collection suffixes are capitalized.

- Unless the pulse sequence is exclusively associated with a specific qMRI application (for example, MP2RAGE), sequence names are not used as suffixes.

- File collection suffixes for qMRI applications attain a clear description of the qMRI method that they relate to in the file collections appendix.

- Hyperlinks to example applications and reference method articles are encouraged whenever possible.

- If it is possible to derive a qMRI application from an already existing file collection suffix by defining a set of logical conditions over the metadata fields, the tables of the deriving the intended qMRI application from an ambiguous file collection and the anatomy data priority levels sections are extended instead of introducing a new suffix.

B.2.1 Application specific notes for anatomy imaging file collection

General notes:

- Some BIDS metadata field values are calculated based on the values of other metadata fields that are not listed as required fields. These fields include: NumberShots. The calculation of the values may depend on the type of the acquisition. These acquisitions include: MP2RAGE and TB1SRGE.

UNIT1 images

Although the UNIT1 image is provided as an output by the acquisition sequence, it is used as an input to offline calculation of a T1map using a dictionary lookup approach. However, complex data is needed for an accurate calculation of the UNIT1 image, which is not commonly provided by the stock sequence. Instead, the magnitude and phase images are exported. Please see the relevant discussion at qMRLab issue #255.

Therefore, the UNIT1 image provided by the scanner is RECOMMENDED to be stored under the anat raw dataset directory along with the MP2RAGE file collection and to be used as the primary input for quantifying a T1map.

If an additional UNIT1 image is calculated offline, then the output is to be stored in the derivatives folder with necessary provenance information.

NumberShots metadata field

Note that the type of NumberShots field can be either a number or an array of numbers.

- If a single number is provided, this should correspond to the number of SlicesPerSlab or ReconMatrixPE. However, in this case, SlicePartialFourier or PartialFourierPE fraction is needed to calculate the number of partitions before and after of the k-space center to calculate a T1 map.

- If before/after calculation is performed during the BIDS conversion of the MP2RAGE data, then the value of NumberShots metadata field can be given as a 1X2 array, with first entry corresponding to before and the second to the after.

Formula:

If NumberShots is an array of numbers such that "NumberShots": [before, after], the values of before and after are calculated as follows:

$$before = SlicesPerSlab * (SlicePartialFourier - 0.5)$$

$$after = SlicesPerSlab/2$$

Other metadata fields

The value of the RepetitionTimeExcitation field is not commonly found in the DICOM files. When accessible, the value of EchoSpacing corresponds to this metadata. When not accessible, 2 X EchoTime can be used as a surrogate.

Further information about other MP2RAGE qMRI protocol fields can be found in the qMRLab documentation.

B.2.2 Radiofrequency (RF) field mapping

Some RF file collections call for the use of special notations that cannot be resolved by by entities that can generalize to other applications. Instead of introducing an entity that is exclusive to a single application, method developers who commonly use these file collections for the MPM application reached the consensus on the use of acq entity to distinguish individual files. These suffixes include: TB1AFI, TB1TFL, TB1RFM, and RB1COR.

TB1SRGE specific notes

Calculation of before and after entries for NumberShots metadata field of TB1SRGE is more involved than that of MP2RAGE. The formula can be found in a reference implementation, which requires information about BaseResolution (that is, image matrix size in PE direction), partial Fourier fraction in the PE direction, number of reference lines for parallel imaging acceleration, and the parallel imaging acceleration factor in PE direction.

TB1EPI specific notes

The flip and echo entities MUST be used to distinguish images with this suffix. The use of flip follows the default convention. However, this suffix defines a specific use case for the echo entity:

Table B.11 TB1EPI entity specifications.

echo-1	echo-2
Lower EchoTime	Higher EchoTime
Spin Echo (SE) image	Stimulated Echo (STE) image

At each FlipAngle, the TB1EPI suffix lists two images acquired at two echo times. The first echo is a spin echo (SE) formed by the pulses alpha-2alpha. However, the second echo in this method is generated in a different fashion compared to a typical MESE acquisition. The second echo is a stimulated echo (STE) that is formed by an additional alpha pulse (that is, alpha-2alpha-alpha).

The FlipAngle value corresponds to the nominal flip angle value of the STE pulse. The nominal FA value of the SE pulse is twice this value.

Note that the following metadata fields MUST be defined in the accompanying JSON files:

Table B.12 TB1EPI metadata specifications.

Field name	Definition
TotalReadoutTime	The effective readout length defined as EffectiveEchoSpacing * PEReconMatrix, with EffectiveEchoSpacing = TrueEchoSpacing / PEacceleration
MixingTime	Time interval between the SE and STE pulses

To properly identify constituents of this particular method, values of the echo entity MUST index the images as follows:

```
sub-01/
└─ fmap/
    ├─ sub-01_echo-1_flip-1_TB1EPI.nii.gz (SE)
    ├─ sub-01_echo-1_flip-1_TB1EPI.json
    ├─ sub-01_echo-2_flip-1_TB1EPI.nii.gz (STE)
    ├─ sub-01_echo-2_flip-1_TB1EPI.json
    ├─ sub-01_echo-1_flip-2_TB1EPI.nii.gz (SE)
    ├─ sub-01_echo-1_flip-2_TB1EPI.json
    ├─ sub-01_echo-2_flip-2_TB1EPI.nii.gz (STE)
    └─ sub-01_echo-2_flip-2_TB1EPI.json
```

TB1AFI specific notes

This method calculates a B1+ map from two images acquired at two interleaved excitation repetition times (TR). Note that there is no entity for the TR that can be used to label the files corresponding to the two repetition times and the definition of repetition time depends on the modality (functional or anatomical) in the specification.

Therefore, to properly identify constituents of this particular method, values of the acq

entity SHOULD begin with either tr1 (lower TR) or tr2 (higher TR) and MAY be followed by freeform entries:

Table B.13 TB1AFI entity specifications.

First TR	Second TR	Use case
_acq-tr1	_acq-tr2	Single acquisition
_acq-tr1Test	_acq-tr2Test	Acquisition Test
_acq-tr1Retest	_acq-tr2Retest	Acquisition Retest

```
sub-01/
└─fmap/
     ├─ sub-01_acq-tr1_TB1AFI.nii.gz
     ├─ sub-01_acq-tr1_TB1AFI.json
     ├─ sub-01_acq-tr2_TB1AFI.nii.gz
     └─ sub-01_acq-tr2_TB1AFI.json
```

TB1TFL and TB1RFM specific notes

These suffixes describe two outputs generated by Siemens tfl_b1_map and rf_map product sequences, respectively. Both sequences output two images. The first image appears like an anatomical image and the second output is a scaled flip angle map.

To properly identify files of this particular file collection, values of the acq entity SHOULD begin with either anat or famp and MAY be followed by freeform entries:

Table B.14 TB1TFL and TB1RFM entity specifications.

Anatomical (like) image	Scaled flip angle map	Use case
_acq-anat	_acq-famp	Single acquisition
_acq-anatTest	_acq-fampTest	Acquisition Test
_acq-anatRetest	_acq-fampRetest	Acquisition Retest

```
sub-01/
└─fmap/
     ├─ sub-01_acq-anat_TB1TFL.nii.gz
     ├─ sub-01_acq-anat_TB1TFL.json
     ├─ sub-01_acq-famp_TB1TFL.nii.gz
     └─ sub-01_acq-famp_TB1TFL.json
```

The example above applies to the TB1RFM suffix as well.

RB1COR specific notes

This method generates a sensitivity map by combining two low resolution images collected by two transmit coils (the body and the head coil) upon subsequent scans with identical acquisition parameters.

To properly identify constituents of this particular method, values of the acq entity SHOULD begin with either body or head and MAY be followed by freeform entries:

Table B.15 RB1COR entity specifications.

Body coil	Head coil	Use case
_acq-body	_acq-head	Single acquisition
_acq-bodyMTw	_acq-headMTw	MTw for MPM
_acq-bodyPDw	_acq-headPDw	PDw for MPM
_acq-bodyT1w	_acq-headT1w	T1w for MPM

```
sub-01/
└── fmap/
    ├── sub-01_acq-body_RB1COR.nii.gz (Body coil)
    ├── sub-01_acq-body_RB1COR.json
    ├── sub-01_acq-head_RB1COR.nii.gz (Head coil)
    └── sub-01_acq-head_RB1COR.json
```

B.3 File collections appendix

Here, some concrete use-cases of entity-linked file collections are listed using descriptive tables, organized by modality.

The tables in this appendix catalog applications where the use of a file collection is required.

Certain entities interlink the files in a file collection through a metadata field. Unlike other common entities (for example run), they require an iteration over different values of the metadata fields they represent. Please keep the following list of linking entities up-to-date with the file collections included in this appendix:

1. Magnetic Resonance Imaging

    (a) echo

    (b) flip

    (c) inv

    (d) mt

(e) part

The current list of file collections are listed below for the anatomy imaging (Table B.16) and fieldmap (Table B.18) applications.

Table B.16 File collections for anatomy imaging data.

Suffix	Linking entities	Application	Description
VFA	flip	Variable flip angle	The VFA method involves at least two spoiled gradient echo (SPGR) of steady-state free precession (SSFP) images acquired at different flip angles. Depending on the provided metadata fields and the sequence type, data may be eligible for DESPOT1, DESPOT2 and their variants (https://doi.org/10.1002/mrm.20314).
IRT1	inv, part	Inversion recovery T1 mapping	The IRT1 method involves multiple inversion recovery spin-echo images acquired at different inversion times (https://doi.org/10.1002/mrm.22497).

Table B.16 continued: File collections for anatomy imaging data.

MP2RAGE	flip, inv, echo, part	Magnetization prepared two gradient echoes	The MP2RAGE method is a special protocol that collects several images at different flip angles and inversion times to create a parametric T1map by combining the magnitude and phase images (https://doi.org/10.1016/j.neuroimage.2009.10.002).
MESE	echo	Multi-echo spin-echo	The MESE method involves multiple spin echo images acquired at different echo times and is primarily used for T2 mapping. Please note that this suffix is not intended for the logical grouping of images acquired using an Echo Planar Imaging (EPI) readout.
MEGRE	echo	Multi-echo gradient-echo	Anatomical gradient echo images acquired at different echo times. Please note that this suffix is not intended for the logical grouping of images acquired using an Echo Planar Imaging (EPI) readout.

Table B.16 continued: File collections for anatomy imaging data.

MTR	mt	Magnetization transfer ratio	This method is to calculate a semi-quantitative magnetization transfer ratio map.
MTS	flip, mt	Magnetization transfer saturation	This method is to calculate a semi-quantitative magnetization transfer saturation index map. The MTS method involves three sets of anatomical images that differ in terms of application of a magnetization transfer RF pulse (MTon or MToff) and flip angle (https://doi.org/10.1002/mrm.21732).

Table B.16 continued: File collections for anatomy imaging data.

MPM	flip, mt, echo, part	Multi-parametric mapping	The MPM approaches (a.k.a hMRI) involves the acquisition of highly-similar anatomical images that differ in terms of application of a magnetization transfer RF pulse (MTon or MToff), flip angle and (optionally) echo time and magnitue/phase parts (https://doi.org/10.3389/fnins.2013.00095). See https://owncloud.gwdg.de/index.php/s/iv2TOQwGy4FGDDZ for suggested MPM acquisition protocols.

Table B.18 File collections for field mapping data.

Suffix	Meta-data relevant entity	Application	Description

Table B.18 continued: File collections for field mapping data.

TB1DAM	flip	Double-angle B1+ mapping	The double-angle B1+ method (https://doi.org/10.1006/jmra.1993.1133) is based on the calculation of the actual angles from signal ratios, collected by two acquisitions at different nominal excitation flip angles. Common sequence types for this application include spin echo and echo planar imaging.
TB1EPI	flip, echo	B1+ mapping with 3D EPI	This B1+ mapping method (https://doi.org/10.1002/mrm.21083) is based on two EPI readouts to acquire spin echo (SE) and stimulated echo (STE) images at multiple flip angles in one sequence, used in the calculation of deviations from the nominal flip angle.

Table B.18 continued: File collections for field mapping data.

TB1AFI	Please see the qMRI appendix.	Actual Flip Angle Imaging (AFI)	This method (https://doi.org/10.1002/mrm.21120) calculates a B1+ map from two images acquired at interleaved (two) TRs with identical RF pulses using a steady-state sequence.
TB1TFL	Please see the qMRI appendix.	Siemens tfl_b1_map	B1+ data acquired using tfl_b1_map product sequence by Siemens based on the method by https://doi.org/10.1002/mrm.22423. The sequence generates one anatomical image and one scaled flip angle map.
TB1RFM	Please see the qMRI appendix.	Siemens rf_map	B1+ data acquired using rf_map product sequence by Siemens.

Table B.18 continued: File collections for field mapping data.

TB1SRGE	flip, inv	SA2RAGE	Saturation-prepared with 2 rapid gradient echoes (SA2RAGE) uses a ratio of two saturation recovery images with different time delays, and a simulated look-up table to estimate B1+ (https://doi.org/10.1002/mrm.23145). This sequence can also be used in conjunction with MP2RAGE T1 mapping to iteratively improve B1+ and T1 map estimation (https://doi.org/10.1371/journal.pone.0069294).
RB1COR	Please see the qMRI appendix.	B1- field correction	Low resolution images acquired by the body coil (in the gantry of the scanner) and the head coil using identical acquisition parameters to generate a combined sensitivity map as described in https://doi.org/10.1002/mrm.26058.

APPENDIX C   SUPPLEMENTARY MATERIALS FOR ARTICLE 3

C.1   Summary statistics

Table C.1 Summary statistics for VENUS and vendor-native acquisitions in the system phantom. Tags 750, PRI and SKY stand for G1, S1 and S2, respectively. The rth label identifies VENUS acquisitions.

Session	RefT1 (mean)	T1 (mean)	T1 (me-dian)	T1 (std)	Num Samples
sub-phantom_ses-rth750	1.99	2.06	2.05	0.08	55
sub-phantom_ses-rth750	1.45	1.47	1.47	0.03	55
sub-phantom_ses-rth750	0.98	0.99	0.99	0.02	55
sub-phantom_ses-rth750	0.71	0.66	0.65	0.02	55
sub-phantom_ses-rth750	0.5	0.44	0.44	0.02	55
sub-phantom_ses-rth750	0.35	0.3	0.3	0.01	55
sub-phantom_ses-rth750	0.25	0.23	0.23	0.01	55
sub-phantom_ses-rth750	0.18	0.19	0.19	0.01	55
sub-phantom_ses-rth750	0.13	0.15	0.15	0.01	55
sub-phantom_ses-rth750	0.09	0.12	0.12	0.01	55
sub-phantom_ses-rthPRI	1.99	2.12	2.12	0.06	55
sub-phantom_ses-rthPRI	1.45	1.46	1.45	0.03	55
sub-phantom_ses-rthPRI	0.98	1.01	1.01	0.02	55
sub-phantom_ses-rthPRI	0.71	0.7	0.7	0.01	55
sub-phantom_ses-rthPRI	0.5	0.48	0.48	0.01	55
sub-phantom_ses-rthPRI	0.35	0.34	0.35	0.01	55
sub-phantom_ses-rthPRI	0.25	0.26	0.26	0.01	55
sub-phantom_ses-rthPRI	0.18	0.22	0.22	0.01	55
sub-phantom_ses-rthPRI	0.13	0.16	0.16	0.01	55
sub-phantom_ses-rthPRI	0.09	0.11	0.11	0	55
sub-phantom_ses-rthSKY	1.99	1.99	1.99	0.08	55
sub-phantom_ses-rthSKY	1.45	1.42	1.43	0.04	55
sub-phantom_ses-rthSKY	0.98	1	1	0.03	55

Table C.1 continued: Summary statistics for VENUS and vendor-native acquisitions in the system phantom.

sub-phantom_ses-rthSKY	0.71	0.7	0.69	0.02	55
sub-phantom_ses-rthSKY	0.5	0.47	0.47	0.01	55
sub-phantom_ses-rthSKY	0.35	0.32	0.32	0.01	55
sub-phantom_ses-rthSKY	0.25	0.25	0.25	0.01	55
sub-phantom_ses-rthSKY	0.18	0.2	0.2	0.01	55
sub-phantom_ses-rthSKY	0.13	0.15	0.15	0.01	55
sub-phantom_ses-rthSKY	0.09	0.1	0.1	0.01	55
sub-phantom_ses-vendor750	1.99	2.6	2.59	0.08	55
sub-phantom_ses-vendor750	1.45	1.7	1.7	0.04	55
sub-phantom_ses-vendor750	0.98	1.16	1.16	0.04	55
sub-phantom_ses-vendor750	0.71	0.77	0.77	0.02	55
sub-phantom_ses-vendor750	0.5	0.51	0.51	0.01	55
sub-phantom_ses-vendor750	0.35	0.38	0.38	0.02	55
sub-phantom_ses-vendor750	0.25	0.27	0.27	0.01	55
sub-phantom_ses-vendor750	0.18	0.23	0.23	0.01	55
sub-phantom_ses-vendor750	0.13	0.19	0.19	0.01	55
sub-phantom_ses-vendor750	0.09	0.15	0.15	0.01	55
sub-phantom_ses-vendorPRI	1.99	2.11	2.08	0.11	55
sub-phantom_ses-vendorPRI	1.45	1.46	1.45	0.06	55
sub-phantom_ses-vendorPRI	0.98	1.01	1	0.04	55
sub-phantom_ses-vendorPRI	0.71	0.69	0.68	0.03	55
sub-phantom_ses-vendorPRI	0.5	0.48	0.47	0.02	55
sub-phantom_ses-vendorPRI	0.35	0.33	0.32	0.01	55
sub-phantom_ses-vendorPRI	0.25	0.24	0.24	0.01	55
sub-phantom_ses-vendorPRI	0.18	0.2	0.19	0.01	55
sub-phantom_ses-vendorPRI	0.13	0.14	0.13	0.01	55
sub-phantom_ses-vendorPRI	0.09	0.09	0.09	0.01	55
sub-phantom_ses-vendorSKY	1.99	2.17	2.15	0.07	55
sub-phantom_ses-vendorSKY	1.45	1.55	1.55	0.04	55
sub-phantom_ses-vendorSKY	0.98	1.07	1.07	0.04	55
sub-phantom_ses-vendorSKY	0.71	0.73	0.73	0.02	55
sub-phantom_ses-vendorSKY	0.5	0.49	0.49	0.02	55

Table C.1 continued: Summary statistics for VENUS and vendor-native acquisitions in the system phantom.

sub-phantom_ses-vendorSKY	0.35	0.33	0.33	0.01	55
sub-phantom_ses-vendorSKY	0.25	0.24	0.24	0.01	55
sub-phantom_ses-vendorSKY	0.18	0.2	0.2	0.01	55
sub-phantom_ses-vendorSKY	0.13	0.14	0.14	0.01	55
sub-phantom_ses-vendorSKY	0.09	0.09	0.09	0.01	55