NeuroLibre
Reproducible Preprints

Leveraging Large Language Models for Interactive Exploration of MRI Research Reproducibility: A Self-Evolving Review

Agah Karakuzu [1,2], **Mathieu Boudreau** [1], **and Nikola Stikov** [1,2,3]

1 NeuroPoly, Polytechnique Montreal, Quebec, Canada **2** Montreal Heart Institute, Montreal, Quebec, Canada **3** Center for Advanced Interdisciplinary Research, Ss. Cyril and Methodius University, Skopje, North Macedonia

DOI: 10.55458/neurolibre.00021

Interactive Version

- Living Preprint

Code

- Technical Screening
- Submitted Repository

Reproducibility Assets

- Repository
- Dataset
- Living Preprint
- Container

Moderator: Pierre Bellec
Screener(s):

- @agahkarakuzu

Submitted: 23 December 2023
Published: 19 June 2025

Disclaimer

This pdf is intended for content registration purposes only! For full access and interactive reading of this publication, please visit the live version.

Abstract

This interactive version accompanies the article published in Magnetic Resonance in Medical Sciences (Karakuzu et al., 2024), providing a reproducible and online-executable implementation designed to realize the self-evolving functionality described in the original work.

Magnetic resonance imaging has progressed significantly with the introduction of advanced computational methods and novel imaging techniques, but their wider adoption hinges on their reproducibility. This concise review synthesizes reproducible research insights from recent MRI articles to examine the current state of reproducibility in neuroimaging, highlighting key trends and challenges.

Unlike the static published version, this interactive platform leverages a custom GPT model designed specifically for automated analysis and synthesis of information pertaining to reproducibility insights, enabling continuous evolution of the review as new literature emerges. Readers can directly engage with the computational analyses, modify parameters, and contribute to the ongoing synthesis, transforming traditional literature review into a dynamic, community-driven resource that adapts and grows with the field.

Introduction

Reproducibility is a cornerstone of scientific inquiry, particularly relevant for data-intensive and computationally demanding fields of research, such as magnetic resonance imaging (MRI) (Stikov et al., 2019). Ensuring reproducibility thus poses a unique set of challenges and necessitates the diligent application of methods that foster transparency, verification, and interoperability of research findings.

While numerous articles have addressed the reproducibility of clinical MRI studies, few have looked at the reproducibility of the MRI methodology underpinning these studies. This is understandable given that the MRI development community is smaller, driven by engineers and physicists, with modest representation from clinicians and statisticians.

However, performing a thorough meta-analysis or a systematic review of these studies in the context of reproducibility presents challenges due to:

- the diversity in study designs across various MRI development subfields, and
- the absence of standardized statistics to gauge reproducibility performance.

Considering these challenges, we opted to conduct a mini-review leveraging the semantic extraction capabilities of the advanced language models. Specifically, we trained a custom **GPT** model using a knowledge base constructed for a selection of articles coupled with web scraping of content pertaining to their reproducibility.

With this mini-review we aim to examine the current landscape of reroducible research practices across various MRI studies, drawing attention to common strategies, tools, and repositories used to achieve reproducible outcomes. We anticipate that this approach provides a living review that can be automatically updated to accommodate the continuously expanding breadth of methodologies, helping us identify commonalities and discrepancies across studies.

Methodology

In distilling reproducibility insights powered by GPT, this review centered on 31 research articles published in the journal Magnetic Resonance in Medicine (MRM), chosen by the editor for their dedication to enhancing reproducibility in MRI. Since 2020, the journal has published interviews with authors of these selected publications, discussing the tools and practices they used to bolster the reproducibility of their findings (available here).

Mapping selected articles in the semantic landscape of reproducibility

We performed a literature search to identify where these studies fall in the broader literature of reproducible neuroimaging. To retrieve articles dedicated to reproducibility in MRI, we utilized the Semantics Scholar API (Fricke, 2018) with the following query terms on November 23, 2023:

```
(code | data | open-source | github | jupyter ) & ((MRI & brain) | (MRI neuroimaging))
& reproducib~.
```

Among `1098` articles included in the Semantic Scholar records, SPECTER vector embeddings (Cohan et al., 2020) were available for `612` articles, representing the semantics of publicly accessible content in abstracts and titles. For these articles, the high-dimensional semantic information captured by the word embeddings was visualized using UMAP (McInnes et al., 2018) Figure 1. This visualization allowed the inspection of the semantic clustering of the articles, facilitating a deeper understanding of their contextual placement within the reproducibility landscape. In addition, the following diagram illustrates the hierarchical clustering of the selected studies in the broader literature:

Creating a knowledge base for a custom GPT

We created a custom GPT model, designed specifically to assist in the analysis and synthesis of information pertaining to the `31` reproducible research insights. The knowledge base of this retrieval-augmented generation framework incorporates GPT-4 summaries of the abstracts from `31` MRM articles, merged with their respective MRM Highlights interviews, as well as the titles and keywords associated with each article (refer to Appendix A). This compilation was assembled via API calls to OpenAI on November 23, 2023, using the `gpt-4-1106-preview` model.

This specialized GPT, named `RRInsights`, is tailored to process and interpret the provided data in the context of reproducibility, for the system prompts please see Appendix B.

Results

Contextual placement of the selected articles in the landscape of reproducibility

The MRI systems cluster was predominantly composed of articles published in MRM, with only two publications appearing in a different journal (Adebimpe et al., 2022; Tilea et al., 2009). Additionally, this cluster was sufficiently distinct from the rest of the reproducibility literature, as can be seen by the location of the dark red dots Figure 1. Few other selected articles (8/31) were found at the intersection of the MRI systems, deep learning, and data/workflows clusters, which in total spans 103 articles. Since the custom GPT model was trained on the 31 selected MRM articles (red dots), Figure 1 serves as a map for inferring the topics where RRInsights is more likely to be context-aware.

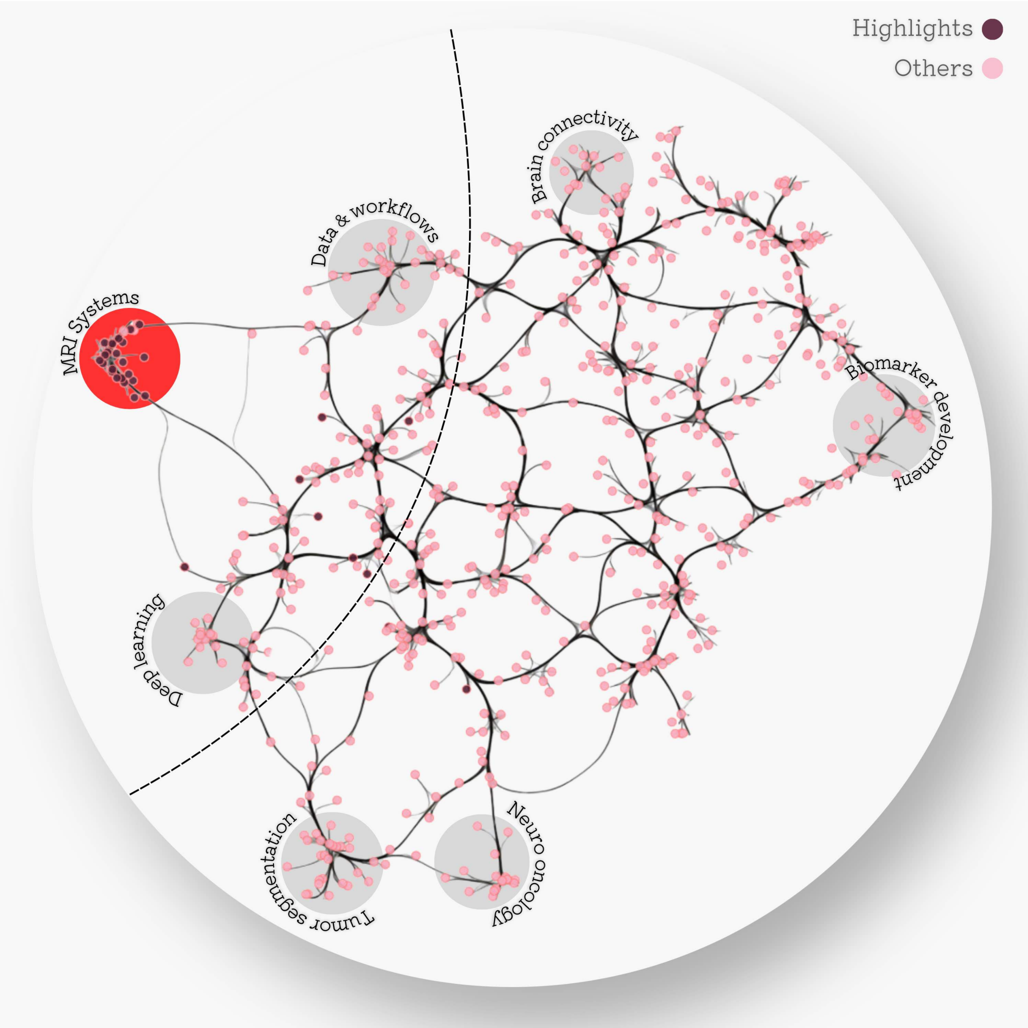

Figure 1: Edge-bundled connectivity of the 612 articles identified by the literature search. A notable cluster is formed by most of the MRM articles that were featured in the reproducible research insights (purple nodes), particularly in the development of MRI systems. Few other selected articles fell at the intersection of MRI systems, deep learning, and workflows. Notable clusters for other studies (pink) are annotated by gray circles.

Custom GPT for reproducibility insights

Through its advanced natural language processing capabilities, RRInsights can efficiently analyze the scoped literature, extract key insights, and generate comprehensive overviews of the research papers focusing on MRI technology. The custom GPT is available at https://chat.openai.com/g/g-5uDwBInx4-rrinsights (requires subscription as of May 2024).

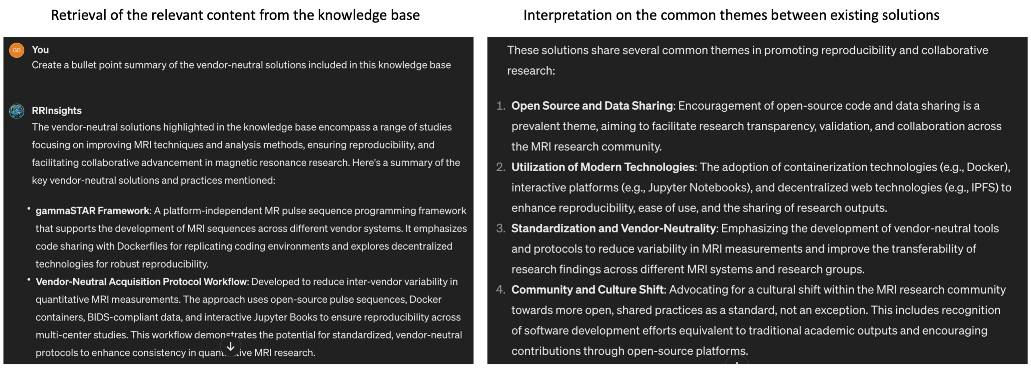

Figure 2: An example user interaction with the RRInsights custom GPT. The model is capable of fetching the studies concerned with the requested content (i.e., vendor-neutral solutions and provide summaries highlighting thematic similarities between them, particularly focusing on reproducibility aspects.

GPT-powered summary of the reproducible magnetic resonance neuroimaging

Most MRI development is done on commercial systems using proprietary hardware and software. Peeking inside the black boxes that generate the images is non-trivial, but it is essential for promoting reproducibility in MRI.

Quantitative MRI articles are powerful showcases of reproducible research practices, as they usually come with fitting models that can be shared on public code repositories. The applications range from MR spectroscopy (Clarke et al., 2021; Songeon et al., 2023; Wilson, 2021) to ASL (Woods et al., 2022), diffusion MRI (Cai et al., 2021; Tristan-Vega et al., 2022), CEST (Huang et al., 2022), magnetization transfer (Assländer et al., 2022; Boudreau, Karakuzu, Boré, et al., 2024; Malik et al., 2020; Rowley et al., 2021), B1 mapping (Delgado et al., 2020) and relaxometry (Balbastre et al., 2022; Boudreau, Karakuzu, Cohen-Adad, et al., 2024; Hafyane et al., 2018; Kapre et al., 2020; Keenan et al., 2025; Y. Lee et al., 2019; Whitaker et al., 2020).

Transparent reconstruction and analysis pipelines (Maier et al., 2021) are also prominently featured in the reproducible research insights, including methods for real-time MRI (Zhao et al., 2021), parallel imaging (Hess et al., 2021), large-scale volumetric dynamic imaging (Ong et al., 2020), pharmacokinetic modeling of DCE-MRI (Ahmed & Levesque, 2020), phase unwrapping (Dymerska et al., 2021), hyperpolarized MRI (Tustison et al., 2021), Dixon imaging (Rydén et al., 2020) and X-nuclei imaging (McCallister et al., 2021). Deep learning is increasingly present in the reproducibility conversation, as MRI researchers are trying to shine a light on AI-driven workflows for phase-focused applications (Cole et al., 2021), CEST (Huang et al., 2022), diffusion-weighted imaging (Barbieri et al., 2020), myelin water imaging (J. Lee et al., 2020), B1 estimation (Abbasi-Rad et al., 2021), and tissue segmentation (Estrada et al., 2020).

Reproducibility of MRI hardware is still in its infancy, but a recent study integrated RF coils with commercial field cameras for ultrahigh-field MRI, exemplifying the coupling of hardware advancements with software solutions. The authors shared the design CAD files, performance data, and image reconstruction code, ensuring that hardware innovations can be reproduced and utilized by other researchers (Gilbert et al., 2022).

Finally, vendor-neutral pulse sequences are putting interoperability and transparency at the center of the reproducibility landscape. Pulseq and gammaSTAR are vendor-neutral platforms enabling the creation of MRI pulse sequences that are compatible with three major MRI vendors (Cordes et al., 2020; Layton et al., 2017). In addition, VENUS is an end-to-end vendor-neutral workflow that was shown to reduce inter-vendor variability in quantitative MRI measurements of myelin, thereby strengthening the reproducibility of quantitative MRI research and facilitating multicenter clinical trials (Karakuzu et al., 2020; Karakuzu, Biswas, et al., 2022).

Data sharing

There is a growing number of studies providing access to raw imaging data, pre-processing pipelines, and post-analysis results. Repositories like Zenodo, XNAT, and the Open Science Framework (OSF) serve as vital resources for housing and curating MRI data. Data sharing is also made easier thanks to unified data representations, such as the ISMRM raw data format (Inati et al., 2017) for standardizing k-space data, and the Brain Imaging Data Structure (BIDS) for organizing complex datasets (Gorgolewski et al., 2016) and their derivatives (Karakuzu, Appelhoff, et al., 2022).

Code sharing

Software repositories such as GitHub and GitLab are making it easier to centralize processing routines and to adopt version control, unit tests and other robust software development practices. The introduction of tools for automated QA processes, as seen in the development of platforms like PreQual for DWI analysis (Cai et al., 2021), signifies an emphasis on interoperability and standardization.

The increasing adoption of containerization and virtual environments makes workflows transparent and easy to execute. Tools like Docker and Singularity are used to package computing environments, making them portable and reproducible across different systems. Studies employing these tools enable MRI researchers to replicate computational processing pipelines without dealing with dependency issues in local computational environments (Cordes et al., 2020; Estrada et al., 2020; Karakuzu, Biswas, et al., 2022).

The rise of machine learning and artificial intelligence in MRI necessitates rigorous evaluation to ensure reproducibility. Studies that use deep learning are beginning to supplement their methodological descriptions with the open-source code, trained models, and simulation tools that underpin their algorithms. Algorithms such as DeepCEST, developed for B1 inhomogeneity correction at 7T, showcase how clinical research can be improved by reproducible research practices (Huang et al., 2022). Sharing these algorithms allows others to perform direct comparisons and apply them to new datasets.

Vendor-neutrality

Finally, pulse-sequence and hardware descriptions are slowly entering the public domain (Cordes et al., 2020; Gilbert et al., 2022; Karakuzu et al., 2020; Layton et al., 2017). For a long time MRI vendors have been reluctant to open up their systems (Stikov & Karakuzu, 2023), but standardized phantoms (Stupic et al., 2021) are creating benchmarks that require transparency and reproducibility. This is particularly relevant for quantitative MRI applications, where scanner upgrades and variabilities across sites are a major hurdle to wider clinical adoption (Boudreau, Karakuzu, Boré, et al., 2024; Keenan et al., 2019; Y. Lee et al., 2019).

Dissemination

Reproducibility is also bolstered by interactive documentation and tools such as Jupyter Notebooks, allowing for dynamic presentation and hands-on engagement with data and methods. Platforms incorporating such interactive elements are being utilized with greater frequency, providing real-time demonstration of analysis techniques and enabling peer-led validation.

Resources such as MRHub, MRPub, [Open Source Imaging] (https://www.opensourceimaging.org/projects/) and NeuroLibre (Harding et al., 2023; Karakuzu, DuPre, et al., 2022) serve as a gateway to a wide range of tools and tutorials that promote reproducibility in MRI. The curation of these resources is essential for ensuring that publications featuring Jupyter Notebooks and R Markdown files (Trisovic et al., 2022) remain executable and properly archived Karakuzu (2025).

Discussion and Future Directions

The progress towards reproducibility in MRI research points to a distinct cultural shift in the scientific community. The move towards open-access publishing, code-sharing platforms, and data repositories reflects a concerted effort to uphold the reproducibility of complex imaging studies. Adopting containerization technologies, pushing for standardization, and consistently focusing on quality assurance are key drivers that will continue to improve reproducibility standards in MRI research.

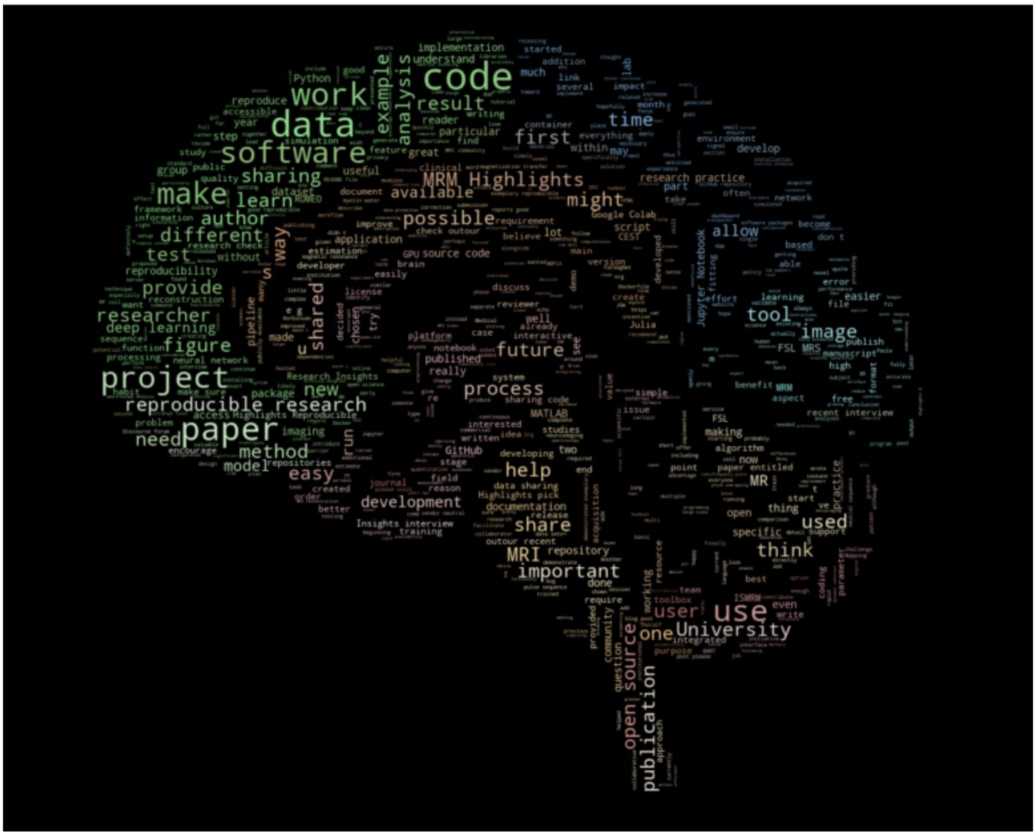

Figure 3: Word cloud generated from the articles included in this review, highlighting the concepts and vocabulary that is driving reproducibility in MRI.

Figure 3 is a word cloud generated from the articles included in this review, highlighting the concepts and vocabulary that is driving reproducibility in MRI. As can be seen from the figure, the components of reproducibility in MRI research are multifaceted, integrating not just data and code, but also the analytical pipelines and hardware configurations. The shift towards comprehensive sharing is motivated both by a scientific ethic of transparency and the practical need for rigorous validation of complex methodologies (Boudreau et al., 2021, 2022).

However, this shift is not without challenges (Niso et al., 2022). Variations in data acquisition and analysis methodologies limit cross-study comparisons. Sensitivity to software and hardware versions can impede direct reproducibility. Privacy concerns and data protection regulations can be barriers to data sharing, particularly with clinical images.

While challenges persist, steps are taken by individual researchers and institutions to prioritize reproducibility. Moving forward, the MRI community should work collectively to overcome barriers, institutionalize reproducible practices, and constructively address data sharing concerns to further the discipline's progress. Beyond their benefit in establishing a research culture that is more aligned with the computational demands of modern neuroimaging, end-to-end reproducible workflows offer a powerful technical solution needed to bring metrological rigor to MRI measurements (Cashmore et al., 2021; Karakuzu et al., 2025). One drawback of this approach is increased external dependencies (Warrington et al., 2023), which can be mitigated by the use of data-driven, portable and modular computational pipelines (Di Tommaso et al., 2017; Karakuzu et al., 2019). Executor agnostic nature of these pipelines also allows for the use of different computational backends, such as cloud computing and high-performance computing resources. This capability is particularly relevant for decoupling the technical complexity of post-acquisition processing from the clinical workflow to create workflows that are more amenable to clinical adoption (Dupuis et al., 2024).

Time and again, the literature shows that aligning the acquisitions (i.e., making the protocols similar to the utmost extent possible) do not suffice to achieve multicenter agreement, neither for radiomic (and other morphological) features (Klontzas, 2024) nor for quantitative applications (Bauer et al., 2010). Absent MRI measurements performed within a metrological framework using well-defined standards, solutions such as physics-driven corrections (Chhetri et al., 2021) or statistical harmonization approaches (Warrington et al., 2023) can still benefit research studies that are based on existing datasets. Nevertheless, surface-level addressing of discrepancies stemming from the complex interplay between MRI physics and biological systems will be insufficient for developing robust calibration strategies, which remain the only clinically acceptable method to characterize and mitigate hardware-based variability. As quantitative computed-tomography (Skornitzke et al., 2022) and ultrasound (Hoferer et al., 2023) are coming to the forefront of developing calibration strategies for non-invasive tissue characterization, MRI should also claim its place in the metrological framework. A prerequisite for this is to recognize that a useful measurement aims to establish error bounds in the pursuit of normative values for physiologically relevant parameters (Cashmore et al., 2021), rather than modeling the normative values while unaddressed sources of variability persist.

The initiatives and tools identified in this review serve as a blueprint for future studies to replicate successful practices, safeguard against bias, and accelerate neuroscientific discovery. As MRI research continues to advance, upholding the principles of reproducibility will be essential to maintaining the integrity and translational potential of its findings.

We also hope that our methodology in generating this review will pave the way for future studies that leverage large language models to create unique literature insights. In particular, we believe that the RRInsights GPT can serve as a blueprint for generating a scoping review (Mak & Thomas, 2022) and inspire other scientists to experiment with the format of scientific publications in the age of AI.

References

Abbasi-Rad, S., O'Brien, K., Kelly, S., Vegh, V., Rodell, A., Tesiram, Y., Jin, J., Barth, M., & Bollmann, S. (2021). Improving FLAIR SAR efficiency at 7T by adaptive tailoring of adiabatic pulse power through deep learning B1+ estimation. *Magn. Reson. Med.*, *85*(5), 2462–2476.

Adebimpe, A., Bertolero, M., Dolui, S., Cieslak, M., Murtha, K., Baller, E. B., Boeve, B., Boxer, A., Butler, E. R., Cook, P., Colcombe, S., Covitz, S., Davatzikos, C., Davila, D. G., Elliott, M. A., Flounders, M. W., Franco, A. R., Gur, R. E., Gur, R. C., … Satterthwaite, T. D. (2022). ASLPrep: A platform for processing of arterial spin labeled

MRI and quantification of regional brain perfusion. *Nat. Methods*, *19*(6), 683–686. https://doi.org/10.1038/s41592-022-01458-7

Ahmed, Z., & Levesque, I. R. (2020). Pharmacokinetic modeling of dynamic contrast-enhanced MRI using a reference region and input function tail. *Magnetic Resonance in Medicine*, *83*(1), 286–298. https://doi.org/10.1002/mrm.27913

Assländer, J., Gultekin, C., Flassbeck, S., Glaser, S. J., & Sodickson, D. K. (2022). Generalized bloch model: A theory for pulsed magnetization transfer. *Magn. Reson. Med.*, *87*(4), 2003–2017. https://doi.org/10.58530/2022/2707

Balbastre, Y., Aghaeifar, A., Corbin, N., Brudfors, M., Ashburner, J., & Callaghan, M. F. (2022). Correcting inter-scan motion artifacts in quantitative R1 mapping at 7T. *Magn. Reson. Med.*, *88*(1), 280–291.

Barbieri, S., Gurney-Champion, O. J., Klaassen, R., & Thoeny, H. C. (2020). Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted MRI. *Magn. Reson. Med.*, *83*(1), 312–321. https://doi.org/10.1002/mrm.27910

Bauer, C. M., Jara, H., Killiany, R., Initiative, A. D. N., & others. (2010). Whole brain quantitative T2 MRI across multiple scanners with dual echo FSE: Applications to AD, MCI, and normal aging. *Neuroimage*, *52*(2), 508–514. https://doi.org/10.1016/j.neuroimage.2010.04.255

Boudreau, M., Karakuzu, A., Boré, A., Pinsard, B., Zelenkovski, K., Alonso-Ortiz, E., Boyle, J., Bellec, L., & Cohen-Adad, J. (2024). Longitudinal reproducibility of brain and spinal cord quantitative MRI biomarkers. *Imaging Neuroscience*. https://doi.org/10.1162/imag_a_00409

Boudreau, M., Karakuzu, A., Cohen-Adad, J., Bozkurt, E., Carr, M., Castellaro, M., Concha, L., Doneva, M., Dual, S. A., Ensworth, A., & others. (2024). Repeat it without me: Crowdsourcing the T1 mapping common ground via the ISMRM reproducibility challenge. *Magnetic Resonance in Medicine*. https://doi.org/10.1002/mrm.30111

Boudreau, M., Poline, J.-B., Bellec, P., & Stikov, N. (2021). On the open-source landscape of PLOS computational biology. *PLoS Comput. Biol.*, *17*(2), e1008725. https://doi.org/10.1371/journal.pcbi.1008725

Boudreau, M., Stikov, N., & Jezzard, P. (2022). On the open-source landscape of magnetic resonance in medicine. *Magn. Reson. Med.*, *88*(4), 1495–1497.

Cai, L. Y., Yang, Q., Hansen, C. B., Nath, V., Ramadass, K., Johnson, G. W., Conrad, B. N., Boyd, B. D., Begnoche, J. P., Beason-Held, L. L., Shafer, A. T., Resnick, S. M., Taylor, W. D., Price, G. R., Morgan, V. L., Rogers, B. P., Schilling, K. G., & Landman, B. A. (2021). PreQual: An automated pipeline for integrated preprocessing and quality assurance of diffusion weighted MRI images. *Magn. Reson. Med.*, *86*(1), 456–470. https://doi.org/10.1002/mrm.28678

Cashmore, M. T., McCann, A. J., Wastling, S. J., McGrath, C., Thornton, J., & Hall, M. G. (2021). Clinical quantitative MRI and the need for metrology. *The British Journal of Radiology*, *94*(1120), 20201215. https://doi.org/10.1259/bjr.20201215

Chhetri, G., McPhee, K. C., Wilman, A. H., Initiative, A. D. N., & others. (2021). Bloch modelling enables robust T2 mapping using retrospective proton density and T2-weighted images from different vendors and sites. *Neuroimage*, *237*, 118116. https://doi.org/10.1016/j.neuroimage.2021.118116

Clarke, W. T., Stagg, C. J., & Jbabdi, S. (2021). FSL-MRS: An end-to-end spectroscopy analysis package. *Magn. Reson. Med.*, *85*(6), 2950–2964.

Cohan, A., Feldman, S., Beltagy, I., Downey, D., & Weld, D. S. (2020). *SPECTER: Document-level representation learning using citation-informed transformers*. https://doi.org/10.18653/v1/2020.acl-main.207

Cole, E., Cheng, J., Pauly, J., & Vasanawala, S. (2021). Analysis of deep complex-valued convolutional neural networks for MRI reconstruction and phase-focused applications. *Magn. Reson. Med.*, *86*(2), 1093–1109. https://doi.org/10.1002/mrm.28733

Cordes, C., Konstandin, S., Porter, D., & Günther, M. (2020). Portable and platform-independent MR pulse sequence programs. *Magn. Reson. Med.*, *83*(4), 1277–1290. https://doi.org/10.1002/mrm.28020

Delgado, P. R., Kuehne, A., Periquito, J. S., Millward, J. M., Pohlmann, A., Waiczies, S., & Niendorf, T. (2020). B1 inhomogeneity correction of RARE MRI with transceive surface radiofrequency probes. *Magn. Reson. Med.*, *84*(5), 2684–2701.

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. https://doi.org/10.1038/nbt.3820

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. https://doi.org/10.1371/journal.pcbi.1009651

Dupuis, A., Boyacioglu, R., Keenan, K. E., & Griswold, M. A. (2024). Real-time automated quality control for quantitative MRI. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 1–11. https://doi.org/10.1007/s10334-024-01205-3

Dymerska, B., Eckstein, K., Bachrata, B., Siow, B., Trattnig, S., Shmueli, K., & Robinson, S. D. (2021). Phase unwrapping with a rapid opensource minimum spanning tree algorithm (ROMEO). *Magn. Reson. Med.*, *85*(4), 2294–2308. https://doi.org/10.1002/mrm.28563

Estrada, S., Lu, R., Conjeti, S., Orozco-Ruiz, X., Panos-Willuhn, J., Breteler, M. M. B., & Reuter, M. (2020). FatSegNet: A fully automated deep learning pipeline for adipose tissue segmentation on abdominal dixon MRI. *Magn. Reson. Med.*, *83*(4), 1471–1483. https://doi.org/10.1002/mrm.28163

Fricke, S. (2018). Semantic scholar. *J. Med. Libr. Assoc.*, *106*(1), 145. https://doi.org/10.5195/jmla.2018.280

Gilbert, K. M., Dubovan, P. I., Gati, J. S., Menon, R. S., & Baron, C. A. (2022). Integration of an RF coil and commercial field camera for ultrahigh-field MRI. *Magn. Reson. Med.*, *87*(5), 2551–2565. https://doi.org/10.1002/mrm.29130

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., Handwerker, D. A., Hanke, M., Keator, D., Li, X., Michael, Z., Maumet, C., Nichols, B. N., Nichols, T. E., Pellman, J., … Poldrack, R. A. (2016). The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. *Sci Data*, *3*, 160044. https://doi.org/10.1038/sdata.2016.44

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. https://doi.org/10.1101/343079

Harding, R. J., Bermudez, P., Bernier, A., Beauvais, M., Bellec, P., Hill, S., Karakuzu, A., Knoppers, B. M., Pavlidis, P., Poline, J.-B., & others. (2023). The canadian open neuroscience platform—an open science framework for the neuroscience community. *PLOS Computational Biology*, *19*(7), e1011230. https://doi.org/10.1371/journal.pcbi.1011230

Hess, A. T., Dragonu, I., & Chiew, M. (2021). Accelerated calibrationless parallel transmit mapping using joint transmit and receive low-rank tensor completion. *Magn. Reson. Med.*, *86*(5), 2454–2467. https://doi.org/10.1002/mrm.28880

Hoferer, I., Jourdain, L., Girot, C., Benatsou, B., Leguerney, I., Cournede, P., Marouf, A., Hoarau, Y., Lassau, N., & Pitre-Champagnat, S. (2023). New calibration setup for quantitative DCE-US imaging protocol: Toward standardization. *Medical Physics*, *50*(9), 5541–5552. https://doi.org/10.1002/mp.16362

Huang, J., Lai, J. H. C., Tse, K.-H., Cheng, G. W. Y., Liu, Y., Chen, Z., Han, X., Chen, L., Xu, J., & Chan, K. W. Y. (2022). Deep neural network based CEST and AREX processing: Application in imaging a model of alzheimer's disease at 3 T. *Magn. Reson. Med.*, *87*(3), 1529–1545. https://doi.org/10.1002/mrm.29044

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., Xue, H., Campbell-Washburn, A. E., Sørensen, T. S., & Hansen, M. S. (2017). ISMRM raw data format: A proposed standard for MRI raw datasets. *Magn. Reson. Med.*, *77*(1), 411–421. https://doi.org/10.1002/mrm.26089

Kapre, R., Zhou, J., Li, X., Beckett, L., & Louie, A. Y. (2020). A novel gamma GLM approach to MRI relaxometry comparisons. *Magn. Reson. Med.*, *84*(3), 1592–1604. https://doi.org/10.1002/mrm.28192

Karakuzu, A. (2025). Toward a woven literature: Open-source infrastructure for reproducible publishing. *NeuroLibre Reproducible Preprints*, 41. https://doi.org/10.55458/neurolibre.00041

Karakuzu, A., Appelhoff, S., Auer, T., Boudreau, M., Feingold, F., Khan, A. R., Lazari, A., Markiewicz, C., Mulder, M., Phillips, C., Salo, T., Stikov, N., Whitaker, K., & Hollander, G. de. (2022). qMRI-BIDS: An extension to the brain imaging data structure for quantitative magnetic resonance imaging data. *Sci Data*, *9*(1), 517. https://doi.org/10.1038/s41597-022-01571-4

Karakuzu, A., Biswas, L., Cohen-Adad, J., & Stikov, N. (2022). Vendor-neutral sequences and fully transparent workflows improve inter-vendor reproducibility of quantitative MRI. *Magn. Reson. Med.*, *88*(3), 1212–1228. https://doi.org/10.1002/mrm.29292

Karakuzu, A., Blostein, N., Caron, A. V., Boré, A., Rheault, F., Descoteaux, M., & Stikov, N. (2025). Rethinking MRI as a measurement device through modular and portable pipelines. *Magnetic Resonance Materials in Physics, Biology and Medicine*, 1–17. https://doi.org/10.1007/s10334-025-01245-3

Karakuzu, A., Boudreau, M., Duval, T., Boshkovski, T., Leppert, I., Cabana, J.-F., Gagnon, I., Beliveau, P., Pike, G., Cohen-Adad, J., & Stikov, N. (2020). qMRLab: Quantitative MRI analysis, under one umbrella. *J. Open Source Softw.*, *5*(53), 2343. https://doi.org/10.21105/joss.02343

Karakuzu, A., Boudreau, M., Duval, T., Leppert, I., Boshkovski, T., Cohen-Adad, J., & Stikov, N. (2019). The qMRLab workflow: From acquisition to publication. *Proceedings of ISMRM*.

Karakuzu, A., Boudreau, M., & Stikov, N. (2024). Reproducible research practices in magnetic resonance neuroimaging: A review informed by advanced language models. *Magnetic Resonance in Medical Sciences*, *23*(3), 252–267. https://doi.org/10.2463/mrms.rev.2023-0174

Karakuzu, A., DuPre, E., Tetrel, L., Bermudez, P., Boudreau, M., Chin, M., Poline, J.-B., Das, S., Bellec, P., & Stikov, N. (2022). *NeuroLibre : A preprint server for full-fledged reproducible neuroscience*. https://doi.org/10.31219/osf.io/h89js

Keenan, K. E., Gimbutas, Z., Dienstfrey, A., & Stupic, K. F. (2019). Assessing effects of scanner upgrades for clinical studies. *J. Magn. Reson. Imaging*, *50*(6), 1948–1954. https://doi.org/10.1002/jmri.26785

Keenan, K. E., Tasdelen, B., Javed, A., Ramasawmy, R., Rizzo, R., Martin, M. N., Stupic, K. F., Seiberlich, N., Campbell-Washburn, A. E., & Nayak, K. S. (2025). T1 and T2 measurements across multiple 0.55 t MRI systems using open-source vendor-neutral sequences. *Magnetic Resonance in Medicine*, *93*(1), 289–300. https://doi.org/10.1002/mrm.30281

Klontzas, M. E. (2024). Radiomics feature reproducibility: The elephant in the room. In *European Journal of Radiology* (p. 111430). Elsevier. https://doi.org/10.1016/j.ejrad.2024.111430

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. *Magn. Reson. Med.*, *77*(4), 1544–1552. https://doi.org/10.1002/mrm.26235

Lee, J., Lee, D., Choi, J. Y., Shin, D., Shin, H.-G., & Lee, J. (2020). Artificial neural network for myelin water imaging. *Magn. Reson. Med.*, *83*(5), 1875–1883.

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.

Maier, O., Baete, S. H., Fyrdahl, A., Hammernik, K., Harrevelt, S., Kasper, L., Karakuzu, A., Loecher, M., Patzig, F., Tian, Y., & others. (2021). CG-SENSE revisited: Results from the first ISMRM reproducibility challenge. *Magnetic Resonance in Medicine*, *85*(4), 1821–1839. https://doi.org/10.1002/mrm.28569

Mak, S., & Thomas, A. (2022). Steps for conducting a scoping review. *Journal of Graduate Medical Education*, *14*(5), 565–567. https://doi.org/10.4300/jgme-d-22-00621.1

Malik, S. J., Teixeira, R. P. A. G., West, D. J., Wood, T. C., & Hajnal, J. V. (2020). Steady-state imaging with inhomogeneous magnetization transfer contrast using multiband radiofrequency pulses. *Magn. Reson. Med.*, *83*(3), 935–949. https://doi.org/10.1002/mrm.27984

McCallister, A., Chung, S. H., Antonacci, M., Z Powell, M., Ceppe, A. S., Donaldson, S. H., Lee, Y. Z., Branca, R. T., & Goralski, J. L. (2021). Comparison of single breath hyperpolarized 129 xe MRI with dynamic 19 F MRI in cystic fibrosis lung disease. *Magn. Reson. Med.*, *85*(2), 1028–1038.

McInnes, L., Healy, J., & Melville, J. (2018). *UMAP: Uniform manifold approximation and projection for dimension reduction*. https://arxiv.org/abs/1802.03426

Niso, G., Botvinik-Nezer, R., Appelhoff, S., De La Vega, A., Esteban, O., Etzel, J. A., Finc, K., Ganz, M., Gau, R., Halchenko, Y. O., Herholz, P., Karakuzu, A., Keator, D. B., Markiewicz, C. J., Maumet, C., Pernet, C. R., Pestilli, F., Queder, N., Schmitt, T., … Rieger, J. W. (2022). Open and reproducible neuroimaging: From study inception to publication. *Neuroimage*, *263*, 119623. https://doi.org/10.31219/osf.io/pu5vb

Ong, F., Zhu, X., Cheng, J. Y., Johnson, K. M., Larson, P. E. Z., Vasanawala, S. S., & Lustig, M. (2020). Extreme MRI: Large-scale volumetric dynamic imaging from continuous non-gated acquisitions. *Magn. Reson. Med.*, *84*(4), 1763–1780. https://doi.org/10.1002/mrm.28235

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 B1+ 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

Rydén, H., Berglund, J., Norbeck, O., Avventi, E., Sprenger, T., Niekerk, A. van, & Skare, S. (2020). RARE two-point dixon with dual bandwidths. *Magn. Reson. Med.*, *84*(5), 2456–2468. https://doi.org/10.1002/mrm.28293

Skornitzke, S., Vats, N., Kopytova, T., Tong, E. W. Y., Hofbauer, T., Weber, T. F., Rehnitz, C., Stackelberg, O. von, Maier-Hein, K., Stiller, W., & others. (2022). Asynchronous calibration of quantitative computed tomography bone mineral density assessment for opportunistic osteoporosis screening: Phantom-based validation and parameter influence evaluation. *Scientific Reports*, *12*(1), 20729. https://doi.org/10.1038/s41598-022-24546-2

Songeon, J., Courvoisier, S., Xin, L., Agius, T., Dabrowski, O., Longchamp, A., Lazeyras, F., & Klauser, A. (2023). In vivo magnetic resonance 31 P-Spectral analysis with neural networks: 31P-SPAWNN. *Magn. Reson. Med.*, *89*(1), 40–53.

Stikov, N., & Karakuzu, A. (2023). The relaxometry hype cycle. *Front. Physiol.*, *14*, 1281147. https://doi.org/10.3389/fphys.2023.1281147

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

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. *Magn. Reson. Med.*, *86*(3), 1194–1211. https://doi.org/10.1002/mrm.28779

Tilea, B., Alberti, C., Adamsbaum, C., Armoogum, P., Oury, J. F., Cabrol, D., Sebag, G., Kalifa, G., & Garel, C. (2009). Cerebral biometry in fetal magnetic resonance imaging: New reference data. *Ultrasound Obstet. Gynecol.*, *33*(2), 173–181. https://doi.org/10.1002/uog.6276

Trisovic, A., Lau, M. K., Pasquier, T., & Crosas, M. (2022). A large-scale study on research code quality and execution. *Scientific Data*, *9*(1), 60. https://doi.org/10.1038/s41597-022-01143-6

Tristan-Vega, A., Paris, G., Luis-Garcia, R. de, & Aja-Fernandez, S. (2022). Accurate free-water estimation in white matter from fast diffusion MRI acquisitions using the spherical means technique. *Magn. Reson. Med.*, *87*(2), 1028–1035. https://doi.org/10.1002/mrm.28997

Tustison, N. J., Altes, T. A., Qing, K., He, M., Miller, G. W., Avants, B. B., Shim, Y. M., Gee, J. C., Mugler, J. P., 3rd, & Mata, J. F. (2021). Image- versus histogram-based considerations in semantic segmentation of pulmonary hyperpolarized gas images. *Magn. Reson. Med.*, *86*(5), 2822–2836. https://doi.org/10.1002/mrm.28908

Warrington, S., Ntata, A., Mougin, O., Campbell, J., Torchi, A., Craig, M., Alfaro-Almagro, F., Miller, K. L., Morgan, P. S., Jenkinson, M., & others. (2023). A resource for development and comparison of multimodal brain 3 t MRI harmonisation approaches.

Imaging Neuroscience, *1*, 1–27. https://doi.org/10.1162/imag_a_00042

Whitaker, S. T., Nataraj, G., Nielsen, J.-F., & Fessler, J. A. (2020). Myelin water fraction estimation using small-tip fast recovery MRI. *Magn. Reson. Med.*, *84*(4), 1977–1990. https://doi.org/10.1002/mrm.28259

Wilson, M. (2021). Adaptive baseline fitting for 1H MR spectroscopy analysis. *Magn. Reson. Med.*, *85*(1), 13–29. https://doi.org/10.1002/mrm.28385

Woods, J. G., Wong, E. C., Boyd, E. C., & Bolar, D. S. (2022). VESPA ASL: VElocity and SPAtially selective arterial spin labeling. *Magn. Reson. Med.*, *87*(6), 2667–2684. https://doi.org/10.1002/mrm.29159

Zhao, Z., Lim, Y., Byrd, D., Narayanan, S., & Nayak, K. S. (2021). Improved 3D real-time MRI of speech production. *Magn. Reson. Med.*, *85*(6), 3182–3195. https://doi.org/10.1002/mrm.28651