QUALITY ASSURANCE OF THE FEDERAL INTERAGENCY TRAUMATIC BRAIN INJURY RESEARCH (FITBIR) MRI DATABASE TO ENABLE INTEGRATED MULTI-SITE ANALYSIS

Adam M. Saunders[1], Michael E. Kim[1], Gaurav Rudravaram[1], Elyssa M. McMaster[1], Chloe Scholten[2], Simon Vandekar[3], Tonia S. Rex[3], François Rheault[4], Bennett A. Landman[1]

[1]Vanderbilt University, TN, USA
[2]University of Calgary, AB, Canada
[3]Vanderbilt University Medical Center, TN, USA
[4]Université de Sherbrooke, QC, Canada

ABSTRACT

The Federal Interagency Traumatic Brain Injury Research (FITBIR) database is a centralized data repository for traumatic brain injury (TBI) research. It includes over 45,529 magnetic resonance images (MRI) from 6,211 subjects (9,229 imaging sessions) across 26 studies with heterogeneous organization formats, contrasts, acquisition parameters, and demographics. In this work, we organized all available structural and diffusion MRI from FITBIR along with relevant demographic information into the Brain Imaging Data Structure. We analyzed whole-brain mean fractional anisotropy, mean diffusivity, total intracranial volume, and the volumes of 132 regions of interest using UNesT segmentations. There were 4,868 subjects (7,035 sessions) with structural MRI and 2,666 subjects (3,763 sessions) with diffusion MRI following quality assurance and harmonization. We modeled profiles for these metrics across ages with generalized additive models for location, scale, and shape (GAMLSS) and found significant differences in subjects with TBI compared to controls in volumes of 54 regions of the brain ($q < 0.05$, likelihood ratio test with false discovery rate correction).

Index Terms— traumatic brain injury, magnetic resonance imaging, diffusion MRI, imaging informatics

1. INTRODUCTION

Traumatic brain injury (TBI) encompasses a wide variety of presentations and outcomes, often classified into mild, moderate and severe based on the Glasgow coma scale [1]. The Glasgow coma scale is unevenly correlated with outcomes and does not capture the full patient experience [2]. Coarse ratings can obscure subgroups of TBI presentation, leading to efforts to characterize TBI on sensitive and multi-dimensional frameworks such as clinical presentation, blood biomarkers, imaging, and modifying factors (CBI-M) [3].

Structural and diffusion-weighted magnetic resonance imaging (MRI) provide insight into volumetric and microstructural changes in the brain following TBI. Even in mild TBI, the brain undergoes volumetric changes that correlate with outcomes [4]. Diffusion tensor metrics like fractional anisotropy (FA) and mean diffusivity (MD) are predictive of TBI outcomes [5].

The search for more sensitive classification systems for TBI severity has highlighted the need for large-scale data repositories like the FITBIR database in the United States [6] and the Collaborative European NeuroTrauma Effectiveness Research in TBI (CENTER-TBI) database [7]. Here, we curate and analyze diffusion tensor metrics and brain region volumes across all MRI available from FITBIR (Fig. 1).

2. METHODS

We retrieved all available structural and diffusion MRI from FITBIR. We converted all images to NIfTI format with dcm2niix [8] and reorganized available data into the Brain

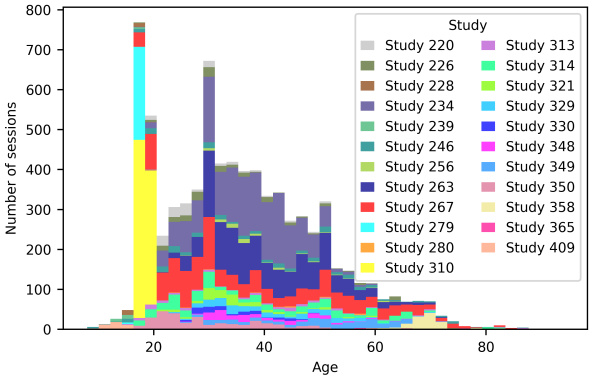

Fig. 1. We unify 6,211 subjects (9,229 sessions) across 26 studies by organizing the images and relevant demographics and metadata into BIDS format. (Sessions from studies with age reported are shown here.)

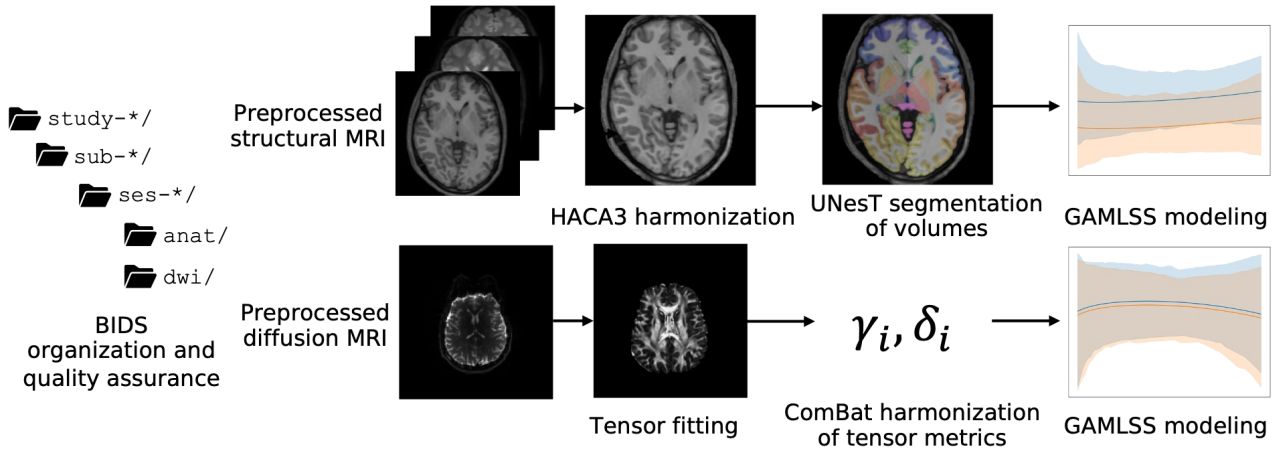

Fig. 2. We harmonized the structural MRI using HACA3 and calculated the volumes of 132 regions of interest. For the diffusion MRI, we first calculated whole-brain mean FA, mean MD, and total intracranial volume (TICV) and harmonized these metrics with ComBat. We used generalized additive models for location, scale, and shape (GAMLSS) to model these metrics across ages 15-90 years in patients with TBI versus controls.

Imaging Data Structure (BIDS) [9]. We recorded common data elements in each study to store age, sex, race, ethnicity, case/control status, Glasgow coma scale score, scanner type, scanner ID, and location of scanner as available.

Next, we visualized a montage of slices from each image and performed quality assurance with AutoQA [10]. We removed images with 2D slices, extremely limited fields of view, and severe anatomical deformations or artifacts. We included studies containing participants with a reported age, sex, and case/control status and sessions with diffusion MRI or a T1-weighted image available (Fig. 2).

For the structural MRI, we harmonized the images as in Saunders et al. [11]. Briefly, we used SMORE super-resolution [12], N4 bias correction [13], and rigid registration to MNI space. We applied HACA3 harmonization to fuse information from all structural MRI and harmonize contrasts to a single T1-weighted image [14]. We calculated the volumes for 132 regions of interest (ROIs) using UNesT [15].

We preprocessed all diffusion MRI sessions using PreQual, which includes denoising, normalization, susceptibility and eddy current nonlinearity correction [16]. We combined multiple acquisitions with separately acquired b_0 images or reverse phase encoding images where possible. In acquisitions with different echo times per shell, we only used the shell nearest to a b-value of 1000 s/mm^2.

To process the diffusion MRI, we used nf-neuro, a set of neuroimaging pipelines and workflows based on the Nextflow workflow manager [17]. We skull-stripped the preprocessed T1-weighted images with SynthStrip [18] and applied rigid registration to the diffusion image. For the diffusion MRI, we resampled to 1 mm isotropic resolution, skull-stripped using the mean b_0 image, corrected for free water contributions to the diffusion signal [19], and fit tensors to the signal. As the diffusion tensor model assumes Gaussian diffusion, we only analyzed sessions that have at least 12 b-values between 500 s/mm^2 and 1500 s/mm^2, and we ignored volumes with higher b-values. We calculated TICV, mean FA, and mean MD within all brain tissue (excluding ventricles). To harmonize the resulting metrics from diffusion MRI, we used ComBat by considering the study as site along with age, sex, and case/control status as covariates [20]. We ran quality assurance after all processing with AutoQA.

3. RESULTS AND DISCUSSION

A total of 4,868 subjects (7,035 sessions, 19 studies) passed inclusion criteria and quality assurance for structural MRI, while there were 2,666 subjects (3,763 sessions, 13 studies) with diffusion MRI and a corresponding T1-weighted image. The mean FA, mean MD, and TICV demonstrated significant differences between studies that ComBat harmonization reduced (Fig. 3).

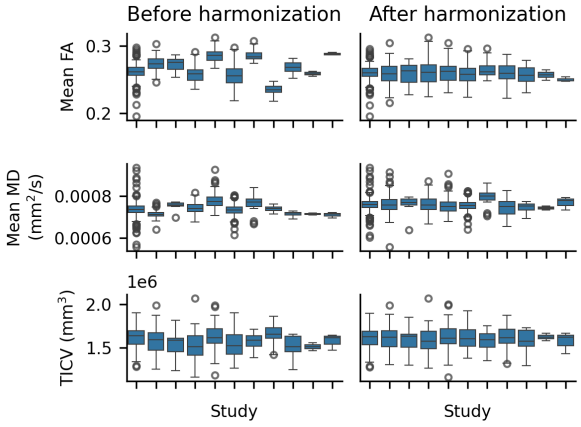

Fig. 3. There are significant differences in metrics even among the controls from FITBIR studies with age, sex, and case/control status reported ($p < 0.05$ for each metric, ANOVA). ComBAT reduces biases across studies ($p < 0.05$ for mean MD only).

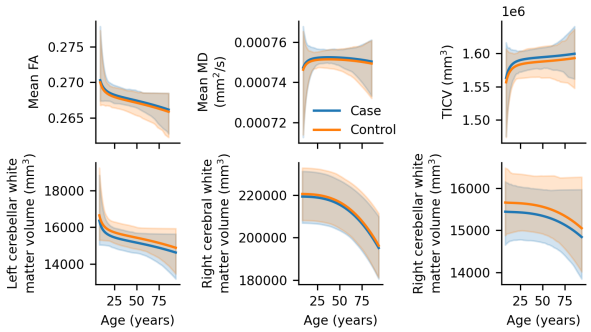

Fig. 4. Mean FA, mean MD, and TICV were not significantly different between TBI and controls, while 54 out of 132 ROI volumes have significant differences, including the white matter regions shown (bottom row, $q < 0.05$, likelihood ratio test with FDR correction).

To describe trends across age in these metrics in patients with TBI and controls, we fit generalized additive models for location, scale, and shape (GAMLSS) curves to the ROI volumes and diffusion tensor metrics based on age, sex and case/control status using the generalized Gamma distribution family [21]. We rejected outliers more than five standard deviations from the mean value for each study (296 measurements from structural MRI and 18 measurements from diffusion MRI). Upon visual inspection, many outliers were due to errors in skull-stripping and segmentation. We bootstrapped 95th percentile confidence intervals and ran a likelihood ratio test using a null model of curves fit without case/control status. We used Benjamini-Yekutieli multiple comparisons correction with a false discovery rate (FDR) of 0.05 (Fig. 4). We found no significant difference in the models for mean FA, mean MD, and TICV, whereas there were significant differences in 54 out of 132 ROIs from structural MRI. These ROIs included the cerebellum, brain stem, subcortical gray matter, insular cortex, and many regions along the longitudinal fissure, all largely symmetric across the hemispheres of the brain.

The use of large-scale retrospective structural and diffusion MRI in this study highlights the need for consistent data collection, organization, and processing, especially with heterogeneous pathologies. Only 77% of structural MRI sessions and 62% of diffusion MRI sessions were suitable for analysis, largely due to missing demographic information such as age, sex or case/control status (~10% of all sessions) and poor image quality or severe anatomical deformations (~10% of structural and ~20% of diffusion sessions). Several studies had missing diffusion gradient information that was unrecoverable (315 sessions) or T1-weighted images collected days apart from diffusion MRI, rendering it difficult to link information from multiple contrasts (81 sessions).

Harmonization helped account for the large intra-study variability in derived metrics. HACA3 allowed for harmonization of structural MRI without an explicit definition of site, but it has not been routinely applied or validated for raw diffusion MRI. Instead, we performed harmonization of diffusion MRI on derived metrics with ComBat. Along with other considerations for careful distribution matching, ComBat requires an explicit mapping of site [22], [23]. While we considered each study as a site, FITBIR contains several consortium-level studies with data from many institutions, and many institutions within FITBIR did not collect data from healthy controls as a reference.

We can curate a useful subset of data from FITBIR for harmonization and further analysis. Careful organization, processing, and quality assurance can account for variability in data quality, while harmonization can account for intra-site variation. With large-scale imaging available for harmonization and analysis from repositories like FITBIR, we can investigate complex and varied conditions like TBI with high statistical power.

4. ACKNOWLEDGMENTS

This work was supported by DoD grant HT94252410563, NIH 1R01EB017230, NIH P50HD103537, the Alzheimer's Disease Sequencing Project Phenotype Harmonization Consortium that is funded by NIA (U24 AG074855, U01 AG068057 and R01 AG059716), the Canadian Institutes of Health Research, the Canadian Neurodevelopmental Research Training Platform, and the Hotchkiss Brain Institute. The Vanderbilt Institute for Clinical and Translational Research is funded by the National Center for Advancing Translational Sciences Clinical Translational Science Award Program, Award Number 5UL1TR002243-03. This work was conducted in part using the resources of the Advanced Computing Center for Research and Education at Vanderbilt University, Nashville, TN. Data and/or research tools used in the preparation of this manuscript were obtained and analyzed from the controlled access datasets distributed from the DoD- and NIH-supported Federal Interagency Traumatic Brain Injury Research (FITBIR) Informatics Systems. FITBIR is a collaborative biomedical informatics system created by the Department of Defense and the National Institutes of Health to provide a national resource to support and accelerate research in TBI. Dataset identifiers: FITBIR-STUDY0000220, 226, 228, 234, 239, 246, 256, 263, 267, 279, 280, 310, 313, 314, 321, 329, 330, 348, 349, 350, 358, 363, 364, 365, 409, and 411. This manuscript reflects the views of the authors and may not reflect the opinions or views of the DOD, NIH, or of the Submitters submitting original data to FITBIR Informatics System. We used generative AI to debug, edit, and autocomplete code. We take accountability for the review of all work generated by AI.

5. REFERENCES

[1] G. Teasdale and B. Jennett, "ASSESSMENT OF COMA AND IMPAIRED CONSCIOUSNESS: A Practical Scale," *The Lancet*, vol. 304, no. 7872, pp. 81–84, July 1974, doi: 10.1016/S0140-6736(74)91639-0.

[2] H. M. Lindsey, C. B. Hodges, K. M. Greer, E. A. Wilde, and T. L. Merkley, "Diffusion-Weighted Imaging in Mild Traumatic Brain Injury: A Systematic Review of the Literature," *Neuropsychol. Rev. 2021 331*, vol. 33, no. 1, pp. 42–121, Mar. 2021, doi: 10.1007/S11065-021-09485-5.

[3] G. T. Manley *et al.*, "A new characterisation of acute traumatic brain injury: the NIH-NINDS TBI Classification and Nomenclature Initiative," *Lancet Neurol.*, vol. 24, no. 6, pp. 512–523, June 2025, doi: 10.1016/S1474-4422(25)00154-1/ATTACHMENT/D37573D6-B714-4F3C-B861-388783C05568/MMC1.PDF.

[4] E. D. Bigler, "Volumetric MRI Findings in Mild Traumatic Brain Injury (mTBI) and Neuropsychological Outcome," *Neuropsychol. Rev.*, vol. 33, no. 1, pp. 5–41, Mar. 2023, doi: 10.1007/s11065-020-09474-0.

[5] S. Richter *et al.*, "Predicting recovery in patients with mild traumatic brain injury and a normal CT using serum biomarkers and diffusion tensor imaging (CENTER-TBI): an observational cohort study," *eClinicalMedicine*, vol. 75, Sept. 2024, doi: 10.1016/j.eclinm.2024.102751.

[6] H. J. Thompson, M. S. Vavilala, and F. P. Rivara, "Common Data Elements and Federal Interagency Traumatic Brain Injury Research Informatics System for TBI Research," *Annu. Rev. Nurs. Res.*, vol. 33, no. 1, pp. 1–11, May 2015, doi: 10.1891/0739-6686.33.1.

[7] A. I. R. Maas *et al.*, "Collaborative European NeuroTrauma Effectiveness Research in Traumatic Brain Injury (CENTER-TBI): A Prospective Longitudinal Observational Study," *Neurosurgery*, vol. 76, no. 1, p. 67, Jan. 2015, doi: 10.1227/NEU.0000000000000575.

[8] X. Li, P. S. Morgan, J. Ashburner, J. Smith, and C. Rorden, "The first step for neuroimaging data analysis: DICOM to NIfTI conversion," *J. Neurosci. Methods*, vol. 264, pp. 47–56, May 2016, doi: 10.1016/j.jneumeth.2016.03.001.

[9] S. Bao *et al.*, "Integrating the BIDS Neuroimaging Data Format and Workflow Optimization for Large-Scale Medical Image Analysis," *J. Digit. Imaging*, vol. 35, no. 6, pp. 1576–1589, Dec. 2022, doi: 10.1007/s10278-022-00679-8.

[10] M. E. Kim *et al.*, "Scalable quality control on processing of large diffusion-weighted and structural magnetic resonance imaging datasets," *PLOS One*, vol. 20, no. 8, p. e0327388, Aug. 2025, doi: 10.1371/journal.pone.0327388.

[11] A. M. Saunders *et al.*, "Phenotype discovery of traumatic brain injury segmentations from heterogeneous multi-site data," in *SPIE Medical Imaging: Image Processing*, 2026. [Online]. Available: https://arxiv.org/abs/2511.03767

[12] C. Zhao, B. E. Dewey, D. L. Pham, P. A. Calabresi, D. S. Reich, and J. L. Prince, "SMORE: A Self-Supervised Anti-Aliasing and Super-Resolution Algorithm for MRI Using Deep Learning," *IEEE Trans. Med. Imaging*, vol. 40, no. 3, pp. 805–817, Mar. 2021, doi: 10.1109/TMI.2020.3037187.

[13] N. J. Tustison *et al.*, "N4ITK: Improved N3 Bias Correction," *IEEE Trans. Med. Imaging*, vol. 29, no. 6, pp. 1310–1320, June 2010, doi: 10.1109/TMI.2010.2046908.

[14] L. Zuo *et al.*, "HACA3: A unified approach for multi-site MR image harmonization," *Comput. Med. Imaging Graph.*, vol. 109, p. 102285, Oct. 2023, doi: 10.1016/j.compmedimag.2023.102285.

[15] X. Yu *et al.*, "UNesT: Local spatial representation learning with hierarchical transformer for efficient medical segmentation," *Med. Image Anal.*, vol. 90, p. 102939, Dec. 2023, doi: 10.1016/J.MEDIA.2023.102939.

[16] L. Y. Cai *et al.*, "PreQual: An automated pipeline for integrated preprocessing and quality assurance of diffusion weighted MRI images," *Magn. Reson. Med.*, vol. 86, no. 1, pp. 456–470, July 2021, doi: 10.1002/mrm.28678.

[17] A. Valcourt Caron *et al.*, "nf-neuro: Developing maintainable pipelines for seamless processing of diffusion MRI," in *Proceedings of ISMRM Workshop on 40 Years of Diffusion: Past, Present & Future Perspectives*, 2025.

[18] A. Hoopes, J. S. Mora, A. V. Dalca, B. Fischl, and M. Hoffmann, "SynthStrip: skull-stripping for any brain image," *NeuroImage*, vol. 260, p. 119474, Oct. 2022, doi: 10.1016/J.NEUROIMAGE.2022.119474.

[19] O. Pasternak, N. Sochen, Y. Gur, N. Intrator, and Y. Assaf, "Free water elimination and mapping from diffusion MRI," *Magn. Reson. Med.*, vol. 62, no. 3, pp. 717–730, Sept. 2009, doi: 10.1002/mrm.22055.

[20] J.-P. Fortin *et al.*, "Harmonization of multi-site diffusion tensor imaging data," *NeuroImage*, vol. 161, pp. 149–170, Nov. 2017, doi: 10.1016/j.neuroimage.2017.08.047.

[21] R. A. Rigby and D. M. Stasinopoulos, "Generalized Additive Models for Location, Scale and Shape," *J. R. Stat. Soc. Ser. C Appl. Stat.*, vol. 54, no. 3, pp. 507–554, June 2005, doi: 10.1111/j.1467-9876.2005.00510.x.

[22] P.-M. Jodoin *et al.*, "ComBAT Harmonization for diffusion MRI: Challenges and Best Practices," May 19, 2025, *arXiv*: arXiv:2505.14722. doi: 10.48550/arXiv.2505.14722.

[23] M. E. Kim *et al.*, "Empirical assessment of the assumptions of ComBat with diffusion tensor imaging," *J. Med. Imaging*, vol. 11, no. 02, Apr. 2024, doi: 10.1117/1.JMI.11.2.024011.