## Title

Image data harmonization tools for the analysis of post-traumatic epilepsy development in preclinical multisite MRI studies.



## Authors

Bhagavatula, Sweta
Cabeen, Ryan
Harris, Neil
et al.

## Publication Date

2023-09-01

## DOI

10.1016/j.eplepsyres.2023.107201

Peer reviewed

HHS Public Access

Author manuscript

*Epilepsy Res*. Author manuscript; available in PMC 2024 September 01.

Published in final edited form as:

*Epilepsy Res*. 2023 September ; 195: 107201. doi:10.1016/j.eplepsyres.2023.107201.

# Image Data Harmonization Tools for the Analysis of Posttraumatic Epilepsy Development in Preclinical Multisite MRI Studies

**Sweta Bhagavatula**[1], **Ryan Cabeen**[1], **Neil G. Harris**[4], **Olli Gröhn**[2], **David K. Wright**[3], **Rachael Garner**[1], **Alexis Bennett**[1], **Celina Alba**[1], **Aubrey Martinez**[1], **Xavier Ekolle Ndode-Ekane**[2], **Pedro Andrade**[2], **Tomi Paananen**[2], **Robert Ciszek**[2], **Riikka Immonen**[2], **Eppu Manninen**[2], **Noora Puhakka**[2], **Jussi Tohka**[2], **Mette Heiskanen**[2], **Idrish Ali**[3], **Sandy R. Shultz**[3], **Pablo M. Casillas-Espinosa**[3], **Glenn R. Yamakawa**[3], **Nigel C. Jones**[3], **Matthew R. Hudson**[3], **Juliana C. Silva**[3], **Emma L. Braine**[3], **Rhys D. Brady**[3], **Cesar E. Santana-Gomez**[4], **Gregory D. Smith**[4], **Rick Staba**[4], **Terence J O'Brien**[3], **Asla Pitkänen**[2], **Dominique Duncan**[1] **EpiBioS4Rx investigators**

[1]USC Stevens Neuroimaging and Informatics Institute, Keck School of Medicine of USC, University of Southern California, Los Angeles, CA;

[2]A.I. Virtanen Institute for Molecular Sciences, University of Eastern Finland, Kuopio, Finland;

[3]Departments of Neuroscience and Neurology, Central Clinical School, Alfred Health, Monash University, Melbourne, Victoria, Australia;

[4]Department of Neurology, David Geffen School of Medicine at University of California at Los Angeles, Los Angeles, CA

## Abstract

Preclinical MRI studies have been utilized for the discovery of biomarkers that predict posttraumatic epilepsy (PTE). However, these single site studies often lack statistical power due to limited and homogeneous datasets. Therefore, multisite studies, such as the Epilepsy Bioinformatics Study for Antiepileptogenic Therapy (EpiBioS4Rx), are developed to create large, heterogeneous datasets that can lead to more statistically significant results. EpiBioS4Rx collects preclinical data internationally across sites, including the United States, Finland, and Australia. However, in doing so, there are robust normalization and harmonization processes that are required to obtain statistically significant and generalizable results. This work describes the tools and procedures used to harmonize multisite, multimodal preclinical imaging data acquired by EpiBioS4Rx. There were four main harmonization processes that were utilized, including file

format harmonization, naming convention harmonization, image coordinate system harmonization, and diffusion tensor imaging (DTI) metrics harmonization. By using Python tools and bash scripts, the file formats, file names, and image coordinate systems are harmonized across all the sites. To harmonize DTI metrics, values are estimated for each voxel in an image to generate a histogram representing the whole image. Then, the Quantitative Imaging Toolkit (QIT) modules are utilized to scale the mode to a value of one and depict the subsequent harmonized histogram. The standardization of file formats, naming conventions, coordinate systems, and DTI metrics are qualitatively assessed. The histograms of the DTI metrics were generated for all the individual rodents per site. For inter-site analysis, an average of the individual scans was calculated to create a histogram that represents each site. In order to ensure the analysis can be run at the level of individual animals, the sham and TBI cohort were analyzed separately, which depicted the same harmonization factor. The results demonstrate that these processes qualitatively standardize the file formats, naming conventions, coordinate systems, and DTI metrics of the data. This assists in the ability to share data across the study, as well as disseminate tools that can help other researchers to strengthen the statistical power of their studies and analyze data more cohesively.

## 1. INTRODUCTION

Traumatic brain injury (TBI) is typically the result of an external force to the brain that leads to several pathologies, including posttraumatic epilepsy (PTE). While epilepsy is a common neurological disorder that affects millions, the particular cases of PTE are quite varied and present symptoms several years after the TBI, making it difficult to study the particular pathophysiology of the disease (Cabeen et al., 2020; Immonen et al., 2019). In order to better understand these outcomes, there is a search for noninvasive imaging biomarkers of TBI and PTE. The Epilepsy Bioinformatics Study for Antiepileptogenic Therapy (EpiBioS4Rx) is one study that aims to identify epileptogenic biomarkers after TBI. Within the EpiBioS4Rx project, there are preclinical and clinical components that are conducted over various research centers based in the United States, Finland, and Australia. Out of the preclinical components, the first aspect of EpiBioS4Rx utilizes lateral fluid percussion injury (FPI) in rodents to discover and analyze biomarkers for PTE (Ndode-Ekane et al., 2019; Weiler et al., 2022). Each site collecting preclinical data obtains magnetic resonance imaging (MRI) scans of these rodents at four time points: 2 days, 9 days, 1 month, and 5 months after the injury (Vespa et al., 2019). Along with the establishment of biomarkers for PTE in a rodent model, a secondary project of EpiBioS4Rx aims to generate a preclinical trial that is based on the biomarkers that are discovered by the primary project. This secondary project will not only analyze the effects of pharmacological therapeutics on PTE, but it will also include an additional site, the Albert Einstein College of Medicine, which leads to larger datasets and more diverse samples.

For EpiBioS4Rx, there has already been previous analysis on the harmonization of data acquisition across the sites, focusing on variations in MRI scanners and noise characteristics (Ciszek et al., 2019; Immonen et al., 2019; Pitkänen et al., 2019). The purpose of this paper is to present and discuss the development of several frameworks and pipelines that harmonize aspects of the MRI imaging data after they are acquired from each site of the EpiBioS4Rx preclinical study (Immonen et al., 2019). It will also present

these harmonization techniques as a way of depicting the influences, or lack thereof, on harmonization protocols, such as TBI. Therefore, the specific focus was to create pipelines that are able to harmonize sham and TBI rodent data simultaneously. It was necessary to generate a pipeline for EpiBioS4Rx, and other preclinical multisite TBI studies, that addresses the additional harmonizing and preprocessing challenges that arise with TBI data. TBI introduces varying factors, such as lesion size, changes in cortical thickness, and tissue destruction (Xiong et al., 2013). One specific factor that is important to harmonize across TBI and sham rodents, despite the potential changes that occur with TBI, is diffusion tensor imaging (DTI) metrics. DTI is an MRI technique that focuses on detecting pathological tissue damage. It is intended to detect pathological alterations in unhealthy tissue before the changes appear in conventional imaging (Homos et al., 2017). This is accomplished by detecting changes in diffusion and the directionality of diffusion within tissues due to cellular barriers. One specific DTI metric is fractional anisotropy (FA), which measures the degree of the directionality of diffusion (Pierpaoli et al., 1996). FA has been shown to be an important feature in TBI analysis for PTE (Gupta et al., 2005). There are several factors that can alter these parameters, including eddy currents, imperfections in gradient calibrations, motion artifacts, the time of day, and temperature (Hasan et al., 2014; Kozak et al., 2009; Liu et al., 2010; Thomas et al., 2018; Truong et al., 2011). These variations can also lead to broader issues, such as image quality. Due to these variations, it is important to implement harmonization processes that standardize these metrics and allow for inter-site analysis. This paper will present these pipelines so they can be implemented for current and future preclinical multisite MRI studies.

Recently, there has been a rise in multisite clinical studies, specifically for neuroimaging. Various normalization and harmonization processes have been developed by multisite studies and other consortia to overcome the challenges of data variation and thereby obtain the statistical power necessary to provide meaningful insight into a number of neuropsychiatric disorders, emphasizing the relevance and importance of these pipelines (Bell et al., 2022; Koike et al., 2021; Treiman, 2019; Wrobel et al., 2020). For instance, studies in the Strategic Research Program for Brain Science (SRPBS) consortium utilized harmonization methods that removed solely measurement bias to predict brain features in neuropsychiatric patients (Yamashita et al., 2020, 2019). Another multisite study, the Japanese Alzeheimer's Disease Neuroimaging Initiative, or J-ADNI, was also able to utilize harmonization of protocol and procedures in acquiring data to improve the ability to predict the progression of cognitively normal elderly patients, mild cognitive impairment (MCI), and mild Alzheimer's Disease (AD) (Beheshti et al., 2017; Iwatsubo et al., 2018). Additionally, the Adolescent Brain Cognitive Development study has benefitted from harmonizing imaging methods and assessments across 21 sites to understand brain structure and function with regards to adolescent development and addiction (Casey et al., 2018). Therefore, through several studies, it has been shown that harmonization techniques are necessary to produce reliable and effective analysis for neuroimaging studies (Schnack et al., 2004; Yu et al., 2018).

While clinical multisite MRI studies, such as these, are extremely beneficial and have been becoming relatively common, there is almost complete lack of preclinical multisite MRI studies. Preclinical MRI studies introduce more avenues for research, however, there

are specific features of the preclinical multisite MRI study that must be addressed when compared to one that is clinical (Brockmann et al., 2007). Preclinical multisite studies typically have a greater variation in MRI field strength and other hardware configurations. For instance, the range of field strengths in the preclinical portion of EpiBioS4Rx is 4.7 to 9.4 T (Gao et al., 2015; Gsell et al., 2020). Furthermore, when compared to clinical settings, pulse sequences are less standardized (Moffat et al., 2004). Therefore, it is common to have in-house modifications of pulse sequences to accommodate for differences in hardware in an optimal way. This poses a significant additional challenge compared to clinical settings.

The normalization of metrics has led to the ability to store all the preclinical data in one repository, which has been a major step in increasing the efficiency of the analysis. In addition to making access to data more seamless and rapid, building these data tools is important for EpiBioS4Rx as it increases the statistical power of the study and leads to more cohesive and robust analysis of the data. Not only are these tools beneficial for this study, but also sharing them on public platforms, such as GitHub, for communal access can assist other multi-site preclinical projects' analysis to obtain more statistically significant results.

## 2. METHODS

### 2.1 Overview of Harmonization Processes

Preclinical MRI data was acquired from three preclinical EpiBioS4Rx sites: University of Eastern Finland, University of Melbourne/Florey Institute, and University of California, Los Angeles. The data set from University of Eastern Finland contains 43 rodents; 32 TBI and 11 sham. The University of Melbourne/Florey Institute data set contains 41 rodents; 31 TBI and 10 sham. Finally, the University of California, Los Angeles data set contains 43 rodents; 33 TBI and 10 sham. These values are presented in Figure 1. Four main harmonization processes were applied to the MRI scans across all sites: file format harmonization, naming convention harmonization, image coordinate system harmonization, and diffusion tensor imaging (DTI) metrics harmonization. These processes are visualized in Figure 2. The software primarily used in these processes includes the Quantitative Imaging Toolkit (QIT) (Cabeen et al., 2018). This package includes several computational tools that are particularly helpful for the visualization and analysis for neuroimaging data (https://cabeen.io/qitwiki/).

### 2.2 File Format Harmonization

As the MRI scans are uploaded across the sites, there is variation in the format that the data is stored and the naming of the sequences. However, the imaging data produced by a Bruker MRI provides a "method" file (Doss et al., 2022). This is a metadata file that provides information about the session and acquisition of the data, including information about the scanner and the imaging sequence. The challenge with harmonizing files from the Bruker scanner is that the "method" file is in a proprietary format, so it is difficult to read and interpret. Therefore, we created a Python tool, ParseBruker, to parse and summarize Bruker "method" files, making them easier to understand and work with. This tool, which is available on GitHub, is able to ensure that the "method" files are converted to a JSON output and formatted output based on a used-specified pattern (https://github.com/cabeen/parsebruker). Each method file comprises a header, header attributes, parameters,

and parameter attributes. The header and parameters are key value pairs: the header provides information about the session, such as date and environment, whereas the parameters are regarding the scanner protocol. The attributes are additional details based on the header and parameters, such as spatial resolution and diffusion gradient b-vectors and b-values. Once the method file is interpreted, the imaging data can then be sorted based on the variations in sequences, for example. This is the primary tool that is utilized in converting the Bruker files to a more easily readable NIfTI format that can be further used for cross-site analysis.

### 2.3 Naming Convention Harmonization

Along with the file format variation across sites, there are also different naming conventions that are utilized. Currently, there is no standardized naming that is widely used among rodent data, unlike human data which often leverages the Brain Imaging Data Structure, or BIDS (Gorgolewski et al., 2016). Human imaging sequences are slightly more standardized, whereas rodent imaging is not due to its experimental nature. Although BIDS is increasingly being used in the scope of preclinical MRI studies, it is not as ubiquitous as a standardized naming convention should be (Ioanas et al., 2022; Ioanas and Rudin, 2019; Kim et al., 2021). The main issue is that each imaging center has a unique way of sequencing the rodent data, which is a problem that is common for all preclinical multisite studies. Additionally, the community lacks many publicly-available converters to maintain consistent conventions across studies. There is a converter available, known as BrkRaw, which is intended to convert large raw datasets into an organized structure according to BIDS (https://github.com/BrkRaw/brkraw). However, there are still issues with identifying scans due to the various naming conventions associated with the raw data. This still leads to the challenge of not being able to compare data across sites, especially at specific time points. Therefore, we created bash scripts to address these challenges. The bash scripts, EpibiosAuxBrukerImport.sh and EpibiosAuxBrukerCommon.sh, is available to the public on GitHub (https://github.com/cabeen/epibios-mri-rat). These scripts rename the files in a consistent way using string pattern matching rules, which are manually defined, to identify the files from the diffusion weighted imaging (DWI), multi-gradient echo (MGE), specifically T2-star, and magnetization transfer (MT) scans. This harmonizes the protocol names for all images in the study: dwi, dwilow, mge, mtlow, mthigh, and rare, leading to the naming convention across all sites to be standardized so that the data can be more easily shared and comprehended for all researchers in the study.

### 2.4 Coordinate System Harmonization

The next step that is essential to processing data across several sites is the harmonization of the coordinate system of the images. The conversion from Bruker format to NIFTI can often lead to differences in image coordinate systems and orientations, which vary across scans and sites. To ensure that all the images' coordinate systems are harmonized, we applied bash scripts to scans across all sites that utilize the BrkRaw tool and the novel ParseBruker tool. This process is depicted in the bash script, EpibiosAuxBrukerImport.sh, and the beginning of the pipeline in the EpibiosAuxProcess.makefile, which is available to the public on GitHub (https://github.com/cabeen/epibios-mri-rat). These steps ensure that all images are processed in the same way, avoiding any changes in coordinate systems across scans.

### 2.5 DTI Metrics Harmonization

The DTI metrics, specifically fractional anisotropy (FA), were also harmonized across animals and sites (Fortin et al., 2017). Harmonization of the FA values was used as a way to represent the harmonization of DTI metrics as FA is the most common and comparable factor (Fox et al., 2012). While there are others, such as mean diffusivity, it is assumed that FA values are sufficient to present the effects of harmonization.

In the EpiBioS4Rx dataset, every DW image had an FA value that was estimated for each voxel in that image. We used a global statistical approach to standardize these FA values. We first generated a histogram from the signal intensities within voxels isolated using brain extraction, and we then identified the peak value (or mode) of the histogram as a feature to normalize across individuals. This process was implemented in the QIT module VolumeHistogram. To harmonize the data, a QIT module, VolumeHarmonize, computes the mode of the data, then applies a global scaling factor that brings the peak to a value of one, and generates a subsequent histogram for quality control. This is due to the fact that normal appearing tissue is the most common and also has a value of one. The mode is utilized rather than the mean to avoid bias by lesion size. Once a histogram has been generated for each animal within each site, these histograms can be averaged to produce a histogram that is representative of the site.

Despite the harmonization process, it is still possible that outliers exist due to various factors, including severe artifacts or errors in skull stripping. It is also possible that harmonizing individual rodents could lead to global effects being ignored. So, in order to understand the local and global effects, the harmonization process was conducted and compared within sites and across sites. Also, this process was run on sham and TBI rodents separately to assess whether the harmonization factor differs for TBI.

## 3. RESULTS

These harmonization processes standardize file formats, naming conventions, coordinate systems, and DTI metrics, which are qualitatively assessed. The raw file names and harmonized file names, which are standardized across all rodents and sites, are depicted in Figure 3. The histograms were generated for all the FA values in every individual rodent's whole brain scan across all sites. The example of Monash University is shown in Figure 4, where every time point (2 days, 9 days, 1 month, and 5 months after the injury) and all sham and TBI rodents are presented. Through a qualitative assessment, there is a clear, visual decrease in variability after the harmonization process. For inter-site analysis, the individual histograms were averaged to depict a histogram that is representative of each site. In Figure 5, the whole brain rodent scans of the sham and TBI cohorts were analyzed separately. Each site-specific histogram includes either the sham or TBI rodents at all four time points. The raw and harmonized data for each is also shown.

## 4. DISCUSSION & CONCLUSION

These results qualitatively indicate that the harmonization processes were successful in normalizing aspects of the data post-acquisition. This is especially useful for the analysis

of FA values across sham and TBI. TBI introduces several factors to the data, including variations in lesion sizes and tissue destruction, which could lead to alterations in FA values. It may be thought that this could influence the harmonization factor. However, despite all the varying effects of TBI, based on the visual representation of the data, TBI does not influence the harmonization factor. The harmonization process was as effective at normalizing FA values in TBI rodents as it had in sham rodents, ensuring that the analysis can be run at the level of the individual animals.

Within EpiBioS4Rx, compared to the harmonization processes in the clinical aspect of the project, the preclinical project requires some different methods. For instance, BIDS exists for the standardized and organized naming and file structure for human data (Gorgolewski et al., 2016). However, the development of a naming convention specifically for the preclinical component of the EpiBioS4Rx study was necessary. By harmonizing this, along with the other factors discussed, data can be more easily shared among all researchers in the preclinical study. Additionally, when running analysis for the discovery of biomarkers for PTE, it is much less challenging if the file formats, naming convention, coordinate systems, and DTI metrics are all standardized. This standardization allows researchers to seamlessly choose rodent cases from each site and apply similar pipelines. Beyond this particular study, building these data tools is important as it can be utilized by other preclinical multisite MRI studies to make data more useful and expedite research (Duncan et al., 2019). As these methods increase the statistical power of the study and lead to more cohesive and robust analysis, they can be applied to other studies that may have similar variations prior to harmonization.

Future directions for this project include adapting this preliminary harmonization pipeline to other projects of EpiBioS4Rx. As mentioned previously, there is an additional preclinical project to analyze the impact of pharmacological treatments on PTE, which includes data from an additional site at the Albert Einstein College of Medicine. However, this data was obtained from the Agilent MRI scanner. This inclusion in the future, while it will lead to a larger, more diverse dataset, will also require more harmonization processes to account for differences in scanner type. Also, when conducting these harmonization methods, it will be beneficial to create a pipeline that can include all pre-processing steps for all four sites through one replicable script. Future analysis that includes more sites can additionally focus on any impacts in harmonization that may arise due to variation in lesion extent and location that is based on protocol and acquisition.

Some limitations of this study include only developing harmonization protocols from MRI data that was obtained from Bruker scanners. While these pipelines are intended to serve as a preliminary framework for future work, it is acknowledged that the pipeline is only applied to one scanner. Preclinical scanners have a large variability in bore sizes, gradient performance, and radiofrequency (RF)-coils to the level that two preclinical scanners are seldom alike. This leads to an increased need for more harmonization protocols between scanners in order to conduct statistically significant inter-site analysis. As mentioned previously, future EpiBioS4Rx work aims to develop effective data-analysis methods that include other scanners, such as the Agilent MRI scanner. Another limitation, specifically for the analysis of DTI metrics, is that there was only a qualitative assessment of FA

values as a representation for the effects of harmonization. Although there was a clear qualitative standardization of FA values through the harmonization protocol, it may have been beneficial to include other DTI metrics that are quantitatively assessed.

While it is acknowledged that these scripts and tools are currently specific for the EpiBioS4Rx study, it is believed that their applicability will be useful for several other studies. Other studies can utilize these tools and develop similar scripts to process and structure their data with similar formats, naming conventions, and coordinate systems. Also, it is useful for other studies to recognize how TBI, despite all the variations it introduces, does not impact the harmonization factor when normalizing FA values. The dissemination of this will hopefully lead to more collaboration and analysis of preclinical multisite MRI data that can be used to improve the understanding of PTE and its implications.

## REFERENCES

Beheshti I, Maikusa N, Daneshmand M, Matsuda H, Demirel H, Anbarjafari G, 2017. Classification of Alzheimer's Disease and Prediction of Mild Cognitive Impairment Conversion Using Histogram-Based Analysis of Patient-Specific Anatomical Brain Connectivity Networks. JAD 60, 295–304. 10.3233/JAD-161080 [PubMed: 28800325]

Bell TK, Godfrey KJ, Ware AL, Yeates KO, Harris AD, 2022. Harmonization of multi-site MRS data with ComBat. NeuroImage 257, 119330. 10.1016/j.neuroimage.2022.119330 [PubMed: 35618196]

Brockmann MA, Kemmling A, Groden C, 2007. Current issues and perspectives in small rodent magnetic resonance imaging using clinical MRI scanners. Methods 43, 79–87. 10.1016/j.ymeth.2007.07.001 [PubMed: 17720566]

Cabeen RP, Immonen R, Harris NG, Grohn O, Smith G, Manninen E, Garner R, Duncan D, Pitkanen A, Toga AW, 2020. A Computational Diffusion MRI Framework for Biomarker Discovery in a Rodent Model of Post-Traumatic Epileptogenesis, in: 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI). Presented at the 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), IEEE, Iowa City, IA, USA, pp. 1013–1017. 10.1109/ISBI45749.2020.9098575

Cabeen RP, Laidlaw DH, Toga AW, 2018. Quantitative Imaging Toolkit: Software for Interactive 3D Visualization, Data Exploration, and Computational Analysis of Neuroimaging Datasets. ISMRM-ESMRMB Abstracts 12–14.

Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, Soules ME, Teslovich T, Dellarco DV, Garavan H, Orr CA, Wager TD, Banich MT, Speer NK, Sutherland MT, Riedel MC, Dick AS, Bjork JM, Thomas KM, Chaarani B, Mejia MH, Hagler DJ, Daniela Cornejo M, Sicat CS, Harms MP, Dosenbach NUF, Rosenberg M, Earl E, Bartsch H, Watts R, Polimeni JR, Kuperman JM, Fair DA, Dale AM, 2018. The Adolescent Brain Cognitive Development (ABCD) study: Imaging acquisition across 21 sites. Developmental Cognitive Neuroscience 32, 43–54. 10.1016/j.dcn.2018.03.001 [PubMed: 29567376]

Ciszek R, Ndode-Ekane XE, Gomez CS, Casillas-Espinosa PM, Ali I, Smith G, Puhakka N, Lapinlampi N, Andrade P, Kamnaksh A, Immonen R, Paananen T, Hudson MR, Brady RD, Shultz SR, O'Brien TJ, Staba RJ, Tohka J, Pitkänen A, 2019. Informatics tools to assess the success of procedural harmonization in preclinical multicenter biomarker discovery study on post-traumatic epileptogenesis. Epilepsy Res 150, 17–26. 10.1016/j.eplepsyres.2018.12.010 [PubMed: 30605864]

Doss KKM, Mion PE, Kao Y-CJ, Kuo T-T, Chen J-C, 2022. Performance Evaluation of a PET of 7T Bruker Micro-PET/MR Based on NEMA NU 4–2008 Standards. Electronics 11, 2194. 10.3390/electronics11142194

Duncan D, Vespa P, Pitkänen A, Braimah A, Lapinlampi N, Toga AW, 2019. Big data sharing and analysis to advance research in post-traumatic epilepsy. Neurobiol Dis 123, 127–136. 10.1016/j.nbd.2018.05.026 [PubMed: 29864492]

Fortin J-P, Parker D, Tunç B, Watanabe T, Elliott MA, Ruparel K, Roalf DR, Satterthwaite TD, Gur RC, Gur RE, Schultz RT, Verma R, Shinohara RT, 2017. Harmonization of multi-site diffusion tensor imaging data. NeuroImage 161, 149–170. 10.1016/j.neuroimage.2017.08.047 [PubMed: 28826946]

Fox RJ, Sakaie K, Lee JC, Debbins JP, Liu Y, Arnold DL, Melhem ER, Smith CH, Philips MD, Lowe M, Fisher E, 2012. A Validation Study of Multicenter Diffusion Tensor Imaging: Reliability of Fractional Anisotropy and Diffusivity Values. AJNR Am J Neuroradiol 33, 695–700. 10.3174/ajnr.A2844 [PubMed: 22173748]

Gao Y, Chen Y, Ma D, Jiang Y, Herrmann KA, Vincent JA, Dell KM, Drumm ML, Brady-Kalnay SM, Griswold MA, Flask CA, Lu L, 2015. Preclinical MR fingerprinting (MRF) at 7 T: effective quantitative imaging for rodent disease models: HIGH-FIELD PRECLINICAL MRF. NMR Biomed. 28, 384–394. 10.1002/nbm.3262 [PubMed: 25639694]

Gorgolewski KJ, Auer T, Calhoun VD, Craddock RC, Das S, Duff EP, Flandin G, Ghosh SS, Glatard T, Halchenko YO, Handwerker DA, Hanke M, Keator D, Li X, Michael Z, Maumet C, Nichols BN, Nichols TE, Pellman J, Poline JB, Rokem A, Schaefer G, Sochat V, Triplett W, Turner JA, Varoquaux G, Poldrack RA, 2016. The brain imaging data structure, a format for organizing and describing outputs of neuroimaging experiments. Sci Data 3, 160044. 10.1038/sdata.2016.44 [PubMed: 27326542]

Gsell W, Molinos C, Correcher C, Belderbos S, Wouters J, Junge S, Heidenreich M, Velde GV, Rezaei A, Nuyts J, Cawthorne C, Cleeren F, Nannan L, Deroose CM, Himmelreich U, Gonzalez AJ, 2020. Characterization of a preclinical PET insert in a 7 tesla MRI scanner: beyond NEMA testing. Phys. Med. Biol 65, 245016. 10.1088/1361-6560/aba08c [PubMed: 32590380]

Gupta RK, Saksena S, Agarwal A, Hasan KM, Husain M, Gupta V, Narayana PA, 2005. Diffusion Tensor Imaging in Late Posttraumatic Epilepsy. Epilepsia 46, 1465–1471. 10.1111/j.1528-1167.2005.01205.x [PubMed: 16146442]

Hasan KM, Moeller FG, Narayana PA, 2014. DTI-based segmentation and quantification of human brain lateral ventricular CSF volumetry and mean diffusivity: Validation, age, gender effects and biophysical implications. Magnetic Resonance Imaging 32, 405–412. 10.1016/j.mri.2014.01.014 [PubMed: 24582546]

Homos MD, Ali MT, Osman MF, Nabil DM, 2017. DTI metrics reflecting microstructural changes of normal appearing deep grey matter in multiple sclerosis. The Egyptian Journal of Radiology and Nuclear Medicine 48, 1005–1008. 10.1016/j.ejrnm.2017.04.012

Immonen R, Smith G, Brady RD, Wright D, Johnston L, Harris NG, Manninen E, Salo R, Branch C, Duncan D, Cabeen R, Ndode-Ekane XE, Gomez CS, Casillas-Espinosa PM, Ali I, Shultz SR, Andrade P, Puhakka N, Staba RJ, O'Brien TJ, Toga AW, Pitkänen A, Gröhn O, 2019. Harmonization of pipeline for preclinical multicenter MRI biomarker discovery in a rat model of post-traumatic epileptogenesis. Epilepsy Res 150, 46–57. 10.1016/j.eplepsyres.2019.01.001 [PubMed: 30641351]

Ioanas H-I, Rudin M, 2019. BIDS-formatted example mouse brain data for SAMRI. 10.5281/ZENODO.3831124

Ioanas H-I, Saab BJ, Rudin M, 2022. Whole-brain opto-fMRI map of mouse VTA dopaminergic activation reflects structural projections with small but significant deviations. Transl Psychiatry 12, 60. 10.1038/s41398-022-01812-5 [PubMed: 35165257]

Iwatsubo T, Iwata A, Suzuki K, Ihara R, Arai H, Ishii K, Senda M, Ito K, Ikeuchi T, Kuwano R, Matsuda H, Japanese Alzheimer's Disease Neuroimaging Initiative, Sun C, Beckett LA, Petersen RC, Weiner MW, Aisen PS, Donohue MC, Alzheimer's Disease Neuroimaging Initiative, 2018. Japanese and North American Alzheimer's Disease Neuroimaging Initiative studies: Harmonization for international trials. Alzheimer's & Dementia 14, 1077–1087. 10.1016/j.jalz.2018.03.009

Kim E, Censo DD, Baraldo M, Simmons C, Rosa I, Randall K, Ballard C, Dickie B, Williams S, Killick R, Cash D, 2021. In vivo multi-parametric manganese-enhanced MRI for detecting senile plaques in rodent models of Alzheimer's disease. 10.18112/OPENNEURO.DS003463.V1.0.1

Koike S, Tanaka SC, Okada T, Aso T, Yamashita A, Yamashita O, Asano M, Maikusa N, Morita K, Okada N, Fukunaga M, Uematsu A, Togo H, Miyazaki A, Murata K, Urushibata Y, Autio J, Ose T, Yoshimoto J, Araki T, Glasser MF, Van Essen DC, Maruyama M, Sadato N, Kawato M, Kasai

K, Okamoto Y, Hanakawa T, Hayashi T, 2021. Brain/MINDS beyond human brain MRI project: A protocol for multi-level harmonization across brain disorders throughout the lifespan. NeuroImage: Clinical 30, 102600. 10.1016/j.nicl.2021.102600 [PubMed: 33741307]

Kozak L, Bango M, Szabo M, Rudas G, Vidnyanszky Z, Nagy Z, 2009. Using diffusion MRI for measuring the temperature of cerebrospinal fluid within the lateral ventricles. Acta Paediatrica. 10.1111/j.1651-2227.2009.01528.x

Liu Z, Wang Y, Gerig G, Gouttard S, Tao R, Fletcher T, Styner M, 2010. Quality control of diffusion weighted images, in: Liu BJ, Boonn WW (Eds.), . Presented at the SPIE Medical Imaging, San Diego, California, USA, p. 76280J. 10.1117/12.844748

Moffat BA, Hall DE, Stojanovska J, McConville PJ, Moody JB, Chenevert TL, Rehemtulla A, Ross BD, 2004. Diffusion imaging for evaluation of tumor therapies in preclinical animal models. MAGMA 17, 249–259. 10.1007/s10334-004-0079-z [PubMed: 15580371]

Ndode-Ekane XE, Santana-Gomez C, Casillas-Espinosa PM, Ali I, Brady RD, Smith G, Andrade P, Immonen R, Puhakka N, Hudson MR, Braine EL, Shultz SR, Staba RJ, O'Brien TJ, Pitkänen A, 2019. Harmonization of lateral fluid-percussion injury model production and post-injury monitoring in a preclinical multicenter biomarker discovery study on post-traumatic epileptogenesis. Epilepsy Res 151, 7–16. 10.1016/j.eplepsyres.2019.01.006 [PubMed: 30711714]

Pierpaoli C, Jezzard P, Basser PJ, Barnett A, Di Chiro G, 1996. Diffusion tensor MR imaging of the human brain. Radiology 201, 637–648. 10.1148/radiology.201.3.8939209 [PubMed: 8939209]

Pitkänen A, O'Brien TJ, Staba R, 2019. Preface - Practical and theoretical considerations for performing a multi-center preclinical biomarker discovery study of post-traumatic epileptogenesis: lessons learned from the EpiBioS4Rx consortium. Epilepsy Res 156, 106080. 10.1016/j.eplepsyres.2019.01.007 [PubMed: 30685321]

Schnack HG, Van Haren NEM, Hulshoff Pol HE, Picchioni M, Weisbrod M, Sauer H, Cannon T, Huttunen M, Murray R, Kahn RS, 2004. Reliability of brain volumes from multicenter MRI acquisition: A calibration study. Hum. Brain Mapp 22, 312–320. 10.1002/hbm.20040 [PubMed: 15202109]

Thomas C, Sadeghi N, Nayak A, Trefler A, Sarlls J, Baker CI, Pierpaoli C, 2018. Impact of time-of-day on diffusivity measures of brain tissue derived from diffusion tensor imaging. NeuroImage 173, 25–34. 10.1016/j.neuroimage.2018.02.026 [PubMed: 29458189]

Treiman DM, 2019. The potential and challenge of multicenter preclinical research. Epilepsy Res 156, 106152. 10.1016/j.eplepsyres.2019.106152 [PubMed: 31229310]

Truong TK, Chen N, Song AW, 2011. Dynamic correction of artifacts due to susceptibility effects and time-varying eddy currents in diffusion tensor imaging. NeuroImage 57, 1343–1347. 10.1016/j.neuroimage.2011.06.008 [PubMed: 21689763]

Vespa PM, Shrestha V, Abend N, Agoston D, Au A, Bell MJ, Bleck TP, Blanco MB, Claassen J, Diaz-Arrastia R, Duncan D, Ellingson B, Foreman B, Gilmore EJ, Hirsch L, Hunn M, Kamnaksh A, McArthur D, Morokoff A, O'Brien T, O'Phelan K, Robertson CL, Rosenthal E, Staba R, Toga A, Willyerd FA, Zimmermann L, Yam E, Martinez S, Real C, Engel J, 2019. The epilepsy bioinformatics study for anti-epileptogenic therapy (EpiBioS4Rx) clinical biomarker: Study design and protocol. Neurobiol Dis 123, 110–114. 10.1016/j.nbd.2018.07.025 [PubMed: 30048805]

Weiler M, Casseb RF, de Campos BM, Crone JS, Lutkenhoff ES, Vespa PM, Monti MM, EpiBioS4Rx Study Group, 2022. Evaluating denoising strategies in resting-state functional magnetic resonance in traumatic brain injury (EpiBioS4Rx). Hum Brain Mapp 43, 4640–4649. 10.1002/hbm.25979 [PubMed: 35723510]

Wrobel J, Martin ML, Bakshi R, Calabresi PA, Elliot M, Roalf D, Gur RC, Gur RE, Henry RG, Nair G, Oh J, Papinutto N, Pelletier D, Reich DS, Rooney WD, Satterthwaite TD, Stern W, Prabhakaran K, Sicotte NL, Shinohara RT, Goldsmith J, 2020. Intensity warping for multisite MRI harmonization. NeuroImage 223, 117242. 10.1016/j.neuroimage.2020.117242 [PubMed: 32798678]

Xiong Y, Mahmood A, Chopp M, 2013. Animal models of traumatic brain injury. Nat Rev Neurosci 14, 128–142. 10.1038/nrn3407 [PubMed: 23329160]

Yamashita A, Sakai Y, Yamada T, Yahata N, Kunimatsu A, Okada N, Itahashi T, Hashimoto R, Mizuta H, Ichikawa N, Takamura M, Okada G, Yamagata H, Harada K, Matsuo K, Tanaka SC, Kawato M, Kasai K, Kato N, Takahashi H, Okamoto Y, Yamashita O, Imamizu H, 2020. Generalizable

brain network markers of major depressive disorder across multiple imaging sites. PLoS Biol 18, e3000966. 10.1371/journal.pbio.3000966 [PubMed: 33284797]

Yamashita A, Yahata N, Itahashi T, Lisi G, Yamada T, Ichikawa N, Takamura M, Yoshihara Y, Kunimatsu A, Okada N, Yamagata H, Matsuo K, Hashimoto R, Okada G, Sakai Y, Morimoto J, Narumoto J, Shimada Y, Kasai K, Kato N, Takahashi H, Okamoto Y, Tanaka SC, Kawato M, Yamashita O, Imamizu H, 2019. Harmonization of resting-state functional MRI data across multiple imaging sites via the separation of site differences into sampling bias and measurement bias. PLoS Biol 17, e3000042. 10.1371/journal.pbio.3000042 [PubMed: 30998673]

Yu M, Linn KA, Cook PA, Phillips ML, McInnis M, Fava M, Trivedi MH, Weissman MM, Shinohara RT, Sheline YI, 2018. Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. Hum Brain Mapp 39, 4213–4227. 10.1002/hbm.24241 [PubMed: 29962049]

## Highlights

- Harmonizing preclinical MRI scans across sites provides large datasets for analysis
- Novel scripts and tools harmonize file formats, names, and image coordinate systems
- QIT modules generated scaled histograms to represent the DTI metrics harmonization
- DTI metrics analysis depicted that TBI does not influence the harmonization factor
- Harmonization leads to data sharing and increasing statistical strength of studies

| University of Eastern Finland | | | University of Melbourne/Florey Institute | | | University of California, Los Angeles | | |
|---|---|---|---|---|---|---|---|---|
| SHAM | TBI | TOTAL | SHAM | TBI | TOTAL | SHAM | TBI | TOTAL |
| 11 | 32 | 43 | 10 | 31 | 41 | 10 | 33 | 43 |

**Figure 1.**

The preclinical MRI data was acquired from three preclinical EpiBioS4Rx sites: University of Eastern Finland, University of Melbourne/Florey Institute, and University of California, Los Angeles. The figure includes the size of each site's sham rodent, TBI rodent, total rodent data set.

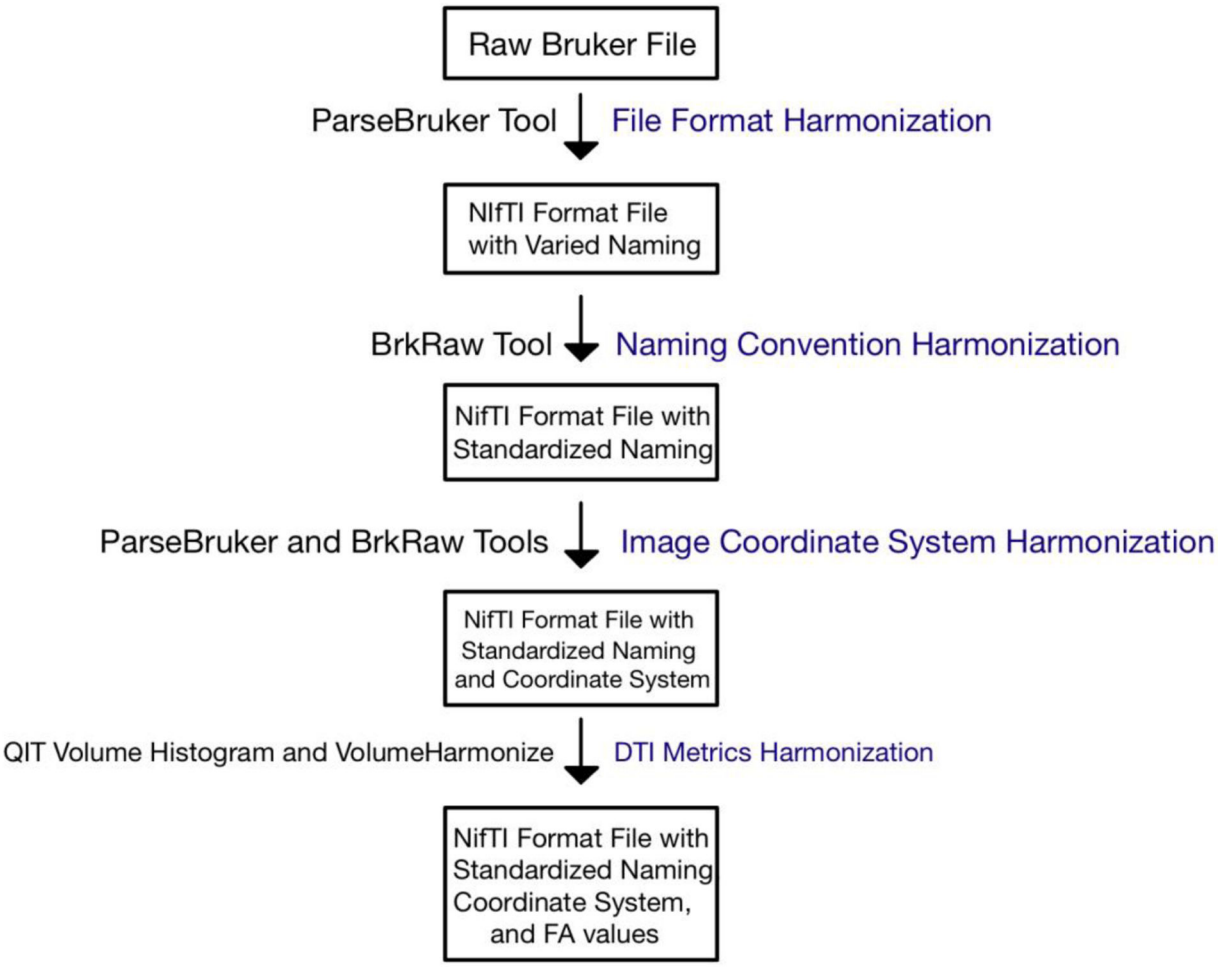


**Figure 2.**

Visual stepwise representation of the harmonization, including the main inputs and outputs, along with the corresponding primary tools utilized.


**Figure 3.**
Image of the raw file names and the filenames after the naming convention harmonization process. This process was repeated for every rodent across the three sites

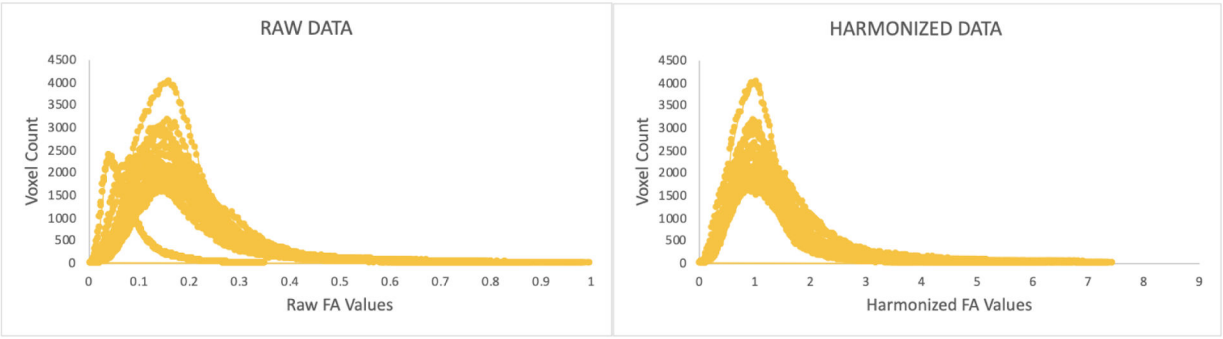


**Figure 4.**

Histograms of the FA values of whole brain scans of rodents, sham and TBI, in the Monash University dataset. All four time points (2 days, 9 days, 1 month, and 5 months after the injury) are presented. The left side depicts the raw histograms and the right side shows the histograms after the harmonization of the DTI metrics.

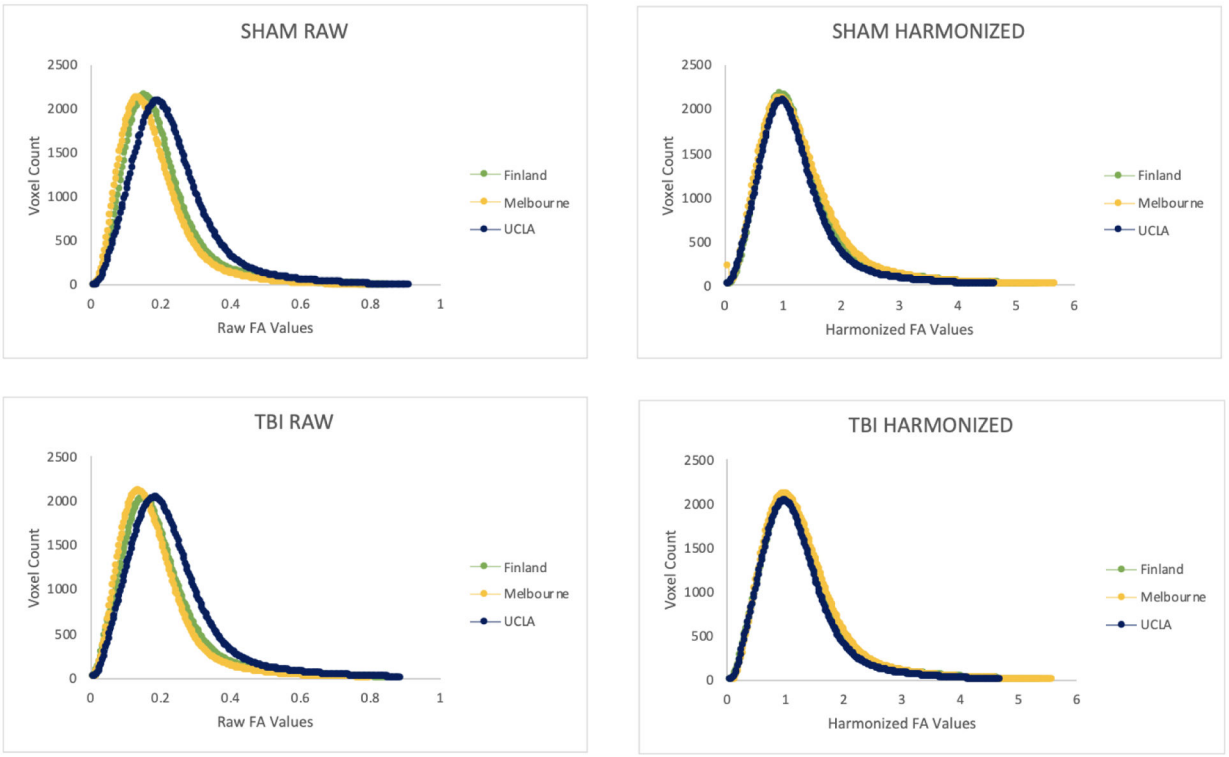


**Figure 5.**

Histograms of FA values of the whole brain rodent scans of the three sites for sham and TBI rodents depicting the raw data and harmonized data. In each histogram, all four time points are presented. The sham and TBI histograms were generated separately to show that regardless of sham vs. TBI or the site, the harmonization factor is the same.