JOURNAL OF UNDERGRADUATE RESEARCH

Evaluating Spatial Filtering on Diffusion MRI Data Harmonization in Parkinsonism

Madelyn Corliss

University of Florida

Faculty Mentor: Dr. David Vaillancourt, Health and Human Performance

Abstract

Background: Parkinsonism is an umbrella term encompassing several disease pathologies that share common motor symptoms. The most prevalent diagnosis is Parkinson's disease, followed by multiple system atrophy, and progressive supranuclear palsy. Early detection and differentiation between types of Parkinsonism remain an issue in clinical practice.
Objective: MRI has the potential to aid the diagnosis of Parkinsonisms. A major hurdle is combining and harmonizing the data across different MRI vendors. The objective of this study was to determine if a full width half maximum gaussian spatial filter helps harmonize data sets collected from different scanners.
Methods: Using 17 different MRI scanners, data was collected from 1,002 subjects. First, the data were spatially filtered using different sizes (no filter, 2mm, 4mm, 6mm). Data were then preprocessed and transformed into Montreal Neurological Institute (MNI) space. Next, support vector machine learning tested the training and validation accuracy of predicting diagnosis at each spatial filter setting.
Results: The training and validation data for weighted sensitivity, specificity, and accuracy were similar for all filter conditions. Differences between the weighted sensitivity, specificity, and accuracy of the training groups for all filter sizes were less than 0.1 and less than 0.2 for validation groups.
Conclusions: Training and validation predictions did not differ across spatial filters, suggesting the accuracy of the algorithm is robust at different spatial filter sizes. In conclusion, the size of the spatial filter applied to diffusion MRI data does not result in a change in the outcome of the machine learning approach.

Keywords: diffusion MRI, harmonization, parkinsonism, spatial filtering

Introduction

Parkinsonism is a broad umbrella term that encompasses several different diseases and pathologies that share common motor symptoms. The three most prevalent types include Parkinson's disease (PD), multiple system atrophy (MSA), and progressive supranuclear palsy (PSP). The introduction of a non-invasive biomarker for the different types of Parkinsonism can help improve the diagnosis accuracy (Hughes et al., 2002). Finding an algorithm that distinguishes between typical (PD) and atypical (MSA and PSP) Parkinsonism could help decrease the number of misdiagnoses and hopefully allow for earlier medical intervention. The main interest is in evaluating magnetic resonance imaging (MRI) as a potential biomarker for

differential diagnosis of PD and atypical Parkinsonism. As a first step towards that goal, a spatial filtering procedure was tested to determine if this method can enhance the harmonization of the MRI datasets from different MRI scanners.

This project reanalyzed data from another study (Archer et al., 2019) which included imaging from over 1000 subjects from 17 different MRI machines. Different spatial filter conditions (no filter, 2mm, 4mm, 6mm) were applied to these images before performing free-water calculation on them. The researcher hypothesized that the weighted sensitivity, specificity, and accuracy would all improve with increasing filter size. With 6mm being the largest filter size tested, it was predicted to be the best filtering method.

Most studies that have examined diffusion MRI data are single-site studies. Incorporating multiple sites, cohorts, and scanners into the same analysis pipeline would allow for further generalizability and scalability. Of 190 studies published in the *American Journal of Neuroradiology*, Siemens accounts for 38% of scanners used followed by 32% General Electric (GE), 26% Philips, and 4% other manufacturers (Ramezanpour et al., 2019). Changes in the volume of some parts of the brain, like the thalamus and globus pallidus (Plitman et al., 2021), have been reported with the use of different manufacturers, emphasizing the importance of normalization of data among different scanners. Still, there are a few studies, such as the Alzheimer's Disease Neuroimaging Initiative (ADNI), that have included Siemens, GE, and Philips even though there has not been a systematic assessment of how to combine these data in quantitative analysis pipelines.

When an MRI is acquired there is a reconstruction algorithm that moves the data from K-space to image space. Most analysis procedures process data in image space. Differences in voxel size and some unique filtering can occur in the reconstruction algorithm. In the current study, the 3dmerge function in the software package Analysis of Functional Neuroimages (AFNI) (Cox, 1996) was used to merge and edit the data sets, such that unique filters were applied to each data file. Our goal was to determine if adding a spatial filter could impact the outcome of a machine learning approach to predicting a diagnosis based on images from participants with PD, MSA, and PSP. A spatial filter, specifically a full width half maximum gaussian filter (Ashburner & Friston, 2001), harmonizes the noise across different datasets collected from different MRI scanners.

Methods

Data sources

This was an international study utilizing 17 different MRI scanners across eight cohorts in the United States, Germany, and Australia (including the University of Florida, Penn State Hershey Medical Center, Medical University Innsbruck, Northwestern University, University of Michigan, Parkinson's Progression Marker's Initiative, and 4 Repeat Tauopathy Neuroimaging Initiative). A total of 1002 participants were tested with 278 being healthy controls, 511 with PD, 84 with MSA, and 129 with PSP. There were 608 male and 394 female subjects with a total average age of 65.05. The average Movement Disorder Society Unified Parkinson's Disease Rating Scale part III (MDS-UPDRS III) scores per group reported controls with a score of 3.42, PD subjects with 30.15, MSA subjects with 51.13, and PSP subjects with 40.65 (Archer et al., 2019). This demonstrates that MSA is the diagnosis with the most severe motor deficits of those being tested.

Patients with Parkinsonism have all received a diagnosis from a movement disorder specialist which is the current gold standard. Control participants self-reported no history of neurological problems. All participants provided written informed consent which was approved by institutional review boards.

Data preprocessing and normalization

The preprocessing and normalization of data followed a previous study by Archer and colleagues (Archer et al., 2019). Custom MATLAB scripts were used to get free-water and free-water corrected fractional anisotropy (FA_T) images for the subjects, then each image was inspected to ensure that the entire brain was clearly visible and that artifacts did not exist. This included inspection for evidence of stroke, warping errors, and more.

Prior to nonlinear registration to Montreal Neurological Institute (MNI) space, a linear registration was performed. This was accomplished using Advanced Normalization Tools (ANTs) (Klein et al., 2009). A rigid transformation was applied to all the images to move them from subject space to MNI space where the template brain was the Human Connectome Project FA template. Linear registrations were followed with nonlinear warping (Symmetric

Normalization in ANTs). The nonlinear transform used the same mean Human Connectome Project FA template.

Regions of interest

All regions and tracts of interest were in MNI space. The region of interest template (Archer et al., 2018) included a total of 17 regions in the basal ganglia, midbrain, thalamus, cortex, and cerebellum. Forty-three white matter tracts were also analyzed using existing tractography templates, which include the sensorimotor area tract template (S-MATT) (Archer et al., 2018), transcallosal tractography template (TCATT), and a cerebellar white matter atlas (van Baarsen et al., 2016).

Machine learning approach

There were four conditions for this study based on the full width at half maximum (FWHM) gaussian filter that was applied to the image. The conditions were no filter, 2mm, 4mm, and 6mm. Scores from the MDS-UPDRS III, sex, and age for the PD, MSA, and PSP subjects were considered.

Diffusion-weighted MRI (dMRI) data for weighted sensitivity, specificity, and accuracy were analyzed. Each combination of variables was used in parts of a support vector machine (SVM) learning algorithm in Python. Participant data were randomly assigned to either training or validation sets and then split into five subgroups for five-fold cross-validation. This trains the algorithm and checks the generalizability of the system's performance.

Results

Data for participants with MSA/PSP vs PD and MSA vs PSP were compared. Training (80% of the data) and validation (20% of the data) were similar for the filter conditions, so this model could transfer to an independent data set. The training data of all four filter sizes for weighted sensitivity, specificity, and accuracy ranged from 0.82-0.94, 0.89-0.96, and 0.86-0.94 respectively. The validation sets had small ranges for weighted sensitivity (0.81-0.96), specificity (0.82-1.00), and accuracy (0.83-0.95) as well.

Figure 1 shows the weighted sensitivity, specificity, and accuracy with no filter applied in both training (A,B,C) and validation (D,E,F) sets. Figures 2, 3, and 4 show the same data with

the 2mm, 4mm, and 6mm filters applied respectively. Figure 5 has the weighted accuracy for all the filter conditions in both training and validation sets.

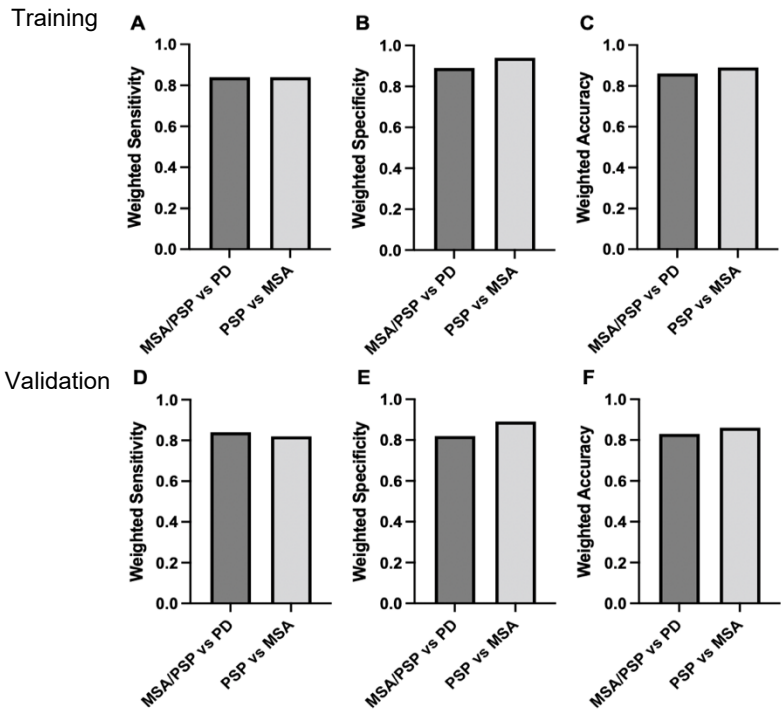

Figure 1. Weighted sensitivity, specificity, and accuracy with no filter applied to the imaging for both training (A,B,C) and validation (D,E,F) data sets.

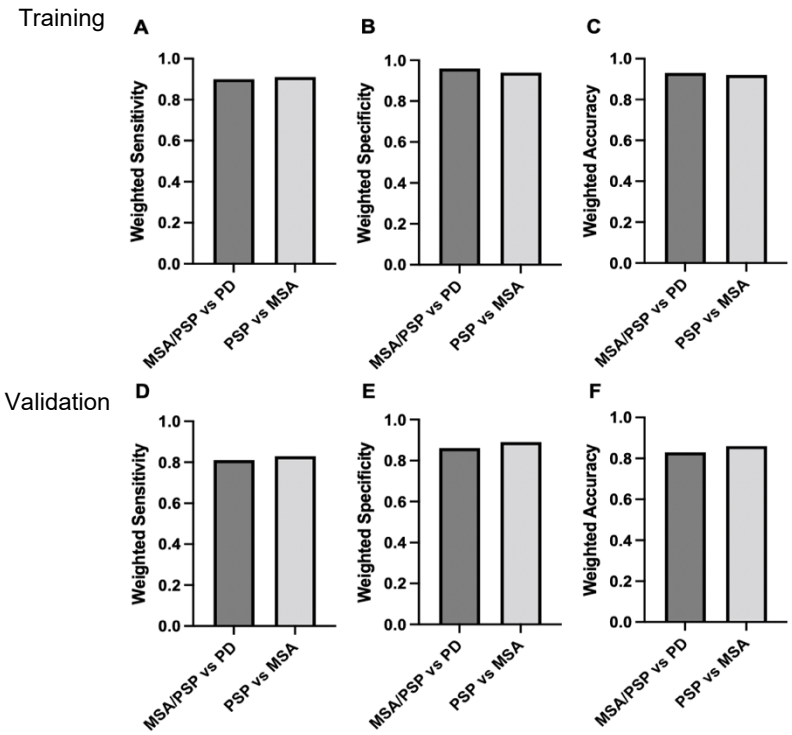

Figure 2. Weighted sensitivity, specificity, and accuracy with the 2mm filter applied to the imaging for both training (A,B,C) and validation (D,E,F) data sets.

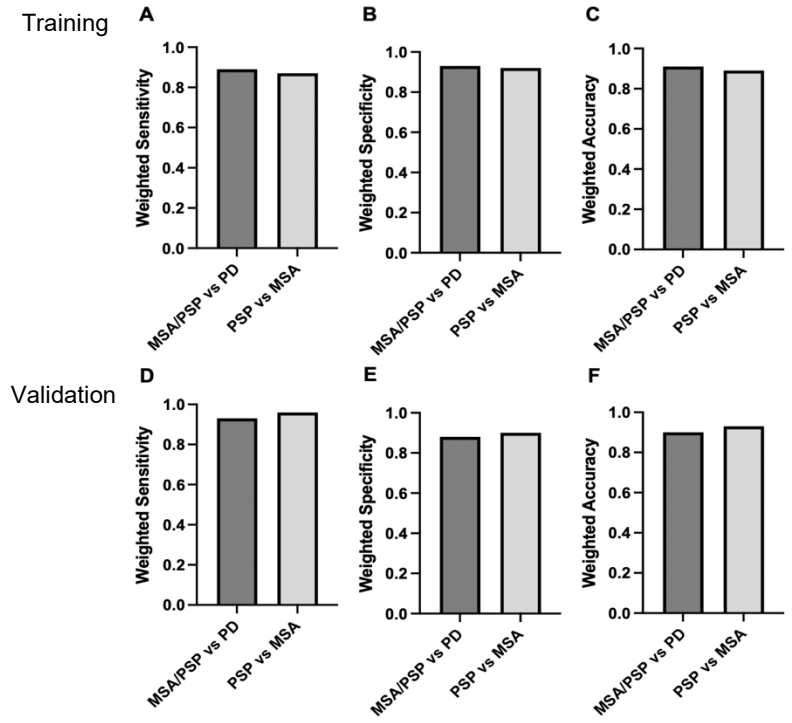

Figure 3. Weighted sensitivity, specificity, and accuracy with the 4mm filter applied to the imaging for both training (A,B,C) and validation (D,E,F) data sets.

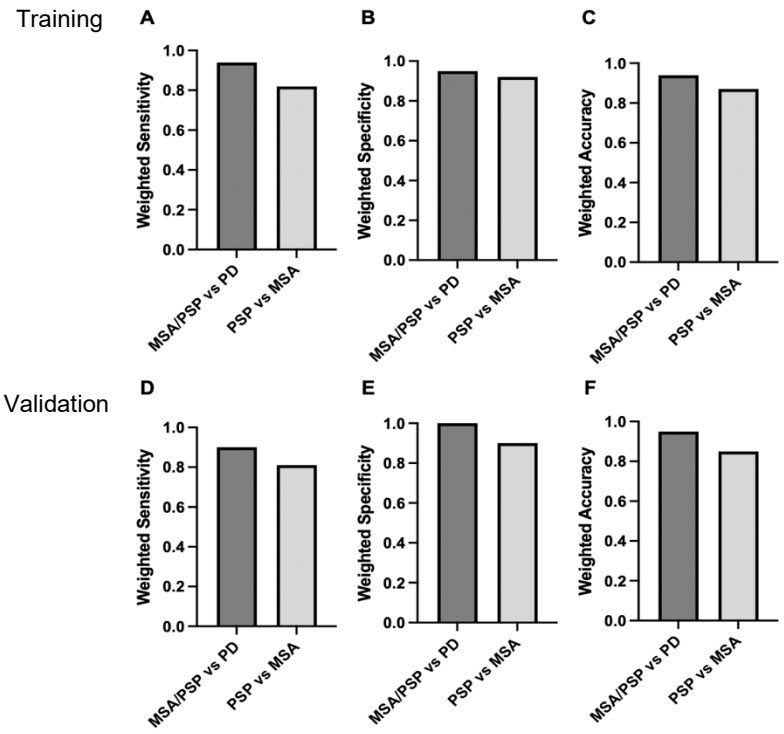

Figure 4. Weighted sensitivity, specificity, and accuracy with the 6mm filter applied to the imaging for both training (A,B,C) and validation (D,E,F) data sets.

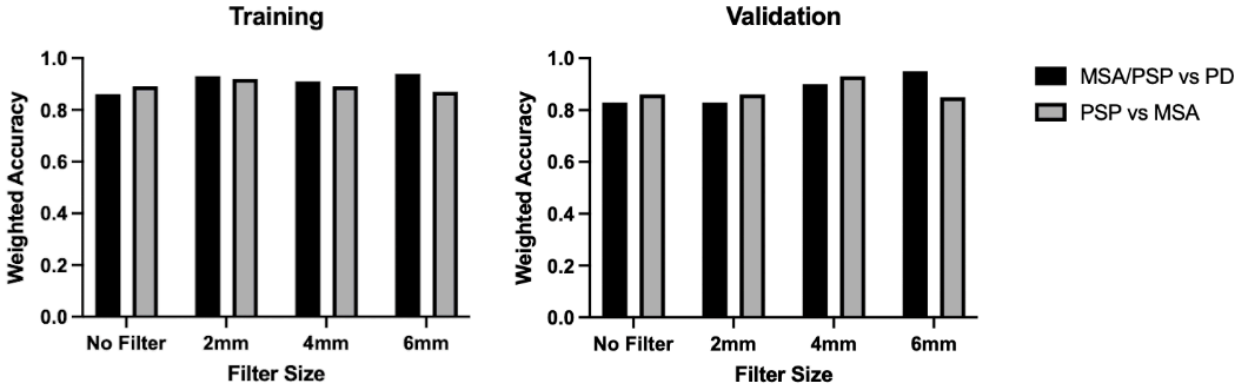

Figure 5. The weighted accuracy for training and validation sets of all four filter conditions.

Discussion

The key finding of this study was that 2mm, 4mm, and 6mm filters did not impact the harmonization of diffusion MRI data. It does not seem that there was a difference in the machine learning approach outcome whether no filter or the 2mm, 4mm, or 6mm filters were applied.

This was shown by similar results (ie. weighted sensitivity, specificity, and accuracy data) for each of the tested filter conditions compared to a control group with no filter applied. The small range of results for both training (less than 0.1) and validation (less than 0.2) groups indicate that the SVM was working well and could be generalized to an independent data set. The implications of this are twofold. First, the differences in MRI scanners need to be considered when looking at data collected from multiple sites. Second, there is still a gap in knowledge about how to accommodate these differences.

There are a few reasons why the filtering method may not impact the data. First, the free-water algorithm already included some normalization processes because the values are between 0 - 1. This algorithm estimates the fractional volume of the extracellular compartment and assigns the fractional volume a value between 0 - 1. It is possible that adding a spatial filter to this algorithm does not improve the multisite nature of the data. Second, when an image is acquired in k-space there are reconstruction algorithms that transform the data to image space which is typically the output of an MRI scanner. The reconstruction algorithm is proprietary information from the company that manufactures the MRI scanner, and thus there can be some filtering that already occurs without the researcher's knowledge. Thus, adding an additional spatial filter may not have an impact if the reconstruction algorithm has a filter built-in. Third, it is possible that the filtering procedure tested here is not sufficient and that another method is needed to improve the harmonization of the imaging data across sites.

A multi-site study using the same scanner software version (VE11C) and the same pulse sequences across sites found that free-water imaging has promise in distinguishing between types of Parkinsonism (Mitchell et al., 2019). The study by Mitchell and colleagues did not find a site effect or group-by-site interactions suggesting that free-water imaging performed on matched scanners is robust. Another multi-site study included data collected across at least 8 different sites to look at progression markers of Parkinson's disease over time. The researchers used a two-compartment model to calculate the signal attenuation before standardizing it to MNI space and found that site did not have an effect on their results (Burciu et al., 2017). Finding a way to harmonize inter-site variability is an ongoing problem due to different scanner manufacturers, software versions, and pulse sequences. The ComBat method (Johnson et al., 2007) is one example of a procedure that has been used for harmonization of diffusion and other imaging data

from across sites. ComBat is a tool that limits inter-site variability by removing batch effects, but it does require considerable data from a site to model and covary the site differences. This method has less use in a clinical setting when substantial a priori data does not exist. Although ComBat is typically used for genomic data, it has been reportedly used on diffusion tensor MRI imaging successfully (Fortin et al., 2017). Additionally, another study used FWHM gaussian filters of sizes 2-30mm (with increments of 2mm) to harmonize functional MRI data and reported an ideal filter size of 8mm (Mikl et al., 2008). However, the sample size (20 participants) was notably smaller than that of this study, only included healthy adults, and was performed at a single site.

In conclusion, the present study has determined that using a 2mm, 4mm, or 6mm full width half maximum spatial filter is not sufficient to harmonize data from different MRI scanners. Future work could look at the impact of 8mm or larger FWHM filters on diffusion MRI data. Since there was no effect from the filter sizes examined in this study, it may be possible that an optimal filter not tested here would optimize harmonization in a future study.

Acknowledgements

I would like to extend a huge thank you to my mentor Dr. David Vaillancourt along with Dr. Wei-en Wang, Dr. David Arpin, and the entire Laboratory for Rehabilitation Neuroscience at the University of Florida for all of the support during this project.

References

Archer, D. B., Bricker, J. T., Chu, W. T., Burciu, R. G., Mccracken, J. L., Lai, S., Coombes, S. A., Fang, R., Barmpoutis, A., Corcos, D. M., Kurani, A. S., Mitchell, T., Black, M. L., Herschel, E., Simuni, T., Parrish, T. B., Comella, C., Xie, T., Seppi, K., … Vaillancourt, D. E. (2019). Development and Validation of the Automated Imaging Differentiation in Parkinsonism (AID-P): A Multi-Site Machine Learning Study. *The Lancet. Digital Health*, *1*(5), e222–e231. https://doi.org/10.1016/s2589-7500(19)30105-0

Archer, D. B., Vaillancourt, D. E., & Coombes, S. A. (2018). A Template and Probabilistic Atlas of the Human Sensorimotor Tracts using Diffusion MRI. *Cerebral Cortex (New York, NY)*, *28*(5), 1685–1699. https://doi.org/10.1093/cercor/bhx066

Ashburner, J., & Friston, K. J. (2001). Why voxel-based morphometry should be used. *NeuroImage*, *14*(6), 1238–1243. https://doi.org/10.1006/nimg.2001.0961

Burciu, R. G., Ofori, E., Archer, D. B., Wu, S. S., Pasternak, O., McFarland, N. R., Okun, M. S., & Vaillancourt, D. E. (2017). Progression marker of Parkinson's disease: A 4-year multi-site

imaging study. *Brain: A Journal of Neurology*, *140*(8), 2183–2192. https://doi.org/10.1093/brain/awx146

Cox, R. W. (1996). AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers and Biomedical Research, an International Journal*, *29*(3), 162–173. https://doi.org/10.1006/cbmr.1996.0014

Fortin, J.-P., Parker, D., Tunç, B., Watanabe, T., Elliott, M. A., Ruparel, K., Roalf, D. R., Satterthwaite, T. D., Gur, R. C., Gur, R. E., Schultz, R. T., Verma, R., & Shinohara, R. T. (2017). Harmonization of multi-site diffusion tensor imaging data. *NeuroImage*, *161*, 149–170. https://doi.org/10.1016/j.neuroimage.2017.08.047

Hughes, A. J., Daniel, S. E., Ben-Shlomo, Y., & Lees, A. J. (2002). The accuracy of diagnosis of parkinsonian syndromes in a specialist movement disorder service. *Brain: A Journal of Neurology*, *125*(Pt 4), 861–870. https://doi.org/10.1093/brain/awf080

Johnson, W. E., Li, C., & Rabinovic, A. (2007). Adjusting batch effects in microarray expression data using empirical Bayes methods. *Biostatistics (Oxford, England)*, *8*(1), 118–127. https://doi.org/10.1093/biostatistics/kxj037

Klein, A., Andersson, J., Ardekani, B. A., Ashburner, J., Avants, B., Chiang, M.-C., Christensen, G. E., Collins, D. L., Gee, J., Hellier, P., Song, J. H., Jenkinson, M., Lepage, C., Rueckert, D., Thompson, P., Vercauteren, T., Woods, R. P., Mann, J. J., & Parsey, R. V. (2009). Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *NeuroImage*, *46*(3), 786–802. https://doi.org/10.1016/j.neuroimage.2008.12.037

Mikl, M., Marecek, R., Hlustík, P., Pavlicová, M., Drastich, A., Chlebus, P., Brázdil, M., & Krupa, P. (2008). Effects of spatial smoothing on fMRI group inferences. *Magnetic Resonance Imaging*, *26*(4), 490–503. https://doi.org/10.1016/j.mri.2007.08.006

Mitchell, T., Archer, D. B., Chu, W. T., Coombes, S. A., Lai, S., Wilkes, B. J., McFarland, N. R., Okun, M. S., Black, M. L., Herschel, E., Simuni, T., Comella, C., Xie, T., Li, H., Parrish, T. B., Kurani, A. S., Corcos, D. M., & Vaillancourt, D. E. (2019). Neurite orientation dispersion and density imaging (NODDI) and free-water imaging in Parkinsonism. *Human Brain Mapping*, *40*(17), 5094–5107. https://doi.org/10.1002/hbm.24760

Plitman, E., Bussy, A., Valiquette, V., Salaciak, A., Patel, R., Cupo, L., Béland, M.-L., Tullo, S., Tardif, C. L., Rajah, M. N., Near, J., Devenyi, G. A., & Chakravarty, M. M. (2021). The impact of the Siemens Tim Trio to Prisma upgrade and the addition of volumetric navigators on cortical thickness, structure volume, and 1H-MRS indices: An MRI reliability study with implications for longitudinal study designs. *NeuroImage*, *238*, 118172. https://doi.org/10.1016/j.neuroimage.2021.118172

Ramezanpour, S., Jalilianhasanpour, R., & Yousem, D. M. (2019). Vendors Used in CT and MRI Neuroradiology Research. *American Journal of Neuroradiology*. https://doi.org/10.3174/ajnr.A6180

van Baarsen, K. M., Kleinnijenhuis, M., Jbabdi, S., Sotiropoulos, S. N., Grotenhuis, J. A., & van Cappellen van Walsum, A. M. (2016). A probabilistic atlas of the cerebellar white matter. *NeuroImage*, *124*(Pt A), 724–732. https://doi.org/10.1016/j.neuroimage.2015.09.014