Whole-brain R1 predicts manganese exposure and biological effects in welders

David A. Edmondson[1,2,3], **Chien-Lin Yeh**[1,2], **Sébastien Hélie**[4], **Ulrike Dydak**[1,2]

[1.]School of Health Sciences, Purdue University, West Lafayette, IN

[2.]Department of Radiology and Imaging Sciences, Indiana University School of Medicine, Indianapolis, IN

[3.]Imaging Research Center, Cincinnati Children's Hospital Medical Center, Cincinnati, OH

[4.]Department of Psychological Sciences, Purdue University, West Lafayette, IN

Abstract

Manganese (Mn) is a neurotoxicant that, due to its paramagnetic property, also functions as a magnetic resonance imaging (MRI) T1 contrast agent. Previous studies in Mn toxicity have shown that Mn accumulates in the brain, which may lead to parkinsonian symptoms. In this article, we trained support vector machines (SVM) using whole-brain R1 (R1 = 1/T1) maps from 57 welders and 32 controls to classify subjects based on their air Mn concentration ($[Mn]_{Air}$), Mn brain accumulation ($ExMn_{Brain}$), gross motor dysfunction (UPDRS), thalamic GABA concentration ($GABA_{Thal}$), and total years welding. R1 was highly predictive of $[Mn]_{Air}$ above a threshold of 0.20 mg/m^3 with an accuracy of 88.8% and recall of 88.9%. R1 was also predictive of subjects with $GABA_{Thal}$ having less than or equal to 2.6 mM with an accuracy of 82% and recall of 78.9%. Finally, we used an SVM to predict age as a method of verifying that the results could be attributed to Mn exposure. We found that R1 was predictive of age below 48 years of age with accuracies ranging between 75% and 82% with recall between 94.7% and 76.9% but was not predictive above 48 years of age. Together, this suggests that lower levels of exposure (< 0.20 mg/m^3 and < 18 years of welding on the job) do not produce discernable signatures whereas higher air exposures and subjects with more total years welding produce signatures in the brain that are readily identifiable using SVM.

Keywords

Manganese; Machine Learning; R1; Occupational Exposure; GABA; Welding; Magnetic resonance imaging; neuroimaging

2. INTRODUCTION

Manganese (Mn) is a neurotoxic metal that is a common constituent of welding fume which thousands of welders inhale every day in the manufacturing industry. Symptoms of excessive Mn exposure include bradykinesia, dystonia, postural tremor and rigidity (Guilarte and Gonzales, 2015; Racette et al., 2012; Tsuboi et al., 2007). These symptoms do not recede upon cessation of exposure (Guilarte, 2013), nor are symptoms alleviated with the standard treatment for Parkinson's disease, levodopa (Olanow, 2004). Therefore, preventing symptoms is the primary goal when it comes to risk assessment. The primary method of risk assessment comes from exposure monitoring through air sampling where 0.02 mg/m^3 was recommended by the American Conference of Governmental Industrial Hygienists as the threshold limit value for Mn (Ward et al., 2017). However, much debate continues to surround at what level the limit should be placed. While having a low limit is a conservative method of minimizing risk, it is unknown whether this level is low enough to prevent symptoms from eventually occurring or too low thus creating financial hardship for industry with little to no benefit to workers. Bailey et al. (2018) suggest a possible occupational exposure level of 0.1 up to 0.14 mg/m^3 for respirable Mn based on studies that they considered adequately accounted for exposure levels in the workplace (Bailey et al., 2018).

Since Mn toxicity symptoms are neurological, understanding how Mn affects the brain is of vital importance. Therefore, knowing how much Mn accumulates in the brain after a given exposure would be ideal. When Mn is inhaled into the lungs, it is readily absorbed into the bloodstream (Leggett, 2011) and into the pulmonary veins. Inhaled Mn therefore bypasses the first pass effect of the liver and thus the majority of inhaled Mn is available for transport into the brain. There are two primary barriers that Mn must then pass before entering the brain, the blood brain barrier (BBB) and the blood CSF (cerebral spinal fluid) barrier (BCB). Mn can cross the BBB while bound to transferrin via transferrin transporters and ZIP8 transporters (Aschner and Aschner, 1991). Mn can also penetrate the BCB via divalent metal transporter 1 (DMT1) (Yokel, 2009). Mn competes with Iron (Fe) for DMT1, but it has been found that there is preferential uptake for Mn (versus Fe) in the choroid plexus, site of the BCB (Bornhorst et al., 2012).

Once Mn crosses a barrier into the brain, it disperses throughout via a variety of physical and molecular interactions (Bock et al., 2008). It can be taken into glial cells through transporters and into neurons through Ca2+ voltage-gated transporters (Bedenk et al., 2018; Leuze et al., 2012). Once in cells, Mn tends to be taken into endosomes or the mitochondria (Borg and Cotzias, 1958). Mn can also travel anterograde along axons, a feature that has been utilized to great effect for neuronal tracing using Mn-enhanced MRI, or MEMRI. Similarly, we can use the MRI relaxation rate R1, which is proportional to Mn in the region, to measure Mn uptake, accumulation, and dispersion in the brain (Yeh et al., 2016). This has been done in studies involving rats (Lehallier et al., 2012), monkeys (Dorman, 2006; Park et al., 2007; Shinotoh et al., 1995), and humans (Edmondson et al., 2019; Lee et al., 2018, 2015; Lewis et al., 2016; Long et al., 2015; Ma et al., 2018), but understanding the relationship between exposure and uptake is still ongoing. For instance, in monkeys, there was a strong correlation and linear relationship between R1 and Mn accumulation in the various regions of the brain (Dorman, 2006). However, in humans, R1 has been found to be

higher in subjects with higher exposure versus controls with lower exposure, but the relationship between R1 and exposure is non-linear relationship (Lee et al., 2015; Ma et al., 2018) and may suggest an exposure threshold below which Mn may not impact R1, in other words: a detection limit. Together, this all may suggest that a direct measure of Mn content in the brain, rather than exposure, may be a better metric for assessing risk. To do this, a biological model could be employed to estimate Mn in the brain. Biological modeling is used to simulate how a specific xenobiotic will disperse throughout the body. By accounting for this, a well-designed model could then estimate how much Mn will be in a part of the body at a given time after exposure.

Mn accumulation in the brain may be the direct cause for characteristic motor dysfunction in Mn-induced parkinsonism. We have shown that neurotransmitter concentration, specifically γ-aminobutyric acid (GABA), the primary inhibitory neurotransmitter in the central nervous system, is altered with Mn exposure. Specifically, GABA is higher in the thalamus in groups exposed to higher levels of Mn (Ma et al., 2018) and changes proportionately with levels of Mn exposure (Edmondson et al., 2019).

Because Mn can disperse throughout the brain, we took a whole-brain approach to measuring R1. In this study, we employed a support vector machine (SVM) pipeline utilizing acquired whole-brain R1 MRI data. This pipeline generates SVM models from R1 data which can be used to predict classes within different target variables (variables that we wish to predict classes within). Each target variable was then binarized based on different value thresholds. This method is advantageous because rather than focusing on *a priori* regions within the brain, we can use data from the whole brain to find identifiable patterns that are predictive of target variables. These patterns may then provide further insight into the effects of low-level Mn exposure that would otherwise be missed.

We hypothesize that whole-brain R1 will be predictive of air Mn exposure ($[Mn]_{Air}$, the amount of respirable Mn particulate in the air at the welder's workplace measured using personal air sampling equipment) and excess brain Mn ($ExMn_{Brain}$, the amount of Mn accumulated in the brain due to exposure calculated from $[Mn]_{Air}$, hours worked per day, breaths per hour, and amount of air breathed). Additionally, we hypothesize that whole-brain R1 can predict biological outcomes such as thalamic GABA ($GABA_{Thal}$) concentration and motor dysfunction, as measured by the Unified Parkinson's Disease Rating Scale (UPDRS). Finally, to account for commonly used covariates, we will also focus on total years welding and age as target variables in order to compare performance of these models to other models predicting Mn exposure and biological outcome models.

3. METHODS

3.1. Data Acquisition

The hypotheses were tested by re-analyzing data previously collected in our laboratory at Purdue University (Ma et al., 2018; Ward et al., 2017). Specifically, eighty-nine subjects (57 welders, 32 controls) had been recruited in previous studies directly from a local manufacturing plant. All subjects were male. All subjects gave informed consent at that time. The methods used for data acquisition (Relaxometry, Spectroscopy, UPDRS, and

Exposure Assessment) from these studies is briefly summarized below. Details can be obtained in the original articles cited above. Methods specific to this study are elaborated below.

3.2. Whole-Brain R1 Mapping

Magnetic resonance imaging (MRI) was performed on a 3T GE Signa MRI scanner using an 8-channel head coil. A high-resolution 3D T1-weighted structural image (FSPGR, TR/TE = 6.26/2.67 ms, resolution: 1 x 1 x 1 mm^3) was taken for use in segmentation. Segmentation of the brain was performed using Freesurfer (surfer.nmr.mgh.harvard.edu) which parsed the brain into 192 separate regions of interest (ROI) as masks. Each ROI consists of volume pixels, or voxels, which denote a 3D volume in the images. Two of the ROIs included unlabeled voxels and were thus removed from analysis, leaving 190 ROIs. To produce T1 relaxation time maps, we acquired a spoiled gradient echo imaging sequence (SPGR, TR/TE = 6.36/1.76 ms, resolution: 1 x 1 x 2 mm^3) using two different flip angles (3°, 17°). A 3D T1 relaxation time map was generated using the images from these two flip angles (Christensen et al., 1974). To correct for inhomogeneity in the radiofrequency field, we acquired an inversion recovery SPGR with the same parameters as the SPGR images (Deoni, 2007). R1 in each voxel was then calculated as inverse T1 relaxation time where R1 = 1/T1. We used the segmented masks to extract R1 values for each voxel within each ROI. To best describe the distribution of R1 within each ROI, five statistics were selected: median, variance, skew, 10-percentile, and 90-percentile.

3.3. Target Variables

3.3.1. Air Mn Concentration—Air Mn concentration ($[Mn]_{Air}$) in units of mg/m^3 was determined by performing personal air sampling of respirable Mn on subjects in the workplace, as detailed in Ma et al., 2018 and Ward et al., 2017. Personal air sampling was obtained over 8-hour workdays and averaged to create an 8-hr time-weighted average of $[Mn]_{Air}$. $[Mn]_{Air}$ only consisted of the respirable Mn component and was averaged per worksite in the factory. Air filters on SKC aluminum cyclones with a cut-point of 4 μm were used for sampling. Filters were placed inside the welding helmet for welders and over the shoulder for control subjects.

3.3.2. Excess Brain Mn—A biokinetic model (Figure 1) was used to determine the amount of Mn accumulation due to occupational exposure in the brain above normal amounts, termed *Excess Brain Mn* ($ExMn_{Brain}$). The model was based on a previously published model (Leggett, 2011) and implemented with in-house Python 3 scripts using the SciPy module for calculating ordinary differential equations, *odeint*, with $[Mn]_{Air}$ as the input. To determine how much Mn was inhaled each day, the $[Mn]_{Air}$ was used to calculate the total amount of Mn inhaled over the 8-hour workday using mean values for breaths per hour and amount of air breathed (U.S. EPA, 2011). To simplify the model, we assumed that all respirable Mn that entered the lungs was absorbed into the blood. Hence, this amount of Mn was assumed to be 100% taken into the blood compartment. Using each subject's work-history, a timeline of Mn accumulation was created with the final data point representing the amount of Mn in the brain on the day of scanning. We calculated Mn intake based on an 8-hr per day, 7-day workweek because the welders in our cohort were known to work more than

the typical 40-hour work week, sometimes up to 7 days a week and 10 hours a day. These values were then used as inputs to the biokinetic model. Specifically, excess brain Mn was then extracted according to the differential equation:

$$\frac{dN}{dt} = \lambda_{Blood \rightarrow Brain} * N_{Blood} - \lambda_{Brain \rightarrow Blood} * N_{Brain}$$

where N is the amount of Mn in the compartment and λ is the rate of transfer from one compartment to the other. For any given time, this equation allows us to calculate the amount of Mn in the brain accounting for the amount of Mn entering the brain ($\lambda_{Blood \rightarrow Brain} * N_{Blood}$) and the amount leaving ($-\lambda_{Brain \rightarrow Blood} * N_{Brain}$). Based on the Leggett 2011 model, we used $\lambda_{Blood \rightarrow Brain} = 1$ d^{-1} and $\lambda_{Brain \rightarrow Blood} = 0.00462$ d^{-1}. These values correspond to half-lives of 0.693 days and 150 days, respectively.

3.3.3.   Thalamic GABA—Prior studies (Edmondson et al., 2019; Long et al., 2014; Ma et al., 2018) have shown that thalamic GABA (GABA$_{Thal}$) concentrations are correlated with Mn exposure. Therefore, we included GABA$_{Thal}$ as a target variable. GABA$_{Thal}$ concentrations were obtained from magnetic resonance spectroscopy (MRS) performed in previous studies. MEGA-PRESS localization (TR/TE = 2000/68 ms, 256 Averages) (Mullins et al., 2014) was used on a volume of interest (VOI) placed over the right thalamus (25mm x 30mm x 25mm) in each subject. Spectra were quantified using LCModel V6.3-1B (Provencher, 1993) with a basis set generated by density matrix simulation using GABA coupling constants from Kaiser et al. (2007). Spectra were phase and frequency corrected using a water reference. GABA concentrations are CSF-corrected and are reported in units of mM. To obtain CSF-corrected values, segmentation was performed using SPM8 (Wellcome Department of Imaging Neuroscience, London, United Kingdom). The VOI was segmented into its tissue components, specifically white matter, gray matter, and CSF.

3.3.4.   UPDRS—One method to assess motor dysfunction is with the Unified Parkinson's Disease Rating Scale Part III (UPDRS) (Goetz et al., 2008), a scale commonly used to measure Parkinson's disease progression. To assess whether R1 can predict motor dysfunction, we included UPDRS as a target variable. During participation in a prior study, a certified neurologist tested each subject and rated them based on their motor performance in tasks measuring symptoms such as tremor and bradykinesia. Higher scores represented worse performances.

3.3.5.   Age and Total Years Welding—During their participation in prior studies, both welders and controls filled out a comprehensive work history which included their age as well as any welding experience they had. Values for age and total years welding came from these self-reported answers. Welders and controls were classified as such based on their current job position and whether their occupation required any welding or not. Total welding years is a proxy for lifetime exposure while age is included as a target variable to account for variability in our dbata not due to Mn exposure.

3.3.6.   Correlation Tests—To assess directional relationships and collinearity between target variables, correlation tests between all target variables were performed using R.

Spearman's correlation was used for comparisons with ordinal variables (UPDRS Score) while Pearson's correlation was used for comparisons between continuous variables (GABA$_{Thal}$, [Mn]$_{Air}$, age, years welding, and ExMn$_{Brain}$). We adjusted for multiple comparisons using Benjamini-Hochberg correction, or the false discovery rate (FDR). An adjusted p-value of 0.05 was considered statistically significant.

3.4. Machine Learning

In this study, target variables included: air Mn exposure ([Mn]$_{Air}$), excess brain Mn (ExMn$_{Brain}$), total welding years, age, thalamic GABA (GABA$_{Thal}$), and UPDRS score. For each target variable, models were run across thresholds that were equally distributed and selected based on the range of values while ensuring that at least one subject was in the minority group. Performance of each model was then evaluated using leave one-out cross-validation.

3.4.1. Model Development—A support vector machine (SVM) was created for each target variable, threshold, and statistic. Briefly, SVM models estimate a hyperplane boundary created by maximizing the distance between the data points that are nearest to the hyperplane (support vectors) and the hyperplane. Due to the relatively low sample size, a linear kernel was chosen to avoid overfitting the data and to ensure results remain readily interpretable.

All models were built using the Python machine learning package, Scikit-Learn (Pedregosa et al., 2012). First, missing R1 values due to improper segmentation were imputed with median values using *SimpleImputer* followed by standardization using *StandardScaler*. This method removes the mean and scales to unit variance on each feature (ROI) where each sample score is calculated as a *z*-score:

$$z = \frac{x - u}{s}$$

where *x* is the sample, *s* is the standard deviation of all samples, and *u* is the mean of all samples.

Principle component analysis (PCA) was then performed on the standardized R1 values in each of the 190 brain ROIs to reduce dimensionality and prevent random and structured noise from biasing the results. PCA is performed by calculating the covariance matrix of the data and then extracting the eigenvalues and eigenvectors that explain up to a set threshold of variance in the data. Each eigenvector contains the weights to produce a linear combination of each ROI while the normalized eigenvalues are the percent of total variance explained by the eigenvector. For this study, we chose 90% variance as the cut-off for number of PCs used to then transform the R1 data, removing much of the noise in the data.

SVM was implemented using the Scikit-Learn module *LinearSVC* which utilizes the LIBLINEAR library (Fan, Rong-En, 2008). Without appropriate precautions, an SVM model will be biased towards the majority class to maximize accuracy of the model potentially leading to always predicting the majority class. To measure how far away from

equal distribution of classes, a 95% confidence interval was calculated using a test of proportions. We accounted for unequal distribution of groups by implementing a penalty term, *C*, that adjusts for the imbalance of the classes:

$$C = \frac{n}{(m * q)}$$

where *n* is the number of samples, *m* is the number of classes and *q* is the number of subjects in the class.

3.4.2. Model Testing & Scoring—For each SVM model created for a target variable, statistic, and threshold, leave one-out cross validation (LOO-CV) was used to measure model performance. LOO-CV is performed by creating the model with N-1 subjects, and then using the model to predict the left-out subject. Classes were determined by the threshold where Class 0 was less than or equal to the threshold and Class 1 was greater than the threshold. A confusion matrix for each model was created by calculating the true positives (TP), true negatives (TN), false positives (FP), and false negatives (FN). All models were then combined in a process called "bagging" where multiple machine learning models are combined to improve overall performance. In this case, class predictions from all five statistic SVM models are determined and then the overall prediction is chosen based on a majority rule. We call this our "combined" model.

SVM models were scored using CV accuracy $\left(\frac{TP+TN}{TP+TN+FP+FN}\right)$ and recall $\left(\frac{TP}{TP+FN}\right)$.

Accuracy was chosen because it accounts for all correct class 0 predictions, whereas the other metrics do not. Recall was chosen to therefore account for how well the model can identify class 1 subjects out of all class 1 subjects available. Superior model performance is considered from a combination of high accuracy with high recall. Accuracy depicts the total number of correct predictions and describes how correct the model is when accounting for all possible outcomes while recall focuses on the total number of correct predictions for class 1 out of all possible class 1 available (i.e. the proportion of subjects with Mn levels above threshold that are correctly identified) and only measures how well a model can identify one group, in our case, subjects with high Mn exposure. Finally, due to the inherent class imbalance with different thresholds used, we performed a test of proportions to assess the distribution of errors between FP and FN.

4. RESULTS AND DISCUSSION

4.1. Demographics

As detailed earlier, data for this study was acquired from previous studies. We used data from 52 welders and 37 controls. Welders and controls were approximately age-matched at 41 years old. On average, welders worked in environments with 0.14 mg/m^3 Mn whereas controls worked in environments with 0.004 mg/m^3. All target variables are summarized in Table 1.

4.2. Principle Component Analysis

We used PCA to reduce the number of features needed to perform later analyses and effectively eliminate noise in the data. With the assumption that variability in our R1 data is due to Mn deposition changing the R1 (Dorman, 2006), PCA provides insight as to which regions of the brain had the most variation across all subjects. PCA results are summarized in Table 2. Calculated statistics from R1 distributions in ROIs (median, 10-percentile, and 90-percentile) went from using all 190 ROIs as features to only needing less than 7% total features to represent 90% of the variance in the data.

The principle components (PCs) generated by PCA are linear combinations of each ROI with a weight assigned based on the amount of total data variability that can be explained by the ROI. Some ROIs were more prevalent amongst the different statistics than others within the top 3 PCs. The superior temporal sulcus was a heavily weighted feature in both the 10-percentile and median whereas the parietal-occipital sulcus was heavily weighing in the 10-percentile and 90-percentile. White matter, including the corpus collosum and cerebral white matter, were also prominently featured in the second PC for median, variance, skew, and 90-percentile.

Conceptually, higher Mn accumulation results in more voxels with high values of R1, leading to a shift in the voxel distribution of R1 to the right. Therefore, the structures with greater variance in R1 could be due to greater differences in Mn uptake in these regions across all subjects. Noticeably, the nuclei of the basal ganglia are absent, the region of the brain commonly mentioned as susceptible to Mn accumulation. However, R1 in the basal ganglia has been found either to not be associated with Mn exposure (Edmondson et al., 2019; Ma et al., 2018), or it is after a threshold level of exposure is taken into account (Lee et al., 2018, 2015). Therefore, at lower levels of exposure, the overall variability between subjects in the basal ganglia may be low, thus these regions not having high weightings in PCA. So, while the basal ganglia may be a susceptible region of the brain to Mn, it might not be the best location to obtain an exposure biomarker. Rather, in our study, we found cortical regions and white matter in the brain to have more variability across all subjects, thus possibly having more information about Mn accumulation in the brain.

4.3. Predicting Measures of Mn Exposure

We tested a series of SVM models to measure how well R1 predicts groups of different target variables: air Mn concentration ($[Mn]_{Air}$), excess brain Mn ($ExMn_{Brain}$), total years welding, age, UPDRS score, and thalamic GABA ($GABA_{Thal}$). All SVM model performances are summarized in Figure 2. SVM models with accuracy (A) and recall (R) greater than or equal to 75% are presented in Table 3.

A combined SVM model predicting whether a subject is welder or control was 59.6% accurate and had recall (56.1%) slightly better than chance. However, three of the individual statistic models were more accurate and had higher recall: skew (A: 68.5%, R: 71.2%), variance (A: 69.7%, R: 79.2%) and 90-percentile (A: 57.3%, R: 69.6%). The combined model likely underperformed due to the median model's performance (A: 40.4%, R: 55%). The higher performing recall shows that while the model was predicting welders

appropriately, the sufficiently lower accuracy suggests the model still had trouble differentiating between controls and welders, which may be due to the wide disparity in Mn exposure levels amongst welders.

In the models predicting $[Mn]_{Air}$, we see that Mn exposure levels may influence how accurate the models can be. Especially at higher exposure levels, the model performs remarkably well where the combined models predicting $[Mn]_{Air}$ performed best with thresholds between 0.20 and 0.24 mg/m^3. In this range, accuracy was 88.8% while recall ranged from 88.9% to 75%. While the classes were imbalanced at these higher threshold (Figure 3) the prediction errors were distributed evenly (Figure 4).

We hypothesized that R1 would be more predictive of an estimate of Mn accumulation in the brain using a biokinetic model, $ExMn_{Brain}$. Contrary to our hypothesis, the combined models predicting $ExMn_{Brain}$ did not outperform the models predicting $[Mn]_{Air}$. $ExMn_{Brain}$ models performed best with thresholds between 6.5 and 8.5 mg/m^3, however while accuracy approached 80%, recall hovered around 60%. Therefore, it could be concluded that air Mn exposure at levels around 0.20 mg/m^3 provide stronger, identifiable signatures of R1 that differentiate above and below that threshold. While the SVM model predicting $ExMn_{Brain}$ was relatively accurate with accuracies greater than 70% above levels of 4mg, recall was around 40%, suggesting that the model was reasonably good at identifying subjects below the threshold, but could only get 2 out of 5 subjects greater than the threshold.

According to the biokinetic model used in this study, if a new welder breathed in a time-weighted average of 0.20 mg/m^3 of Mn every day for 365 days, they would have approximately 5 mg $ExMn_{Brain}$, whereas a welder exposed to 0.16 mg/m^3 would have approximately 4 mg $ExMn_{Brain}$. Because the biokinetic model takes into account previous exposures, these levels vary, which may be why accuracies and recall are different at these levels between the two models.

However, the biokinetic model may not be detailed enough to cover all the complexities of how Mn disperses through the brain. Mn which is bound to a protein or stored in an endosome will impact R1 significantly less than Mn that is free (Troughton et al., 2004). Therefore, a biological model may need to take this into account and only calculate for the free Mn in the brain. A physiologically-based pharmacokinetic model (PBPK) has been developed that accounts for the different states of Mn, but only in select regions of the brain (Schroeter et al., 2011). Therefore, when performing this study, it was determined that this may not be representative of the whole brain and so we chose to use a less complicated model that considered the brain as one compartment.

Because other Mn toxicity studies using MRI have used some metric of time spent welding as a proxy for exposure (Lee et al., 2015; Lewis et al., 2016), we trained a model to see if we could predict total years welding as well. Amongst our welders, total years welding was not correlated with $[Mn]_{Air}$ (Pearson's r = 0.11, adj. p > 0.05) or $ExMn_{Brain}$ (Pearson's r = 0.24, adj. p > 0.05). However, the combined models predicting total years welding had high accuracy and recall at the later thresholds of 19 (A: 85.4%, R: 84.6%) and 20 years (A: 85.4%, R: 84.6%). Total years welding was strongly correlated with age in welders

(Pearson's r = 0.75, adj. p < 0.0001). The combined model targeting age performed well with accuracy approaching 80% with recall greater than 80% between thresholds of 42 and 48 years. Once the threshold reached 50, recall dropped drastically until it reached 0% at 56 years, meaning all errors were false negatives. This may be because the majority of the older subjects were welders (16 out of 21) with higher levels of cumulative Mn exposure that influenced R1 and thus affected the predictability of their ages. This was an opposite response compared to the total welding years models where accuracy and recall improved substantially as the threshold was raised to higher levels. This indicates that R1 may be a better predictor of age for subjects that do not have enough Mn exposure to differentiate them from others in the same age class. However, cumulative Mn exposure may impose variability that the age model cannot classify correctly.

4.4. Predicting Measures of Biological Effect

We also looked at how whole-brain R1 might be able to predict biological effects from Mn exposure, focusing on two endpoints: neurochemical changes and motor dysfunction. As stated before, thalamic GABA ($GABA_{Thal}$) was chosen based on previous studies indicating that GABA levels changed based on Mn exposure while UPDRS Scores were chosen because they are the standard for assessing motor dysfunction in parkinsonian patients.

The combined models targeting $GABA_{Thal}$ performed reasonably consistent across all thresholds with accuracies ranging between 63% and 81%, however a few of the thresholds performed remarkably strong. While the model appears to perform strongly at 0.8 mM, the groups were heavily weighted towards those above 0.8 mM and the model performed many Type I errors as depicted in figure 4. Nonetheless, the model did perform well at higher thresholds, specifically 2.1 mM (A: 80.9%, R: 79.2 %) and 2.6 mM (A: 82%, R: 78.9%). A careful look at the plot (Figure 2) shows that the model had consistent accuracy and recall between 2.0 mM and 2.6 mM. While at this range there were more subjects that fell below the threshold, the model did err consistently between the two groups.

UPDRS scores were correlated with age (Spearman's = 0.45, p = 0.002), so it is not surprising that these models performed similarly to the age models, where recall dropped to 0% as the higher thresholds were tested. However, unlike the age models, UPDRS did not have strong prediction values at the lower scores, with the exception of the initial thresholds of 0 and 1. At these thresholds, though, the classes were significantly unbalanced as were the prediction errors, suggesting that the high accuracy and recall were artificially inflated by class proportion. Therefore, in this cohort, R1 can safely be presumed to not be predictive of UPDRS scores.

5. CONCLUSION

The whole-brain approach to data analysis used in this study elucidated how Mn accumulates in the entire brain, including white matter, cortical regions and the basal ganglia. Our study shows that R1 may be a viable predictor for assessing current Mn exposure and the resulting effects in neurochemistry as shown by $GABA_{Thal}$. R1 was reasonably predictive of $GABA_{Thal}$ above chance across all levels of $GABA_{Thal}$. Because of the relationship between $GABA_{Thal}$ and $[Mn]_{Air}$ (Pearson's r = 0.38, p < 0.05), $GABA_{Thal}$

might be considered a candidate biomarker for determining whether someone had recent Mn exposure. However, a relatively low correlation does not readily suggest predictability of $GABA_{Thal}$ from $[Mn]_{Air}$, or vice-versa. Future work may also want to look at using similar models to predict other symptomatic outcomes, such as cognitive and fine motor dysfunction.

We have also shown that whole-brain R1 might only predict exposures greater than 0.20 mg/m^3, which is about 10x higher than the current threshold limit value of 0.02 mg/m^3, suggesting that R1 may not be sensitive enough to predict changes in the brain at lower exposure levels. We also showed that there is a divergence of predictability between age and total years welding. Even though both target variables were correlated with one another, R1 was more predictive of welders with more total years welding than age. This study shows that there is a difference between older subjects and older subjects with many years of welding experience. Therefore, it is likely Mn alters the brain in a signature way that is identified using machine learning.

As discussed earlier, R1 did not sufficiently predict $ExMn_{Brain}$ as calculated from the biokinetic model used in this study, which may be due to our model's incomplete accounting of bound and free Mn. Therefore, further efforts should be made using a model that incorporates these states. However, current models do not take into account cortical regions of the brain, which we have shown are predictive of $[Mn]_{Air}$. So, while the Schroeter et al. 2011 model would be ideal, it may need to be modified to include other regions of the brain. Additionally, while not tested here, it is possible that whole-brain R1 may be able to predict the amount of Mn accumulation in one specific location in the brain, as could be simulated in the Schroeter et al. 2011 model.

While our study was largely successful, it was also limited by a couple factors. First, all subjects in the study cohort were male. While this is representative of the locations where members of the cohort worked, this limits generalizability to the general population. Second, for a study using machine learning, while 89 subjects are enough for a reliable model, more subjects would potentially allow for more generalizable models. Additionally, due to the number of subjects we had, we were limited by the machine learning algorithms available for use thus potentially preventing us from achieving higher model accuracy and recall.

In conclusion, this study employed machine learning to assess the effects of Mn on the whole brain and could be easily adapted to measuring the effects of other toxicants. Even though the imaging preprocessing tools we used are brain-specific (e.g. freesurfer), other tools could be used to assess the effects of toxicants on the liver using a similar data analysis method we used here. However, as can be seen in our results, groups can become significantly unequal. Because sample sizes are much smaller at higher thresholds, more data is needed to further improve these models.

ACKNOWLEDGEMENTS

7. REFERENCES

Aschner M, Aschner JL, 1991 Manganese Neurotoxicity : Cellular Effects and Blood-Brain Barrier Transport. Neurosci. Biobehav. Rev 15, 333–340.

Bailey LA, Kerper LE, Goodman JE, 2018 Derivation of an occupational exposure level for manganese in welding fumes. Neurotoxicology 64, 166–176. 10.1016/j.neuro.2017.06.009 [PubMed: 28624528]

Bedenk BT, Almeida-Corrêa S, Jurik A, Dedic N, Grünecker B, Genewsky AJ, Kaltwasser SF, Riebe CJ, Deussing JM, Czisch M, Wotjak CT, 2018 Mn2+dynamics in manganese-enhanced MRI (MEMRI): Cav1.2 channel-mediated uptake and preferential accumulation in projection terminals. Neuroimage 169, 374–382. 10.1016/j.neuroimage.2017.12.054 [PubMed: 29277401]

Bock NA, Paiva FF, Nascimento GC, Newman JD, Silva AC, 2008 Cerebrospinal fluid to brain transport of manganese in a non-human primate revealed by MRI. Brain Res 1198, 160–170. 10.1016/j.brainres.2007.12.065 [PubMed: 18243167]

Borg DC, Cotzias GC, 1958 Manganese Metabolism in Man: Rapid Exchange of MN56 with Tissue as Demonstrated by Blood Clearance and Liver Uptake. J. Clin. Invest 37, 1269–1278. 10.1172/JCI103714 [PubMed: 13575525]

Bornhorst J, Wehe C. a, Hüwel S, Karst U, Galla H-J, Schwerdtle T, 2012 Impact of manganese on and transfer across blood-brain and blood-cerebrospinal fluid barrier in vitro. J. Biol. Chem 287, 17140–51. 10.1074/jbc.M112.344093 [PubMed: 22457347]

Christensen KA, Grant DM, Schulman EM, Walling C, 1974 Optimal determination of relaxation times of fourier transform nuclear magnetic resonance. Determination of spin-lattice relaxation times in chemically polarized species. J. Phys. Chem 78, 1971–1977. 10.1021/j100612a022

Deoni SCL, 2007 High-resolution T1 mapping of the brain at 3T with driven equilibrium single pulse observation of T1 with high-speed incorporation of RF field inhomogeneities (DESPOT1-HIFI). J. Magn. Reson. Imaging 26, 1106–1111. 10.1002/jmri.21130 [PubMed: 17896356]

Dorman DC, 2006 Correlation of Brain Magnetic Resonance Imaging Changes with Pallidal Manganese Concentrations in Rhesus Monkeys Following Subchronic Manganese Inhalation. Toxicol. Sci 92, 219–227. 10.1093/toxsci/kfj209 [PubMed: 16638924]

Edmondson DA, Ma RE, Yeh C-L, Ward E, Snyder S, Azizi E, Zauber SE, Wells EM, Dydak U, 2019 Reversibility of neuroimaging markers influenced by lifetime occupational manganese exposure. Toxicol. Sci. 1, 1–13. 10.1093/toxsci/kfz174

Fan Rong-En, et al., 2008 LIBLINEAR: A library for large linear classification. J. Mach. Learn. Res 9, 1871–1874.

Goetz CG, Tilley BC, Shaftman SR, Stebbins GT, Fahn S, Martinez-Martin P, Poewe W, Sampaio C, Stern MB, Dodel R, Dubois B, Holloway R, Jankovic J, Kulisevsky J, Lang AE, Lees A, Leurgans S, LeWitt PA, Nyenhuis D, Olanow CW, Rascol O, Schrag A, Teresi JA, van Hilten JJ, LaPelle N, Agarwal P, Athar S, Bordelan Y, Bronte-Stewart HM, Camicioli R, Chou K, Cole W, Dalvi A, Delgado H, Diamond A, Dick JP, Duda J, Elble RJ, Evans C, Evidente VG, Fernandez HH, Fox S, Friedman JH, Fross RD, Gallagher D, Goetz CG, Hall D, Hermanowicz N, Hinson V, Horn S, Hurtig H, Kang UJ, Kleiner-Fisman G, Klepitskaya O, Kompoliti K, Lai EC, Leehey ML, Leroi I, Lyons KE, McClain T, Metzer SW, Miyasaki J, Morgan JC, Nance M, Nemeth J, Pahwa R, Parashos SA, Schneider JSJS, Schrag A, Sethi K, Shulman LM, Siderowf A, Silverdale M, Simuni T, Stacy M, Stern MB, Stewart RM, Sullivan K, Swope DM, Wadia PM, Walker RW, Walker R, Weiner WJ, Wiener J, Wilkinson J, Wojcieszek JM, Wolfrath S, Wooten F, Wu A, Zesiewicz TA, Zweig RM, 2008 Movement Disorder Society-Sponsored Revision of the Unified Parkinson's Disease Rating Scale (MDS-UPDRS): Scale presentation and clinimetric testing results. Mov. Disord. 23, 2129–2170. 10.1002/mds.22340 [PubMed: 19025984]

Guilarte TR, 2013 Manganese neurotoxicity: new perspectives from behavioral, neuroimaging, and neuropathological studies in humans and non-human primates. Front. Aging Neurosci 5, 23 10.3389/fnagi.2013.00023 [PubMed: 23805100]

Guilarte TR, Gonzales KK, 2015 Manganese-induced parkinsonism is not idiopathic Parkinson's disease: Environmental and genetic evidence. Toxicol. Sci. 146, 204–212. 10.1093/toxsci/kfv099 [PubMed: 26220508]

Kaiser LG, Young K, and Matson GB (2007). Elimination of spatial interference in PRESS-localized editing spectroscopy. Magn. Reson. Med 58, 813–818. [PubMed: 17899586]

Lee EY, Flynn MR, Du G, Lewis MM, Fry R, Herring AH, Van Buren E, Van Buren S, Smeester L, Kong L, Yang Q, Mailman RB, Huang X, 2015 T1 relaxation rate (R1) indicates nonlinear mn accumulation in brain tissue of welders with low-level exposure. Toxicol. Sci 146, 281–289. 10.1093/toxsci/kfv088 [PubMed: 25953701]

Lee EY, Flynn MR, Lewis MM, Mailman RB, Huang X, 2018 Welding-related brain and functional changes in welders with chronic and low-level exposure. Neurotoxicology 64, 50–59. 10.1016/j.neuro.2017.06.011 [PubMed: 28648949]

Leggett RW, 2011 A biokinetic model for manganese. Sci. Total Environ 409, 4179–86. 10.1016/j.scitotenv.2011.07.003 [PubMed: 21802707]

Lehallier B, Coureaud G, Maurin Y, Bonny J-M, 2012 Effects of manganese injected into rat nostrils: implications for in vivo functional study of olfaction using MEMRI. Magn. Reson. Imaging 30, 62–9. 10.1016/j.mri.2011.08.009 [PubMed: 22055859]

Leuze C, Kimura Y, Kershaw J, Shibata S, Saga T, Chuang K-H, Shimoyama I, Aoki I, 2012 Quantitative measurement of changes in calcium channel activity in vivo utilizing dynamic manganese-enhanced MRI (dMEMRI). Neuroimage 60, 392–9. 10.1016/j.neuroimage.2011.12.030 [PubMed: 22227885]

Lewis MM, Flynn MR, Lee EY, Van Buren S, Van Buren E, Du G, Fry RC, Herring AH, Kong L, Mailman RB, Huang X, 2016 Longitudinal T1 relaxation rate (R1) captures changes in short-term Mn exposure in welders. Neurotoxicology 57, 39–44. 10.1016/j.neuro.2016.08.012 [PubMed: 27567731]

Long Z, Dyke JP, Ma R, Huang CC, Louis ED, Dydak U, 2015 Reproducibility and effect of tissue composition on cerebellar γ-aminobutyric acid (GABA) MRS in an elderly population. NMR Biomed 28, 1315–1323. 10.1002/nbm.3381 [PubMed: 26314380]

Long Z, Li X-RR, Xu J, Edden RAEE, Qin W-PP, Long L-LL, Murdoch JB, Zheng W, Jiang Y-MM, Dydak U, 2014 Thalamic GABA Predicts Fine Motor Performance in Manganese-Exposed Smelter Workers. PLoS One 9, e88220 10.1371/journal.pone.0088220 [PubMed: 24505436]

Ma RE, Ward EJ, Yeh C-L, Snyder S, Long Z, Gokalp Yavuz F, Zauber SE, Dydak U, 2018 Thalamic GABA levels and occupational manganese neurotoxicity: Association with exposure levels and brain MRI. Neurotoxicology 64, 30–42. 10.1016/j.neuro.2017.08.013 [PubMed: 28873337]

Mullins PG, Mcgonigle DJ, O'Gorman RL, Puts NAJJ, Vidyasagar R, Evans CJ, Edden RAEE, Brookes MJ, Garcia A, Foerster BR, Petrou M, Price D, Solanky BS, Violante IR, Williams S, Wilson M, 2014 Current practice in the use of MEGA-PRESS spectroscopy for the detection of GABA. Neuroimage 86, 43–52. 10.1016/j.neuroimage.2012.12.004 [PubMed: 23246994]

Olanow CW, 2004 Manganese-Induced Parkinsonism and Parkinson's Disease. Ann. N. Y. Acad. Sci 1012, 209–223. 10.1196/annals.1306.018 [PubMed: 15105268]

Park JD, Chung YH, Kim CY, Ha CS, Yang SO, Khang HS, Yu IK, Cheong HK, Lee JS, Song C-W, Kwon IH, Han JH, Sung JH, Heo JD, Choi BS, Im R, Jeong J, Yu IJ, 2007 Comparison of high MRI T1 signals with manganese concentration in brains of cynomolgus monkeys after 8 months of stainless steel welding-fume exposure. Inhal. Toxicol 19, 965–71. 10.1080/08958370701516108 [PubMed: 17849280]

Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, Blondel M, Müller A, Nothman J, Louppe G, Prettenhofer P, Weiss R, Dubourg V, Vanderplas J, Passos A, Cournapeau D, Brucher M, Perrot M, Duchesnay É, 2012 Scikit-learn: Machine Learning in Python. J. Mach. Learn. Res 12, 2825–2830. 10.1007/s13398-014-0173-7.2

Provencher SW, 1993 Estimation of metabolite concentrations from localized in vivo proton NMR spectra. Magn. Reson. Med 30, 672–679. 10.1002/mrm.1910300604 [PubMed: 8139448]

Racette BA, Aschner M, Guilarte TR, Dydak U, Criswell SR, Zheng W, 2012 Pathophysiology of manganese-associated neurotoxicity. Neurotoxicology 33, 881–6. 10.1016/j.neuro.2011.12.010 [PubMed: 22202748]

Schroeter JD, Nong A, Yoon M, Taylor MD, Dorman DC, Andersen ME, Clewell HJ, 2011 Analysis of manganese tracer kinetics and target tissue dosimetry in monkeys and humans with multi-route

physiologically based pharmacokinetic models. Toxicol. Sci 120, 481–498. 10.1093/toxsci/kfq389 [PubMed: 21205636]

Shinotoh H, Snow BJ, Hewitt K. a, Pate BD, Doudet D, Nugent R, Perl DP, Olanow W, Calne DB, 1995 MRI and PET studies of manganese-intoxicated monkeys. Neurology 10.1212/WNL.45.6.1199

Troughton JS, Greenfield MT, Greenwood JM, Dumas S, Wiethoff AJ, Wang J, Spiller M, McMurry TJ, Caravan P, 2004 Synthesis and evaluation of a high relaxivity manganese(II)-based MRI contrast agent. Inorg. Chem 43, 6313–6323. 10.1021/ic049559g [PubMed: 15446878]

Tsuboi Y, Uchikado H, Dickson DW, 2007 Neuropathology of Parkinson's disease dementia and dementia with Lewy bodies with reference to striatal pathology. Park. Dis. Mov. Disord 13, S221–S224. 10.1016/S1353-8020(08)70005-1

U.S. EPA, 2011 Exposure Factors Handbook 2011 Edition (Final Report) Washington, D.C 10.1016/b978-0-12-803125-4.00012-2

Ward EJ, Edmondson DA, Nour MM, Snyder S, Rosenthal FS, Dydak U, 2017 Toenail Manganese: A Sensitive and Specific Biomarker of Exposure to Manganese in Career Welders. Ann. Work Expo. Heal 62, 101–111. 10.1093/annweh/wxx091

Yeh C-L, Ward EJ, Ma R, Snyder S, Schmidt-Wilcke T, Dydak U, 2016 P125 Whole-brain r1 mapping of manganese in welders - visualisation of increased mn levels in the brain. Occup. Environ. Med 73, A161 LP-A161. 10.1136/oemed-2016-103951.442

Yokel R. a, 2009 Manganese flux across the blood-brain barrier. Neuromolecular Med 11, 297–310. 10.1007/s12017-009-8101-2 [PubMed: 19902387]
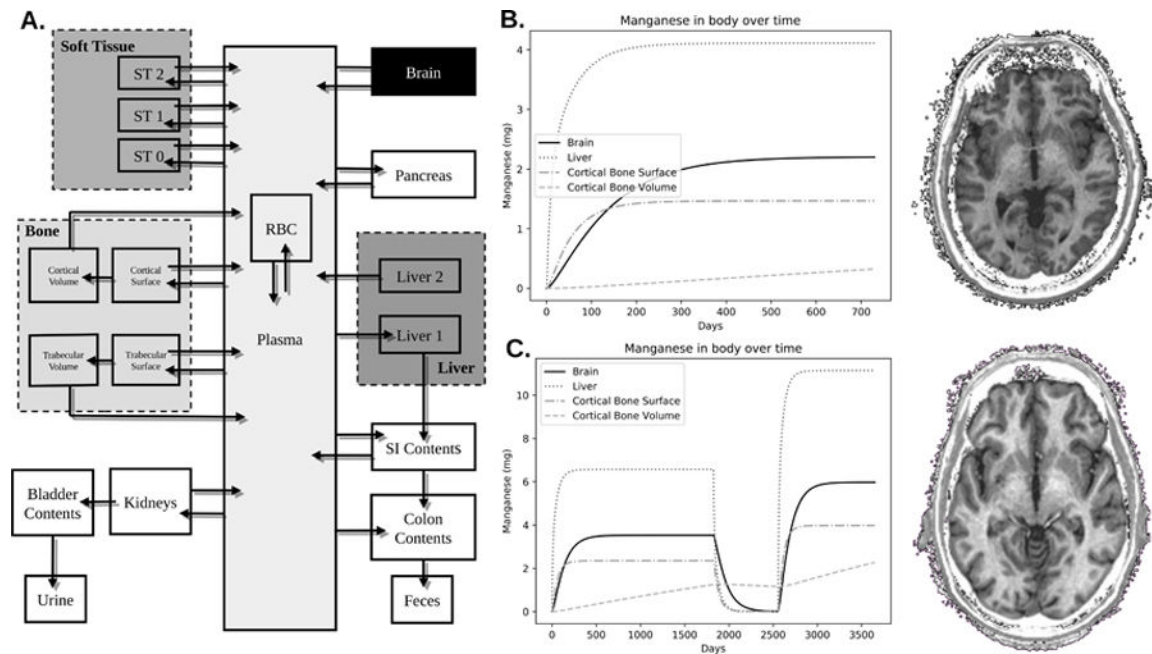
Figure 1. Mn Biokinetic Model.

A) This model is derived from a previously published model. Respirable Mn was assumed to be transported into the plasma compartment with 100% efficiency. The model then describes through a series of compartments how Mn disperses itself throughout the body. This study used values of excess brain Mn that was calculated from the brain compartment. **B)** Shows the trend of Mn accumulation in 6 regions of the body: brain, cortical bone surface, cortical bone volume, trabecular bone surface, trabecular bone volume, and blood plasma. This welder worked for 2 years with an average air Mn exposure concentration of 0.088 mg/m³. **C)** Another welder worked in three jobs over 10 years. The first job was for 5 years with an air concentration of 0.14 mg/m³. The second job was welding aluminum, and thus no Mn exposure. The third job lasted for 3 years prior to his participation in our study with an exposure level of 0.24 mg/m³. As can be seen by the R1 images on the right, taken from the same location and with the same contrast, the subject in C has much brighter white matter – indicative of higher R1 due to Mn accumulation.
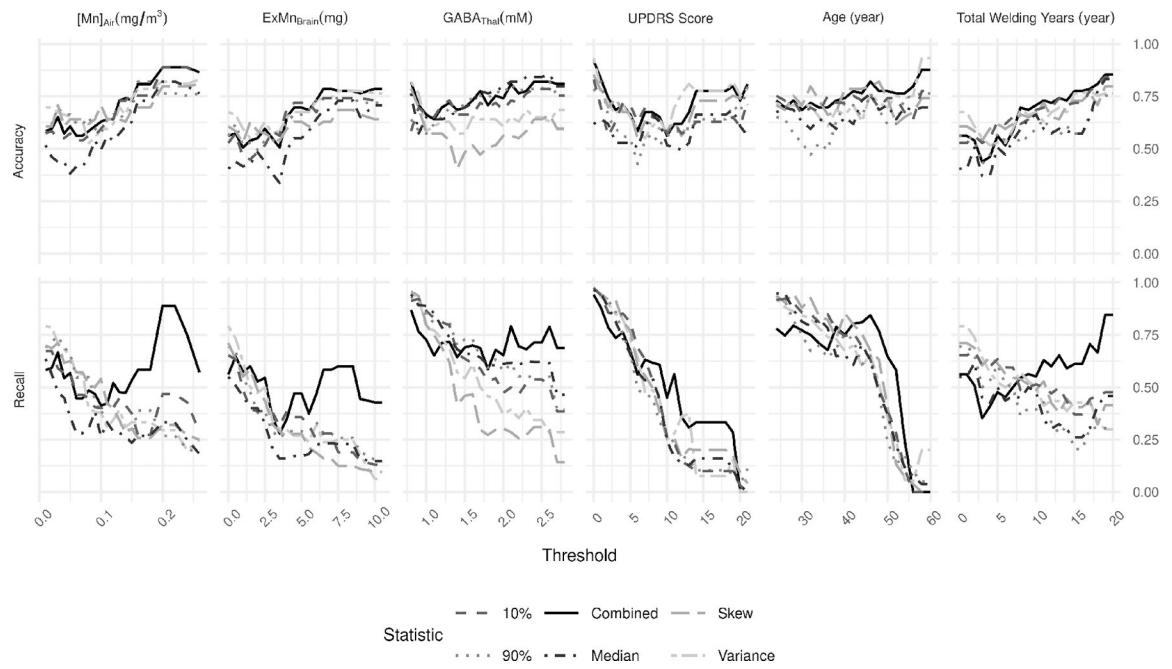
Figure 2. SVM Model Performance.

For each target variable, accuracy (top row) and recall (bottom row) are presented across all thresholds. The combined (solid line) model represents the model performance combining predictions from the other five statistics and using a majority rule to determine its prediction.
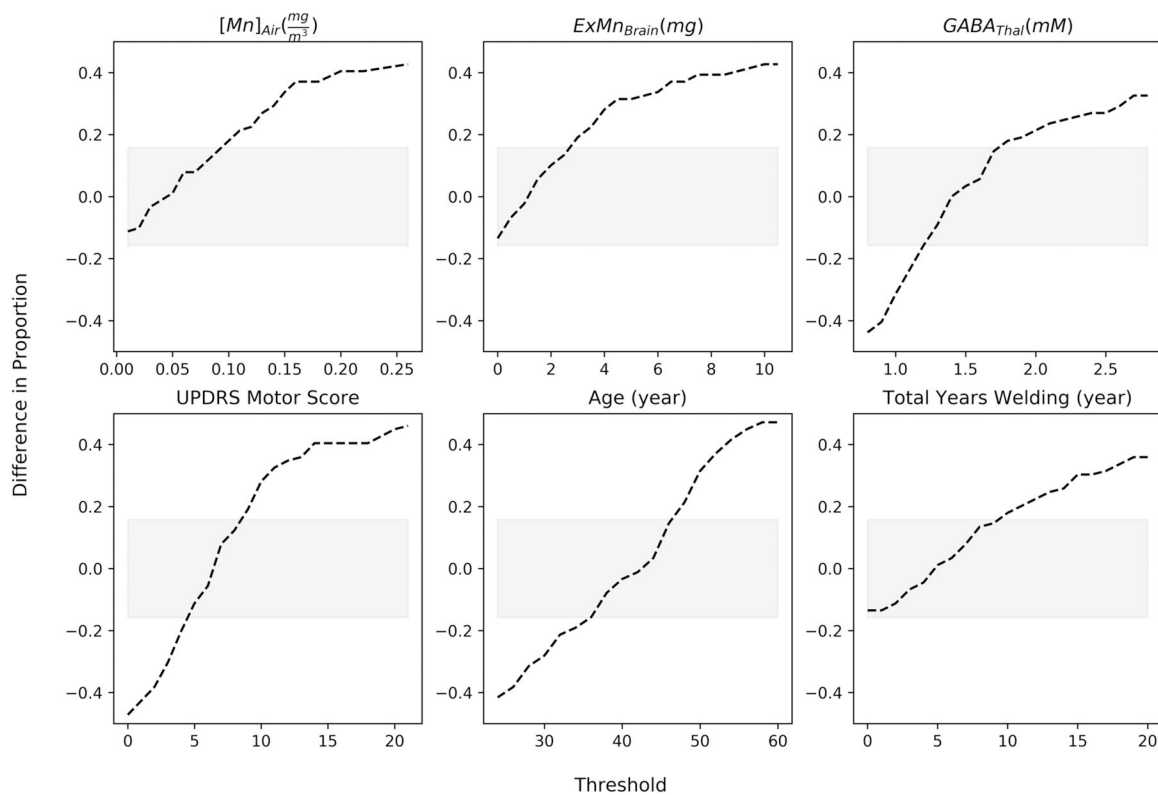
Figure 3. Proportionality of Observed Class Distributions.

The dotted-line represents the difference in proportions between both classes being used in a model at a specific threshold. When the proportion is negative, there are more subjects in the class representing greater than the threshold while when the proportion is positive, there are more subjects in the class representing less than or equal to the threshold. The grey region represents the 95% confidence interval for classes that are equally split. In general, the thresholds move from having greater numbers in the class representing greater than the threshold to having greater numbers in the class less than or equal to the threshold.
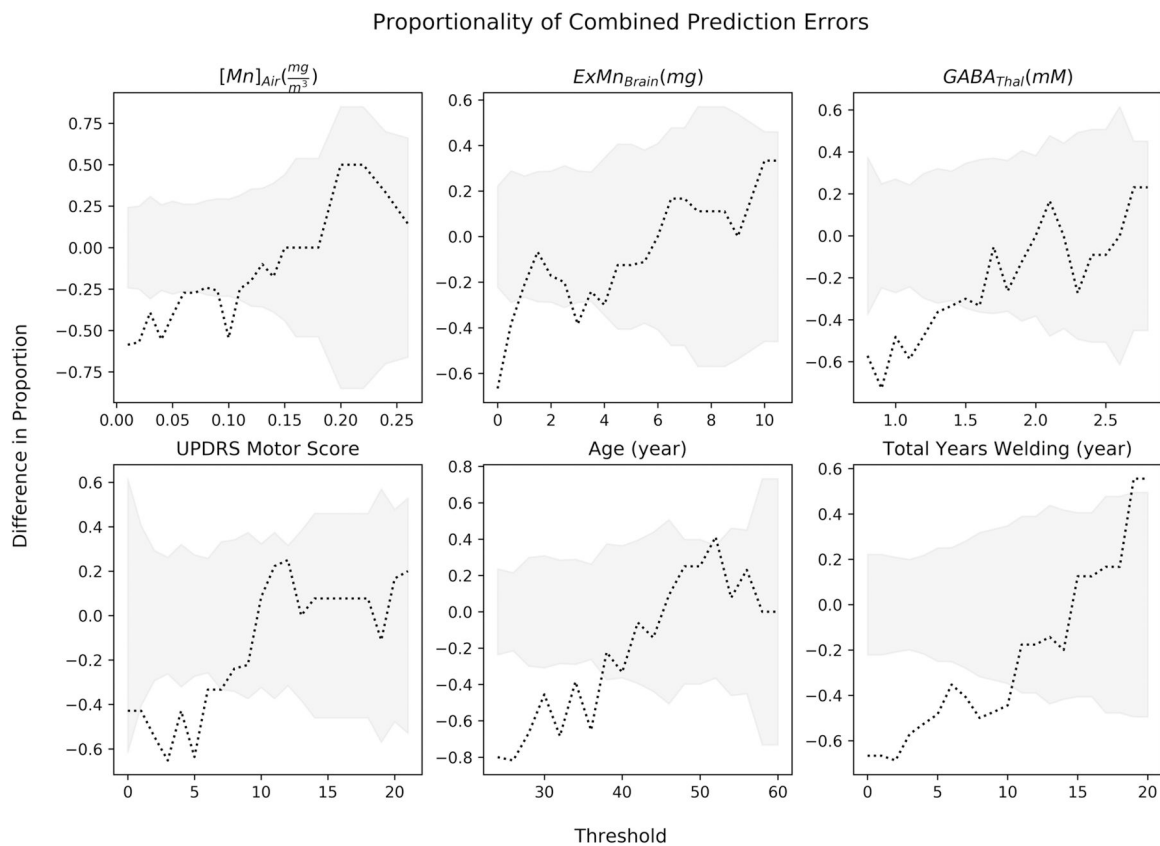
Figure 4. Proportionality of Combined Prediction Errors.
Wald confident intervals were calculated for each threshold and is depicted in grey. The dotted line is the difference in proportion of combined model prediction errors. A negative difference in proportion represents the model incorrectly predicting less in the class less than or equal to the threshold, thus more false positive predictions. A positive difference in proportions represents more false negative predictions.

Table 1

Demographics and Statistics (Mean [SD], Range)

	Welders	Controls
N	52	37
$[Mn]_{Air}$ (mg/m^3)	0.137 [0.134], 0 – 0.716	0.004 [0.005], 0 – 0.0252
$ExMn_{Brain}$ (mg)	3.86 [3.84], 0.018 – 20.60	0
$GABA_{Thal}$ (mM)	1.55 [0.60], 0.66 – 3.42	1.37 [0.56], 0.693 – 2.63
UPDRS	8.4 [5.2], 0 – 22	7.2 [5.9], 0 – 28
Age (Years)	41 [10], 21 – 61	41 [11], 20 – 61
Total Welding Years (Years)	12.85 [8.6], 1.66 – 37	0

Abbreviations: $[Mn]_{Air}$: amount of respirable Mn particulate in the air at the welder's workplace; $ExMn_{Brain}$: amount of Mn accumulated in the brain; $GABA_{Thal}$: concentration of GABA in the Thalamus; UPDRS: Unified Parkinson's Disease Rating Scale.

Table 2

Results for Principle Component Analysis of Whole-Brain R1

PCs (90%)		PC1		PC2		PC3	
		%Var	Regions[*]	%Var	Regions[*]	% Var	Regions[*]
MEDIAN	6	77%	Superior temporal sulcus [†]	6.4%	Ventricles Corpus Collosum	2.8%	Superior parietal gyrus Superior occipital gyrus
VARIANCE	25	41.6%	Cerebral White Matter, Parieto-occipital, cingulate-marginalis	11%	Cuneus Precuneus Corpus Collosum	4.9%	Cerebellum White Matter Cerebellum Cortex Brainstem
SKEW	49	12.3%	Superior circular insula sulcus Posterior lateral fissure	7.8%	Cerebral White Matter Precentral superior sulcus Postcentral Sulcus	5%	Superior Frontal Gyrus Cerebellum Cortex Temporal Medial Parahip gyrus
10-PERCENTILE	13	72.6%	Parieto-occipital sulcus Temporal Sulcus Cingulate	4.8%	Ventricles Blood Vessels	2.5%	Precentral Gyrus Postcentral Gyrus Cerebellar White Matter
90-PERCENTILE	13	63.8%	Parietal Occipital Sulcus Temporal sulcus Hippocampus	9.6%	Cuneus gyrus Precuneus Gyri Corpus Collosum	3.9%	Ventricles Paracentral gyrus, sulcus Optic Chiasm

*
Regions presented had high weightings in both right and left hemisphere

†
While superior temporal sulcus is the highest weighted, the difference amongst regions varied by less than 1.

Table 3

Summary of high performing SVM models (0.75 Accuracy, Recall)

	THRESHOLD	STATISTIC	ACCURACY	RECALL
[MN]$_{AIR}$	(mg/m^3)			
	0.2	Combined	0.888	0.889
	0.22	Combined	0.888	0.889
	0.24	Combined	0.888	0.75
GABA$_{THAL}$	(mM)			
	0.8	Median	0.764	0.944
	0.8	Variance	0.82	0.947
	0.8	Skew	0.787	0.958
	0.8	Combined	0.82	0.869
	0.9	Variance	0.764	0.929
	2.1	Combined	0.809	0.792
	2.6	Combined	0.82	0.789
UPDRS SCORE	(score)			
	0	Variance	0.933	0.976
	0	Skew	0.854	0.974
	0	10%	0.831	0.974
	0	90%	0.775	0.972
	0	Combined	0.921	0.943
	1	Variance	0.787	0.932
	1	Skew	0.775	0.944
	1	90%	0.82	0.947
	1	Combined	0.831	0.88
AGE	(years)			
	24	10%	0.753	0.917
	28	Skew	0.753	0.947
	32	Skew	0.798	0.926
	36	10%	0.753	0.849
	40	Skew	0.787	0.854
	42	Skew	0.787	0.8
	42	10%	0.753	0.761
	42	Combined	0.775	0.804
	44	Skew	0.787	0.756
	44	Combined	0.764	0.81
	46	Combined	0.82	0.844
	48	Combined	0.775	0.769
TOTAL WELDING YEARS	(years)			
	19	Combined	0.854	0.846
	20	Combined	0.854	0.846