JOURNAL OF NEUROTRAUMA 38:848–857 (April 1, 2021)

DOI: 10.1089/neu.2020.7322

# Acute Post-Concussive Assessments of Brain Tissue Magnetism Using Magnetic Resonance Imaging

Kevin M. Koch,[1] Andrew S. Nencka,[1] Brad Swearingen,[1] Anne Bauer,[2] Timothy B. Meier,[3] and Michael McCrea[3]


## Abstract

Recent studies have demonstrated the promising capabilities of magnetic resonance imaging (MRI)-based quantitative susceptibility maps (QSM) in producing biomarkers of brain injury. The present study aims to further explore acute QSM changes in athletes after sports concussion and investigate prognostication capabilities of QSM-derived imaging metrics. The QSM were derived from neurological MRI data acquired on a cohort ($n = 78$) of concussed male American football athletes within 48 h of injury. The MRI-derived QSM values in subcortical gray and white matter compartments after concussion showed differences relative to a matched uninjured control group (white matter: $z = 3.04$, $p = 0.002$, subcortical gray matter: $z = -2.07$, $p = 0.04$). Subcortical gray matter QSM MRI measurements also correlated strongly with duration of symptoms ($\rho = -0.46$, $p = 0.002$) within a subcohort of subjects who had symptom durations for at least one week ($n = 39$). The acute QSM MRI metrics showed promising prognostication capabilities, with subcortical gray matter compartment QSM values yielding a mean classification area under the curve performance of 0.78 when predicting symptoms of more than two weeks in duration. The results of the study reproduce previous acute post-concussion group QSM findings and provide promising initial prognostication capabilities of acute QSM measurements in a post-concussion setting.


**Keywords:** MRI; quantitative susceptibility mapping, sports concussion; susceptibility weighted imaging; tissue magnetism; traumatic brain injury

## Introduction

The severity, mechanisms, and outcome risk factors of traumatic brain injury (TBI) are diverse and can vary widely from one injury to another.[1] To roughly differentiate injuries, TBI are categorized using the Glasgow Coma Scale (GCS) score into mild, moderate, and severe injury categories.[2] Although mild TBI (mTBI) rarely places injured patients in emergent danger, it is well-documented that a substantial fraction of those with mTBI show persistent symptoms and functional deficits.[3–5] Recent large cohort studies have reported that only a third of patients with mTBI achieve full recovery by 12 months post-injury.[6]

Sport-related concussions (SRC) generally fall within the mildest spectrum of TBI severity. Because of the poor sensitivity of GCS scores to SRC severity, alternative measures, such as the Sport Concussion Assessment Tool (SCAT) are utilized as first-stage clinical assessments of acute injury severity.[7] Despite the relatively minor acute severity of SRC injuries, increasing public awareness of the potential neurological impact from individual concussive events and longer-term risks from repeated injury events have amplified efforts to understand pathophysiological changes induced by such injuries. Recent research studies have also been leveraged to inform best practices in determining post-SRC return-to-play timelines.[8]

Advanced diagnostic imaging, including magnetic resonance imaging (MRI) and computed tomography (CT), have undergone substantial exploration as probes for post-SRC changes in brain structure and function. The physiological changes induced by the SRC subcategory of mTBI are difficult to detect with qualitative interpretation of MRI or CT.[9] Advanced quantitative MRI (qMRI), however, which moves beyond the limits of qualitative morphological MRI analysis, has shown promise as an investigative probe of post-concussive pathophysiological changes in brain tissue. As exemplified by the growing utility of qMRI in cancer research,[10] the sensitivity of MRI to local tissue environments and physiological processes can be a valuable tool in pathology quantification.

Most previous qMRI studies of SRC have focused on diffusion tensor imaging,[11,12] diffusion kurtosis imaging tensor imaging,[13] arterial spin labeling,[14] and functional connectivity.[15,16]

Departments of [1]Radiology and [3]Neurosurgery, Medical College of Wisconsin, Milwaukee, Wisconsin, USA.
[2]Data Science Group, The New York Times, New York, New York, USA.

Quantitative Susceptibility Mapping (QSM) is a relatively newer qMRI metric that has recently shown sensitivity to mTBI.[17–19]

As a quantitative extension of routinely applied susceptibility-weighted MRI (SWI) sequences,[20] QSM utilizes raw MRI signals extracted from SWI MRI acquisitions to estimate an approximated isotropic magnetic susceptibility tensor for each imaged tissue voxel.[21] The utility of QSM as a diagnostic and research tool hinges on its sensitivity to physiological and structural changes in tissue biomagnetism.

The biomagnetism of brain tissue can reveal changes in biological components such as ferretin, hemosiderin, water content, myelin, cerebral blood volume, and calcium,[21,22] Along with conventional SWI, QSM MRI has been used previously to identify regions of focal tissue damage in complicated mTBI on a cohort of military personnel.[23] In addition, preliminary group studies of post-injury QSM have shown group magnetism changes in particular gray and white matter regions.[18,17] White-matter QSM measurements after SRC have also shown promising correlations with injury severity metrics on an individual subject level.[17]

The present study seeks to expand on previous analyses of QSM in athletes after sports concussion and focus on acute QSM measurements. Specifically, this study has three general aims. First, it seeks to reproduce previously reported group differences in acute post-injury QSM values relative to control groups in a larger injury cohort. Second, this study explores correlations of acute QSM measurements with conventional symptom checklist scores. Third, a measure of injury severity (self-reported symptom duration) is used to explore prognostication capabilities of acute QSM measurements, both as a stand-alone measurement and in combination with symptom checklist scores.

## Methods

### Subjects

Male high school and collegiate American football players were recruited within a three-year regional prospective study of SRC in athletes. The criteria for exclusion from the study included contraindications or injuries that would prevent participation in the study protocol, current psychotic disorders or reported narcotic use, and/or a history or suspicion of conditions associated with cognitive impairments (e.g., epilepsy, moderate-to-severe TBI). Adult participants and parents of minors provided written informed consent, while minors completed written assent. The study was approved by the Medical College of Wisconsin Institutional Review Board.

Certified athletic trainers and/or team physicians trained in sports medicine performed the initial diagnosis of concussion injuries. The definition of concussion was derived from the Centers for Disease Control and Prevention HEADS UP educational initiative: ''An injury resulting from a forceful bump, blow, or jolt to the head that results in rapid movement of the head and causes a change in the athlete's behavior, thinking, physical functioning, or the following symptoms: headache, nausea, vomiting, dizziness/balance problems, fatigue, difficulty sleeping, drowsiness, sensitivity to light/noise, blurred vision, memory difficulty, and difficulty concentrating.''

Injured subjects enrolled within the study were imaged with an advanced MRI protocol within 48 h after injury. The SCAT-3 symptom checklist, the Standardized Assessment of Concussion (SAC), and the Balance Error Scoring System (BESS) were also collected at baseline (on initial subject enrollment in study) and at the MRI visit.

Male control subjects were matched on level (high school vs. college), school, team, age, estimated pre-morbid intelligence, race, handedness, concussion history, and position.

Additional exclusion criteria for present analysis included the contingency that QSM raw data were saved successfully during the MRI examination and that the reconstructed QSM passed quality control checks. After these image quality assurance checks, which excluded subjects with significant motion, registration failures, or QSM processing failures, this study included a total of 153 QSM MRI datasets (78 acute injury subjects, and 75 matched control subjects). Complete study cohort demographics are summarized in Table 1. All methods utilized within the study were performed in accordance with relevant guidelines and regulations.

### QSM MRI

MRI was performed on a 3T MRI system (Discovery MR 750, GE Healthcare, Waukesha, WI) using a 32ch Nova Medical (Wilmington, MA) head receive array. The QSM were acquired at 0.75 mm × 0.96 mm in-plane resolution across 2 mm axial slices. Phase-sensitive MRI acquired at four echo times (10.4,17.4, 24.4,31.4) msec and a repetition time of 58.6 msec were used to construct magnetic field perturbation maps. The total acquisition time for each QSM MRI dataset was 4 min.

The MRI-based magnetic perturbation scalar maps of the longitudinal component the perturbed magnetic fields were derived from the phase-sensitive MRI data. After background field removal from the perturbation maps,[24] susceptibility inversion was performed using an adapted localized processing formulation[25] of the Morphology-Enabled Dipole Inversion (MEDI) algorithm (MEDI Toolbox update version date 11/06/2017).[26] Relevant parameters utilized in the MEDI algorithm are lambda = 1250 (regularization) and zero padding of 50% of the field of view. The QSM computations were performed using Matlab (MathWorks, Natick, MA) Dell PowerEdge R730 servers running Red Hat Enterprise Linux 6.7 on Intel Xeon E5-2620v3 CPUs.

### Statistical analyses

Individual and group analyses were enabled via subject image registrations to a common coordinate space. A 2-mm isotropic Montreal Neurological Institute (MNI) neurological template space was utilized.[27] Registrations were performed using software contained within the FSL suite.[28] Specifically, the *FLIRT* function in FSL was utilized using nine degrees of freedom and trilinear interpolation to develop transformation parameters. Warping of the

Table 1. Characteristics of the Study Cohort

| | *Injured* (n = *78 males)* | *Control* (n = *75 males)* | *z* | p |
|---|---|---|---|---|
| Age (y) | 17.95 (17.63–18.27) | 18.43 (18.05–18.81) | -1.53 | 0.13 |
| Height (in) | 71.79 (71.18–72.40) | 71.19 (70.62–71.76) | 1.41 | 0.16 |
| Weight (lb) | 215.99 (205.11–226.86) | 205.59 (196.47–214.70) | 1.09 | 0.27 |
| Y.I.S (y) | 8.12 (7.41–8.84) | 8.08 (7.45–8.71) | -0.01 | 0.99 |
| P.SRC | 0.64 (0.43–0.85) | 0.52 (0.33–0.71) | 0.80 | 0.42 |

Y.I.S, years in sport; P.SRC, number of previous sports-related concussions; *p,* value from Mann-Whitney *U* test; z, approximated z-score from *U* test.

images to the template space utilized the *applywarp* function to apply the computed transformations. Manual verification of template registrations was performed by the study team. No registration failures were identified in the template registration pipeline that necessitated further exclusion of any QSM datasets from the analysis cohort.

Brain regions of sufficient QSM stability within the anatomic template space were identified by computing coefficients of variation (CV) across a control cohort ($n = 68$ examinations) from an independent study that utilized identical acquisition parameters. A stability mask was then constructed from this CV map at a threshold of CV <0.8, which is a QSM stability threshold that has been utilized within previous published studies.[17]

Regional neurological segmentations extracted from the Johns Hopkins University and Harvard MNI space atlases[29,30] were to define white and subcortical gray matter (SCGM) global compartments. The utilized compartments were refined using a physics-based masking operation that restricted white matter as $\chi$ <0.03 ppm[31] and SCGM as $\chi$ >0.05 ppm susceptibility thresholds. This additional compartment masking was performed on a mean composite QSM volume computed from the aforementioned independent control cohort. Mean susceptibility values were computed within each gray and white matter region of interest for each subject and yielded a pair of QSM-MRI based biomarker estimates for each subject.

Reported QSM values were referenced to a cerebral spinal fluid (CSF) baseline, which was computed as the mean CSF compartment susceptibility value averaged across the control subject cohort.

Group comparisons of the acute injured and control cohorts were independently performed for the two QSM biomarkers and three clinical metrics. Normality tests on each distribution were computed using the Jarque-Bera approach.[32] Given the resulting prevalence of non-normal behavior of distributions within the study data, non-parametric statistical tests were utilized. Statistical differences between groups were performed using Mann-Whitney *U* tests. Effect sizes of differences between the groups were also estimated using estimated z-scores from the *U* test results.

Within the injured group, correlations were computed between each acute predictive metric (two QSM MRI compartments and three clinical metrics) and self-reported symptom duration. The symptom duration was determined through reviewer questioning at subject visits or follow-ups, whereby subjects were asked to recall the date at which post-injury symptoms (dizziness, headache, balance, confusion, lack of focus, etc.) had resolved fully. To assess the strength of the biomarker correlations as a function of injury severity, correlations were computed for three categories of symptom duration, whereby subjects with shorter durations of symptoms are progressively excluded from analysis. These categories were defined as: (1) all injuries, (2) injuries with symptom duration of at least one week, and (3) injuries with symptom duration of at least two weeks. The Spearman $\rho$ was used to assess correlations between each metric and reported symptom duration.

Cross correlations of the QSM and clinical measures were computed at an individual level within the injured cohort utilizing the full analysis complement of subject markers: acute QSM (white matter and SCGM) and acute clinical measures (SCAT-3, SAC, BESS; differences from baseline scores). Again, the Spearman $\rho$ was used to compute correlations.

Subcompartment QSM-MRI analyses of the white matter and SCGM compartments (based on Johns Hopkins and Harvard MNI atlas segmentations) were performed using individualized z-score analysis. Mean subregional QSM-MRI metrics were computed from each subcompartment ($j$) within each individual subject ($i$), yielding the indexed parameter, $I_i^j$. In addition, means and standard deviations of the means were computed within of each subcompartment of the control cohort, yielding respective parameters, $C_{mean}{}^j$ and $C_{std}{}^j$. Subcompartment ($j$) z-scores were then computed for each subject ($i$), according to:

$$zScore_i^j = \frac{I_i^j - C_{mean}^j}{C_{std}^j} \quad (1)$$

For each subject, the maximum subcompartment absolute z-score was then computed:

$$j_i^{maxZ} = max_j\left(zScore_i^j\right) \quad (2)$$

Finally, the incidence rate of maximum subcompartment z-score was computed across the injured subject cohort of size $n_{inj}$:

$$Max\ Rate^j = \frac{\sum_i j_i^{maxZ}}{n_{inj}} \quad (3)$$

This process was repeated for each of the aforementioned injury category cohorts.

Predictive modeling of symptom duration was performed using binomial logistic regression. The predictors utilized in the regression models included the acute QSM MRI metrics and clinical measures (reported as differences from baseline scores). Regression models were computed using the individual (univariate) and combined (multi-variate) predictors of self-reported symptomatic durations. After the aforementioned correlation analysis, models were trained for two binary classifications: symptom duration of at least oe week (Category 2 injuries) and symptom duration of at least two weeks (Category 3 injuries) in length.

Before logistic regression modeling, predictor values for each metric were standardized through z-score transformations (mean of zero and a standard deviation of one). In addition, algorithm bias was mitigated by dividing each predictor data into training and testing cohorts using stratified k-fold cross-validation with k = 10, yielding a 10% validation holdout in each fold. Model training was class balanced[33] by weighting the dataset by class population ratios. For example, in training a set of 61 null and 13 injured cases, the injured cases were weighted by a factor of 61/37.

The results of these logistic models were used to perform receiver operating characteristic (ROC) analysis[34] of the specificity and sensitivity of the various markers to classify symptom duration within subjects. The area under the curve (AUC) of a ROC plot (sensitivity vs. 1-specificity) was used as a measure of performance of the predictive models. For a given model, an AUC of 1.0 indicates a perfect sensitivity and specificity of the classifier, whereas 0.5 indicates a random classification. Mean and 95% confidence intervals were computed for each of the logistic regression models, calculated from each fold of the k-fold cross-validation procedure.

All statistical and analysis procedures were performed using the Matlab Statistics and Machine Learning Toolbox (MathWorks, Natick, MA).

## Results

Table 1 provides general characteristics of the study cohort. Using a significance threshold of 0.05, none of the presented metrics showed any significant difference between the injured and control groups.

Acute QSM MRI and clinical checklist metrics for both study cohorts are depicted graphically in Figure 1. Within each plot, the central line indicates the median of the distribution, and the box indicates the interquartile range. Whiskers in each plot extend to the most extreme non-outlier distribution points not considered outliers, while outliers are plotted individually. Group-wise statistical tests of the QSM and checklist metrics are summarized in Table 2.

Several relevant observations within the MRI QSM results are noted. First, both derived acute QSM MRI tissue compartment metrics (white matter and SCGM) within the injured group showed

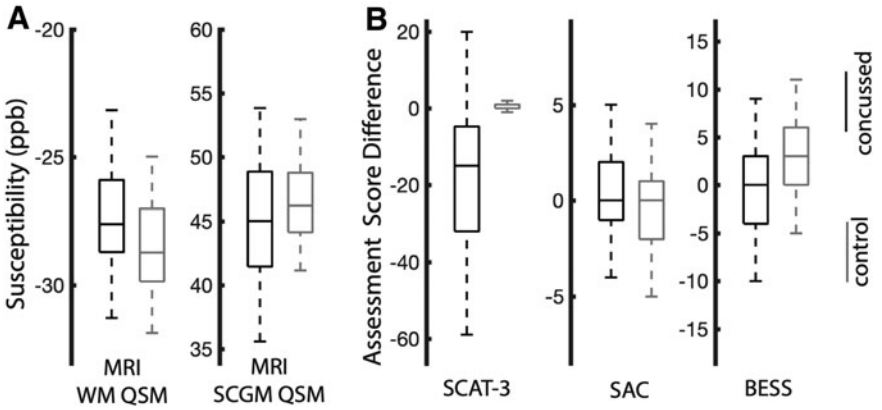


**FIG. 1.** Distribution box-plots for acute (within 48 h) (**A**) quantitative susceptibility maps (QSM) magnetic resonance imaging (MRI) measurements and (**B**) clinical assessment scores (from baseline). Distribution plots for the matched control groups are also displayed for each metric. WM, white matter; SCGM, subcortical gray matter; SCAT, Sport Concussion Assessment Tool; SAC, Standardized Assessment of Concussion; BESS, Balance Error Scoring System.

strong acute differences from the matched control group. Second, the effective z-scores of the observed effects showed opposite signs of QSM changes in the white matter and subcortical gray matter compartments. Relative to controls, the white matter compartment in the injured group shows a roughly 1 ppb acute increase in magnetic susceptibility relative to controls, whereas subcortical gray matter shows a roughly 2 ppb decrease.

As expected, the SCAT-3 and BESS checklist scores (measured acutely post-injury and relative to pre-season baselines) showed strong differences relative to the control group.

Correlation analysis of the acute QSM MRI and clinical metrics with injury outcomes (self-reported symptom duration) are presented in Table 3. The Spearman $\rho$ and associated hypothesis test $p$ values are reported for pairwise correlations between each metric against self-reported symptom duration for three categories of injured subjects: (1) all injured subjects, (2) those subjects with self-reported symptoms of at least one week, and (3) at least 2 weeks. A key result from this analysis is that two of the clinical metric scores (SCAT-3 and SAC) were the only metrics showing correlation trends with the inclusion of all injuries.

These clinical metrics, however, showed poorer correlations when the lesser injuries (>1 week of symptom duration) were excluded from the analysis. For subjects with more substantial injuries, the BESS metric and QSM biometrics showed notable correlation trends with symptom duration. For the most severe injuries (>2 weeks of symptoms), none of the individual acute metrics showed correlations with the symptom duration.

Table 2. Statistical Group Tests of Acute Quantitative Susceptibility Maps Magnetic Resonance Imaging and Clinical Checklist Metrics

| *QSM MRI* | | | | *Clinical metrics* | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| WM | | SCGM | | SCAT-3 | | SAC | | BESS | |
| z | *p* | z | *p* | z | *p* | z | *p* | z | *p* |
| 3.04 | 0.002 | -2.07 | 0.038 | -7.81 | < 0.001 | 1.35 | 0.178 | -3.92 | < 0.001 |

QSM, quantitative susceptibility maps; MRI, magnetic resonance imaging; WM, white matter; SCGM, subcortical gray matter; SCAT-3, Sport Concussion Assessment Tool; SAC, Standardized Assessment of Concussion; BESS, Balance Error Scoring System.

Estimated z-scores and $p$ values from Mann-Whitney $U$ tests between the concussed and control groups are displayed for each metric.

Figure 2 provides representative images of MRI-based susceptibility maps across brain tissues. The QSM MRI for two axial slices demonstrates changes in the white matter (A) and subcortical gray matter compartments (B). Column i provides mean control QSM MRI masked within respective tissue compartments (row A, white matter, row B, subcortical gray matter). The utilized QSM tissue masks are derived from template-space atlases and externally derived QSM stability maps. Column ii provides masked difference QSM MRI from an individual injured subject relative to the control QSM mean. The injured subject displayed in column ii had a reported symptom duration of 34 days.

Figure 3 provides scatter plots of QSM values for the white matter (row A) and gray matter (row B) compartments for Category 2 injuries (those beyond one week of symptom duration). Trends are clearly identified for both white matter (row A, $\rho = 0.32$, $p = 0.05$) and SCGM (row B, $\rho = -0.46$, $p = 0.002$).

Cross-correlation of the analyzed metrics (available in Supplementary Table 1) showed strong correlations between the QSM compartment metrics ($\rho = -0.47$), as well as SCAT3 and BESS metrics ($\rho = -0.37$). These trends were observed in the Category 1 and 2 injured groups, but not in the Category 3 (most severe) injury group.

Table 3. Correlation of Acute Clinical and Quantitative Susceptibility Maps Metrics against Self-Reported Symptom Durations

| *Metric* | *Cat. 1: All injuries* | | | *Cat. 2: Symptoms ≥1 week* | | | *Cat. 3: Symptoms ≥2 weeks* | | |
|---|---|---|---|---|---|---|---|---|---|
| QSM MRI | $\rho$ | p | n | $\rho$ | p | n | $\rho$ | p | n |
| WM | 0.03 | 0.81 | 74 | 0.32 | 0.05 | 39 | -0.003 | 0.99 | 13 |
| SCGM | -0.16 | 0.18 | 74 | -0.46 | <0.01 | 39 | 0.12 | 0.70 | 13 |
| Checklists | | | | | | | | | |
| SCAT-3 | 0.38 | <0.01 | 73 | 0.07 | 0.67 | 39 | 0.08 | 0.79 | 13 |
| SAC | -0.22 | 0.06 | 72 | 0.06 | 0.71 | 37 | 0.05 | 0.87 | 11 |
| BESS | -0.18 | 0.15 | 70 | -0.37 | 0.03 | 35 | 0.29 | 0.42 | 10 |

Cat., Category; QSM, quantitative susceptibility maps; MI, magnetic resonance imaging; WM, white matter; SCGM, subcortical gray matter; SCAT, Sport Concussion Assessment Tool; SAC, Standardized Assessment of Concussion; BESS, Balance Error Scoring System.

Spearman $\rho$, hypothesis test $p$ values, and cohort sizes for each correlation ($n$) are reported.

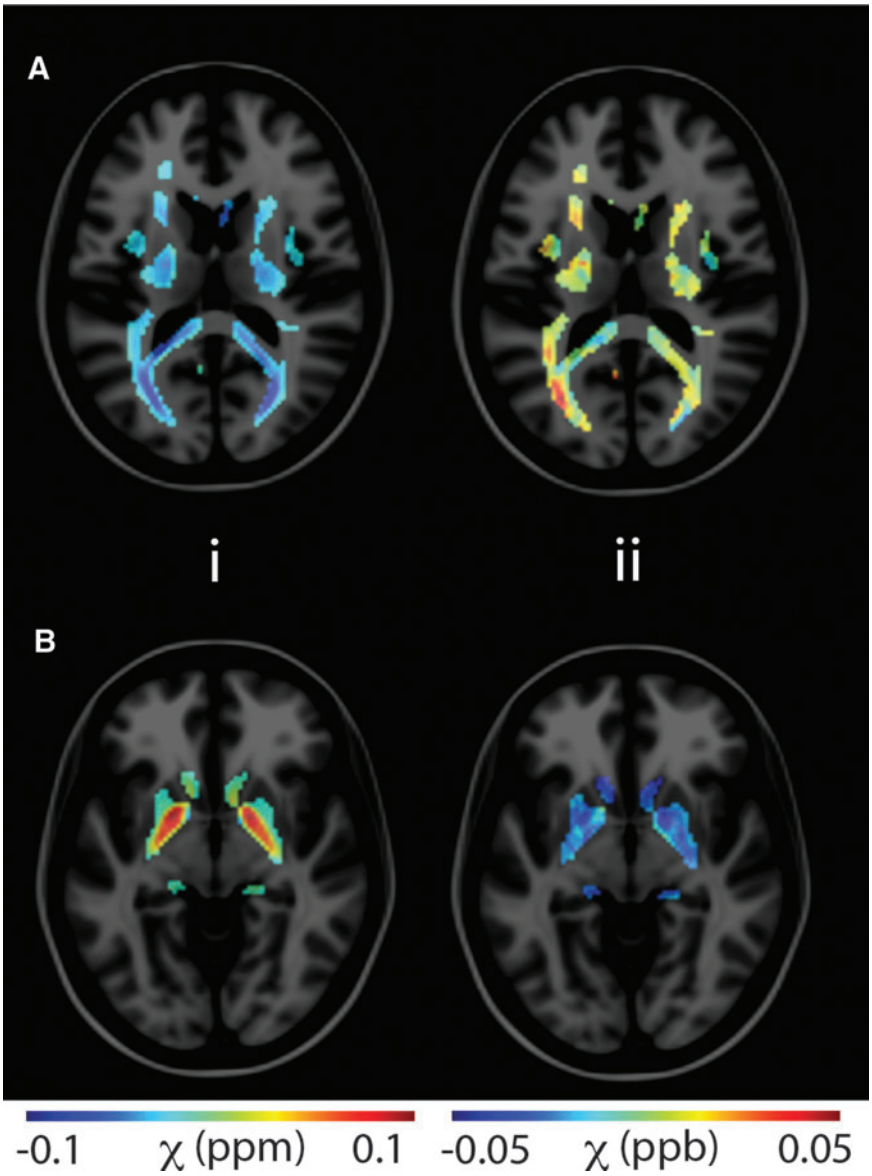


**FIG. 2.** Representative quantitative susceptibility maps (QSM) for two axial slices demonstrating changes in the white matter (**A**) and subcortical gray matter compartments (**B**). Column i displays mean control QSM masked within respective tissue compartments. Column ii provides masked difference QSM from an injured subject relative to the control QSM mean. Masks are derived from MNI atlases and externally derived QSM stability maps. Color image is available online.

To explore the effects of acute post-injury QSM MRI changes within white matter and SCGM tissue subcompartments, QSM MRI z-scores for individual injured subjects were computed relative to control values. The results of this analysis are graphically depicted in Figure 4. Individual subject z-scores for each self-reported symptom injury Category (1,2,3) are depicted across each subregion within the white matter (A) and SCGM (B) tissue compartments. For each subcompartment, bar graphs represent the rate at which each region showed the maximal injury z-score within a subject. The aim of this analysis was to identify whether any subcompartments strongly impacted concussion-induced QSM MRI changes on an individual level.

The z-scores within the injured white matter (A) or gray matter (B) compartments did not show a clear or substantial preference to any subcompartments for any of the injury severity cohort categories. In general, the respective positive and negative z-score trends for the white and gray matter subcompartments increased in magnitude with injury category.

The prognostication capabilities of acute QSM MRI and clinical metrics in predicting post-concussive symptom duration are summarized through the multi-variate logistic regression modeling results presented in Table 4. As described further in the Methods section of this report, ROCs were constructed from each derived logistic regression model. The AUC metric, also known as the discrimination coefficient, of the constructed models was utilized as the primary performance metric in evaluation of the computed models.

Figure 5 provides an example ROC curve of the univariate acute QSM MRI SCGM Category 3 model (Fig. 4A), along with the multi-variate model (Fig. 4B) derived from all acute QSM MRI and clinical metrics for Category 3 injuries.

For each constructed model, the mean AUC of k-fold cross-validation results ($k = 10$) along with the 95% confidence intervals

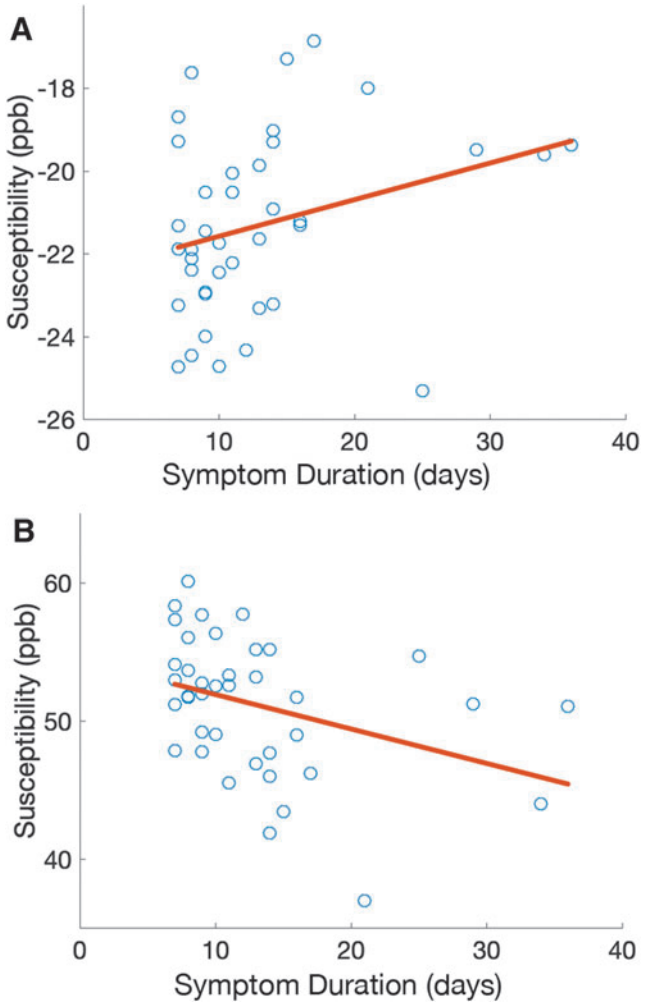


**FIG. 3.** Scatter plots of quantitative susceptibility maps (QSM) values for the white matter (**A**) and subcortical gray matter (**B**) compartments plotted against self-reported symptom duration for each injured subject. Color image is available online.

are reported for the two modeled injured group categories. Performance is reported as both recall (performance on the training dataset) and validation (performance on a reduced test dataset).

It is clear that the predictive model performs substantially better for the more severe injury category (>2 weeks of symptoms). Of note, the univariate model constructed using the QSM subcortical gray matter compartment demonstrated reasonable validation AUC performance (0.78). The use of both QSM metrics and clinical assessment scores yielded a similar prognostic validation performance. For the lower injury category, the QSM metrics did not show the same level of prognostic capabilities (validation AUC <0.6), whereas the clinical assessment score multi-variate model was able to perform reasonably (validation AUC = 0.70). Relative to their training performance, the constructed models generally showed reduced validation performance, which is an expected consequence of the modest dataset size and suggests a limited level of overfitting.

## Discussion

Previous studies have demonstrated preliminary trends of QSM MRI changes after mTBI.[18,17] In the present study, the quantitative trends identified in these studies have been reproduced in acute QSM MRI measurements. In particular, increases in magnetic susceptibility have been identified in white matter tissue, while decreases were found in the subcortical gray matter.

Compared with previous studies of QSM in mTBI, the present analysis has provided more emphasis on individualized correlations of QSM MRI values with injury severity metrics (symptom duration). Koch and associates[17] demonstrated a correlation of QSM MRI white matter values with a SRC injury severity metric (return-to-play delays). In the present study, this finding was reproduced using a different metric (symptom duration) in subjects with more than one week of symptoms. Subcortical gray matter values were similarly shown to anticorrelate with symptom duration.

Individualized z-score analysis of subregions within injured subjects provided another unique viewpoint into QSM MRI changes after SRC. This analysis is particularly important, because the QSM MRI tissue changes are not expected to be spatially homogeneous across a large spectrum of different SRC injuries. By examining z-score changes across multiple subregions, localized individualized QSM MRI changes can be visualized. As depicted in Figure 4, the QSM MRI changes induced by SRC become more apparent with injuries yielding longer symptom duration. Of interest, no particular white or subcortical gray matter subcompartments showed a clear increase in sensitivity to injury. This result is not unexpected, because the nature of post-SRC pathophysiology is not hypothesized to be locally focused when evaluated across a broad spectrum of SRC events. Future work may use specific injury classifications, such as head impact measurements, to explore further localized post-injury QSM MRI effects.

Utilizing this study's injury cohort for training of prognostic logistic regression models yielded modest predictive capabilities, with the subcortical gray matter compartment showing the most promise for Category 3 injuries (symptoms for at least two weeks) [(validation AUC: 0.78). This validation performance was similar to that of the combined use of acute QSM MRI and clinical metrics for this injury inclusion category. None of the analyzed metrics (individually or in combination) showed exceptional prognostic capabilities for less severe injuries (symptom duration less than two weeks). The unilateral SCAT-3 and combined use of all (clinical and QSM MRI) metrics, however, yielded modest prognostic validation performance (SCAT-3: validation AUC = 0.67, all metrics: validation AUC = 0.63).

The results of this preliminary prognostication modeling effort suggest that combined QSM MRI and clinical metrics could provide useful validation or identification of more severe SRC cases and could aid in the guidance of return-to-play decisions for these athletes. The incorporation of additional biometrics, such as alternative advanced MRI metrics or blood markers, could also improve the sensitivity of these preliminary prognostic models to lower grades of SRC severity.

While the prognostic logistic regression models demonstrate better capabilities for more severe injuries with longer symptom durations, continuous metric correlations (Table 3) showed better agreement with reported symptom duration when more moderate injuries are exclusively included in the analysis. An explanation for this discrepancy is the broader modeling capabilities of the logistic regression approach, whereby continuous variable labels (symptom durations) are converted into binary classifiers (symptomatic or asymptomatic) for each analysis cohort. This suggests that the more severe injuries do not linearly correlate with any of the analyzed biometrics and reported symptom duration, while more mild injuries do show trends of linear correlation with some metrics.

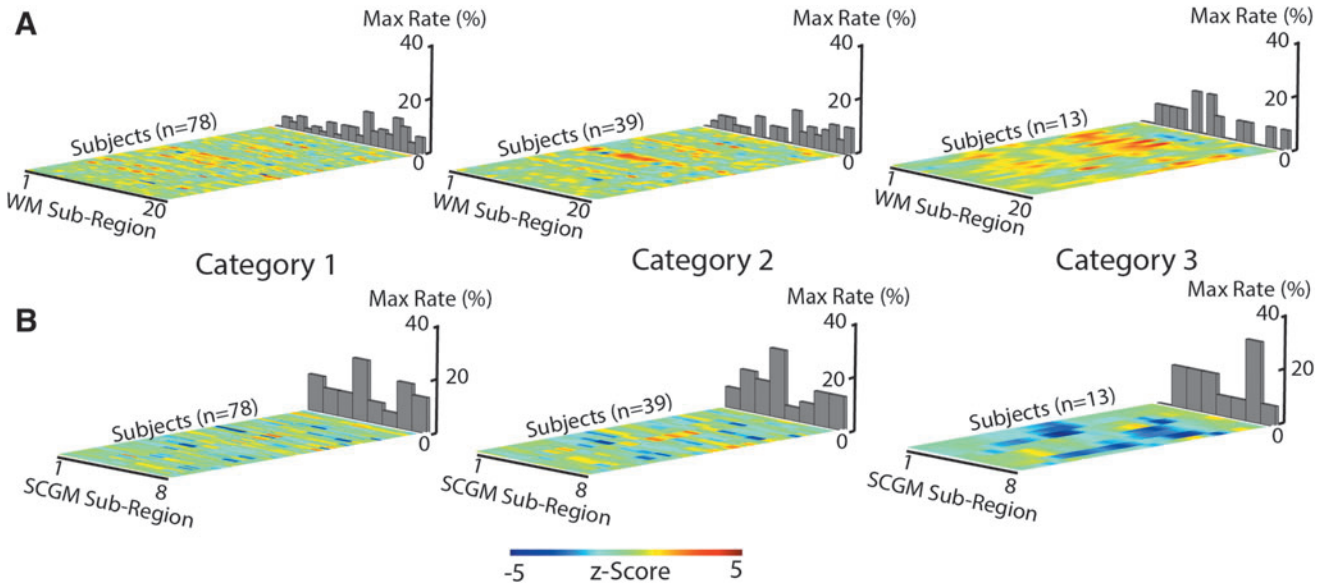


**FIG. 4.** Z-score maps across concussed subjects for each of the three injury categories (1,2,3) and subregions within the (**A**) white matter (WM) and (**B**) subcortical gray matter (SCGM) tissue compartments. Bar graphs indicate incidence rate at which a given subregion provided the maximum z-score for the given subject. The utilized subregion index labels are as follows. WM: 1 = L. Anterior thalamic radiation, 2 = R. Anterior thalamic radiation, 3 = L. Corticospinal tract, 4 = R. Corticospinal tract, 5 = L. Cingulum (cingulate gyrus), 6 = R. Cingulum (cingulate gyrus), 7 = L. Cingulum (hippocampus), 8 = R. Cingulum (hippocampus), 9 = Forceps major, 10 = Forceps minor, 11 = L. Inferior fronto-occipital fasciculus, 12 = R. Inferior fronto-occipital fasciculus, 13 = L. R. Inferior longitudinal fasciculus, 14 = R. Inferior longitudinal fasciculus, 15 = L. Superior longitudinal fasciculus, 16 = R. Superior longitudinal fasciculus, 17 = L. Uncinate fasciculus, 18 = R. Uncinate fasciculus, 19 = L. Superior longitudinal fasciculus (temporal), 20 = R. Superior longitudinal fasciculus (temporal). SCGM: 1 = L. Thalamus, 2 = L. Caudate, 3 = L. Putamen, 4 = L. Pallidum, 5 = R. Thalamus, 6 = R. Caudate, 7 = R. Putamen, 8 = R. Pallidum. Color image is available online.

Previous reports of QSM MRI findings after mTBI have provided limited speculative context on the biophysical cause of the observed effects. The consistent reproduction of these QSM MRI effects warrants further investigation of the pathophysiological source.

Within the white matter, increased water content is a probable explanation for the post-injury tissue magnetism increases that have been identified in other studies[17,18] and reproduced in the present study. Crude estimates utilizing known tissue magnetism properties suggests that the magnetism changes observed in the present study reflect between a 1.5–3% increase in water volume—which could be indicative of mild diffuse edema in the white matter. While demyelination and iron increases could also cause susceptibility increases, these potential causes are not known to be acute physiological white matter effects after mTBI.

In the subcortical gray matter, calcium ion influx and cerebral blood volume reduction have been measured in acute TBI phases.[35,36] Again, crude biophysical estimates using the tissue magnetism changes observed in this study suggest potential CBV reductions of 4%, calcium ion concentration increases of 8%, or an antisymmetrical superposition of these two parameters. Unfortunately, the effects of these two subcortical gray matter injury effects on QSM MRI values cannot be decoupled without additional diagnostic information.

While the validation of these hypothesized post-injury biophysical sources of QSM MRI changes are beyond the scope of this

Table 4. Performance of Logistic Regression Models Trained Using Acute SRC Metrics

| | *Cat. 2: Symptoms ≥1 week* | | *Cat. 3: Symptoms ≥2 weeks* | |
|---|---|---|---|---|
| | *Training* | *Validation* | *Training* | *Validation* |
| Clin-SCAT | 0.688 (0.667–0.709) | 0.672 (0.439–0.905) | 0.672 (0.657–0.687) | 0.558 (0.360–0.757) |
| Clin-SAC | 0.610 (0.596–0.624) | 0.522 (0.370–0.675) | 0.558 (0.531–0.585) | 0.417 (0.192–0.641) |
| Clin-BESS | 0.537 (0.517–0.558) | 0.461 (0.213–0.710) | 0.670 (0.651–0.689) | 0.733 (0.543–0.924) |
| QSM-WM | 0.542 (0.531–0.554) | 0.511 (0.335–0.688) | 0.697 (0.676–0.718) | 0.717 (0.471–0.962) |
| QSM-GM | 0.523 (0.510–0.536) | 0.422 (0.211–0.634) | 0.756 (0.739–0.772) | 0.783 (0.597–0.970) |
| QSM-ALL | 0.596 (0.578–0.614) | 0.500 (0.308–0.692) | 0.765 (0.750–0.779) | 0.633 (0.410–0.857) |
| Clin–All | 0.676 (0.661–0.690) | 0.589 (0.476–0.702) | 0.618 (0.604–0.632) | 0.520 (0.305–0.735) |
| All Metrics | 0.701 (0.681–0.721) | 0.633 (0.483–0.783) | 0.872 (0.853–0.890) | 0.760 (0.598–0.922) |

Cat., Category; SCAT, Sport Concussion Assessment Tool; SAC, Standardized Assessment of Concussion; BESS, Balance Error Scoring System; QSM, quantitative susceptibility maps; WM, white matter; GM, gray matter.

AUC metrics are reported as means (95% confidence interval) for the 100 bootstrapped models computed for each metric. In addition, precision recall and validation performance results are reported.

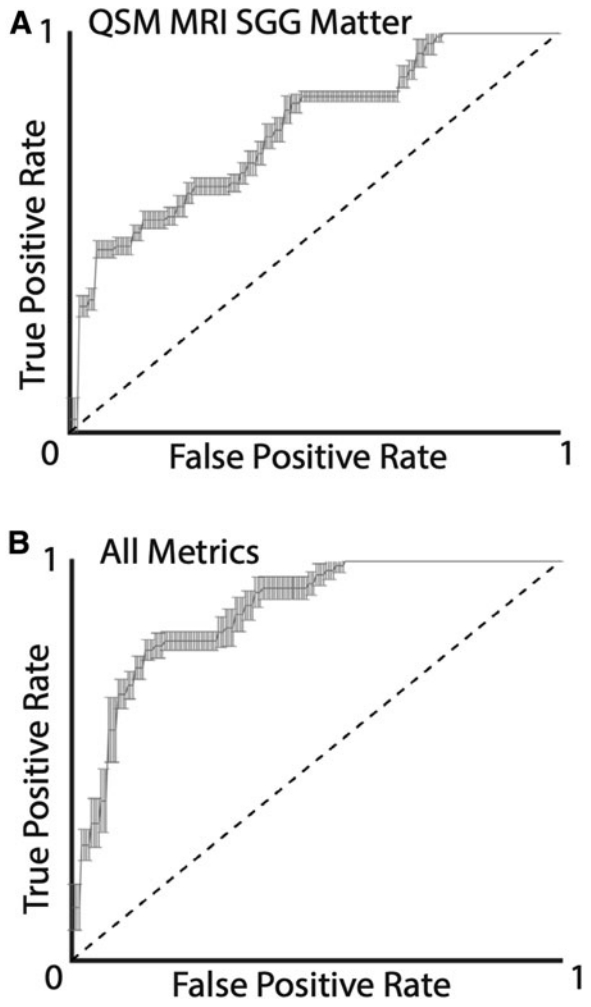


**FIG. 5.** Receiver-operator curves of logistic regression models for Category 3 injuries (symptom duration >2 weeks) using (**A**) univariate acute quantitative susceptibility maps (QSM) magnetic resonance imaging (MRI) subcortical gray matter (SCGM) values as a predictor and (**B**) multi-variate acute QSM MRI and symptom checklist metrics as predictors.

study, future studies could perform more in-depth analyses that may help illustrate the biophysical sources.

The present study utilized well-established logistic regression approaches to prognostic modeling. Given the limited cohort size, modest overfitting of trained models was observed. More sophisticated prognostication models, leveraging regularization of the applied cost functions, can be applied to reduce the observed over-fitting.

The present study had several limitations that may be overcome in future post-injury analyses of QSM MRI after mTBI. First, despite a very large recruitment pool (1318 athletes), acute QSM MRI data were acquired only on 13 subjects with injuries inducing symptoms for more than two weeks. As a result, the majority of the 78 injured subjects with acute QSM MRI had relatively mild injuries. This reduced the statistical power of both group and individualized correlation analysis.

Given the limited cohort size, modest overfitting of trained logistic regression models was observed. More sophisticated prognostication models, leveraging regularization of the applied cost functions can be applied to reduce the observed overfitting. It is anticipated that larger studies with larger cohorts of more severe injuries will help to provide greater confidence in the preliminary trends observed in this study. In particular, the use of more advanced prognostication models on larger datasets may also enable the use of individual subcompartment QSM values in building models composed of larger predictor arrays.

The choice of MRI QSM inversion algorithm offers a second opportunity for improving the results of the current study. The construction of QSM is a well-known technical challenge that has undergone more than a decade of algorithmic development. The methods utilized in the present study leverage well-known methods that are optimized to reduce persistent artifacts in QSM. Despite these efforts, imperfect QSM estimation inevitably added to variances with the group and individualized analyses. This source of error was a likely contributor to the poor individualized QSM MRI correlation with symptom duration for weaker injuries. In particular, as seen in Table 3, Category 2 injuries showed the best correlations, which are likely explained by the balance of injury severity (>1 week of symptom duration) and remaining cohort size ($n = 39$) in this group.

## Conclusion

This study of QSM MRI in an acute post-concussive setting has provided two notable results. First, it reproduced previously reported acute post-concussion QSM findings in a larger injury cohort. Specifically, findings of increased global white matter susceptibility and decreased global subcortical gray matter susceptibility in the injured cohort, relative to a matched control group, were reproduced. In addition, this study has provided preliminary evidence that acute QSM measurements may hold value for prognostication in a post-concussion setting. Future work will investigate the use of deep-learning applications, which recently have seen rapid progress from several groups,[37–39] in post-TBI QSM MRI reconstruction. The use of these new algorithms may allow for improved sensitivity of QSM measurements to lower grades of mTBI.

## Acknowledgments

Supported by the Department of Defense Broad Agency Announcement for Extramural Medical Research through award number W81XWH-14-1-0561.

The REDCap electronic database and the Adult Translational Research Unit used for this project were supported by the National Center for Advancing Translational Sciences, NIH, award number UL1TR001436.

This research was completed, in part, with computational resources and technical support provided by the Research Computing Center at the Medical College of Wisconsin.

## Authors' Contributions

K.M.K., A.S.N., and M.M. conceived the experiment(s). B.S. K.M.K. performed post-processing of data. A.B. provided consultation and review of statistical analyses. K.M.K., A.S.N, and T.B.M. performed statistical analyses. K.M.K. wrote manuscript. All authors reviewed the manuscript.

## Funding Information

Supported by the Department of Defense Broad Agency Announcement for Extramural Medical Research through award number W81XWH-14-1-0561.

The REDCap electronic database and the Adult Translational Research Unit used for this project were supported by the National Center for Advancing Translational Sciences, NIH, award number UL1TR001436.

This research was completed, in part, with computational resources and technical support provided by the Research Computing Center at the Medical College of Wisconsin.

## Author Disclosure Statement

No competing financial interests exist.

## Supplementary Material

Supplementary Table S1

## References

1. Faden, A.I. (2001). Neuroprotection and traumatic brain injury: the search continues. Arch. Neurol. 58, 1553–1555.
2. Teasdale, G., and Jennett, B. (1974). Assessment of coma and impaired consciousness: a practical scale. Lancet 2, 81–84.
3. Borg, J., Holm, L., Peloso, P.M., Cassidy, J.D., Carroll, L., Von Holst, H., Paniak, C., and Yates, D. (2004). Non-surgical intervention and cost for mild traumatic brain injury: results of the WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. J. Rehabil. Med. Suppl. 43, 76–83.
4. Ponsford, J., Willmott, C., Rothwell, A., Cameron, P., Kelly, A., Nelms, R., and Curran, C. (2002). Impact of early intervention on outcome following mild head injury in adults. J. Neurol. Neurosurg. Psychiatry 73, 330–332.
5. Prince, C., and Bruhns, M. (2017). Evaluation and treatment of mild traumatic brain injury: the role of neuropsychology. Brain Sci. 7, 105–114.
6. McMahon, P.J., Hricik, A., Yue, J.K., Puccio, A.M., Inoue, T., Lingsma, H.F., Beers, S.R., Gordon, W.A., Valadka, A.B., Manley, G.T., Okonkwo, O.O., and TRACK-TBI Investigators. (2014). Symptomatology and functional outcome in mild traumatic brain injury: results from the prospective TRACK-TBI study. J. Neurotrauma 31, 26–33.
7. Echemendia, R.J., Meeuwisse, W., McCrory, P., Davis, G.A., Putukian, M., Leddy, J., Makdissi, M., Sullivan, S.J., Broglio, S.P., Raftery, M., Schneider, K., Kissick, J., McCrea, M., Dvorak, J., Sills, A.K., Aubry, M., Engebretsen, L., Loosemore, M., Fuller, G., Kutcher, J., Ellenbogen, R, Guskiewicz, K., Patricios, J., and Herring, H. (2017). The Sport Concussion Assessment Tool 5th edition (SCAT5): background and rationale. Br. J. Sports Med. 51, 848–850.
8. McCrea, M., Guskiewicz, K., Randolph, C., Barr, W.B., Hammeke, T.A., Marshall, S.W., and Kelly, J.P. (2009). Effects of a symptom-free waiting period on clinical outcome and risk of reinjury after sport-related concussion. Neurosurgery 65, 876–882.
9. McCrory, P., Meeuwisse, W., Dvorak, J., Aubry, M., Bailes, J., Broglio, S., Cantu, R.C., Cassidy, D., Echemendia, R.J., Castellani, R.J., Davis, G.A., Ellenbogen, R., Emery, C., Engebretsen, L., Feddermann-Demont, N., Giza, C.C., Guskiewicz, K.M., Herring, S., Iverson, G.L., Johnston, K.M., Kissick, J., Kutcher, J., Leddy, J.J., Maddocks, D., Makdissi, M., Manley, G.T., Michael McCrea, Meehan, W.P., Nagahiro, S., Jon Patricios, J., Putukian, M., Schneider, K.J., Sills, A., Tator, C.H., Turner, M., and Vos, P.E. (2017). Consensus statement on concussion in sport—the 5th international conference on concussion in sport held in Berlin, October 2016. Br. J. Sports Med. 51, 838–847.
10. Clarke, L.P., Nordstrom, R.J., Zhang, H., Tandon, P., Zhang, Y., Redmond, G., Farahani, K., Kelloff, G., Henderson, L., Shankar, L., Deye, J., Capala, J., and Jacobs, P. (2014). The quantitative imaging network: NCI's historical perspective and planned goals. Transl. Oncol. 7, 1–4.
11. Shenton, M., Hamoda, H., Schneiderman, J., Bouix, S., Pasternak, O., Rathi, Y., Vu, M.A., Purohit, M.P., Helmer, K., Koerte, I., Lin, A.P., Westin, C.F., Kikinis, R., Kubicki, M., Stern, R.A., and Zafonte, R. (2012). A review of magnetic resonance imaging and diffusion tensor imaging findings in mild traumatic brain injury. Brain Imaging Behav. 6, 137–192.
12. Wallace, E.J., Mathias, J.L., and Ward, L. (2018). Diffusion tensor imaging changes following mild, moderate and severe adult traumatic brain injury: a meta-analysis. Brain Imaging Behav. 12, 1607–1621.
13. Lancaster, M., Olson, D., McCrea, M., Nelson, L., LaRoche, A.A., and Muftuler, L.T. (2016). Acute white matter changes following sport-related concussion: a serial diffusion tensor and diffusion kurtosis tensor imaging study. Hum. Brain Mapp. 37, 3821–3834.
14. Wang, Y., Nelson, L., LaRoche, A., Pfaller, A., Nencka, A.S., Koch, K.M., and McCrea, M.A. (2016). Cerebral blood flow alterations in acute sport-related concussion. J. Neurotrauma 33, 1227–1236.
15. Mayer, A., Bellgowan, P., and Hanlon, F. (2015). Functional magnetic resonance imaging of mild traumatic brain injury. Neurosci. Biobehav. Rev. 49, 8–18.
16. McCrea, M., Meier, T., Huber, D., Ptito, A., Bigler, E., Debert, C.T., Manley, G., Menon, D., Chen, J.K., Wall, R., Schneider, K.J., and Mcallister, T. (2017). Role of advanced neuroimaging, fluid biomarkers and genetic testing in the assessment of sport-related concussion: a systematic review. Br. J. Sports Med. 51, 919–929.
17. Koch, K.M., Meier, T.B., Karr, R., Nencka, A.S., Muftuler, L.T., and McCrea, M. (2018). Quantitative susceptibility mapping after sports-related concussion. Am. J. Neuroradiol. 39, 1215–1221.
18. Lin, H., Liu, H., Tsai, P., Hsu, F., Lu, C.F., Kao, Y.C., Hsieh, W.J., Huang, H.F., Chen, H.L., Blakeley, P., Lee, G.A., and Chen, C.Y. (2017). Quantitative susceptibility mapping in mild traumatic brain injury. Proceedings of the ISMRM, 2395.
19. Schweser, F., Kyyriainen, J., Preda, M., Pitkänen, A., Toffolo, K., Poulsen, A., Donahue, K., Levy, B., and Poulsen, D. (2019). Visualization of thalamic calcium influx with quantitative susceptibility mapping as a potential imaging biomarker for repeated mild traumatic brain injury. Neuroimage 200, 250–258.
20. Haacke, E.M., Cheng, N.Y., House, M.J., Liu, Q., Neelavalli, J., Ogg, R.J., Khan, A., Ayaz, M., Kirsch, W., and Obenaus, A. (2005). Imaging iron stores in the brain using magnetic resonance imaging. Magn. Reson. Imaging 23, 1–25.
21. Wang, Y., and Liu, T. (2015). Quantitative susceptibility mapping (QSM): decoding MRI data for a tissue magnetic biomarker. Magn. Reson. Med. 73, 82–101.
22. Reichenbach, J., Schweser, F., Serres, B., and Deistung, A. (2015). Quantitative susceptibility mapping: concepts and applications. Clin. Neuroradiol. 25, Suppl. 2, 225–230.
23. Liu, W., Soderlund, K., Senseney, J., Joy, D., Yeh, P., Ollinger, J., Sham, E.B., Liu, T., Wang, Y., Oakes, T.R., and Riedy, G. (2016). Imaging cerebral microhemorrhages in military service members with chronic traumatic brain injury. Radiology 278, 536–545.
24. Sun, H., and Wilman, A. (2014). Background field removal using spherical mean value filtering and Tikhonov regularization. Magn. Reson. Med. 71, 1151–1157.
25. Anderson, C., Nencka, A., Muftuler, T., Schmainda, K., and Koch, K. (2016). Volume parcellated quantitative susceptibility mapping. Proceedings of the ISMRM, 1108.
26. Liu, J., Liu, T., de Rochefort, L., Ledoux, J., Khalidov, I., Chen, W., Tsiouris, A., Wisnieff, C., Spincemaille, P., Prince, M., and Wang, Y. (2012). Morphology enabled dipole inversion for quantitative susceptibility mapping using structural consistency between the magnitude image and the susceptibility map. Neuroimage 59, 2560–2568.
27. Mazziotta, J., Toga, A., Evans, A., Fox, P., Lancaster, J., Zilles, K., Woods, R., Paus, T., Simpson, G., Pike, B., Holmes, C., Collins, L., Thompson, P., MacDonald, D, Iacoboni, M., Schormann, T., Amunts, K., Palomero-Gallagher, N., Geyer, S., Parsons, L., Narr, K., Kabani, N., Le Goualher, G., Feidler, J., Smith, K., Boomsma, D., Poi, H.H., Cannon, T., Kawashima, R., and Mazoyer, B.(2001). A four-dimensional probabilistic atlas of the human brain. J. Am. Med. Inform. Assoc. 8, 401–430.
28. Jenkinson, M., Beckmann, C., Behrens, T., Woolrich, M., and Smith, S. (2012). FSL. Neuroimage 62, 782–790.
29. Desikan, R.S., Ségonne, F., Fischl, B., Quinn, B.T., Dickerson, B.C., Blacker, D., Buckner, R.L., Dale, A.M., Maguire, R.P., Hyman, B.T., Albert, M.S., and Killiany, R.J. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31, 968–980.
30. Mori, S., Wakana, S., Van Zijl, P.C., and Nagae-Poetscher, L. (2005). *MRI Atlas of Human White Matter*. Elsevier: Amsterdam, Netherlands
31. Li, W., Wu, B., Batrachenko, A., Bancroft-Wu, V., Morey, R.A., Shashi, V., Langkammer, C., De Bellis, M.D., Ropele, S., Song, A.W.,

and Liu, C. (2014). Differential developmental trajectories of magnetic susceptibility in human brain gray and white matter over the lifespan. Hum. Brain Mapp. 35, 2698–2713.
32. Jarque, C.M., and Bera, A.K. (1987). A test for normality of observations and regression residuals. Int. Stat. Rev. 55, 163–172.
33. Japkowicz, N., and Stephen, S. (2002). The class imbalance problem: a systematic study. Intel. Data Anal. 6, 429–449.
34. Hanley, J.A., and McNeil, B.J. (1982). The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143, 29–36.
35. Osteen, C.L., Moore, A.H., Prins, M.L., and Hovda, D.A. (2001). Age-dependency of 45calcium accumulation following lateral fluid percussion: acute and delayed patterns. J. Neurotrauma 18, 141–162.
36. Bartnik-Olson, B.L., Holshouser, B., Wang, H., Grube, M., Tong, K., Wong, V., and Ashwal, S. (2014). Impaired neurovascular unit function contributes to persistent symptoms after concussion: a pilot study. J. Neurotrauma 31, 1497–1506.
37. Bollmann, S., Rasmussen, K.G., Kristensen, M., Blendal, R.G., Østergaard, L.R., Plocharski, M., O'Brien, K., Langkammer, C., Janke, A., and Barth, M. (2019). DeepQSM-using deep learning to solve the dipole inversion for quantitative susceptibility mapping. Neuroimage 195, 373–383.
38. Yoon, J., Gong, E., Chatnuntawech, I., Bilgic, B., Lee, J., Jung, W., Ko, J., Jung, H., Setsompop, K., Zaharchuk, G., Kim, E.Y., Pauly, J., and Lee, J. (2018). Quantitative susceptibility mapping using deep neural network: QSMNet. Neuroimage 179, 199–206.
39. Liu, J. and Koch, K. M. (2019). MRI tissue magnetism quantification through total field inversion with deep neural networks. Available at: arxiv.org. Last a ccessed: November 1, 2020.

Address correspondence to:
*Kevin M. Koch, PhD*
*Department of Radiology*
*Medical College of Wisconsin*
*8701 West Watertown Plank Road*
*Milwaukee, WI 53226*
*USA*

*E-mail:* kmkoch@mcw.edu