



Published in final edited form as:

*J Int Neuropsychol Soc.* 2022 August ; 28(7): 687–699. doi:10.1017/S1355617721000928.

## Test–Retest Reliability of a Semi-Structured Interview to Aid in Pediatric Traumatic Brain Injury Diagnosis

Danielle C. Hergert<sup>1</sup>, Veronik Sicard<sup>1</sup>, David D. Stephenson<sup>1</sup>, Sharvani Pabbathi Reddy<sup>1</sup>, Cidney R. Robertson-Benta<sup>1</sup>, Andrew B. Dodd<sup>1</sup>, Edward J. Bedrick<sup>2</sup>, Gerard A. Gioia<sup>3,4</sup>, Timothy B. Meier<sup>5</sup>, Nicholas A. Shaff<sup>1</sup>, Davin K. Quinn<sup>6</sup>, Richard A. Campbell<sup>6</sup>, John P. Phillips<sup>1,7</sup>, Andrei A. Vakhtin<sup>1</sup>, Robert E. Sapien<sup>8</sup>, Andrew R. Mayer<sup>1,6,7,9</sup>

<sup>1</sup>The Mind Research Network/Lovelace Biomedical Research Institute, Albuquerque, NM, USA

<sup>2</sup>Department of Epidemiology and Biostatistics, University of Arizona, Tucson, AZ, USA

<sup>3</sup>Department of Pediatrics and Psychiatry & Behavioral Sciences, George Washington University School of Medicine, Washington, DC, USA

<sup>4</sup>Division of Pediatric Neuropsychology, Children’s National Hospital, Washington, DC, USA

<sup>5</sup>Department of Neurosurgery, Medical College of Wisconsin, Milwaukee, WI, USA

<sup>6</sup>Department of Psychiatry and Behavioral Sciences, University of New Mexico, Albuquerque, NM, USA

<sup>7</sup>Department of Neurology, University of New Mexico, Albuquerque, NM, USA

<sup>8</sup>Department of Emergency Medicine, University of New Mexico Health Sciences Center, Albuquerque, NM, USA

<sup>9</sup>Department of Psychology, University of New Mexico, Albuquerque, NM, USA

### Abstract

**Objective:** Retrospective self-report is typically used for diagnosing previous pediatric traumatic brain injury (TBI). A new semi-structured interview instrument (New Mexico Assessment of Pediatric TBI; NewMAP TBI) investigated test–retest reliability for TBI characteristics in both the TBI that qualified for study inclusion and for lifetime history of TBI.

**Method:** One-hundred and eight-four mTBI (aged 8–18), 156 matched healthy controls (HC), and their parents completed the NewMAP TBI within 11 days (subacute; SA) and 4 months (early chronic; EC) of injury, with a subset returning at 1 year (late chronic; LC).

**Results:** The test–retest reliability of common TBI characteristics [loss of consciousness (LOC), post-traumatic amnesia (PTA), retrograde amnesia, confusion/disorientation] and post-concussion

---

Correspondence and reprint requests to: Andrew R. Mayer, Mind Research Network, Pete and Nancy Domenici Hall, 1101 Yale Blvd. NE, Albuquerque, NM 87106, USA. amayer@mrn.org.

#### SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S1355617721000928>.

#### CONFLICTS OF INTEREST

The authors have nothing to declare.

symptoms (PCS) were examined across study visits. Aside from PTA, binary reporting (present/absent) for all TBI characteristics exhibited acceptable (0.60) test–retest reliability for both Qualifying and Remote TBIs across all three visits. In contrast, reliability for continuous data (exact duration) was generally unacceptable, with LOC and PCS meeting acceptable criteria at only half of the assessments. Transforming continuous self-report ratings into discrete categories based on injury severity resulted in acceptable reliability. Reliability was not strongly affected by the parent completing the NewMAP TBI.

**Conclusions:** Categorical reporting of TBI characteristics in children and adolescents can aid clinicians in retrospectively obtaining reliable estimates of TBI severity up to a year post-injury. However, test–retest reliability is strongly impacted by the initial data distribution, selected statistical methods, and potentially by patient difficulty in distinguishing among conceptually similar medical concepts (i.e., PTA vs. confusion).

### Keywords

Pediatrics; Adolescent; Traumatic brain injury; Brain concussion; Self-reports; Test–retest reliability

---

## INTRODUCTION

Pediatric traumatic brain injury (TBI) represents a large public health concern, with an estimated 750,000 new cases of pediatric mild TBI (mTBI) occurring each year (Zemek et al., 2017). Most TBI diagnostic criteria, such as the World Health Organization or American Congress of Rehabilitation Medicine, require that individuals experience an external force to the head followed by some alteration in mental status (Mayer, Quinn, & Master, 2017). In contrast, other diagnostic criteria such as the Concussion in Sport Group guidelines only require the emergence of one or more post-concussion symptoms to diagnose mTBI, even in the absence of altered mental status (McCrory et al., 2017). Indeed, recent studies highlight the variability that exists in both diagnosing whether a child experienced an mTBI (7.1–98.7% classified as mTBI across 17 definitions; Crowe et al., 2018) or experienced persistent symptomatology (48.7% range across six definitions; Mayer et al., 2020). For this study, the terms mTBI and concussion will be used synonymously.

Although recent blood-based biomarkers (Bazarian et al., 2018) and imaging findings have shown diagnostic potential (Mayer, Kaushal, et al., 2018), TBI, and especially mTBI, is typically diagnosed based on retrospective self-report of injury characteristics such as duration of loss of consciousness (LOC), post-traumatic amnesia (PTA), retrograde amnesia (RGA), and the presence of new symptoms in lieu of medical records (Carroll, Cassidy, Holm, Kraus, & Coronado, 2004; Mayer et al., 2017; McCrory et al., 2017; Menon, Schwab, Wright, & Maas, 2010). Semi-structured interviews improve recall and elicit more accurate information about diagnostic characteristics of TBI relative to self-report measures, thereby improving the validity of self-reported information (Corrigan & Bogner, 2007; Vanderploeg, Groer, & Belanger, 2012). Semi-structured interviews also exhibit good to excellent inter-rater reliability (Corrigan & Bogner, 2007) and improve test–retest reliability of recalled information in adults (Bogner & Corrigan, 2009; Bogner et al., 2017). However, no such

semi-structured interviews currently exist for use in pediatric populations for retrospective diagnosis of TBI beyond acute care settings (Gioia, Collins, & Isquith, 2008).

Test–retest reliability is a critical psychometric property for understanding the stability of a construct over time (Lexell & Downham, 2005). Eighty-five percent of military service members were consistent in reporting whether or not they experienced altered mental status (i.e., confusion/disorientation), LOC, or if they remembered the head injury up to 1-year post-deployment (Nelson et al., 2015). A civilian TBI (Sherer et al., 2015) study showed that most participants self-reported longer LOC (66%) and PTA (84%) durations than medical records, potentially impacting diagnostic accuracy (e.g., moderate-to-severe TBI). Similarly, athletes with orthopedic injuries are more accurate in reporting on global information, such as whether or not they had an injury, rather than reporting specifics, such as injury location and even diagnoses (Gabbe, Finch, Bennell, & Wajswelner, 2003; Vanderlei et al., 2017). Indeed, several factors may affect reliability, such as time since injury, biases in recall (e.g., good old days), and the severity of the head injury and impact on memory consolidation.

The aim of the current study was therefore to assess the test–retest reliability for the retrospective assessment of self-reported TBI history through a newly developed, pediatric-focused, semi-structured interview (New Mexico Assessment of Pediatric Traumatic Brain Injury; NewMAP TBI; see Appendix A). The NewMAP TBI purposefully incorporated the Post-Concussion Symptom Inventory (PCSI; Sady, Vaughan, & Gioia, 2014), a validated pediatric-specific assessment of symptomatology following trauma, to avoid establishing a new assessment of common PCS. The psychometric properties of TBI characteristics (i.e., number of head injuries, LOC, PTA, RGA, alterations of mental status, and PCS) obtained during a semi-structured interview were assessed for both the instance of mTBI resulting in study enrollment (hereafter referred to as the Qualifying mTBI, injured group only) and lifetime history of TBI (hereafter referred to as Remote TBIs, for both injured and healthy control [HC] groups).

We hypothesized that test–retest reliability coefficients would be higher for binary (presence *vs.* absence) and categorical characterization of injury characteristics relative to self-reported durations (Gabbe et al., 2003; Sherer et al., 2015; Vanderlei et al., 2017). Multiple statistical techniques were used to quantify test–retest reliability due to non-normality (Mayer et al., 2020) and the different distributions of data types (binary *vs.* categorical *vs.* continuous). Finally, PCS are typically greatest at the time of injury and decrease with recovery (Eisenberg, Meehan, & Mannix, 2014; Yeates et al., 2009). Thus, as a preliminary check of the construct validity of self-reported symptoms, we hypothesized that PCS burden shortly after the injury would be higher than symptom burden on both the day of formal assessment (approximately 1 week post-injury in the current study) and when compared to retrospectively rated symptoms (1 month prior to injury).

## METHODS

### Participants

Children and adolescents with mTBI (8–18 years) were consecutively recruited from local emergency department and urgent care settings between July 2016 and February 2020

for this prospective cohort study. Inclusion criteria were based on the American Congress of Rehabilitation Medicine (Kay et al., 1993) and Concussion in Sport Group guidelines (McCroory et al., 2017), with the former being used as the higher threshold and the latter as the lower threshold for sustaining an mTBI. Specifically, all mTBI participants experienced head trauma and one of the following: an alteration in mental status, LOC (if present)  $\geq 30$  min, PTA (if present)  $\geq 24$  hr, or at least two new PCS. Glasgow Coma Scale Scores, if available, were  $\geq 13$ . All Qualifying mTBI were diagnosed by qualified medical personnel at the time of hospital admission and confirmed by the senior author (ARM). Matched HC were recruited from the local community through flyers and word of mouth.

The larger parent study is ongoing. Participants were included in this study (Figure 1) if they had completed the first two study visits by the end of February 2020. Exclusion criteria for the mTBI group were: (1) a history of (a) a neurological diagnosis other than TBI, (b) previous moderate or severe TBI with  $>30$  min LOC, (c) developmental disorder (i.e., autism spectrum disorder or intellectual disability), (d) any psychiatric disorders other than adjustment disorder, or (e) substance abuse/dependence; (2) a non-English-speaking child/adolescent or guardian; or (3) a positive urine-based drug screen at any of the three visits. HC had identical exclusion criteria as well as: (1) diagnosis of attention-deficit/hyperactivity disorder or a learning disability and (2) sustained mTBI within 6 months of their initial visit. All participants provided informed consent or assent according to institutional guidelines at the University of New Mexico School of Medicine. All study procedures were completed in accordance with the Helsinki Declaration.

## Materials and Procedures

Trained staff members administered the NewMAP TBI (see Appendix A and Figure 2) in tandem to both the parent and child/adolescent at subacute (SA:  $7.43 \pm 2.17$  days post-injury; range=1–11), early chronic (EC:  $130.66 \pm 15.15$  days post-injury; range=101–194), and late chronic (LC:  $371.17 \pm 34.23$  days post-injury; range=238–489) post-injury phases in an outpatient setting. Staff read the interview prompts and recorded participants' verbatim responses electronically. The NewMAP TBI is conceptually similar to existing semi-structured interviews for adults, such as the Ohio State University TBI-ID, the Virginia Commonwealth University retrospective concussion diagnostic interview, blast version, and an adult version of this interview (Corrigan & Bogner, 2007; Mayer, Hanlon, et al., 2018; Walker et al., 2015), but incorporated pediatric-specific scales (i.e., PCSI) and language. The first step of the NewMAP TBI involved a global recording of the four most recent potential head and/or neck injuries. Information obtained during the semi-structured interview includes a narrative of the circumstances related to the head injury, approximate date and year of the injury, type of injury classified based on previously published categories (e.g., fall vs. motor vehicle crash, Haarbauer-Krupa et al., 2018), whether or not the injury was sport- or recreation-related (Rice, 2008), and details about injury severity (i.e., LOC, PTA, RGA, confusion/disorientation). The trained technician then determined whether the injury met minimal diagnostic criteria for mTBI (McCroory et al., 2017). If criteria were met, both the parent and child/adolescent completed a more thorough assessment of the PCS experienced acutely following the injury using a slightly modified version of the PCSI (i.e., temporal flag of "Yesterday and Today" removed) for adolescents (13–18 years old),

hereafter referred to as the acute-PCSI measurement. When more than one TBI was reported at any of the visits, two raters (authors DH and VS) separately read each description and indicated which reported injuries were the same across study visits.

In addition to the semi-structured interview, all children/adolescents and parents independently completed separate and age-appropriate modified child/adolescent and parent versions of the PCSI self-report questionnaire (see Supplemental Materials) to rate PCS retrospectively 1 month prior to the SA visit (R-PCSI) and for the day of the initial visit (SA-PCSI). The PCSI was modified with author (GG) permission as follows: (1) the version of the PCSI for 13–18 years old (adolescent version) was also administered to 12-year-old participants to equate the response scale across the 1-year follow-up and (2) all references to an injury were removed from questionnaires and instructions to avoid response bias in HC. Participants who were 8–11 years old completed the older child self-report form. The following formula (Barlow, Crawford, Brooks, Turley, & Mikrogianakis, 2015) was used to equate symptom burden for analyses given the different number of items between parent and child/adolescent versions of the PCSI (child range: 0–21; parent range: 0–20): PCSI symptom burden scores = (Total PCSI/Total Possible PCSI Score) × 100.

Participants also completed a short version of the Test of Memory and Malinger (TOMMe10) as an independent measure of performance validity (Denning, 2012; see Figure 2 for study procedures).

### Statistical Analyses

T-test and chi-square tests compared demographic variables and the number of individuals scoring below recommended cut offs on a measure of performance validity (TOMMe10) between mTBI and HC, as well as those who completed three *versus* two visits. Four main TBI characteristics were examined for both Qualifying and Remote TBIs: LOC, PTA, RGA, and confusion/disorientation. Analyses of Qualifying TBI data were limited to mTBI participants to better control for time since injury, while Remote TBI analyses included both groups. Each TBI characteristic was examined using a range of precision: (1) self-reported duration of time (continuous converted to minutes), (2) a conversion of self-reported duration to categorical variables, and (3) binary endorsement for present or absent. Categorical classifications for LOC, PTA, RGA, and confusion/disorientation were identical to a previous adult study (Mayer, Hanlon, et al., 2018) and based on previously published (Lezak, 1995; Ruff & Jurica, 1999) categorical classifications (see Table 1).

The distributions for LOC, PTA, RGA, confusion/disorientation, number of symptoms (range=0–21), and symptom burden (range=0–126) were zero-inflated and positively skewed. Data were, therefore, log-transformed following the addition of a constant. Intraclass correlation coefficients (Shrout & Fleiss, 1979) were used to examine test–retest reliability [two-way random effects, absolute agreement, single measurement; ICC(2,1)] for continuous self-report variables. Gwet’s AC<sub>1</sub> estimation evaluated test–retest reliability of binary variables (Gwet, 2008). Gwet’s AC<sub>2</sub> estimation evaluated test–retest reliability of categorical variables and the reported number of Remote TBIs. All reliability estimates were categorized as poor (< 0.39), fair (0.40–0.59), good (0.60–0.74), or excellent (> 0.75) based on previous guidelines (Cicchetti, 2001), with reliability coefficients of > 0.60 (i.e.,

good or excellent ranges) operationally defined to be acceptable. Because estimates of test–retest reliability (i.e., Gwet’s  $AC_1/AC_2$  and ICCs) are not mathematically equivalent, reliability coefficients for all variables (binary, categorical, and continuous) and algorithms are presented for thoroughness in Supplemental Materials.

Secondary analyses compared the acute-PCSI (derived from the NewMAP TBI) to the R-PCSI and SA-PCSI using generalized estimating equations (GEE) with a negative binomial distribution to establish preliminary evidence for construct validity (i.e., highest symptom burden at the time of injury). These analyses were limited to participants aged 12 and older to minimize the impact of parent–child version differences in scoring (Mayer et al., 2020).

## RESULTS

### Participants

One-hundred and eight-four mTBI and 156 statistically matched (age/sex) HC were included in the study. HC parents had significantly higher educational attainment than mTBI parents. Additionally, fewer HC scored below the recommended cut off on a measure of performance validity (See Table 2). There were no differences in demographic or severity of injury characteristics between mTBI completing all three *versus* two visits (all  $p$ 's > 0.10; See Supplemental Table 1). Eleven HC reported a previous Remote TBI at the SA visit. Ten mTBI reported new head injuries that occurred between visits. New injuries sustained after the first visit were not included in the following reliability analyses. Importantly, some mTBI reported LOC or PTA outside the mild severity range for their Qualifying or Remote TBI during the NewMAP TBI administration, along with other anomalies (see Supplemental Materials for full details). These data points were purposefully included in the analyses given the primary test–retest reliability aim.

### Qualifying TBI Analyses

The primary mechanisms of injury (Haarbauer-Krupa et al., 2018) for the Qualifying TBI were motor vehicle crashes (24.7%), falls (30.8%), strikes by objects (15.4%), strikes by people (18.7%), assaults (4.4%), bicycle accidents (5.5%), and other types of injuries (0.5%), with 61% sport- or recreation-related. Binary LOC, RGA, and confusion/disorientation variables had good to excellent reliabilities (See Tables 1 and 3), while binary PTA variables had fair reliabilities. For categorical variables, LOC, PTA, and RGA variables exhibited good to excellent test–retest reliabilities across visits, while categorical confusion/disorientation variables had fair reliabilities. Continuous LOC variables had good reliability from SA to EC, excellent reliability from EC to LC, and fair reliability from SA to LC. Conversely, continuous PTA, RGA, and confusion/disorientation variables had poor to fair reliabilities. Finally, the acute-PCSI number of symptoms and symptom burden scores exhibited fair to good reliabilities across all visits. Supplemental Table 2 provides reliability coefficients (Gwet’s statistics, ICC and transformed ICC) for all combinations of TBI characteristics and algorithms for the Qualifying TBI.

Secondary analyses revealed the expected, significant linear trend in symptom burden for both parent (Wald  $\chi^2=115.81$ ;  $p<0.001$ ) and adolescent (Wald  $\chi^2=91.22$ ;  $p<0.001$ ) PCSI



ratings (see Figure 3). Specifically, the highest PCSI symptom burden scores (all  $p$ 's<0.001) were reported on the acute-PCSI ( $32.89 \pm 18.58$ ) followed by the SA-PCSI ( $20.49 \pm 20.39$ ) and R-PCSI ( $7.97 \pm 14.36$ ) for parent-rated version. Similarly, the acute-PCSI had the highest symptom burden scores ( $32.66 \pm 18.52$ ) relative to SA-PCSI ( $22.14 \pm 20.35$ ) and R-PCSI ( $14.53 \pm 18.75$ ; all  $p$ 's<0.001) for the adolescent-rated version.

### The Effect of Different Parents/Guardians

One potential source of variability is different parents or guardians completing the NewMAP TBI with the child/adolescent across multiple visits. However, test–retest reliabilities (Supplemental Table 3) for the reduced sample with different parent raters ( $n=30$ ) were generally consistent with the overall sample. Exceptions included continuous LOC, which was reduced across all time points from the good to fair range from SA to EC, fair to poor range for SA to LC, and excellent to fair range from EC to LC. Continuous PTA reliability coefficients also were reduced from fair to poor from SA to EC and from EC to LC. Lastly, acute-PCSI number of symptoms and symptom burden scores had good reliability from SA to EC and EC to LC; however, reliability was poor for acute-PCSI number of symptoms while symptom burden had fair reliability from SA to LC.

### Remote TBIs Analyses

Test–retest reliability coefficients for the number of Remote TBIs were excellent across the three visits (See Table 3) for the overall sample as most individuals (85.8%) reported no Remote TBI history. For the subset of individuals who reported a history of Remote TBI at any of the three visits, the test–retest reliability was good from SA to EC, fair from SA to LC, and excellent from EC to LC (Table 3). Notably, of the 82 Remote TBIs that were reported during at least one of the visits, 36 (44%) were not consistently reported across each visit. In examining the reliability of Remote TBIs characteristics that were consistently reported, all binary variables had good to excellent reliabilities. Categorical LOC, PTA, and RGA all had excellent reliabilities, while confusion/disorientation variables had fair to good reliability. Compared to reliabilities for the Qualifying TBI analyses, reliabilities for continuous variables for LOC decreased across visits, with fair reliability from SA to EC, poor reliability from SA to LC, and good reliability from EC to LC. For PTA, the reliability of SA to EC and SA to LC was similar to the Qualifying TBI analyses; however, the reliability from EC to LC improved to the good range. Both RGA and confusion/disorientation continuous variables had similar reliability to the Qualifying TBI, except reliability improved from SA to LC from poor to good for RGA and poor to excellent for confusion/disorientation. Reliability also improved from fair to good for RGA for EC to LC. The reliability of acute-PCSI number of symptoms and symptom burden scores at the time of injury was fair to excellent across visits (See Table 3). Supplemental Table 4 provides reliability coefficients (Gwet's statistics, ICC and transformed ICC) for all combinations of TBI characteristics and algorithms for the Remote TBI.

### Impact of Reliability Algorithms

The statistical properties (zero inflation, skew) of the TBI characteristics could potentially impact reliability, and estimates of test–retest reliability (i.e., Gwet's  $AC_1/AC_2$  and ICCs) are not mathematically equivalent. Therefore, reliability estimates were compared across

algorithms to determine the impact of these factors on the results. Specifically, Supplemental Tables 2 and 4 indicate that Gwet's  $AC_1$  and  $AC_2$  produced the same values for binary variables as expected. For both the Qualifying and Remote TBI, Gwet's  $AC_2$  typically resulted in the highest value among all algorithms, although this was likely due to differences in the algorithms (e.g., Gwet's  $AC_2$  vs. ICC; see the following paragraph) as well as reflecting increased reliability for the transformed categorical data. Specifically, when the continuous variables were converted into categories, ICC values increased. When using a base 10 logarithmic transformation on the continuous data with an added constant of 1, ICC values also increased if the original distribution was skewed and zero inflated (e.g., LOC, PTA). When the original distribution was normal (acute-PCSI data), log-transformed and original data produced similar ICCs.

Calculations on simulated data were therefore conducted to determine how estimated reliabilities for both ICC and Gwet's  $AC_1/AC_2$  can be moderated by different data types (binary, categorical, and continuous), which have different distributions and varying degrees of "true" correlations. For continuous distributions, values were randomly sampled from a bivariate normal distribution across varying levels of correlation strength (i.e., the ICC; 0.0–0.9 in increments of 0.1). This sample distribution was then transformed by converting all negative values to zeros, thus mirroring the obtained data from the semi-structured interview, in which half of the bivariate sample report zeros. Simulation of categorical variables was performed in two different ways with zeroes from the previously created distribution remaining as zeroes and the remaining values split into four categories as either an even distribution or a distribution weighted toward lower ordinal categories.

Simulation of binary variables was achieved by splitting the original bivariate normal distribution below (0) and above (1) the medians for the X and Y values. Supplemental Table 5 shows how reliability estimates were impacted by applying different types of transformations to the bivariate normal distribution sample. The calculations showed that converting zero-inflated continuous data into a binary variable attenuated the ICC. However, the categorization of the zero-inflated data gives results comparable to the zero-inflated continuous data. For binary data,  $AC_1$  equaled  $AC_2$ .  $AC_2$  is much larger than ICC for categorized data when ICC was small or moderate. The  $AC_2$  increased with the value of ICC but not as rapidly as the ICC based on categorized scores.

## DISCUSSION

Several diagnostic schemes (e.g., ACRM, WHO) predominantly rely on TBI characteristics to determine the severity level (e.g., mild vs. moderate; Carroll et al., 2004; Mayer et al., 2017; McCrory et al., 2017; Menon et al., 2010), with recent expert consensus panels basing diagnosis predominantly on post-injury symptom burden for mTBI (McCrory et al., 2017). However, there are currently no validated semi-structured interviews for assessing pediatric TBI. Only a handful of studies have evaluated the psychometric properties of recalled information for commonly reported TBI characteristics in adults (Bogner et al., 2017; King et al., 1997; Mayou, Black, & Bryant, 2000; McMillan, Jongen, & Greenwood, 1996; Nelson et al., 2015; Roberts, Spitz, Mundy, & Ponsford, 2019; Roberts, Spitz, &



Ponsford, 2016; Sherer et al., 2015), and no studies have examined these factors in pediatric TBI.

Current results indicated that for both Qualifying and Remote TBIs, binary reporting (present/absent) of confusion/disorientation, LOC, and RGA had acceptable test–retest reliabilities, whereas the binary reliability for PTA was below the acceptable range. In contrast, the majority of TBI characteristics did not achieve acceptable test–retest reliabilities when durations were rated on a continuous basis. LOC was the only exception, which achieved acceptable reliability in half of the time points assessed for both Qualifying and Remote TBIs when data were log-transformed to account for skew (see Table 3). Importantly, LOC, RGA, and PTA all had acceptable test–retest reliabilities when continuous self-report data were separated into more discrete categories based on previously published criteria (Lezak, 1995; Ruff & Jurica, 1999), whereas categorical ratings of disorientation/confusion were below the acceptable range (see discussion on distribution properties). Test–retest reliabilities for TBI characteristics were not strongly affected when children/adolescents were accompanied by different parents at each appointment. Thus, current results suggest that continuous self-report may contribute to inconsistency in the diagnosis of pediatric TBI severity, which has recently been highlighted as a major clinical concern (Crowe et al., 2018).

Although preliminary support for the construct validity of retrospective ratings of PCS severity was obtained (acute injury > subacute period > pre-injury period), the recall of PCS burden (number of symptoms and total burden) exhibited acceptable reliability in only approximately half of the instances as a function of time post-injury for both Remote and Qualifying TBIs. Symptoms mimicking PCS frequently occur in healthy individuals as well as other non-TBI-related conditions such as learning disabilities, attention deficit hyperactive disorder, and depression (Iverson & Lange, 2003; Iverson et al., 2015). This increases the challenges of correctly recalling when symptoms truly occurred post-injury relative to a variety of documented recall biases (Caplan et al., 2016; Gunstad & Suhr, 2001; Lange, Iverson, & Rose, 2010; Ruff, Iverson, Barth, Bush, & Broshek, 2009). Notably, the time between appointments (e.g., 4 months or 1 year) was associated with slight decreases in reliability coefficients. This is consistent with previous studies that report much higher test–retest reliabilities across 1–6 week reporting windows (Mailer, Valovich-McLeod, & Bay, 2008; Mayer et al., 2020; Sady et al., 2014).

Current results also suggest that the underlying statistical properties (zero inflation, skew) of the TBI characteristics and the choice of reliability algorithm may have a large impact on test–retest reliability coefficients. Specifically, as exhibited in Table 1, most TBI characteristics have a zero inflated (i.e., high percentage of cases with “None” responses), positively skewed distribution. The ICC is the most commonly used metric to evaluate test–retest reliability but is designed for continuous, normally distributed data (Lexell & Downham, 2005). Log transformations of the continuous self-report data improved reliability in some instances (i.e., LOC, PTA) due to reduced skew of the data. The ICC, however, is not appropriate for estimating reliabilities of binary or categorical responses, as these data require an approach that takes into account the overall agreement of endorsing a particular category across time (Gwet, 2008). Therefore, the statistical properties of the

TBI characteristic being assessed should drive the choice of the appropriate algorithm for determining test–retest reliability coefficients.

Similar to past studies of mTBI (Castile, Collins, McIlvain, & Comstock, 2012; Delaney, Lacroix, Leclerc, & Johnston, 2002; Guskiewicz et al., 2003), confusion/disorientation was the most commonly endorsed TBI characteristic (82.4% of the sample) followed by LOC (49.4%), PTA (37.8%), and RGA (15.0%). These statistical properties may explain why confusion/disorientation’s test–retest reliability was highest when it was binarized, whereas reliability for other variables improved following categorical transformations. Similarly, calculations based on numerical simulations indicated that the rate of convergence between the ICC and Gwet’s statistics were partially determined by the initial distributions of the data (frequencies of responses in different categories), illustrating the complex interplay that occurs between the data distribution, algorithm choice, and reliability coefficients.

Another potential explanation for differences in reliability coefficients is that participants may not understand nuances associated with different TBI characteristics, and/or may conflate similar symptoms. It can be challenging to distinguish PTA from attention deficits, confusion, and/or psychogenic amnesia or dissociation post-injury from a lay perspective (Corrigan & Bogner, 2007; Ruff et al., 2009; Stuss et al., 1999; Vanderploeg et al., 2012). Similarly, disorientation (e.g., “Where are you right now?”) can also occur as a result of RGA (Ruff et al., 2009). Furthermore, prior to the full recovery of continuous, episodic memory, individuals may have isolated memories (e.g., brief memory of being transported to the hospital, but no memory again until hospital discharge) that may be forgotten over time, which can lead to inconsistent reporting of PTA (Roberts et al., 2019). The term “post-traumatic confusional state” has been proposed instead of “post-traumatic amnesia” to encompass all of these medical conditions (Kristman et al., 2014; Stuss et al., 1999). In contrast, a LOC, or lack thereof, may be more obvious (bystanders) or memorable to nonmedical personnel. Previous studies using semi-structured interviews show adequate test–retest reliability of LOC in adult civilian (Bogner et al., 2017) and prison (Bogner & Corrigan, 2009) populations with moderate to severe TBI when asked if LOC exceeded 30 min. Self-reported Remote TBI history, including whether or not individuals sustained LOC, has been shown to have adequate test–retest reliability in older adults as well (Wilmoth et al., 2018). However even in adults, test–retest reliability for LOC duration within the mild range has not been thoroughly investigated (Corrigan, Yang, Singichetti, Manchester, & Bogner, 2018).

Finally, most participants did not report a history of Remote TBIs (85.8%), resulting in excellent test–retest reliability for this characteristic across the entire sample. Test–retest reliability was much lower when limiting analyses to individuals who reported a history of Remote TBIs, with a significant portion of the sample failing to report the same number of Remote injuries across all three visits (36/82, 44% of Remote TBIs not reported across all visits). This can be concerning for both clinical work and research in the assessment of multiple TBIs, particularly when trying to determine whether multiple TBIs are associated with worse or long-term impairment (Kerr, Thomas, Simon, McCrea, & Guskiewicz, 2018; Manley et al., 2017).

## Limitations

Although medically diagnosed with mTBI for their Qualifying injury, a small portion of our sample (5.9%) reported LOC or PTA outside the mild severity range during at least one of their visits. These cases were included given the aim of the current manuscript, with similar findings reported in adult samples (Sherer et al., 2015). Because all participants were seen in emergency care settings and diagnosed with mTBI, additional studies are required to examine reliability of self-report among children and adolescents with either more severe TBI or in other medical settings. Furthermore, the number of lifetime TBIs reported in this sample was low, so that results may not generalize to a sample of participants with a high number of lifetime TBIs. Inter-rater reliability could not be assessed as rater information was not recorded. This limitation needs to be addressed in subsequent studies. Also, as discussed in detail in the sections above, there are many statistical perils for trying to establish rates of test–retest reliability among non-normally distributed data types. Additional methodological studies are required to determine the full impact of these limitations. Finally, while the study did include a performance validity measure, there was no measure of symptom validity and therefore information is unavailable regarding potential under- or overreporting of self-reported symptoms.

## Future Directions

Based on current findings, slight alterations will be made to the NewMAP TBI to improve the reliability of self-reported TBI characteristics (updated version presented in Appendix A). This includes providing categories for the duration of TBI characteristics to participants since current results suggest that the exact duration of reporting was generally unreliable. In the instructions for PTA and RGA, additional formal prompts were provided for interviewers (e.g., “What was the first/last memory you remembered after/before the injury? Do you remember being on site of the injury/in the car/in the hospital?”). These questions may elicit a more accurate temporal reporting of PTA and RGA to better help individuals understand the differences among these phenomena relative to confusion (Ruff et al., 2009; Vanderploeg et al., 2012). Additional studies in different pediatric TBI samples (e.g., with premorbid conditions such as the presence of psychiatric or neurodevelopmental disorders, substance use) will be needed to further examine the reliability and validity of the NewMAP TBI under these use cases.

## CONCLUSIONS

In summary, reliable self-report is critical for accurate diagnosis and prognosis of TBI (Asken et al., 2016; Haran et al., 2016; McCrory et al., 2017; Sarmiento, Gioia, Kirkwood, Wade, & Yeates, 2019). Binary or categorical variables of TBI characteristics for both Qualifying and Remote TBIs had acceptable reliabilities, whereas most continuously reported data were at or below traditional levels of acceptable reliability. Although slight reductions based on the time from injury were observed, test–retest reliability of the NewMAP TBI was relatively robust to duration between assessments (4 months vs. 1 year) and different parent/child raters. Current results suggest that discretizing continuous variables may provide the best balance between psychometric properties and clinical information that are needed to diagnose injury severity.

## Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

## Acknowledgments

### FUNDING

This research was supported by grants from the National Institutes of Health to Andrew R. Mayer (grant numbers NIH 01 R01 NS098494-01A1, R01 NS098494-03S1A1, and P30 GM122734). The NIH had no role in study review, data collection and analysis, decision to publish, or preparation of the manuscript.

## REFERENCES

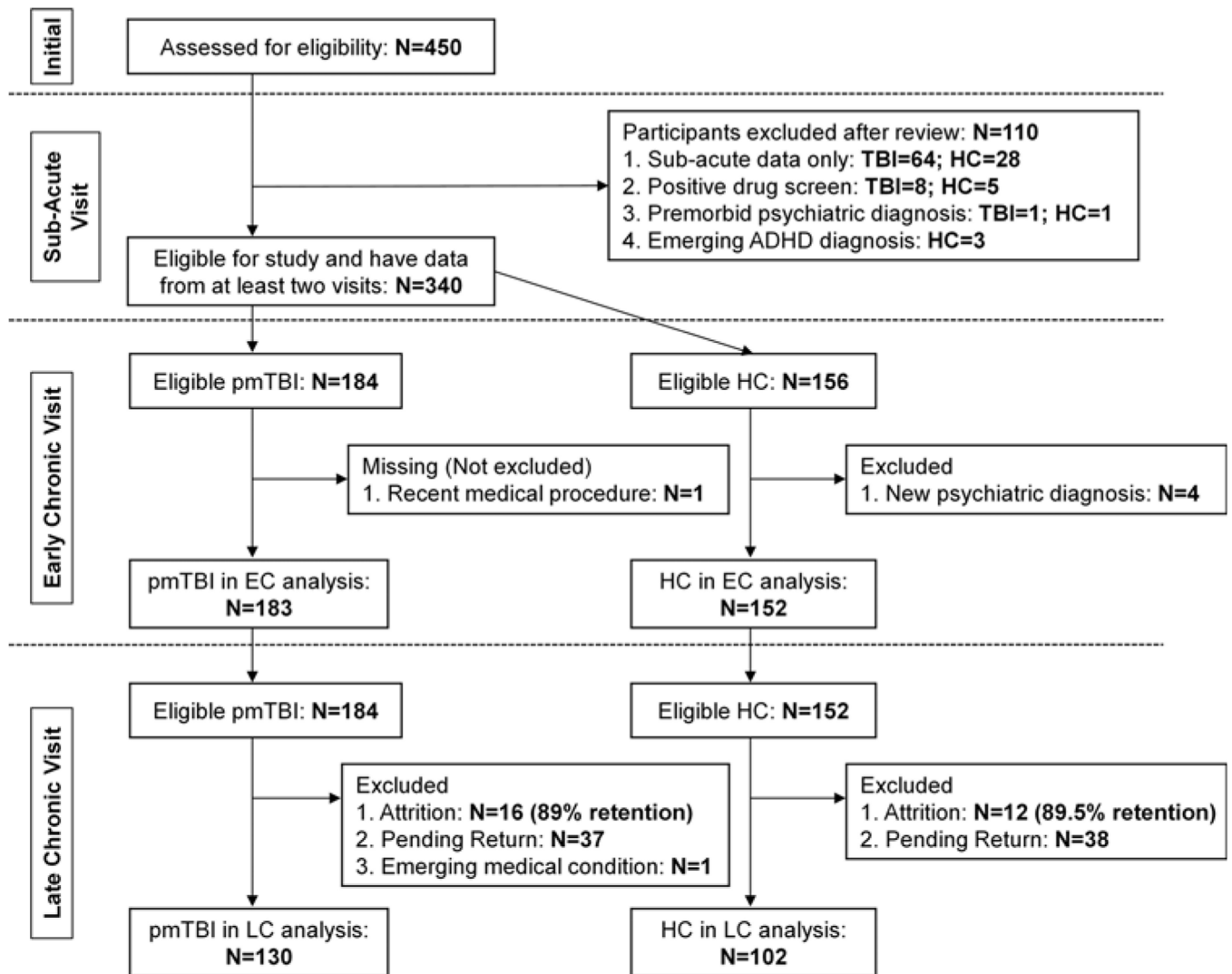
- Asken BM, McCrea MA, Clugston JR, Snyder AR, Houck ZM, & Bauer RM (2016). "Playing Through It": Delayed reporting and removal from athletic activity after concussion predicts prolonged recovery. *Journal of Athletic Training*, 51(4), 329–335. Retrieved from PM:27111584 [PubMed: 27111584]
- Barlow KM, Crawford S, Brooks BL, Turley B, & Mikrogianakis A. (2015). The incidence of postconcussion syndrome remains stable following mild traumatic brain injury in children. *Pediatric Neurology*, 53(6), 491–497. Retrieved from PM:26421987 [PubMed: 26421987]
- Bazarian JJ, Biberthaler P, Welch RD, Lewis LM, Barzo P, Bogner-Flatz V, . . . Jagoda AS (2018). Serum GFAP and UCH-L1 for prediction of absence of intracranial injuries on head CT (ALERT-TBI): A multicentre observational study. *Lancet Neurology*, 17(9), 782–789. Retrieved from PM:30054151 [PubMed: 30054151]
- Bogner J & Corrigan JD (2009). Reliability and predictive validity of the Ohio State University TBI identification method with prisoners. *Journal of Head Trauma Rehabilitation*, 24(4), 279–291. Retrieved from PM:19625867 [PubMed: 19625867]
- Bogner JA, Whiteneck GG, MacDonald J, Juengst SB, Brown AW, Philippus AM, . . . Corrigan JD (2017). Test-retest reliability of traumatic brain injury outcome measures: A traumatic brain injury model systems study. *Journal of Head Trauma Rehabilitation*, 32(5), E1–E16. Retrieved from PM:28195954
- Caplan B, Bogner J, Brenner L, Arciniegas D, Silverberg ND, Iverson GL, . . . Aquino A. (2016). The nature and clinical significance of preinjury recall bias following mild traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 31(6), 388–396. [PubMed: 26580693]
- Carroll LJ, Cassidy JD, Holm L, Kraus J, & Coronado VG (2004). Methodological issues and research recommendations for mild traumatic brain injury: The WHO Collaborating Centre Task Force on Mild Traumatic Brain Injury. *Journal of Rehabilitation Medicine* (43 Suppl.), 113–125. Retrieved from PM:15083875
- Castile L, Collins CL, McIlvain NM, & Comstock RD (2012). The epidemiology of new versus recurrent sports concussions among high school athletes, 2005–2010. *British Journal of Sports Medicine*, 46(8), 603–610. Retrieved from PM:22144000 [PubMed: 22144000]
- Cicchetti DV (2001). The precision of reliability and validity estimates re-visited: Distinguishing between clinical and statistical significance of sample size requirements. *Journal of Clinical and Experimental Neuropsychology*, 23(5), 695–700. Retrieved from PM:11778646 [PubMed: 11778646]
- Corrigan JD & Bogner J. (2007). Initial reliability and validity of the Ohio State University TBI Identification Method. *Journal of Head Trauma Rehabilitation*, 22(6), 318–329. Retrieved from PM:18025964 [PubMed: 18025964]
- Corrigan JD, Yang J, Singichetti B, Manchester K, & Bogner J. (2018). Lifetime prevalence of traumatic brain injury with loss of consciousness. *Injury Prevention*, 24(6), 396–404. [PubMed: 28848057]
- Crowe LM, Hearps S, Anderson V, Borland ML, Phillips N, Kochar A, . . . Babl FE (2018). Investigating the variability in mild traumatic brain injury definitions: A prospective cohort

- study. *Archives of Physical Medicine and Rehabilitation*, 99(7), 1360–1369. Retrieved from PM:29407521 [PubMed: 29407521]
- Delaney JS, Lacroix VJ, Leclerc S, & Johnston KM (2002). Concussions among university football and soccer players. *Clinical Journal of Sport Medicine*, 12(6), 331–338. Retrieved from PM:12466687 [PubMed: 12466687]
- Denning JH (2012). The efficiency and accuracy of the Test of Memory Malingering trial 1, errors on the first 10 items of the test of memory malingering, and five embedded measures in predicting invalid test performance. *Archives of Clinical Neuropsychology*, 27(4), 417–432. Retrieved from PM:22543569 [PubMed: 22543569]
- Eisenberg MA, Meehan WP III, & Mannix R. (2014). Duration and course of post-concussive symptoms. *Pediatrics*, 133(6), 999–1006. Retrieved from PM:24819569 [PubMed: 24819569]
- Gabbe BJ, Finch CF, Bennell KL, & Wajswelner H. (2003). How valid is a self reported 12 month sports injury history? *British Journal of Sports Medicine*, 37(6), 545–547. Retrieved from PM:14665599 [PubMed: 14665599]
- Gioia GA, Collins M, & Isquith PK (2008). Improving identification and diagnosis of mild traumatic brain injury with evidence: psychometric support for the acute concussion evaluation. *Journal of Head Trauma Rehabilitation*, 23(4), 230–242. Retrieved from PM:18650767 [PubMed: 18650767]
- Gunstad J & Suhr JA (2001). “Expectation as etiology” versus “the good old days”: Postconcussion syndrome symptom reporting in athletes, headache sufferers, and depressed individuals. *Journal of the International Neuropsychological Society*, 7(3), 323–333. Retrieved from PM:11311033 [PubMed: 11311033]
- Guskiewicz KM, McCrea M, Marshall SW, Cantu RC, Randolph C, Barr W, . . . Kelly JP (2003). Cumulative effects associated with recurrent concussion in collegiate football players: The NCAA Concussion Study. *JAMA*, 290(19), 2549–2555. [PubMed: 14625331]
- Gwet KL (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, 61(Pt 1), 29–48. Retrieved from PM:18482474 [PubMed: 18482474]
- Haarbauer-Krupa J, Arbogast KB, Metzger KB, Greenspan AI, Kessler R, Curry AE, . . . Master CL (2018). Variations in mechanisms of injury for children with concussion. *The Journal of Pediatrics*, 197, 241–248. Retrieved from PM:29627189 [PubMed: 29627189]
- Haran HP, Bressan S, Oakley E, Davis GA, Anderson V, & Babl FE (2016). On-field management and return-to-play in sports-related concussion in children: Are children managed appropriately? *Journal of Science and Medicine in Sport*, 19(3), 194–199. Retrieved from PM:25772997 [PubMed: 25772997]
- Iverson GL & Lange RT (2003). Examination of “postconcussion-like” symptoms in a healthy sample. *Applied Neuropsychology*, 10(3), 137–144. Retrieved from PM:12890639 [PubMed: 12890639]
- Iverson GL, Silverberg ND, Mannix R, Maxwell BA, Atkins JE, Zafonte R, & Berkner PD (2015). Factors associated with concussion-like symptom reporting in high school athletes. *JAMA Pediatrics*, 169(12), 1132–1140. Retrieved from PM:26457403 [PubMed: 26457403]
- Kay T, Harrington DE, Adams R, Anderson T, Berrol S, Cicerone K, . . . Harley P. (1993). Definition of mild traumatic brain injury. *Journal of Head Trauma Rehabilitation*, 8(3), 86–87.
- Kerr ZY, Thomas LC, Simon JE, McCrea M, & Guskiewicz KM (2018). Association between history of multiple concussions and health outcomes among former college football players: 15-year follow-up from the NCAA Concussion Study (1999–2001). *American Journal of Sports Medicine*, 46(7), 1733–1741. [PubMed: 29620911]
- King N, Crawford S, Wenden F, Moss N, Wade D, & Caldwell F. (1997). Measurement of post-traumatic amnesia: How reliable is it? *Journal of Neurology, Neurosurgery, and Psychiatry*, 62(1), 38–42. [PubMed: 9010398]
- Kristman VL, Borg J, Godbolt AK, Salmi LR, Cancelliere C, Carroll LJ, . . . Cassidy JD (2014). Methodological issues and research recommendations for prognosis after mild traumatic brain injury: Results of the International Collaboration on Mild Traumatic Brain Injury Prognosis. *Archives of Physical Medicine and Rehabilitation*, 95(3 Suppl), S265–S277. Retrieved from PM:24581912. [PubMed: 24581912]

- Lange RT, Iverson GL, & Rose A. (2010). Post-concussion symptom reporting and the “good-old-days” bias following mild traumatic brain injury. *Archives of Clinical Neuropsychology*, 25(5).
- Lexell JE & Downham DY (2005). How to assess the reliability of measurements in rehabilitation. *American Journal of Physical Medicine & Rehabilitation*, 84(9), 719–723. [PubMed: 16141752]
- Lezak MD (1995). Neuropathology for neuropsychologists. In *Neuropsychological Assessment* (pp. 170–276).
- Mailer BJ, Valovich-McLeod TC, & Bay RC (2008). Healthy youth are reliable in reporting symptoms on a graded symptom scale. *Journal of Sport Rehabilitation*, 17(1), 11–20. Retrieved from PM:18270383 [PubMed: 18270383]
- Manley G, Gardner AJ, Schneider KJ, Guskiewicz KM, Bailes J, Cantu RC, . . . Iverson GL (2017). A systematic review of potential long-term effects of sport-related concussion. *British Journal of Sports Medicine*, 51(12), 969–977. Retrieved from PM:28455362 [PubMed: 28455362]
- Mayer AR, Hanlon FM, Claus ED, Dodd AB, Miller B, Mickey J, . . . Hutchison KE (2018). An examination of behavioral and neuronal effects of comorbid traumatic brain injury and alcohol use. *Biological Psychiatry: Cognitive Neuroscience and Neuroimaging*, 3(3), 294–302. Retrieved from PM:29486871 [PubMed: 29486871]
- Mayer AR, Kaushal M, Dodd AB, Hanlon FM, Shaff NA, Mannix R, . . . Meier TB (2018). Advanced biomarkers of pediatric mild traumatic brain injury: Progress and perils. *Neuroscience & Biobehavioral Reviews*, 94, 149–165. doi: 10.1016/j.neubiorev.2018.08.002 [PubMed: 30098989]
- Mayer AR, Quinn DK, & Master CL (2017). The spectrum of mild traumatic brain injury: A review. *Neurology*, 89(6), 623–632. [PubMed: 28701496]
- Mayer AR, Stephenson DD, Dodd AB, Robertson-Benta CR, Pabbathi Reddy S, Shaff NA, . . . Quinn DK (2020). Comparison of methods for classifying persistent post-concussive symptoms in children. *Journal of Neurotrauma*, 37(13), 1504–1511. doi: 10.1089/neu.2019.6805 [PubMed: 31964232]
- Mayou RA, Black J, & Bryant B. (2000). Unconsciousness, amnesia and psychiatric symptoms following road traffic accident injury. *British Journal of Psychiatry*, 177(6), 540–545.
- McCrory P, Meeuwisse W, Dvorak J, Aubry M, Bailes J, Broglio S, . . . Vos PE (2017). Consensus statement on concussion in sport—the 5(th) international conference on concussion in sport held in Berlin, October 2016. *British Journal of Sports Medicine*, 51(11), 838–847. Retrieved from PM:28446457 [PubMed: 28446457]
- McMillan T, Jongen E, & Greenwood R. (1996). Assessment of post-traumatic amnesia after severe closed head injury: Retrospective or prospective? *Journal of Neurology, Neurosurgery, and Psychiatry* 60(4), 422–427. [PubMed: 8774408]
- Menon DK, Schwab K, Wright DW, & Maas AI (2010). Position statement: Definition of traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 91(11), 1637–1640. Retrieved from PM:21044706 [PubMed: 21044706]
- Nelson NW, Anderson CR, Thuras P, Kehle-Forbes SM, Arbisi PA, Erbes CR, & Polusny MA (2015). Factors associated with inconsistency in self-reported mild traumatic brain injury over time among military personnel in Iraq. *British Journal of Psychiatry*, 206(3), 237–244. Retrieved from PM:25614533
- Rice SG (2008). Medical conditions affecting sports participation. *Pediatrics*, 121(4), 841–848. Retrieved from PM:18381550 [PubMed: 18381550]
- Roberts CM, Spitz G, Mundy M, & Ponsford JL (2019). Prospective evaluation of first and last memory reports following moderate to severe traumatic brain injury. *Journal of Clinical and Experimental Neuropsychology*, 41(2), 109–117. [PubMed: 29986640]
- Roberts CM, Spitz G, & Ponsford JL (2016). Comparing prospectively recorded posttraumatic amnesia duration with retrospective accounts. *Journal of Head Trauma Rehabilitation*, 31(2), E71–E77. [PubMed: 26098260]
- Ruff RM, Iverson GL, Barth JT, Bush SS, & Broshek DK (2009). Recommendations for diagnosing a mild traumatic brain injury: A National Academy of Neuropsychology education paper. *Archives of Clinical Neuropsychology*, 24(1), 3–10. Retrieved from PM:19395352 [PubMed: 19395352]
- Ruff RM & Jurica P. (1999). In search of a unified definition for mild traumatic brain injury. *Brain Injury*, 13(12), 943–952. Retrieved from PM:10628500 [PubMed: 10628500]

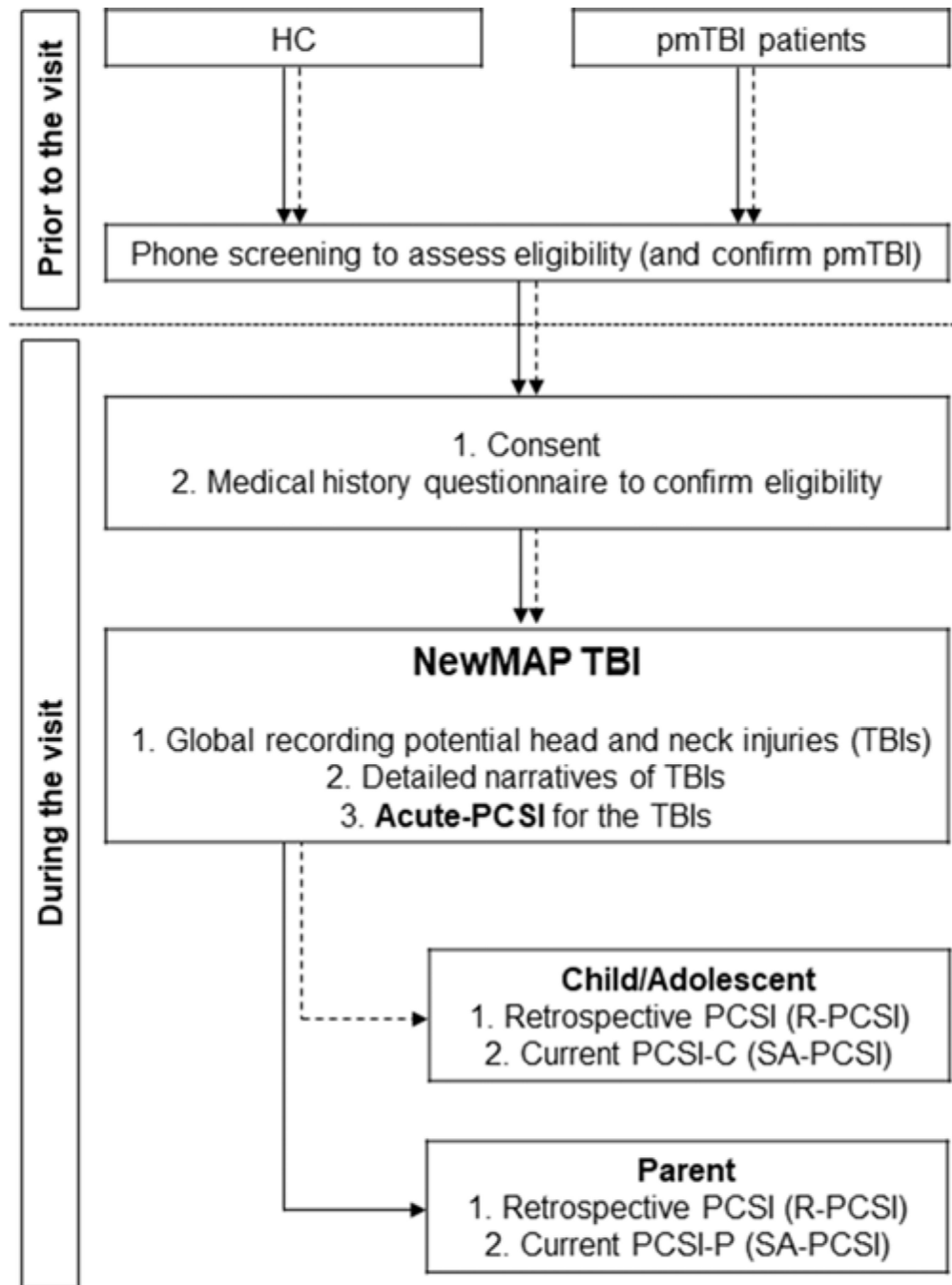


- Sady MD, Vaughan CG, & Gioia GA (2014). Psychometric characteristics of the postconcussion symptom inventory in children and adolescents. *Archives of Clinical Neuropsychology*, 29(4), 348–363. Retrieved from PM:24739735 [PubMed: 24739735]
- Sarmiento K, Gioia GA, Kirkwood MW, Wade SL, & Yeates KO (2019). A commentary for neuropsychologists on CDC's guideline on the diagnosis and management of mild traumatic brain injury among children. *Clinical Neuropsychology*, 1–19. Retrieved from PM:31530221
- Sherer M, Sander AM, Maestas KL, Pastorek NJ, Nick TG, & Li J. (2015). Accuracy of self-reported length of coma and posttraumatic amnesia in persons with medically verified traumatic brain injury. *Archives of Physical Medicine and Rehabilitation*, 96(4), 652–658. Retrieved from PM:25461819 [PubMed: 25461819]
- Shrout PE & Fleiss JL (1979). Intraclass correlations: Uses in assessing rater reliability. *Psychological Bulletin*, 86(2), 420–428. Retrieved from PM:18839484 [PubMed: 18839484]
- Stuss DT, Binns MA, Carruth FG, Levine B, Brandys CE, Moulton RJ, . . . Schwartz ML (1999). The acute period of recovery from traumatic brain injury: Posttraumatic amnesia or posttraumatic confusional state? *Journal of Neurosurgery*, 90(4), 635–643. Retrieved from PM:10193606 [PubMed: 10193606]
- Vanderlei FM, Barbosa DA, Machado AF, Bastos FN, Vanderlei LM, Junior JN, & Pastre CM (2017). Analysis of recall bias of information on soccer injuries in adolescents. *Motriz: Revista de Educação Física*, 23(2), e101777.
- Vanderploeg RD, Groer S, & Belanger HG (2012). Initial developmental process of a VA semistructured clinical interview for TBI identification. *Journal of Rehabilitation Research and Development*, 49(4), 545–556. Retrieved from PM:22773258 [PubMed: 22773258]
- Walker WC, Cifu DX, Hudak AM, Goldberg G, Kunz RD, & Sima AP (2015). Structured interview for mild traumatic brain injury after military blast: Inter-rater agreement and development of diagnostic algorithm. *Journal of Neurotrauma*, 32(7), 464–473. Retrieved from PM:25264909 [PubMed: 25264909]
- Wilmoth K, LoBue C, Clem MA, Reddy R, Hynan LS, Didehbani N, . . . Cullum CM (2018). Consistency of traumatic brain injury reporting in older adults with and without cognitive impairment. *The Clinical Neuropsychologist*, 32(3), 524–529. doi: 10.1080/13854046.2017.1378371 [PubMed: 28988512]
- Yeates KO, Taylor HG, Rusin J, Bangert B, Dietrich A, Nuss K, . . . Jones BL (2009). Longitudinal trajectories of postconcussive symptoms in children with mild traumatic brain injuries and their relationship to acute clinical status. *Pediatrics*, 123(3), 735–743. [PubMed: 19254996]
- Zemek RL, Grool AM, Rodriguez DD, DeMatteo C, Rothman L, Benchimol EI, . . . Macpherson AK (2017). Annual and seasonal trends in ambulatory visits for pediatric concussion in Ontario between 2003 and 2013. *The Journal of Pediatrics*, 181, 222–228.e222. Retrieved from PM:27843008

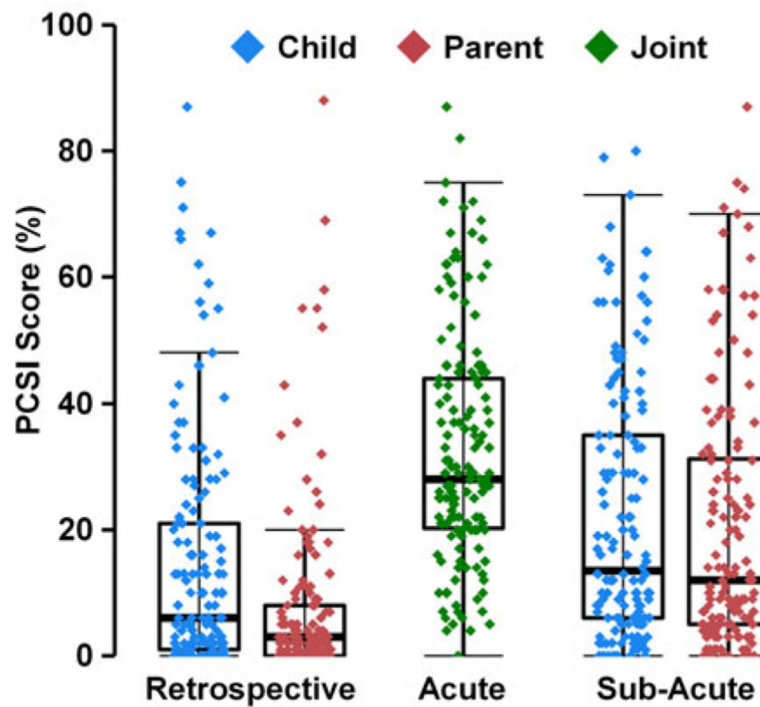


**Fig. 1.**

A flow chart displaying the eligibility, inclusion, and exclusion steps for both healthy controls (HC) and pediatric mild traumatic brain injury (pmTBI) patients for the study. Sample sizes and exclusion criteria are presented separately for the subacute (SA), early chronic (EC), and late chronic (LC) visits.



**Fig. 2.** A flow chart displaying the procedures for a study visit for both healthy controls (HC) and patients with pediatric mild traumatic brain injury (pmTBI) through the subacute visit (SA). Parental interactions are represented by a solid arrow, whereas child/adolescent interactions are represented by a dashed arrow. NewMAP TBI = New Mexico Assessment of Pediatric Traumatic Brain Injury; PCSI = Post-Concussion Symptom Inventory.



**Fig. 3.**

Box and scatter plots demonstrating construct validity for symptom reporting on the NewMAP TBI. The percent of total scores from the Post-Concussion Symptom Inventory (PCSI) scores are reported on the Y-axis. Child and parent ratings were retrospectively acquired for 1 month prior to injury (retrospective) and on the first study visit (subacute). Children and parents jointly completed the semi-structured interview, in which they were asked if the child experienced each PCS following the injury (acute-PCSI). Both parent and child ratings exhibit the expected pattern of acute > subacute > retrospective symptom burden .

Table 1.

TBI characteristics.

Visit	SA		EC		LC	
	Qualifying	Remote	Qualifying	Remote	Qualifying	Remote
<b># of Remote TBIs</b>						
0	-	289 (85.8%)	-	301 (89.3%)	-	205 (88.7%)
1	-	35 (10.4%)	-	26 (7.7%)	-	20 (8.7%)
2	-	9 (2.7%)	-	7 (2.1%)	-	6 (2.6%)
3	-	4 (1.2%)	-	3 (0.9%)	-	0 (0%)
<b>LOC</b>						
<b>Continuous</b>						
Mean ( <i>SD</i> ) in min	2.25 (9.69)	0.62 (1.43)	3.16 (14.71)	4.03 (18.04)	2.38 (7.51)	3.83 (11.95)
Median in min	0	0	0.03	0	0	0.04
<b>Categorical</b>						
None	92 (50.5%)	42 (64.6%)	86 (47.5%)	27 (60%)	67 (52.3%)	14 (46.7%)
< 1 min	35 (19.2%)	9 (13.8%)	38 (21.0%)	5 (11.1%)	23 (18%)	5 (16.7%)
1–5 min	42 (23.1%)	13 (20%)	41 (22.7%)	10 (22.2%)	28 (21.9%)	9 (30.0%)
> 5–30 min	12 (6.6%)	1 (1.5%)	13 (7.2%)	2 (4.4%)	8 (6.3%)	1 (3.3%)
> 30 min	1 (0.5%)	0 (0.0%)	3 (1.7%)	1 (2.2%)	2 (1.6%)	1 (3.3%)
<b>PTA</b>						
<b>Continuous</b>						
Mean ( <i>SD</i> ) in min	250.37 (1165.12)	282.10 (1297.68)	534.40 (3575.81)	304.33 (995.91)	334.00 (1934.97)	1245.50 (5538.62)
Median in min	0	0	0	0	0	0
<b>Categorical</b>						
None	112 (62.2%)	48 (73.8%)	121 (66.9%)	34 (75.6%)	90 (70.3%)	20 (66.7%)
1 sec-< 5 min	4 (2.2%)	1 (1.5%)	8 (4.4%)	0 (0%)	2 (1.6%)	0 (0%)
5 min-< 1 hr	27 (15.0%)	3 (4.6%)	15 (8.3%)	1 (2.2%)	13 (10.2%)	2 (6.7%)
1–24 hr	32 (17.8%)	12 (18.5%)	28 (15.5%)	7 (15.6%)	18 (14.1%)	6 (20%)
> 24 hr	5 (2.8%)	1 (1.5%)	9 (5.0%)	3 (6.7%)	5 (3.9%)	2 (6.7%)
<b>Retrograde Amnesia</b>						
<b>Continuous</b>						

Visit	SA			EC			LC		
	Qualifying	Remote		Qualifying	Remote		Qualifying	Remote	
<b>TBI Type:</b>									
Mean (SD) in min	85.60 (769.9)	94.02 (350.62)	118.50 (494.98)	12.44 (47.79)	83.70 (568.46)	76.63 (280.78)			
Median in min	0	0	0	0	0	0			
<b>Categorical</b>									
None	154 (85.1%)	53 (81.5%)	142 (78.5%)	35 (77.8%)	99 (77.3%)	22 (73.3%)			
1 sec-<5 min	11 (6.1%)	0 (0%)	11 (6.1%)	0 (0%)	8 (6.3%)	2 (6.7%)			
5 min-<1 hr	5 (2.8%)	4 (6.2%)	7 (3.9%)	8 (17.8%)	9 (7.0%)	2 (6.7%)			
1-24 hr	10 (5.5%)	7 (10.9%)	16 (8.8%)	2 (4.4%)	10 (7.8%)	4 (13.3%)			
> 24 hr	1 (0.6%)	0 (0%)	5 (2.8%)	0 (0%)	2 (1.6%)	0 (0%)			
<b>Confusion/Disorientation</b>									
<b>Continuous</b>									
Mean (SD) in min	538.96 (1599.25)	1646.06 (3700.12)	1107.21 (3958.87)	1124.42 (1691.75)	1480.21 (4009.17)	2349.97 (5087.03)			
Median in min	30.00	60.00	30.00	120.00	60.00	180.00			
<b>Categorical</b>									
None	32 (17.7%)	17 (26.6%)	32 (17.7%)	10 (22.2%)	16 (12.5%)	5 (16.7%)			
1 sec-<5 min	22 (12.2%)	4 (6.3%)	19 (10.5%)	2 (4.4%)	12 (9.4%)	2 (6.7%)			
5 min-<1 hr	46 (25.4%)	8 (12.5%)	44 (24.3%)	6 (13.3%)	28 (21.9%)	3 (10.0%)			
1-24 hr	66 (36.5%)	20 (31.3%)	62 (34.3%)	17 (37.8%)	48 (37.5%)	12 (40.0%)			
> 24 hr	15 (8.3%)	15 (23.4%)	24 (13.3%)	10 (22.2%)	24 (18.8%)	8 (26.7%)			
<b>Acute-PCSI # of Symptoms</b>									
Mean (SD)	10.91 (4.80)	6.65 (4.93)	10.86 (5.16)	7.64 (6.83)	11.06 (5.61)	9.60 (6.19)			
Median	10.00	6.00	11.00	5.00	12.00	9.00			
<b>Acute-PCSI Symptom Burden</b>									
Mean (SD)	39.88 (23.06)	19.63 (18.33)	39.67 (24.8)	24.56 (27.36)	40.85 (26.27)	37.37 (60.81)			
Median	35.00	15.00	34.50	12.00	40.00	25.00			

EC = Early Chronic; LC = Late Chronic; LOC = Loss of Consciousness; PCS = Post-Concussive Symptoms; PCSI = Post-Concussion Symptom Inventory; PTA = Post-Traumatic Amnesia; SA = Subacute; SD = Standard Deviation; TBI = Traumatic Brain Injury.

\* Remote TBI analyses include pmTBI patients and HCs that reported a remote TBI at any of the three visits.



**Table 2.**

Demographic information for the full sample.

<b>Demographic Characteristic</b>	<b>pmTBI patients (n=184)</b>	<b>HC (n=156)</b>	<b>Sig.</b>
Female Sex – Count (%)	80 (43.5%)	67 (42.9%)	0.92
Age at Enrollment – Mean ( <i>SD</i> )	13.73 (2.74)	13.69 (2.78)	0.89
Mother Years of Education – Mean ( <i>SD</i> )	14.86 (2.64)	16.43 (1.65)	<0.001
Father Years of Education– Mean ( <i>SD</i> )	14.53 (3.18)	16.73 (2.62)	<0.001
Litigation – Count (%)	9 (4.9%)	–	N/A
LD or ADHD – Count (%)	38 (20.7%)	–	N/A
TOMMe10 < 8 at SA – Count (%)	9 (4.9%)*	1 (0.6%)	0.037

ADHD = Attention-Deficit/Hyperactive Disorder; LD = Learning Disability; HC = Healthy Controls; pmTBI = Pediatric Mild Traumatic Brain Injury; TOMMe10 = Test of Memory Malinger short version.

\* Data are missing for one participant.

**Table 3.**

Test-retest reliability coefficients for TBI characteristics.

TBI Type:	SA to EC		SA to LC		EC to LC	
	Qualifying	Remote	Qualifying	Remote	Qualifying	Remote
# of Reported Remote TBIs – Full Sample <sup>2</sup>	–	0.97	–	0.94	–	0.98
# of Reported Remote TBIs – Only Individuals with a History of Remote TBI (n=59) <sup>2</sup>	–	0.70	–	0.44	–	0.77
LOC Binary <sup>1</sup>	0.67	0.75	0.62	0.60	0.75	1.00
LOC Category <sup>2</sup>	0.87	0.93	0.85	0.86	0.90	0.93
LOC Continuous <sup>3</sup>	0.65	0.53	0.54	0.29	0.80	0.73
PTA Binary <sup>1</sup>	0.56	0.71	0.53	0.66	0.59	0.83
PTA Category <sup>2</sup>	0.74	0.85	0.72	0.77	0.77	0.84
PTA Continuous <sup>3</sup>	0.45	0.54	0.34	0.26	0.44	0.64
RGA Binary <sup>1</sup>	0.68	0.80	0.66	0.88	0.81	0.70
RGA Category <sup>2</sup>	0.87	0.91	0.86	0.93	0.91	0.88
RGA Continuous <sup>3</sup>	0.28	0.36	0.22	0.69	0.59	0.60
Confusion/Disorientation Binary <sup>1</sup>	0.74	0.65	0.77	0.72	0.76	0.79
Confusion/Disorientation Category <sup>2</sup>	0.52	0.50	0.54	0.66	0.53	0.52
Confusion/Disorientation Continuous <sup>3</sup>	0.40	0.43	0.36	0.78	0.44	0.42
Acute-PCSI Number of Symptoms <sup>3</sup>	0.65	0.58	0.52	0.60	0.72	0.80
Acute-PCSI Symptom Burden <sup>3</sup>	0.67	0.59	0.59	0.57	0.67	0.78

EC = Early Chronic; LOC = Loss of Consciousness; LC = Late Chronic; PCSI = Post-Concussion Symptom Inventory; PTA = Post-Traumatic Amnesia; RGA = Retrograde Amnesia; SA = Subacute; TBI = Traumatic Brain Injury.

\* Actual value was negative, but constrained to zero to increase interpretability

<sup>1</sup>Gwet's AC1

<sup>2</sup>Gwet's AC2

$\frac{1}{2}$  Intraclass correlation coefficient (log transformed) + 1.

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript