Published in final edited form as:
Brain Inj. 2024 September 18; 38(11): 880–888. doi:10.1080/02699052.2024.2352524.

Using machine learning to discover traumatic brain injury patient phenotypes: national concussion surveillance system Pilot

Dana Waltzman[a], **Jill Daugherty**[a], **Alexis Peterson**[a], **Angela Lumba-Brown**[b]

[a]Centers for Disease Control and Prevention (CDC), National Center for Injury Prevention and Control (NCIPC), Division of Injury Prevention, Atlanta, Georgia, USA

[b]Department of Emergency Medicine, Stanford University School of Medicine, Stanford, California, USA

Abstract

Objective: The objective is to determine whether unsupervised machine learning identifies traumatic brain injury (TBI) phenotypes with unique clinical profiles.

Methods: Pilot self-reported survey data of over 10,000 adults were collected from the Centers for Disease Control and Prevention (CDC)'s National Concussion Surveillance System (NCSS). Respondents who self-reported a head injury in the past 12 months ($n = 1,364$) were retained and queried for injury, outcome, and clinical characteristics. An unsupervised machine learning algorithm, partitioning around medoids (PAM), that employed Gower's dissimilarity matrix, was used to conduct a cluster analysis.

Results: PAM grouped respondents into five TBI clusters (phenotypes A-E). Phenotype C represented more clinically severe TBIs with a higher prevalence of symptoms and association with worse outcomes. When compared to individuals in Phenotype A, a group with few TBI-related symptoms, individuals in Phenotype C were more likely to undergo medical evaluation (odds ratio [OR] = 9.8, 95% confidence interval[CI] = 5.8–16.6), have symptoms that were not currently resolved or resolved in 8+ days (OR = 10.6, 95%CI = 6.2–18.1), and more likely to report at least moderate impact on social (OR = 54.7, 95%CI = 22.4–133.4) and work (OR = 25.4, 95%CI = 11.2–57.2) functioning.

CONTACT Dana Waltzman, dwaltzman@cdc.gov, Division of Injury Prevention, National Center for Injury Prevention & Control, Centers for Disease Control & Prevention, 4770 Buford Highway, Atlanta, GA 30341.

Disclaimer
The findings and conclusions in this manuscript are those of the authors and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

Disclosure statement
No potential conflict of interest was reported by the author(s).

Supplemental data for this article can be accessed online at https://doi.org/10.1080/02699052.2024.2352524

Conclusion: Machine learning can be used to classify patients into unique TBI phenotypes. Further research might examine the utility of such classifications in supporting clinical diagnosis and patient recovery for this complex health condition.

Keywords

Traumatic brain injury; TBI; machine learning; clustering; phenotypes

Introduction

Clinical manifestations of traumatic brain injury (TBI) are heterogeneous, varying widely from person to person, and often classified as mild, moderate, or severe based on the Glasgow Coma Scale (GCS) score (1). TBI is associated with a broad range of signs and symptoms (2,3). There is a lack of robust, objective criteria and more mild TBI diagnosis relies largely on self-reported symptoms. These mild TBIs have been called 'invisible injuries' given the lack of objective diagnostic criteria and may be unrecognized in certain clinical settings, particularly in multi-trauma cases (4). Variable symptom presentation, response to treatment, and recovery can lead to challenges in managing TBI.

In addition to GCS scoring, three cardinal signs or symptoms of TBI have historically contributed to qualifying the severity of injury including: loss of consciousness, alteration of consciousness/mental state (i.e., feeling disoriented or confused), and post-traumatic amnesia (5). Categorizing TBI patients in this manner fails to account for the complex heterogeneity among individuals and, to date, clinical treatment trials based on this classification system have failed to optimally translate to effective treatment and recovery in the real world (6,7). Patient classification in traditional severity categories may inhibit the discovery of effective therapies that improve outcomes based on more granular clinical profiles (7,8). For these reasons, more refined evidence-based approaches to classification and treatment are needed.

The limitations of the current classification system allow space to find alternative methods that have the potential to stratify TBI patient subpopulations for better targeted treatment. Unsupervised machine learning has been proven as a promising method for discovering patient phenotypes, improving upon the classification and identification of patient subpopulations for several diseases (9,10), including TBI (11–13). The goal of this study was to determine if unsupervised machine learning could identify TBI patient phenotypes with unique clinical and outcome profiles in national survey data.

Methods

Survey

Data from the U.S. Centers for Disease Control and Prevention's pilot National Concussion Surveillance System (NCSS) were analyzed. Detailed NCSS survey methodology has been previously described (14,15) with the goal of developing a TBI tiered case definition to use in self-report surveys and to calculate TBI incidence and prevalence estimates. The NCSS pilot was a random-digit-dial (RDD) telephone survey that used computer-assisted telephone interviewing to collect data. Over 10,000 adults participated in the survey, and data were collected from September 2018 to September 2019. The sampling frame included the non-institutionalized population of males and females aged 18 years and older residing in the 50 states and the District of Columbia.

Measures

Respondents were asked two initial questions regarding head injuries sustained within the past 12 months: 'In the last year, that is since [insert date one year before the interview date], were you examined in a doctor's office, clinic, hospital or elsewhere because of a head injury?' and 'In the last year, that is since [insert date one year before the interview date], did you experience any other injuries to your head that you did not see a doctor about?'

If the respondent answered affirmatively to either of the questions above, they were then asked a series of yes or no questions regarding 12 common postconcussive signs and symptoms in relationship to their reported head injury (being dazed, confused, or having trouble thinking straight; difficulty remembering what happened just before or after injury; loss of consciousness (LOC); nausea or vomiting; headache; dizziness, clumsiness, or balance problems; blurred or double vision; trouble concentrating; difficulty learning or remembering new things; sensitivity to light or noise; change in mood or temperament; and changes in sleep or being more tired than usual). This is in accordance with the previously established TBI case definition for self-report survey data that was derived from Daugherty et al. (14,15). A respondent was defined as sustaining a TBI if at least one of the signs or symptoms were endorsed.

Outcomes associated with the most recent head injury were also examined – these were selected based in part on Daugherty et al. (14,15). which examined the same data source used for this work. They included: whether medical evaluation was sought (yes/no), time to symptom resolution (number of days), and self-reported impact on (a) social and (b) work functioning. Time to symptom resolution was assessed in two ways: experiencing 1 day of symptoms (to indicate lower severity) and still experiencing symptoms at the time of the interview or having had 8+ days of symptoms (to indicate greater severity). For social and work functioning, the response options were based on a Likert scale ('not at all,' 'slightly,' 'moderately,' 'quite a bit,' or 'extremely') and were categorized into a binary variable (not at all/slightly vs. moderately/quite a bit/extremely).

Additionally, demographic information (sex, age, race/ethnicity, education, marital status, and home ownership) was collected as part of the survey.

Analysis

To identify potential TBI phenotypes among those reporting a head injury, respondents were grouped into clusters based upon 12 TBI signs and symptoms. Gower's dissimilarity matrix (16) was computed due to the nature of the binary input data (i.e., presence or absence of each sign or symptom). Unsupervised machine learning was chosen due to the data not being labeled. The unsupervised partitioning around medoids (PAM) algorithm (17) was used to cluster observations. PAM is a partition-based algorithm that selects cluster centers (i.e., medoids) based on actual observations from the data. This is an iterative process, where the final medoid for each cluster is chosen as the data point with the smallest distance to all other data points within the cluster. To determine the optimal number of clusters, the elbow plot and silhouette score were used. Ultimately, the final number of clusters chosen was based on clinical utility (e.g., what generated distinct clinical profiles and what will be generally accepted by the field). The cluster analysis was performed on unweighted data and was run using the 'cluster' package (18) and the cluster visualization was run using the 'Rtsne' package (19) in R software, version 4.1.2 (R Foundation for Statistical Computing, Vienna, Austria).

Descriptive statistics (frequencies and percentages) and bivariate statistics (chi-square tests) were calculated to describe respondent demographics and TBI outcomes by TBI phenotype. Post-hoc tests were conducted, which were pairwise comparisons of column proportions for both the descriptive and bivariate statistics. To determine the association between outcomes and phenotypes, separate logistic regressions were run using the phenotype characterized by the least severity (e.g., Phenotype A ['cluster 1']) as the reference group. This analysis was subset to cases. Associations are presented as odds ratios (ORs) with corresponding 95% confidence intervals (CIs). Analyses were run in SAS 9.4 (SAS Institute, Cary, North Carolina) and accounted for the complex survey design by taking into account weighting, the primary sampling unit, and stratification. Study procedures were approved by the institutional review board of ICF International, Inc. (#FWA00000845) and were consistent with ethical guidelines for human subjects research.

Results

There were 1,364 respondents in the study who reported a head injury in the past 12 months. The sample of respondents with a head injury was composed of 212 non-cases (respondents who had an affirmative response to a head injury but reported no symptoms) and 1,152 respondents with symptomatic head injury that were then classified as a TBI (an affirmative response to a head injury and at least one symptom).

The median number of symptoms reported was four (data not shown). Among respondents who self-reported symptoms, headache was the most common symptom (85%, $n = 980$), followed by confusion (55%, $n = 631$), and balance problems (43%, $n = 501$) (Figure 1). Loss of consciousness (LOC) was reported by 19% ($n = 214$) of respondents with a TBI. Respondents who self-reported *only* headache ($n = 163$) were the most common, followed by respondents who experienced headache and confusion together ($n = 40$) and headache and change in temperament together ($n = 30$). There were 12 respondents who self-reported all 12 signs and symptoms.

The PAM algorithm grouped respondents into five clusters (TBI phenotypes A-E, Table 1 and Supplemental Figure 1S). Phenotype C had the highest prevalence of symptoms (i.e., >50% of respondents in this cluster self-reported 11 out of the 12 symptoms) and included two of the cardinal symptoms (i.e., LOC and difficulty remembering what happened just before/after injury). Additionally, this phenotype had the highest median number of symptoms ($n = 8$, Table 1). Phenotype C respondents had a higher prevalence of medical evaluation (58.1%) for their head injury, a higher prevalence of symptoms that were not currently resolved or resolved in 8+ days (60.5%), greater impact on social and work functioning (Moderately/Quite A Bit/Extremely = 43.2% and 46.9%, respectively), and a lower prevalence of symptoms that resolved in one day (4.5%) compared to other phenotypes (Table 2). Although demographic characteristics were not involved in creating cluster phenotypes, Phenotype C had a higher prevalence of respondents who were a race/ethnicity other than non-Hispanic white (45.6%), had an education of high school or less (52.4%), and were renting a residence (45.5%) when compared to other phenotypes (Table 3).

Phenotype E was composed of all the non-cases (i.e., respondents that reported a head injury with no symptoms) (Table 1). Compared to other phenotypes, Phenotype E had a lower prevalence of those who were evaluated (11.2%) for their head injury and had lower impact on their social and work functioning (Moderately/Quite A Bit/Extremely = 1.2% for both) (Table 2), and a higher prevalence of respondents who were aged 55+ (39.9%) and had a Bachelor's degree or more for education (36.2%) (Table 3).

Phenotype D was the largest cluster and had the second highest median number of symptoms ($n = 5$, Table 1). Phenotype D was most similar to the average respondent in the sample included in the cluster analysis (please refer to the 'Total' column in Tables 2 and 3 to see the overall TBI cluster sample): Phenotype D respondents had the second highest prevalence of medical evaluation (30.7%) for their head injury, symptoms that were not currently resolved or resolved in 8+ days (38.1%), and impact on social and work functioning (Moderately/Quite A Bit/Extremely = 23.1% and 23.6%, respectively) (Table 2). Additionally, similar to the total sample, Phenotype D respondents were 51.8% male, 44.9% were between 18–34 years of age, 69.1% non-Hispanic white, 64.9% had some college education or more, and 56.6% owned a residence (Table 3).

Phenotypes A and B had few symptoms (where >50% of respondents in a cluster self-reported a particular sign or symptom), which included headache alone (Phenotype A) or confusion plus headache (Phenotype B) (Table 1). Correspondingly, Phenotypes A and B had a high prevalence of respondents who had one day of symptoms (48.1% and 35.4%, respectively) and a lower prevalence of respondents who were evaluated (12.4% and 16.7%, respectively) for their head injury, had symptoms that were not currently resolved or resolved in 8+ days (12.7% and 24.0%, respectively), and social (Moderately/Quite A Bit/Extremely = 1.4% and 2.5%, respectively) and work impairment (Moderately/Quite A Bit/Extremely = 3.4% and 5.3%, respectively) (Table 2). However, these phenotypes differed by demographics. Phenotype A had a higher prevalence of respondents who were younger (18–35 years old, 49.8%), while Phenotype B had the highest prevalence of males (65.5%), respondents who were non-Hispanic white (79.1%), and respondents who owned a residence (75.0%) (Table 3).

Table 4 displays the results of the logistic regressions, ordered according to clinical severity (most to least severe). The magnitude of effect (odds ratio) was in the expected direction (e.g., more severe phenotypes demonstrated worse outcomes). Compared to Phenotype A, Phenotype C was associated with higher odds of medical evaluation (OR = 9.8, 95% CI = 5.8–16.6), symptoms that were not currently resolved or resolved in 8+ days (OR = 10.6, 95% CI = 6.2–18.1), higher impacts on social (OR = 54.7, 95% CI = 22.4–133.4) and work (OR = 25.4, 95% CI = 11.2–57.2) functioning, and lower odds of symptoms that resolved in one day (indicator of less severity) (OR = 0.1, 95% CI = 0.03–0.1). Displaying a similar pattern to Phenotype C, Phenotype D was also associated with higher odds of medical evaluation (OR = 3.1, 95% CI = 1.9–5.2) and symptoms that were not currently resolved or resolved in 8+ days (OR = 4.2, 95% CI = 2.6–7.0), higher impacts on social (OR = 21.6, 95% CI = 8.8–52.7) and work (OR = 8.8, 95% CI = 3.9–19.9) functioning, and lower odds of symptoms that resolved in one day (OR = 0.3, 95% CI = 0.2–0.4). Phenotype B was associated with higher odds of symptoms that were not currently resolved or resolved in 8+ days (OR = 2.2, 95% CI = 1.1–4.1) and lower odds of symptoms that resolved in one day (OR = 0.6, 95% CI = 0.4–0.9).

Discussion

In this study, we identified five TBI patient phenotypes using unsupervised machine learning, based solely on self-reported signs and symptoms from national survey data. Each TBI phenotype demonstrated unique clinical characteristics that corresponded to specific differences in the severity of outcomes and demographic profiles. This work directly contributes to a growing foundation of research establishing feasibility of precision diagnostics in TBI. It also supports personalized care with better prognostication of outcomes as compared to traditional classification schemes.

One phenotype (Phenotype C) in this study was characterized by a high number of symptoms and worse outcomes. Additionally, this phenotype also had the highest percentages of respondents with two cardinal symptoms (i.e., LOC and difficulty remembering what happened just before/after injury). Similar to our data, past research has shown that these cardinal symptoms are associated with worse outcomes (e.g., longer functional recovery, increased symptom duration, cognitive impairment, etc.) among TBI patients (20–22). This phenotype was also disproportionately populated by people who were non-Hispanic white, had less education, and rented a residence. This association suggests, as previous literature (23–25) has reported, that TBI may disproportionality impact marginalized groups. Previous studies examining social determinants of health (non-medical factors that influence health outcomes) (26) demonstrate that certain factors (e.g., disability, race, insurance status) can negatively impact an individual's recovery and outcome following TBI (27,28). Recognizing and seeking treatment when a suspected brain injury is sustained are important, and research supports that patients who receive clinical care sooner recover faster (29,30). Even in those who seek care, not all patients with head injury are evaluated for TBI (31,32) and may not be diagnosed or may go untreated (33–35). Future studies could assess and address structural inequalities related to phenotypic diagnosis and subsequent treatment for patients with TBI.

Another distinct cluster demonstrated a phenotype (Phenotype A) where respondents self-reported headache as their only symptom where > 50% of respondents self-reported a particular sign or symptom in that cluster. This finding is consistent with other TBI studies (3,36,37) that also demonstrate a high prevalence of headache among patients. This phenotype likely represents less severe TBIs based on its lower association with worse outcomes (i.e., medical evaluation, symptoms that were not currently resolved or resolved in 8+ days, impact on social and work functioning) and reflects a unique group of patients who may benefit from a streamlined, simple treatment regimen that largely targets headache reduction. However, further research is warranted to ensure resolution of symptoms and no long-term sequelae in this group.

While clinical severity (e.g., GCS or the presence or absence of the cardinal TBI signs/ symptoms) has been demonstrated as strong predictors of worse outcomes after TBI (20,22,38), estimates of these measures are often inaccurate due to recall bias, difficulty in interpretation or collection of these data, and have poor concordance or conflicting associations between them (39,40). Additionally, previous literature (including clinical trials) that categorize individuals with TBI using these standard severity scales have demonstrated limited utility in terms of stratifying TBI patients for effective treatments (6,7). For example, past studies (11,41) that have examined GCS versus machine learning have demonstrated that TBI severity using GCS classification alone may not be optimal or granular enough to capture the complexity of TBI. In comparison to using GCS or to traditional statistics, machine learning can identify non-linear relationships that can reveal meaningful patterns and insights, and create novel representations of clinical profiles that may be better suited to addressing multifaceted and complex health conditions and target treatments (42,43).

Strengths of this work include an innovative machine learning algorithm for classification applying a data-driven approach to identify TBI phenotypes that have a distinct clinical profile. Another strength lies in the use of a large dataset derived from a random-digit-dial (RDD) telephone survey. However, this study does have limitations. Respondents who self-reported a head injury that was then classified as a TBI may not have been clinically evaluated or diagnosed and results should be interpreted with this understanding. However, the data collected as part of this study were intended for surveillance, and limiting the inclusion of individuals who sought care or received a TBI diagnosis would also likely restrict the data to those with 1) a more severe presentation of TBI or 2) ease of accessibility to care (i.e., potentially excluding individuals who are socially disadvantaged) (14). Additionally, self-report data are subject to issues such as recall bias, social desirability bias, and under or over-reporting. Another limitation is that the input data for the clusters were limited to individual signs and symptoms. Other studies (11–13,41) examining TBI phenotypes or endotypes have included biological biomarkers, clinical findings, neuroimaging, laboratory tests, or therapies to create clusters. These data were not available in our study, and future research can consider using machine learning in comprehensive datasets that link biomarkers with symptom profiles and other characteristics to even better inform phenotypes. Prospective controlled studies of phenotype-directed diagnosis and treatment are potentially important next steps for research. Finally, the pilot NCSS survey, from which this study’s data is drawn, was designed to support TBI case ascertainment and measure development. Estimates were weighted with the intent of exploring differences in TBI prevalence derived from different databases, using different case definitions. Weighted estimates from this pilot work were not specifically intended to produce nationally representative estimates of TBI.

Conclusions

Classifying TBI phenotypes using machine learning, as demonstrated here by five distinct phenotypes, may inform next steps in research to focus on their utility in clinical diagnosis and symptom-based treatment for faster patient recovery through a more personalized approach.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

The author(s) reported there is no funding associated with the work featured in this article.

References

1. Teasdale G, Jennett B. Assessment of coma and impaired consciousness: a practical scale. Lancet Neurol. 1974;2(7872):81–84. doi:10.1016/S0140-6736(74)91639-0
2. Centers for Disease Control and Prevention. Symptoms of mild TBI and concussion Atlanta, GA 2021; [Accessed 2021 October 25]. https://www.cdc.gov/traumaticbraininjury/concussion/symptoms.html.
3. Daugherty J, Sarmiento K, Womack LS, Breiding M. Symptom profile of affirmative responses to a self-report concussion question, United States 2019. Brain Inj. 2021;35(11):1413–17. doi:10.1080/02699052.2021.1972340 [PubMed: 34487455]
4. Sharma B, Bradbury C, Mikulis D, Green R. Missed diagnosis of traumatic brain injury in patients with traumatic spinal cord injury. J Rehabil Med. 2014;46(4):370–73. doi:10.2340/16501977-1261 [PubMed: 24549169]
5. American Congress of Rehabilitation Medicine. Definition of mild traumatic brain injury. Developed by the mild traumatic brain injury committee of the head injury interdisciplinary special interest group of the American congress of rehabilitation medicine. J Head Trauma Rehabil. 1993;8(3):86–87. doi:10.1097/00001199-199309000-00010
6. Maas AI, Roozenbeek B, Manley GT. Clinical trials in traumatic brain injury: past experience and current developments. Neurotherapeutics. 2010;7(1):115–26. doi:10.1016/j.nurt.2009.10.022 [PubMed: 20129503]
7. Hawryluk GW, Bullock MR. Past, present, and future of traumatic brain injury research. Neurosurgery Clinics. 2016;27(4):375–96. doi:10.1016/j.nec.2016.05.002 [PubMed: 27637391]
8. Saatman KE, Duhaime AC, Bullock R, Maas AI, Valadka A, Manley GT. Classification of traumatic brain injury for targeted therapies. J Neurotrauma. 2008;25(7):719–38. doi:10.1089/neu.2008.0586 [PubMed: 18627252]
9. Liu Y, Mazumdar S, Bath PA. An unsupervised learning approach to diagnosing Alzheimer's disease using brain magnetic resonance imaging scans. Int J Med Inform. 2023;173:105027. doi:10.1016/j.ijmedinf.2023.105027 [PubMed: 36921480]
10. Kraevsky-Phillips K, Sereika SM, Bouzid Z, Hickey G, Callaway CW, Saba S, Martin-Gill C, Al-Zaiti SS. Unsupervised machine learning identifies symptoms of indigestion as a predictor of acute decompensation and adverse cardiac events in patients with heart failure presenting to the emergency department. Heart Lung. 2023;61:107–13. doi:10.1016/j.hrtlng.2023.05.012 [PubMed: 37247537]
11. Folweiler KA, Sandsmark DK, Diaz-Arrastia R, Cohen AS, Masino AJ. Unsupervised machine learning reveals novel traumatic brain injury patient phenotypes with distinct acute injury profiles and long-term outcomes. J Neurotrauma. 2020;37(12):1431–44. doi:10.1089/neu.2019.6705 [PubMed: 32008422]

12. Qiu H, Zador Z, Lannon M, Farrokhyar F, Duda T, Sharma S. 327 Identification of clinically relevant patient endotypes in traumatic brain injury (TBI) using latent class analysis. Neurosurgery. 2023;69(Supplement_1):48–49. doi:10.1227/neu.0000000000002375_327

13. Rosenblatt CK, Harriss A, Babul AN, Rosenblatt SA. Machine learning for subtyping concussion using a clustering approach. Front Hum Neurosci. 2021;15:716643. doi:10.3389/fnhum.2021.716643 [PubMed: 34658816]

14. Daugherty J, Peterson A, Waltzman D, Breiding M, Chen J, Xu L, DePadilla L, Corrigan JD. Rationale for the development of a traumatic brain injury case definition for the pilot national concussion surveillance system. J Head Trauma Rehabil. 2023;39(2):115–20. doi:10.1097/HTR.0000000000000900

15. Daugherty J, Waltzman D, Breiding M, Peterson A, Chen J, Xu L, Womack LS, DePadilla L, Watson K, Corrigan JD, et al. Refinement of a preliminary case definition for use in traumatic brain injury surveillance. J Head Trauma Rehabil. 2023;39(2):121–39. doi:10.1097/HTR.0000000000000901

16. Gower JC. A General coefficient of similarity and some of its properties. Biometrics. 1971;27(4):857–71. doi:10.2307/2528823

17. Kaufman L, Rousseeuw PJ. Finding groups in data: an introduction to cluster analysis. New York: Wiley; 1990.

18. Maechler M, Rousseeuw P, Struyf A, Hubert M, Hornik K Cluster analysis basics and extensions. R package version 2.1.4. Cran; 2013. https://cran.r-project.org/web/packages/cluster/index.html

19. Krijthe JH, van der Maaten L: Rtsne: T-Distributed stochastic neighbor embedding using Barnes-hut implementation. R package version 0.17. Cran; 2023 https://cran.r-project.org/web/packages/Rtsne/index.html

20. Roy D, Peters ME, Everett AD, Leoutsakos JMS, Yan H, Rao VT, Bechtold K, Sair HI, Van Meter T, Falk H, et al. Loss of consciousness and altered mental state as predictors of functional recovery within 6 months following mild traumatic brain injury. J Neuropsychiatry Clin Neurosci. 2020;32(2):132–38. doi:10.1176/appi.neuropsych.18120379 [PubMed: 31530119]

21. Nelson LD, Furger RE, Ranson J, Tarima S, Hammeke TA, Randolph C, Barr WB, Guskiewicz K, Olsen CM, Lerner EB, et al. Acute clinical predictors of symptom recovery in emergency department patients with uncomplicated mild traumatic brain injury or non-traumatic brain Injuries. J Neurotrauma. 2018;35(2):249–59. doi:10.1089/neu.2017.4988 [PubMed: 29017409]

22. Walker WC, Stromberg KA, Marwitz JH, Sima AP, Agyemang AA, Graham KM, Harrison-Felix C, Hoffman JM, Brown AW, Kreutzer JS, et al. Predicting long-term global outcome after traumatic brain injury: development of a practical prognostic tool using the traumatic brain injury model systems national database. J Neurotrauma. 2018;35(14):1587–95. doi:10.1089/neu.2017.5359 [PubMed: 29566600]

23. Garduño-Ortega O, Li H, Smith M, Yao L, Wilson J, Zarate A, Bushnik T. Assessment of the individual and compounding effects of marginalization factors on injury severity, discharge location, recovery, and employment outcomes at 1 year after traumatic brain injury. Front Neurol. 2022;13:942001. doi:10.3389/fneur.2022.942001 [PubMed: 36090882]

24. Saadi A, Bannon S, Watson E, Vranceanu AM. Racial and ethnic disparities associated with traumatic brain injury across the continuum of care: a narrative review and directions for future research. J Racial Ethn Health Disparities. 2022;9(3):786–99. doi:10.1007/s40615-021-01017-4 [PubMed: 33733427]

25. Peterson AB, Zhou H, Thomas KE. Disparities in traumatic brain injury-related deaths—United States, 2020. J Safety Res. 2022;83:419–26. doi:10.1016/j.jsr.2022.10.001 [PubMed: 36481035]

26. World Health Organization. Social determinants of health; 2023. https://www.who.int/health-topics/social-determinants-of-health#tab=tab_1

27. Whiteneck GG, Cuthbert JP, Corrigan JD, Bogner JA. Risk of negative outcomes after traumatic brain injury: a statewide population-based survey. J Head Trauma Rehabil. 2016;31(1):E43–54. doi:10.1097/HTR.0000000000000141

28. Selassie AW, Pickelsimer EE, Frazier JL, Ferguson PL. The effect of insurance status, race, and gender on ED disposition of persons with traumatic brain injury. Am J Emerg Med. 2004;22(6):465–73. doi:10.1016/j.ajem.2004.07.024 [PubMed: 15520941]

29. Kontos AP, Jorgensen-Wagers K, Trbovich AM, Ernst N, Emami K, Gillie B, French J, Holland C, Elbin RJ, Collins MW, et al. Association of time since injury to the first clinic visit with recovery following concussion. JAMA Neurol. 2020;77(4):435. doi:10.1001/jamaneurol.2019.4552 [PubMed: 31904763]

30. Eagle SR, Puligilla A, Fazio-Sumrok V, Kegel N, Collins MW, Kontos AP. Association of time to initial clinic visit with prolonged recovery in pediatric patients with concussion. J Neurosurg Pediatr. 2020;26(2):165–70. doi:10.3171/2020.2.PEDS2025 [PubMed: 32330895]

31. Ruff RM, Iverson GL, Barth JT, Bush SS, Broshek DK. NAN policy and planning committee. Recommendations for diagnosing a mild traumatic brain injury: a national academy of neuropsychology education paper. Arch Clin Neuropsychol. 2009;24(1):3–10. doi:10.1093/arclin/acp006 [PubMed: 19395352]

32. Womack LS, Breiding MJ, Daugherty J. Concussion evaluation patterns among US adults. J Head Trauma Rehabil. 2022;37(5):303–10. doi:10.1097/HTR.0000000000000756 [PubMed: 35125431]

33. McCrea M, Hammeke T, Olsen G, Leo P, Guskiewicz K. Unreported concussion in high school football players: implications for prevention. Clin J Sport Med. 2004;14(1):13. doi:10.1097/00042752-200401000-00003 [PubMed: 14712161]

34. Meehan W, Mannix R, O’Brien M, Collins M. The prevalence of undiagnosed concussions in athletes. Clin J Sport Med. 2013;23(5):339–42. doi:10.1097/JSM.0b013e318291d3b3 [PubMed: 23727697]

35. Voss JD, Connolly J, Schwab KA, Scher AI. Update on the epidemiology of concussion/mild traumatic brain injury. Curr Pain Headache Rep. 2015;19(7):1–8. doi:10.1007/s11916-015-0506-z

36. Lucas S, Hoffman JM, Bell KR, Dikmen S. A prospective study of prevalence and characterization of headache following mild traumatic brain injury. Cephalalgia. 2014;34(2):93–102. doi:10.1177/0333102413499645 [PubMed: 23921798]

37. Machamer J, Temkin N, Dikmen S, Nelson LD, Barber J, Hwang P, Boase K, Stein MB, Sun X, Giacino J, et al. Symptom frequency and persistence in the first year after traumatic brain injury: a TRACK-TBI study. J Neurotrauma. 2022;39(5–6):358–70. doi:10.1089/neu.2021.0348 [PubMed: 35078327]

38. de Guise E, Leblanc J, Feyz M, Lamoureux J. Prediction of the level of cognitive functional independence in acute care following traumatic brain injury. Brain Inj. 2005;19(13):1087–93. doi:10.1080/02699050500149882 [PubMed: 16286322]

39. Tenovuo O, Diaz-Arrastia R, Goldstein LE, Sharp DJ, van der Naalt J, Zasler ND. Assessing the severity of traumatic brain injury—time for a change? J Clin Med. 2021;10(1):148. doi:10.3390/jcm10010148 [PubMed: 33406786]

40. Nelson LD, Temkin NR, Barber J, Brett BL, Okonkwo DO, McCrea MA, Giacino JT, Bodien YG, Robertson C, Corrigan JD, et al. Functional recovery, symptoms, and quality of life 1 to 5 years after traumatic brain injury. JAMA Netw Open. 2023;6(3):e233660. doi:10.1001/jamanetworkopen.2023.3660 [PubMed: 36939699]

41. Åkerlund CAI, Holst A, Stocchetti N, Steyerberg EW, Menon DK, Ercole A, Nelson DW, Åkerlund C, Amrein K, Andelic N. Clustering identifies endotypes of traumatic brain injury in an intensive care cohort: a CENTER-TBI study. Crit Care. 2022;26(1):228. doi:10.1186/s13054-022-04079-w [PubMed: 35897070]

42. Bergeron MF, Landset S, Maugans TA, Williams VB, Collins CL, Wasserman EB, Khoshgoftaar TM. Machine learning in modeling high school sport concussion symptom resolve. Med Sci Sports Exercise. 2019;51(7):1362–71. doi:10.1249/MSS.0000000000001903

43. Bittencourt NFN, Meeuwisse WH, Mendonça LD, Nettel-Aguirre A, Ocarino JM, Fonseca ST. Complex systems approach for sports injuries: moving from risk factor identification to injury pattern recognition—narrative review and new concept. Br J Sports Med. 2016;50(21):1309–14. doi:10.1136/bjsports-2015-095850 [PubMed: 27445362]

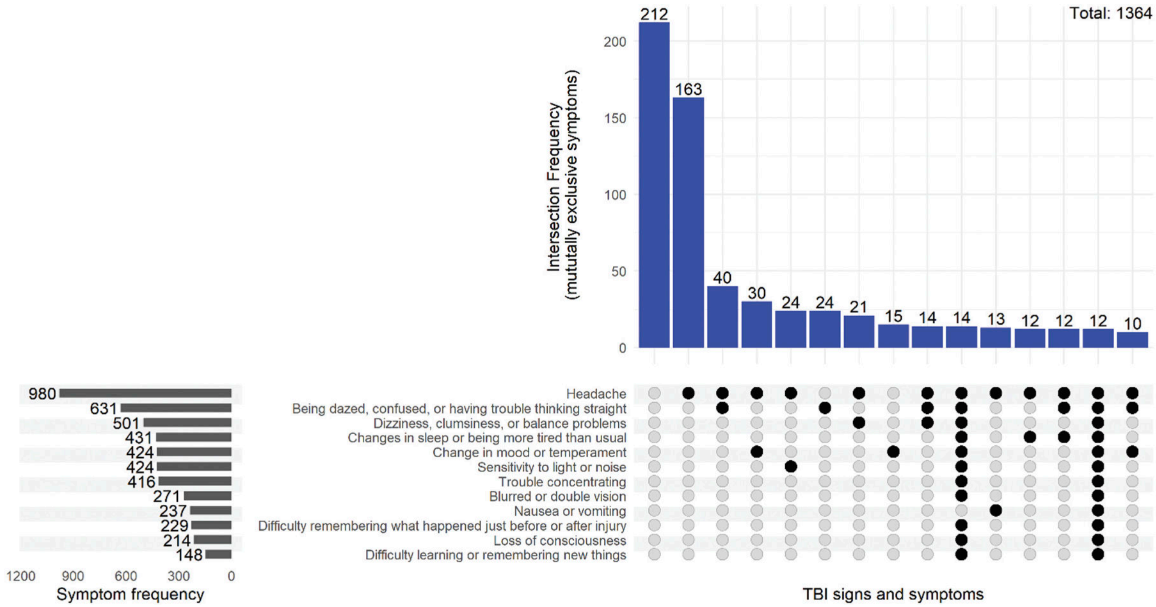

Figure 1.

UpSet plot of self-reported TBI signs and symptoms among adult respondents who sustained a head injury in the past 12 months, National Concussion Surveillance System Pilot, 2018–2019. The blue bars show the number of respondents with each mutually exclusive sign/symptom combination indicated by the black dots. The black bars show the frequency of each symptom.

Table 1.

Clinical phenotype among adults for their most recent head injury#, national concussion surveillance system pilot, 2018–2019.

Sign/Symptom	Phenotype A (cluster 1, *n* = 302)	Phenotype B (cluster 2, *n* = 176)	Phenotype C (cluster 3, *n* = 1298)	Phenotype D (cluster 4, *n* = 1376)	Phenotype E^ [Non-cases] (cluster 3, *n* = 1212)
	%	%	%	%	%
Headache	100	77	80	81	0
Dazed, confused, or trouble thinking straight	0	100	78	59	0
Dizziness, clumsiness, or balance problems	8	17	76	59	0
Trouble concentrating	3	9	65	52	0
Changes in sleep or more tired than usual	5	8	64	56	0
Loss of consciousness	1	10	61	3	0
Difficulty remembering what happened just before/after injury	2	5	61	8	0
Blurred or double vision	5	11	60	16	0
Sensitivity to light or noise	9	8	58	56	0
Change in mood or temperament	12	9	56	55	0
Nausea or vomiting	7	6	51	14	0
Difficulty	0	0	43	6	0
learning or remembering new things					
Number of signs and symptoms					
Mean (standard deviation)	1.5 (0.6)	2.6 (0.9)	7.5 (2.8)	4.8 (1.9)	-
Median (1st quartile, 3rd quartile)	1 (1,2)	3 (2,3)	8 (6,10)	5 (4,6)	-

–Not applicable. Indicates no analysis was conducted because it could not be computed.

#Cell labels for each Phenotype show the percentage of respondents who self-reported a TBI sign or symptom.

^Phenotype E was composed of all the non-cases (i.e., respondents that reported a head injury with no symptoms; these respondents were not considered to have sustained a TBI, according to the aforementioned TBI case definition).

Note: 0% represents a real value and not as ‘missing.’

Table 2.

Respondent outcomes by clinical traumatic brain injury phenotype[#], national concussion surveillance system pilot, 2018–2019.

	Total		PAM clustering (K = 5)											
	[n = 1364]		Phenotype A[1] (cluster 1, n = 302)		Phenotype B[2] (cluster 2, n = 176)		Phenotype C[3] (cluster 3, n = 298)		Phenotype D[4] (cluster 4, n = 376)		Phenotype E[5] [Non-cases] (cluster 1, n = 212)			
Outcome	**N**	**Weighted %**	**N**	**Weighted %**	**N**	**Weighted**	**N**	**Weighted %**	**N**	**Weighted %**	**N**	**Weighted %**	**Test Statistic**	**p-value**
Medically Evaluated													148.7	<0.001
No	919	71.1	252[a]	87.6	137[a]	83.3	109[b]	41.9	241[c]	69.3	180[a]	88.8		
Yes	375	28.9	40[a]	12.4	31[a]	16.7	165[b]	58.1	113[c]	30.7	26[a]	11.2		
Symptoms not resolved or that occurred for 8+ days													99.2	<0.001
No	728	63.6	263[a]	87.3	141[b]	76.0	101[c]	39.5	223[d]	61.9	-	-		
Yes	424	36.4	39[a]	12.7	35[b]	24.0	197[c]	60.5	153[d]	38.1	-	-		
One day of symptoms[6]													131.9	<0.001
No	837	75.0	155[a]	51.9	106[b]	64.6	278[c]	95.5	298[d]	78.6	-	-		
Yes	315	25.0	147[a]	48.1	70[b]	35.4	20[c]	4.5	78[d]	21.4	-	-		
Social Impairment													223.6	<.0001
Not At All/ Slightly	1068	82.6	285[a]	98.6	163[a]	97.5	144[b]	56.8	272[c]	76.9	204[d]	98.8		
Moderately/ Quite A Bit/ Extremely	208	17.4	7[a]	1.4	4[a]	2.5	118[b]	43.2	76[c]	23.1	3[d]	1.2		
Work Impairment													161.2	<.0001
Not At All/ Slightly	1052	80.9	281[a]	96.6	160[a]	94.7	142[b]	53.1	264[c]	76.4	205[d]	98.8		
Moderately/ Quite A Bit/ Extremely	226	19.1	11[a]	3.4	7[a]	5.3	121[b]	46.9	85[c]	23.6	2[d]	1.2		

Abbreviations: PAM = partitioning around medoids.

[a]Phenotype A is characterized by headache (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[b]Phenotype B is characterized by headache and dazed/confused/trouble thinking straight (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[c]Phenotype C is characterized by headache, dazed/confused/trouble thinking straight, Dizziness/clumsiness/balance problems, Trouble concentrating, Changes in sleep/being more tired than usual, Loss of consciousness, Difficulty remembering what happened just before/after injury, Blurred/Double vision, Sensitivity to light/noise, Change in mood/temperament, and Nausea/Vomiting (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[d]Phenotype D is characterized by headache, dazed/confused/trouble thinking straight, Dizziness/clumsiness/balance problems, Trouble concentrating, Changes in sleep/being more tired than usual, Sensitivity to light/noise, and Change in mood/temperament (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[e]Phenotype E is characterized by all the non-cases (i.e., respondents who reported a head injury with no symptoms).

[f]Indicator of less severe injury.

– Not applicable. Indicates no analysis was conducted because it could not be computed.

[#]For post hoc tests, pairwise comparisons of column proportions were computed. Letter superscripts (a, b, c, d) indicate which pairs of columns for a given row are significantly different, $p < 0.05$. If a pair of values are significantly different, the values have different subscript letters assigned to them. For example, for the variable ‘Medically Evaluated,’ for ‘Yes,’ Phenotype A and B are similar (not significantly different from each other) and display the same superscript (‘a’), but Phenotype C is significantly different than both Phenotype A and B and displays a different subscript (‘b’).

Author Manuscript

Author Manuscript

Author Manuscript

Table 3.

Respondent demographics by clinical traumatic brain injury phenotype[#], national concussion surveillance system pilot, 2018–2019.

	Total		PAM clustering (K = 5)											
	(*n* = 1364)		Phenotype A[1] (cluster 1, *n* = 302)		Phenotype B[2] (cluster 2, *n* = 176)		Phenotype C[3] (cluster 3, *n* = 298)		Phenotype D[4] (cluster 4, *n* = 376)		Phenotype E[5] [Non-cases] (cluster 1, *n* = 212)			
Characteristic	**N**	**Weighted %**	**N**	**Weighted %**	**N**	**Weighted**	**N**	**Weighted %**	**N**	**Weighted %**	**N**	**Weighted %**	**Test Statistic**	**p-value**
Sex													10.5	0.03
Female	644	48.1	152[a]	50.4	63[b]	34.5	156[a]	52.9	177[a]	48.2	96[a,b]	46.5		
Male	717	51.9	150[a]	49.6	112[b]	65.5	142[a]	47.1	197[a]	51.8	116[a,b]	53.5		
Age (years)													20.9	0.008
18 – <35	441	42.8	109[a]	49.8	56[a,b]	39.2	89[a,b]	39.5	132[a,b]	44.9	55[b]	35.7		
35 – <55	418	31.4	99[a,b]	30.5	62[a]	36.7	96[a]	35.2	114[a,b]	29.2	47[b]	24.4		
55+	486	25.9	89[a]	19.7	57[a]	24.1	111[a]	25.3	127[a]	25.8	102[b]	39.9		
Race/Ethnicity[6]													20.1	0.001
White, non-Hispanic	1030	65.7	235[a]	65.3	144[b]	79.1	196[c]	54.4	282[a,b]	69.1	173[a,b]	69.8		
Other	334	34.3	67[a]	34.7	32[b]	20.9	102[c]	45.6	94[a,b]	30.9	39[a,b]	30.2		
Education													46.9	<0.001
High school or less	316	37.0	63[a,c]	32.3	33[a,c]	30.0	101[b]	52.4	91[a]	35.1	28[c]	24.1		
Some college	409	38.1	83[a,b]	37.8	60[a]	44.7	96[b]	33.2	114[a,b]	39.3	56[a,b]	39.6		
Bachelor or higher	635	24.9	156[a,c,d]	29.8	81[a,c]	25.3	100[b]	14.4	170[c]	25.6	128[d]	36.2		
Marital Status													6.4	0.17
Married or living with partner	689	49.7	154[a]	46.9	96[a]	55.9	132[a]	44.5	191[a]	52.1	116[a]	54.9		
Other[7]	662	50.3	146[a]	53.1	78[a]	44.1	166[a]	55.5	181[a]	47.9	91[a]	45.1		
Rent or own a residence													15.7	0.003
Rent	489	39.6	106[a,c]	39.4	48[b]	25.0	123[c]	45.5	155[a,c]	43.4	57[a,b]	33.4		

	Total		PAM clustering (K = 5)											
	(*n* = 1364)		Phenotype A[1] (cluster 1, *n* = 302)		Phenotype B[2] (cluster 2, *n* = 176)		Phenotype C[3] (cluster 3, *n* = 298)		Phenotype D[4] (cluster 4, *n* = 376)		Phenotype E[5] [Non-cases] (cluster 1, *n* = 212)			
Characteristic	**N**	**Weighted %**	**N**	**Weighted %**	**N**	**Weighted**	**N**	**Weighted %**	**N**	**Weighted %**	**N**	**Weighted %**	**Test Statistic**	**p-value**
Own	820	60.4	$184^{a,c}$	60.6	122^{b}	75.0	158^{c}	54.5	$210^{a,c}$	56.6	$146^{a,b}$	66.6		

Abbreviations: PAM = partitioning around medoids.

[a]Phenotype A is characterized by headache (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[b]Phenotype B is characterized by headache and dazed/confused/trouble thinking straight (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[c]Phenotype C is characterized by headache, dazed/confused/trouble thinking straight, Dizziness/clumsiness/balance problems, Trouble concentrating, Changes in sleep/being more tired than usual, Loss of consciousness, Difficulty remembering what happened just before/after injury, Blurred/Double vision, Sensitivity to light/noise, Change in mood/temperament, and Nausea/Vomiting (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[d]Phenotype D is characterized by headache, dazed/confused/trouble thinking straight, Dizziness/clumsiness/balance problems, Trouble concentrating, Changes in sleep/being more tired than usual, Sensitivity to light/noise, and Change in mood/temperament (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[e]Phenotype E is characterized by all the non-cases (i.e., respondents who reported a head injury with no symptoms).

[f]Other includes non-Hispanic black, non-Hispanic asian, non-Hispanic other, and Hispanic.

[g]Other includes widowed, divorced, married but separated, and never married.

[#]For post hoc tests, pairwise comparisons of column proportions were computed. Letter superscripts (a, b, c, d) indicate which pairs of columns for a given row are significantly different, $p < 0.05$. If a pair of values are significantly different, the values have different subscript letters assigned to them. For example, for the variable ‘Sex’ for ‘Females,’ Phenotypes A, C, D, and E are similar (not significantly different from each other) and display the same superscript (‘a’). Phenotype B is significantly different than Phenotypes A, C, and D and displays a different subscript (‘b’), but it similar to Phenotype E (which also displays the subscript (‘b’).

Table 4.

Unadjusted odds ratio (OR) estimates for the relationship between clinical traumatic brain injury phenotypes*& and outcomes, national concussion surveillance system Pilot, 2018–2019.

	PAM clustering (K = 5)			
Outcome	**Phenotype C[1] (cluster 3, *n* = 298)**	**Phenotype D[2] (cluster 4, *n* = 376)**	**Phenotype B[3] (cluster 2, *n* = 176)**	**Phenotype A[4] (cluster 1, *n* = 302)**
Symptoms not resolved or that occurred for 8+ days - Yes vs. No				
OR	10.6	4.2	2.2	REF
95% CI	6.2–18.1	2.6–7.0	1.1–4.1	
p-value	<0.001	<0.001	0.02	
One day of symptoms[5] - Yes vs. No				
OR	0.1	0.3	0.6	REF
95% CI	0.03–0.1	0.2–0.4	0.4–0.9	
p-value	<0.001	<0.001	0.03	
Medically Evaluated – evaluated vs. not evaluated				
OR	54.7	3.1	1.4	REF
95% CI	5.8–16.6	1.9–5.2	0.8–2.7	
p-value	<0.001	<0.001	0.28	
Social Impairment- Moderately/Quite A Bit/Extremely vs. Not At All/Slightly				
OR	54.7	21.6	1.8	REF
95% CI	22.4–133.4	8.8–52.7	0.4–7.7	
p-value	<0.001	<0.001	0.42	
Work Impairment- Moderately/Quite A Bit/Extremely vs. Not At All/Slightly				
OR	25.4	8.8	1.6	REF
95% CI	11.2–57.2	3.9–19.9	0.4–6.0	
p-value	<0.001	<0.001	0.47	

Abbreviations: PAM = partitioning around medoids; OR = odds ratio; CI = confidence interval; REF = reference.

[a]Phenotype C is characterized by headache, dazed/confused/trouble thinking straight, Dizziness/clumsiness/balance problems, Trouble concentrating, Changes in sleep/being more tired than usual, Loss of consciousness, Difficulty remembering what happened just before/after injury, Blurred/Double vision, Sensitivity to light/noise, Change in mood/temperament, and Nausea/Vomiting (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[b]Phenotype D is characterized by headache, dazed/confused/trouble thinking straight, Dizziness/clumsiness/balance problems, Trouble concentrating, Changes in sleep/being more tired than usual, Sensitivity to light/noise, and Change in mood/temperament (where >50% of respondents in this cluster self-reported a particular sign or symptom).

[c]Phenotype B is characterized by headache and dazed/confused/trouble thinking straight (where >50% of respondents in this cluster selfreported a particular sign or symptom).

[d]Phenotype A is characterized by headache (where >50% of respondents in this cluster self-reported a particular sign or symptom). indicator of less severe injury.

[*]Phenotypes are ordered according to the clinical severity (most to least severe).

[&]This analysis was subset to cases. Thus, Phenotype E (characterized by the non-cases, i.e., respondents that reported a head injury with no symptoms) was removed from this analysis.