



# HHS Public Access

Author manuscript

*Child Abuse Negl.* Author manuscript; available in PMC 2023 March 01.

Published in final edited form as:

*Child Abuse Negl.* 2022 March ; 125: 105518. doi:10.1016/j.chiabu.2022.105518.

## Screening for pediatric abusive head trauma: Are three variables enough?

**Kent P. Hymel, MD<sup>a</sup>, Wouter Karst, MD<sup>b</sup>, Mark Marinello, MD<sup>c</sup>, Bruce E. Herman, MD<sup>d</sup>, Terra N. Frazier, DO<sup>e</sup>, Christopher L. Carroll, MD<sup>f</sup>, Veronica Armijo-Garcia, MD<sup>g</sup>, Matthew Musick, MD<sup>h</sup>, Kerri Weeks, MD<sup>i</sup>, Suzanne B. Haney, MD, MS<sup>j</sup>, Afshin Pashai, MBA<sup>k</sup>, Ming Wang, PhD<sup>l</sup>**  
**Pediatric Brain Injury Research Network (PediBIRN) Investigators**

<sup>a</sup>Department of Pediatrics, Penn State College of Medicine, Penn State Health Children's Hospital, 600 University Drive, Hershey, Pennsylvania 17033, USA

<sup>b</sup>Department of Forensic Medicine, Netherlands Forensic Institute, PO Box 24044, 2490 AA The Hague, The Netherlands

<sup>c</sup>Department of Pediatrics, Children's Hospital of Richmond at VCU, 1250 East Marshall Street, Richmond, Virginia 23219, USA

<sup>d</sup>Department of Pediatrics, University of Utah School of Medicine, Primary Children's Hospital, 100 North Mario Capecchie Drive, Salt Lake City, Utah 84113, USA

<sup>e</sup>Department of Pediatrics, Children's Mercy Hospital, 2401 Gillham Road, Kansas City, Missouri 64108, USA

<sup>f</sup>Department of Pediatrics, Connecticut Children's Medical Center, 282 Washington Street, Hartford, Connecticut 06106, USA

<sup>g</sup>University of Texas Health Sciences Center at San Antonio, 7703 Floyd Curl Drive, San Antonio, Texas 78229, USA

<sup>h</sup>Department of Pediatrics, Baylor College of Medicine, Texas Children's Hospital, 6621 Fannin Street, Houston, Texas 77030, USA

<sup>i</sup>Department of Pediatrics, University of Kansas School of Medicine, 3243 East Murdoch, Wichita, Kansas 67208, USA

---

**Corresponding Author:** Kent P. Hymel, MD, Penn State Health Milton S. Hershey Medical Center, Department of Pediatrics, Division of Child Abuse Pediatrics, 500 University Drive, MC H085, Hershey PA, 17033, kphymel@gmail.com; Cell phone: 703-674-8989.

**Author Contributions:** Drs. Hymel, Karst, and Wang made substantial contributions to the conception and design of the study. Dr. Wang designed and executed all data analyses. Mr. Pashai facilitated development of the probability calculators. The remaining authors made substantial contributions to the acquisition of data. Dr. Hymel had full access to all the data in the study, drafted the manuscript, and takes responsibility for the integrity of the work as a whole. All authors participated actively in the revision of the manuscript for important intellectual content and have approved the final manuscript as submitted.

**Publisher's Disclaimer:** This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

**Conflicts of Interest:** The authors have no potential, real or perceived, personal or financial, conflicts of interest to report related to this study or manuscript. No honorarium, grant, or other form of payment was given to anyone to produce this manuscript.

<sup>j</sup>Department of Pediatrics, University of Nebraska Medical Center, Children's Hospital and Medical Center, 8200 Dodge Street, Omaha, Nebraska 68114, USA

<sup>k</sup>Department of Health Administration and Policy, George Mason University, 4400 University Drive, Fairfax, Virginia 22030, USA

<sup>l</sup>Department of Public Health Sciences, Penn State College of Medicine, 700 HMC Crescent Road, Hershey, Pennsylvania 17033, USA

## Abstract

**Background:** The PediBIRN 4-variable clinical decision rule (CDR) detects abusive head trauma (AHT) with 96% sensitivity in pediatric intensive care (PICU) settings. Preliminary analysis of its performance in Pediatric Emergency Department settings found that elimination of its fourth predictor variable enhanced screening accuracy.

**Objective:** To compare the AHT screening performances of the “PediBIRN-4” CDR vs. the simplified 3-variable CDR in PICU settings.

**Participants and Settings:** 973 acutely head-injured children <3 years hospitalized for intensive care across 18 sites between February 2011 and March 2021.

**Methods:** Retrospective, secondary analysis of the combined, prospective PediBIRN data sets. AHT definitional criteria and physicians' diagnoses were applied iteratively to sort patients into abusive vs. other head trauma cohorts. Outcome measures of CDR performance included sensitivity, specificity, predictive values, likelihood ratios, ROC AUC, and the correlation between each CDR's patient-specific estimates of AHT probability and the overall positive yield of patients' completed abuse evaluations.

**Results:** Applied accurately and consistently, both CDR's would have performed with sensitivity 93% and negative predictive value 91%. Eliminating the PediBIRN-4's fourth predictor variable resulted in significantly higher specificity (↑d 19%), positive predictive value (↑d 8%), and ROC AUC (↑d 5%), but a 3% reduction in sensitivity. Both CDRs provided patient-specific estimates of abuse probability very strongly correlated with the positive yield of patients' completed abuse evaluations (Pearson's  $r = .95$  and  $.91$ ,  $p = .13$ ).

**Conclusion:** The PediBIRN 3-variable CDR performed with greater AHT screening accuracy than the 4-variable CDR. Both are good predictors of the results of patients' subsequent completed abuse evaluations.

## Keywords

abusive head trauma; child abuse; screening test; clinical decision rule; clinical prediction rule

## INTRODUCTION

To protect young victims of abusive head trauma (AHT) from further abuse, physicians must consider, recognize, evaluate, diagnose, and report suspected abuse. Unfortunately, physicians' responses to child maltreatment have been inconsistent (Jenny, Hymel, Ritzen, Reinert, & Hay, 1999; Lane, & Dubowitz, 2007; Lane, Rubin, Monteith, & Christian, 2002;

Wood et al., 2010; Hymel et al., 2018), the prevalence of AHT has remained consistently high (Hymel et al., 2013; 2014; 2021), and victims of AHT continue to be missed (Jenny, et al., 1999; Letson et al., 2016).

To minimize missed cases, Pediatric Brain Injury Research Network (PediBIRN) investigators conducted sequential, prospective, multicenter studies to derive, validate, and implement a 4-variable clinical decision rule (CDR) for AHT (Hymel et al., 2013; 2014; 2021). Applied as a screening tool, the “PediBIRN-4” CDR directs physicians to complete thorough abuse evaluations on all young, acutely head-injured, “higher risk” patients who present for intensive care with any one or more of its four predictor variables: (1) acute respiratory compromise; (2) bruising of the torso, ear(s), or neck; (3) bilateral or interhemispheric subdural hemorrhage(s) or collection(s); and (4) complex skull fracture(s). Applied accurately and consistently, the PediBIRN-4 will “miss” (stratify as lower risk) approximately 4% of cases (Hymel et al., 2013; 2014; 2015; 2021).

In a recent, external validation study of its potential AHT screening performance in pediatric emergency department (ED) settings (Hymel et al., in press), using data captured by an independent research network, the PediBIRN-4 again demonstrated sensitivity of 0.96 [95% CI: 0.88-0.99], correctly categorizing 75 (96%) of 78 AHT patients as higher risk. Specificity in the ED setting was only 0.29 [95% CI: 0.16-0.46]. Sensitivity analysis of the external ED data set (N=116) revealed that application of a simplified CDR—based solely on the PediBIRN-4’s first three predictor variables—would have increased specificity to 0.84 [95% CI: 0.68-0.93] without compromising sensitivity.

To help confirm or exclude a relative advantage in CDR simplification for application in pediatric intensive care unit (PICU) settings, we conducted a retrospective secondary analysis of the much larger (N=973), combined, PediBIRN derivation, validation, and implementation study data sets (Hymel et al., 2013; 2014; 2021). Our objective was to measure and compare the PICU-based AHT screening performances of the original 4-variable CDR to that of a simplified 3-variable CDR (the “PediBIRN-3”) based solely on the PediBIRN-4’s first three predictor variables (see Table 1). We hypothesized that, applied as directive clinical *decision* rules, the 3-variable CDR would perform with significantly higher specificity, but lower sensitivity. Applied instead as informative clinical *prediction* rules, we hypothesized that both CDRs would provide patient-specific estimates of AHT probability that correlated strongly with the overall positive yield of PediBIRN patients’ subsequent, completed, skeletal surveys and retinal exams.

## MATERIALS AND METHODS

This was a novel, retrospective, secondary analysis of the existing, combined, de-identified PediBIRN data sets. Data were captured between February 2011 and March 2021 in sequential, multicenter studies conducted across 18 North American pediatric intensive care units (PICU). All three PediBIRN studies (Hymel et al., 2013; 2014; 2021) used the same inclusion and exclusion criteria, a priori definitional criteria for AHT (see Table 2), and patient-related data forms. For patients and their families, all three studies were strictly observational. Therefore, at every participating PICU, local institutional review boards

approved study participation with waivers of parental informed consent. The Institutional Review Board at Penn State Health Hershey Medical Center determined that this secondary analysis was not human subject research.

Eligible PediBIRN patients were children under 3 years of age hospitalized for intensive care of acute, closed, traumatic, cranial or intracranial injuries confirmed on initial neuroimaging. Patients with preexisting brain abnormalities and patients whose head injuries resulted from collisions involving motor vehicle(s) were excluded. In all three prior PediBIRN studies: (1) participating PICUs captured complete data regarding >90% of their consecutive eligible patients; (2) prospective study design facilitated capture of uniform demographic, historical, clinical, and radiological data; and (3) data inconsistencies were tracked until resolution (Hymel et al., 2013; 2014; 2021).

To compare 3- vs. 4-variable CDR *screening* performances, we (1) sorted PediBIRN patients iteratively into AHT vs. other head trauma cohorts based on the network's longstanding AHT definitional criteria (see Table 2), and based on PICU and child abuse pediatric (CAP) physicians' final, consensus, diagnostic impressions of definitive/probable AHT vs. other head trauma; (2) applied the PediBIRN 3- and 4-variable CDR's (see Table 1) iteratively to stratify patients as higher vs. lower risk; (3) created four 2x2 contingency tables; (4) populated their cells with counts that sorted patients' as abusive vs. other head trauma and higher vs. lower risk; (5) calculated 3- and 4-variable CDR test characteristics (sensitivity, specificity, predictive values, likelihood ratios, ROC AUC) within 95% confidence intervals; and (6) compared 3- vs. 4-variable CDR test characteristics using the McNemar's test (for sensitivity and specificity), the relative predictive values approach (for predictive values), the regression model approach (for likelihood ratios), or the DeLong's method (for ROC AUC), as appropriate. The analyses were conducted using the package of DTComPair (Stock & Hielscher, 2014) in R version 4.0.2.

To compare 3- vs. 4-variable CDR *prediction* performances, we: (1) applied the network's AHT definitional criteria (see Table 2) to sort patients into abusive vs. other head trauma cohorts; (2) applied the CDRs iteratively to divide the study population into subpopulations that presented for intensive care with unique combinations of each CDR's predictor variables; (3) calculated the percentage of patients sorted as AHT within each subpopulation; (4) assigned that same value, expressed as an estimate of abuse probability, to every patient within that specific subpopulation; (5) calculated the percentage of patients within each subpopulation who revealed findings moderately or highly specific for abuse on their completed skeletal surveys and/or retinal exams; (6) calculated Pearson correlation coefficients, across CDR-defined patient subpopulations, comparing patient specific estimates of abuse probability and the overall positive yield of patients' subsequent, completed, abuse evaluations; and (7) applied Meng, Rosenthal, and Rubin's method (1992) for comparing correlated correlation coefficients.

To provide a measure of clinical context, we also completed a post hoc analysis designed to estimate physicians' relative willingness or reluctance to complete abuse evaluations on patients presenting with the PediBIRN-4's fourth predictor variable (complex skull fractures). Having segregated the PediBIRN-4's higher risk patient population into 15

unique subpopulations, we calculated the proportion of patients in each higher risk subpopulation evaluated with skeletal survey and/or retinal exam, and the overall positive yield of their completed abuse evaluations.

## RESULTS

Significant differences were observed in the demographics of comparative patient groups (see Table 3). Our overall results can be summarized as follows: Applied accurately and consistently as a directive decision rule, the PediBIRN-3 would have performed with greater AHT screening accuracy than the PediBIRN-4 in our study population of 973 young, acutely head-injured patients hospitalized for intensive care. The PediBIRN-3 demonstrated significantly higher specificity (0.56 vs. 0.37,  $p < .001$  and 0.65 vs. 0.44,  $p < .001$ , applying AHT definitional criteria and physicians' final diagnoses, respectively), positive predictive value (0.62 vs. 0.54,  $p < .001$  and 0.74 vs. 0.64,  $p < .001$ ), positive likelihood ratio (2.10 vs. 1.52,  $p < .001$  and 2.69 vs. 1.74,  $p < .001$ ), and ROC AUC (0.83 vs. 0.77,  $p = .002$  and 0.89 vs. 0.82,  $p < .001$ ). These performance enhancements exacted a smaller yet statistically significant reduction in sensitivity (0.96→0.93,  $p = .002$  and 0.98→0.95,  $p = .001$ ). Detailed results are presented in Table 4. Applied instead as evidence-based prediction rules, both CDRs provided patient-specific estimates of abuse probability that were highly correlated with the positive yield of patients' subsequent, completed abuse evaluations (Pearson's  $r = .95$  [95% CI: 0.73-0.99] and .91 [95% CI: 0.76-0.97],  $p = .13$ ; see Table 5 and Figure 1).

Our post hoc analysis revealed that the PediBIRN-4's two higher risk subpopulations that physicians across 18 participating sites evaluated least often for abuse were patients presenting with complex skull fracture(s); with or without acute respiratory compromise; but no bruising of the torso, ear(s), or neck; and no subdural hemorrhage(s) or collection(s). These same two patient subpopulations (see Table 5, data in bold font) also had the lowest positive yields on their completed abuse evaluations—yields on par with patients that the PediBIRN-4 categorized as “lower risk” (see Table 5 and Figure 1, arrow and legend).

## DISCUSSION

The PediBIRN-4's potential AHT screening sensitivity of 96% has been validated in pediatric inpatient (Pfeiffer et al., 2018), intensive care (Hymel et al., 2014), and ED (Hymel et al., in press) settings; applying diverse criteria or methods to define AHT; through analysis of prospective data captured by three independent research networks; and in populations with divergent AHT prevalence. To achieve such high sensitivity, it casts a very wide net. In contrast, the simplified 3-variable CDR casts a smaller net more accurately.

Because missing AHT creates substantial risk (Jenny, et al., 1999), we speculate that many PICU and CAP physicians would opine that casting a wide net is necessary. Given a choice, these physicians would likely deem a 3% reduction in sensitivity to be unacceptable, and would opt to adopt the PediBIRN-4 as their AHT screening tool, even if doing so requires that more “higher risk” patients with non-AHT are evaluated for abuse (see Table 4).

We speculate that those physicians who disagree would cite concern that completing too many “avoidable” abuse evaluations—to miss fewer cases of AHT—increases parental stress

and distrust, increases health care costs, prolongs hospitalizations, and exposes patients to additional risks (false positive results, radiation exposure). Given a choice, these physicians would likely deem a 3% reduction in sensitivity to be a reasonable cost to secure the benefits of a 19-21% increase in specificity and 8-10% increase in PPV (see Table 4).

Our post hoc analysis revealed that many PICU and CAP physicians elected to defer abuse evaluations on their higher risk patients who presented with the PediBIRN-4's fourth predictor variable (complex skull fractures); but no bruising of the torso, ear(s), or neck; and no subdural hemorrhages or collections (see Table 5, data in large bold font). Interestingly, these two higher risk patient subpopulations were also those least likely to reveal corroborating findings of abuse on their subsequent, completed abuse evaluations. In fact, the overall positive yields of completed skeletal surveys and retinal exams in these two "higher risk" patient subpopulations were on par with patients that the PediBIRN-4 categorized as "lower risk" (see Table 5 and Figure 1, arrow and legend). These results support an impression that many physicians have judged correctly that these two patient subpopulations are not at higher risk for abuse. Given a choice, these physicians would likely adopt the simplified 3-variable CDR as their AHT screening tool. Doing so provides assurance that patients who present with any one or more of its three predictor variables will have estimated probabilities of abuse 0.20 (see Table 5 and Figure 1).

Applied to the external pediatric ED data set (N=116), elimination of the PediBIRN-4's fourth predictor variable improved AHT screening performance significantly (Hymel et al., in press). High sensitivity was maintained (0.96 [95% CI: 0.88-0.99]), specificity and positive likelihood ratio increased (0.29 [95% CI: 0.16-0.46] → 0.84 [95% CI: 0.68-0.93] and 1.35 [95% CI: 11.10-1.67] → 6.09 [95% CI: 2.92-12.71], respectively), and negative likelihood ratio decreased (0.13 [95% CI: 0.04-0.46] → 0.05 [95% CI: 0.01-0.14]). Although larger, prospective, validation studies of its performance in ED settings are clearly warranted, these results support an impression that the PediBIRN-3 is the preferred CDR for AHT screening in pediatric ED settings, where high volumes of young children are diagnosed with accidental skull fractures attributed to simple falls.

Looking back, it is interesting to note that the PediBIRN-3 was one of four candidate CDRs considered by PediBIRN investigators in our original, multicenter, CDR derivation study (Hymel et al., 2013). Each performed with sensitivity 0.92. Seeking first and foremost to minimize missed cases of AHT, we opted to cast the widest possible net, and thus adopted the PediBIRN-4 for subsequent validation and implementation (Hymel et al., 2014; 2021). The results of this secondary analysis have returned us "full circle" to reconsider the PediBIRN-3. A screening tool only has value if/when physicians accept and apply it.

The PediBIRN 3- and 4-variable CDRs were developed to inform clinical judgement. They were not designed to replace it. The presenting history, past and family medical history, psychosocial risk assessment, results of tests to confirm or exclude medical mimics, and input from investigators must all be considered. In many cases, these additional data provide clarity. In the remaining cases, when uncertainty regarding the need to evaluate for abuse persists, following the recommendation of a validated, directive, AHT screening tool could lessen practice disparities, missed AHT, and the impacts of physician inexperience and

implicit bias (Jenny, et al., 1999; Lane, et al., 2007, 2010; Letson et al., 2016; Hymel et al., 2018; Wood, et al., 2010).

Clinical care units will now have the option of adopting either the 3- or 4-variable CDR as their AHT screening tool. Applied accurately and consistently as directive AHT screening tools, both perform with higher sensitivity than current AHT screening practices, estimated to be 87% (Letson et al., 2016; Hymel et al., 2015). Physicians who reject the PediBIRN CDRs' directive recommendations can opt instead to apply either CDR as an evidence-based prediction tool (see Table 5 and Figure 1), knowing that both facilitate patient-specific estimation of AHT probability—at or near the time of acute clinical presentation—that is highly correlated with the positive yield of patients' subsequent, completed, abuse evaluations. To apply the PediBIRN-3 and PediBIRN-4 “AHT probability calculators”, visit [www.pedibirn.com](http://www.pedibirn.com).

With this study, the AHT screening performances of the PediBIRN-3 and PediBIRN-4 CDRs have now been validated in both pediatric intensive care and ED settings (Hymel et al., 2014; Hymel et al., in press). Other evidence-based decision rules and prediction tools can/will provide decision support for other consequential decisions regarding possible child physical abuse. Berger et al's PIBIS (Pittsburgh brain injury score) identifies infants at risk for brain injury or AHT who might benefit from neuroimaging (Berger et al., 2016). Pierce et al's BCDRs (bruising clinical decision rules) identify young children with bruising who need evaluation for abuse (Pierce, Kaczor, Aldridge, O'Flynn, & Lorenz, 2010; Pierce et al., 2021). Maguire and colleagues' PredAHT (predicting abusive head trauma) and the PediBIRN-7 are prediction tools that apply the (positive or negative) results of completed abuse evaluations to estimate AHT probability (Cowley, Morris, Maguire, Farewell, & Kemp, 2015; Maguire, Kemp, Lumb, & Farewell, 2011; Hymel et al., 2019).

This study had strengths. The study population was relatively large (N=973). Uniform, complete, patient-specific, clinical, historical, and radiological data were collected prospectively across 18 PICUs. The differences in 3- vs. 4-variable CDR screening performance (see Table 4) were similar using two different methods for sorting AHT vs. other head trauma. We conducted post-hoc analyses that provided insight into the PediBIRN-3's likely acceptability to physicians. Whereas our network's previously published, patient-specific estimates of AHT probability (Hymel et al., 2015) were based on uniform, prospective data from 500 patients, the revised estimates provided in Table 5 are based on equivalent prospective data regarding 973 patients.

The study also had limitations. The study population was limited to young, acutely head-injured patients hospitalized for intensive care. Data regarding PediBIRN CDR performance in other clinical settings is limited (Hymel et al., in press; Pfeiffer et al, 2018). Because there is no gold standard for the diagnosis of AHT, our estimates of CDR test performance and abuse probability are likely inaccurate (see Tables 4 and 5 and Figure 1). Because not every PediBIRN patient underwent skeletal survey and retinal exam, estimates of the overall positive yield of completed abuse evaluations in specific patient subpopulations (see Table 5 and Figure 1) are likely inaccurate as well. Although our results inform estimates of each CDR's performance, substantial additional research is needed to measure, compare,

and contrast their actual acceptability, adoption, and effectiveness as AHT screening and prediction tools, in PICU and in other clinical settings.

## CONCLUSION

The AHT screening performances of the directive PediBIRN 3- and 4-variable CDRs have now been validated in both PICU and pediatric ED settings. Applied as prediction tools, both facilitate early, patient-specific estimation of AHT probability highly predictive of the positive results of patients' subsequent, completed, abuse evaluations. Applied accurately and consistently as a directive decision rule in our PICU study population, the PediBIRN 3 would have performed with greater overall accuracy, but with 3% lower sensitivity. Both CDRs have demonstrated the potential to reduce missed AHT cases below recent estimates. Physicians who make decisions to launch or forgo abuse evaluations must weigh the relative costs vs. benefits of adopting either CDR as an AHT screening tool. PediBIRN investigators welcome independent studies of PediBIRN-3 and PediBIRN-4 effectiveness in more diverse clinical settings.

## Acknowledgments:

The authors would like to acknowledge and thank the remaining PediBIRN investigators who helped to capture the data used in this secondary analysis: Antoinette Laskey, MD, MPH (Primary Children's Medical Center, Salt Lake City, UT); Douglas F. Willson, MD and Robin Foster, MD (The Children's Hospital of Richmond, Richmond, VA); Sandeep K. Narang, MD, JD (University of Texas Health Sciences Center at San Antonio, San Antonio, TX); Deborah A. Pullin, BSN, APRN (Dartmouth-Hitchcock Medical Center, Lebanon, NH); Jeanine M. Graf, MD and Reena Isaac, MD (Texas Children's Hospital, Houston, TX); Kelly Tieves, MD (Children's Mercy Hospital, Kansas City, MO); Edward Truemper, MD (Children's Hospital of Omaha, Omaha, NE); Lindall E. Smith, MD (Wesley Medical Center, Wichita, KS); Renee A. Higgerson, MD and George A. Edwards, MD (Dell Children's Medical Center of Central Texas, Austin, TX); Nancy S. Harper, MD, FAAP and Karl L. Serrao, MD, FAAP, FCCM (Driscoll Children's Hospital, Corpus Christi, TX); Andrew Sirotnak, MD, Joseph Albiez, MD, and Antonia Chiesa, MD (Children's Hospital Colorado, Denver, CO); Christine McKiernan, MD (Baystate Medical Center, Springfield, MA); Mark S. Dias, MD (Penn State College of Medicine, Hershey, PA); Michael Stoiko, MD, Debra Simms, MD, FAAP, and Sarah J. Brown, DO, FACOP, FAAP (Helen DeVos Children's Hospital, Grand Rapids, MI); Amy Ornstein, MD, FRCPC (IWK Health Centre, Halifax, Nova Scotia) AND Phil Hyden, MD (Children's Hospital of Central California, Madera, CA)

## Sources of Funding:

This study was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (grant number P50HD089922) and the Penn State Clinical & Translational Research Institute, Pennsylvania State University CTSA (NIH/CTSA grant number UL1 TR002014). The National Institutes of Health and Pennsylvania State University had no role in the design or conduct of the study; the collection, management, analysis, or interpretation of the data; the preparation, review, or approval of the manuscript; or the decision to submit the manuscript for publication. The content of this study is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health or Pennsylvania State University.

## REFERENCES

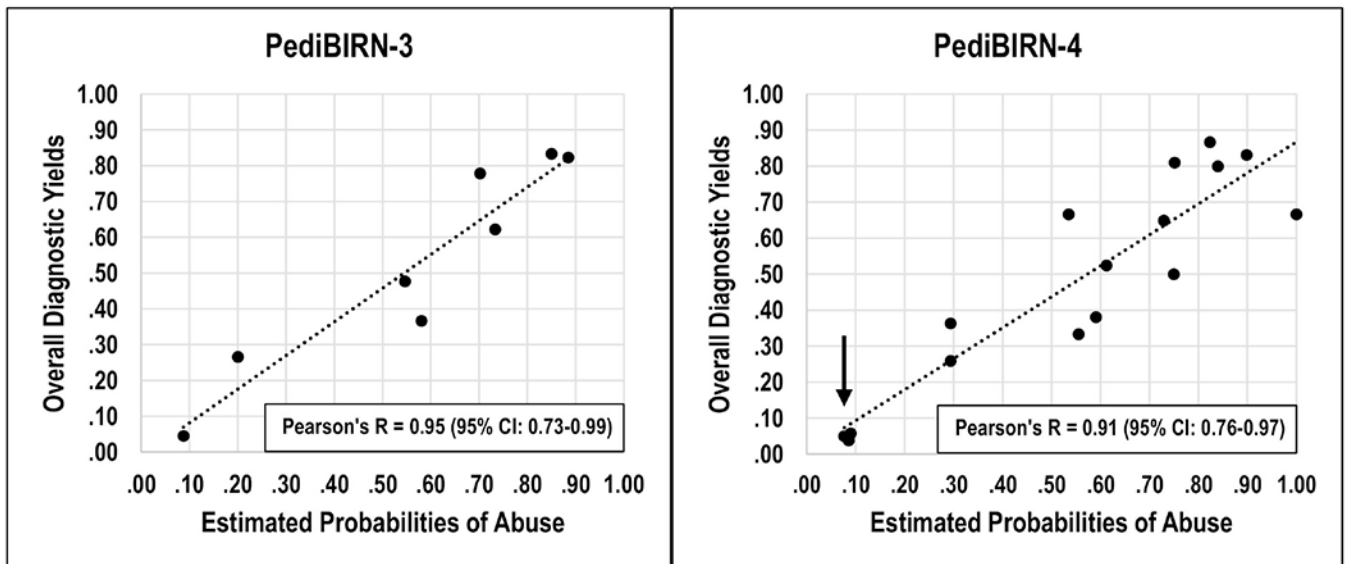
- Berger RP, Fromkin J, Herman B, Pierce MC, Saladino RA, Flom L, ...Kochanek PM (2016). Validation of the Pittsburgh infant brain injury score for abusive head trauma. *Pediatrics*, 138, e20153756. [PubMed: 27338699]
- Cowley LE, Morris CB, Maguire SA, Farewell DM, & Kemp AM (2015). Validation of a prediction tool for abusive head trauma. *Pediatrics*, 136, 290–298. [PubMed: 26216332]
- Hymel KP, Willson DF, Boos SC, Pullin DA, Homa K, Lorenz DJ, ...Armijo-Garcia V (2013). Derivation of a clinical prediction rule for pediatric abusive head trauma. *Pediatric Critical Care Medicine*, 14, 210–220. [PubMed: 23314183]



4. Hymel KP, Armijo-Garcia V, Foster R, Frazier TN, Stoiko M, Christie LM, ... Wang M (2014). Validation of a clinical prediction rule for pediatric abusive head trauma. *Pediatrics*, 134, e1537–1544. [PubMed: 25404722]
5. Hymel KP, Herman BE, Narang SK, Graf JM, Frazier TN, Stoiko M, ... Wang M (2015). Potential impact of a validated screening tool for pediatric abusive head trauma. *The Journal of Pediatrics*, 167,1375–1381. [PubMed: 26477871]
6. Hymel KP, Laskey AL, Crowell KR, Wang M, Armijo-Garcia V, Frazier TN, ... Weeks K (2018). Racial/Ethnic disparities and bias in the evaluation and reporting of abusive head trauma. *The Journal of Pediatrics*, 198,137–143. [PubMed: 29606408]
7. Hymel KP, Wang M, Chinchilli VM, Karst WA, Willson DF, Dias MS, ... Isaac R (2019) Estimating the probability of abusive head trauma after abuse evaluation. *Child Abuse Neglect*, 88, 266–274. [PubMed: 30551063]
8. Hymel KP, Armijo-Garcia V, Musick M, Marinello M, Herman BE, Weeks K, ... Noll J (2021). A cluster randomized trial to reduce missed abusive head trauma in pediatric intensive care settings. *The Journal of Pediatrics*, 236, 260–268. [PubMed: 33798512]
9. Hymel KP, Fingarson AK, Pierce MC, Kaczor K, Makoroff KL, & Wang M External validation of the PediBIRN screening tool for abusive head trauma in pediatric emergency department settings. *Pediatric Emergency Care*, in press.
10. Jenny C, Hymel KP, Ritzen A, Reinert SE, & Hay TC (1999). An analysis of missed cases of abusive head trauma. *Journal of the American Medical Association*, 281, 621–626. [PubMed: 10029123]
11. Lane WG, Rubin DM, Monteith R, & Christian CW (2002). Racial differences in the evaluation of pediatric fractures for physical abuse. *Journal of the American Medical Association*, 288,1603–1609. [PubMed: 12350191]
12. Lane WG, & Dubowitz H (2007). What factors affect the identification and reporting of child abuse-related fractures? *Clinical Orthopaedics and Related Research*, 461, 219–225. [PubMed: 17415005]
13. Letson MM, Cooper JN, Deans KJ, Scribano P, Makoroff KL, ... Berger RP (2016). Prior opportunities to identify abuse in children with abusive head trauma. *Child Abuse Neglect*, 60, 36–45. [PubMed: 27680755]
14. Maguire SA, Kemp AM, Lumb RC, & Farewell DM (2011). Estimating the probability of abusive head trauma: A pooled analysis. *Pediatrics*, 128, e550–e564. [PubMed: 21844052]
15. Meng X.I., Rosenthal R, & Rubin DB (1992). Comparing correlated correlation coefficients. *Psychological Bulletin*, 111,172–175.
16. Pfeiffer H, Smith A, Kemp AM, Cowley LE, Cheek JA, Dalziel SR, ... Babl FE (2018). External validation of the PediBIRN clinical prediction rule for abusive head trauma. *Pediatrics*, 141, e20173674. [PubMed: 29700200]
17. Pierce MC, Kaczor K, Aldridge S, O’Flynn J, & Lorenz DJ (2010), Bruising characteristics discriminating physical child abuse from accidental trauma. *Pediatrics*, 125, 67–74. [PubMed: 19969620]
18. Pierce MC, Kaczor K, Lorenz DL, Bertocci G, Fingarson A, Makoroff K, ... Levanthal JM (2021). Validation of a clinical decision rule to predict abuse in young children based on bruising characteristics. *Journal of the American Medical Association Network Open*, 4, e215832.
19. Stock C, & Hielscher T (2014). DTComPair: comparison of binary diagnostic tests in a paired study design. R package version 1.0.3. URL: <http://CRAN.R-project.org/package=DTComPair> (last referenced September 20, 2021).
20. Wood JN, Hall M, Schilling S, Keren R, Mitra N, & Rubin DM (2010). Disparities in the evaluation and diagnosis of abuse among infants with traumatic brain injury. *Pediatrics*, 126, 408–414. [PubMed: 20713477]

**HIGHLIGHTS**

- Two PediBIRN clinical decision rules (CDR) screen effectively for abusive head trauma.
- The PediBIRN 4-variable clinical decision rule casts a broad net to detect 96% of cases.
- The PediBIRN 3-variable clinical decision rule detects 93% of cases with fewer false positives.
- Both CDRs predict the positive yield of patients' subsequent completed abuse evaluations.



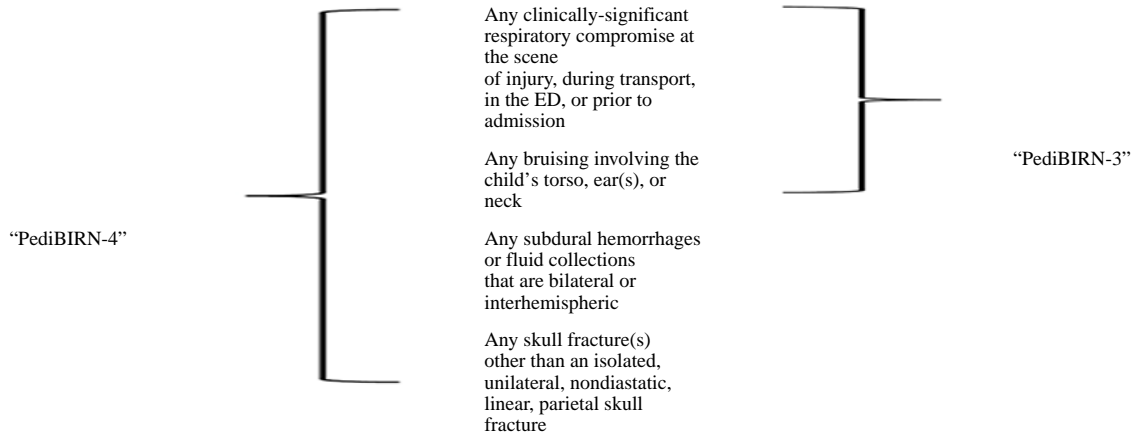
**Figure 1.**

The two plots present the intersections of ‘estimated probabilities of abuse’ and ‘overall diagnostic yield’ across patient subpopulations presenting with different combinations of each clinical decision rule’s predictor variables. The estimated probabilities of abuse were calculated applying definitional criteria for abusive head trauma (see Table 2). The data used to create the plots can be found in Table 5. The two correlations did not differ significantly ( $p = .13$ ). In the plot on the right, the cluster of three data points marked by the arrow presents data regarding the 4-variable decision rule’s “lower risk” patients, and its two “higher risk” patient subpopulations highlighted in large bold font in Table 5.

**Table 1.**

The Pediatric Brain Injury Research Network’s directive 3- and 4-variable clinical decision rules for abusive head trauma.

To minimize missed cases, every acutely head-injured infant or young child hospitalized for intensive care who presents with one or more of these predictor variables should be considered higher risk and thoroughly evaluated for abuse.



Abbreviations: PediBIRN=pediatric brain injury research network

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

**Table 2.**

The PediBIRN network's a priori definitional criteria for pediatric abusive head trauma.

<b>Patients meeting any one or more of these criteria were sorted as AHT.</b>
<ul style="list-style-type: none"> <li>• Primary caregiver<sup>a</sup> admission of abusive acts</li> <li>• Abusive acts by the primary caregiver<sup>a</sup> that were witnessed by an unbiased, independent observer</li> <li>• Specific primary caregiver<sup>a</sup> denial of any head trauma, even though the pre-ambulatory child in his or her care became acutely, clearly and persistently ill with clinical signs subsequently linked to traumatic cranial injuries visible on CT or MR imaging</li> <li>• Primary caregiver<sup>a</sup> account of the child's head injury event that was clearly historically inconsistent with repetition over time</li> <li>• Primary caregiver<sup>a</sup> account of the child's head injury event that was clearly developmentally inconsistent with child's known (or expected) gross motor skills</li> <li>• Two or more categories of extra-cranial injuries considered moderately or highly suspicious for abuse<sup>b</sup></li> </ul>

Abbreviations: AHT=abusive head trauma, CT=computed tomography, MR=magnetic resonance, PediBIRN=pediatric brain injury research network

<sup>a</sup>Defined as the person responsible for the child when he or she was acutely head injured and/or first became clearly and persistently ill with clinical signs subsequently linked to traumatic cranial injuries visible on CT or MR imaging.

<sup>b</sup>Including classic metaphyseal lesion fracture(s) or epiphyseal separation(s); rib fracture(s); fracture(s) of the scapula or sternum; fracture(s) of digits; vertebral body fracture(s), dislocation(s) or fracture(s) of spinous process(es); skin bruising, abrasion(s) or laceration(s) in two or more distinct locations other than knees, shins or elbows; patterned skin bruising or dry contact burn(s); scalding burn(s) with uniform depth, clear lines of demarcation and paucity of splash marks; confirmed intra-abdominal injuries; retinoschisis confirmed by an ophthalmologist; retinal hemorrhages described by an ophthalmologist as dense, extensive, covering a large surface area and/or extending to the ora serrata.

**Table 3.**

Patient demographics in comparative patient cohorts.

	Patients Sorted as Abusive Head Trauma	Patients Sorted as Other Head Trauma	<i>p</i> <.05?
	(n=428)	(n=545)	
<b>Sex</b>			
Male, n (%)	269 (63)	328 (60)	
<b>Age (months)</b>			
Mean (Median, SD, Range)	6.9 (4.0, 7.8, 0-35)	10.5 (7.0, 9.4, 0-36)	
<b>Race</b>			
White or White Hispanic, n (%)	280 (65)	411 (75)	
Black, AA, or Black Hispanic, n (%)	92 (21)	71 (13)	
Other, n (%)	56 (13)	63 (12)	
<b>Ethnicity</b>			
Hispanic or Latino, n (%)	96 (22)	149 (27)	
Not Hispanic or Latino, n (%)	290 (68)	370 (68)	
Unknown or Other, n (%)	42 (10)	26 (5)	
	Patients Diagnosed with Abusive Head Trauma	Patients Diagnosed with Other Head Trauma	<i>p</i> <.05?
	(n=496)	(n=477)	
<b>Sex</b>			
Male, n (%)	321 (65)	276 (59)	
<b>Age (months)</b>			
Mean (Median, SD, Range)	7.6 (4.0, 8.1, 0-35)	10.3 (7.0, 9.5, 0-36)	
<b>Race</b>			
White or White Hispanic, n (%)	327 (66)	364 (76)	
Black, AA, or Black Hispanic, n (%)	94 (19)	69 (14)	
Other, n (%)	75 (15)	44 (9)	
<b>Ethnicity</b>			
Hispanic or Latino, n (%)	113 (23)	132 (28)	
Not Hispanic or Latino, n (%)	337 (68)	323 (68)	
Unknown or Other, n (%)	46 (9)	22 (5)	

Abbreviations: AA=African American, SD=standard deviation

**Table 4.**

2x2 contingency tables comparing the test characteristics of the PediBIRN network’s 3- and 4-variable clinical decision rules.

<u>Applying the PediBIRN-4 and AHT Definitional Criteria</u>			<u>Applying the PediBIRN-3 and AHT Definitional Criteria</u>			<u>p values</u>
	<u>AHT</u>	<u>Other</u>		<u>AHT</u>	<u>Other</u>	
<i>Higher Risk</i>	409	343	<i>Higher Risk</i>	399	242	
<i>Lower risk</i>	19	202	<i>Lower risk</i>	29	303	
	<i>Value</i>	<i>(95% CI)</i>		<i>Value</i>	<i>(95% CI)</i>	
Sensitivity	.96	(.94-.98)	Sensitivity	.93	(.91-.96)	.002
Specificity	.37	(.33-.41)	Specificity	.56	(.51-.60)	<.001
PPV	.54	(.51-.58)	PPV	.62	(.58-.66)	<.001
NPV	.91	(.88-.95)	NPV	.91	(.88-.94)	.901
(+) LR	1.52	(1.42-1.62)	(+) LR	2.10	(1.90-2.31)	<.001
(-) LR	.12	(.08-.19)	(-) LR	.12	(.09-.17)	.902
ROC AUC	.77	(.74-.80)	ROC AUC	.83	(.80-.85)	.002

<u>Applying the PediBIRN-4 and Physicians’ Final Diagnoses</u>			<u>Applying the PediBIRN-3 and Physicians’ Final Diagnoses</u>			<u>p values</u>
	<u>AHT</u>	<u>Other</u>		<u>AHT</u>	<u>Other</u>	
<i>Higher Risk</i>	484	268	<i>Higher Risk</i>	472	169	
<i>Lower risk</i>	12	209	<i>Lower risk</i>	24	308	
	<i>Value</i>	<i>(95% CI)</i>		<i>Value</i>	<i>(95% CI)</i>	
Sensitivity	.98	(.96-.99)	Sensitivity	.95	(.93-.97)	.001
Specificity	.44	(.39-.48)	Specificity	.65	(.60-.69)	<.001
PPV	.64	(.61-.68)	PPV	.74	(.70-.77)	<.001
NPV	.95	(.92-.98)	NPV	.93	(.90-.96)	.107
(+) LR	1.74	(1.60-1.88)	(+) LR	2.69	(2.38-3.04)	<.001
(-) LR	.06	(.03-.10)	(-) LR	.07	(.05-.11)	.142
ROC AUC	.82	(.80-.85)	ROC AUC	.89	(.87-.91)	<.001

Abbreviations: AHT=abusive head trauma, CI=confidence interval, PediBIRN=pediatric brain injury research network; LR=likelihood ratio, NPV=negative predictive value, PPV=positive predictive value, ROC AUC=area under the receiver operating characteristics curve

**Table 5.**

The PediBIRN 4- and 3-variable clinical decision rules' patient-specific estimates of abusive head trauma probability, and the overall positive yield of patients' completed abuse evaluations, in subpopulations distinguished by unique combinations of each decision rule's predictor variables.

The specific combination of predictor variables present on admission		F	S	B	N	n	Estimate of AHT probability, applying definitional criteria <sup>a</sup>	Patients diagnosed with AHT	Estimate of AHT probability, applying physicians' diagnoses	Patients evaluated with skeletal survey and/or retinal exam	Patients with positive skeletal survey and/or retinal exam	Overall positive yield of completed abuse evaluations <sup>a</sup>
R	B	S	B	N	n	p (95% CI)	n	p (95% CI)	n (%)	n	p (95% CI)	
+	+	+	+	25	21	.84 (.64-.95)	24	.96 (.80-1.00)	25 (100%)	20	.80 (.59-.93)	
+	+	○	+	79	71	.90 (.81-.96)	78	.99 (.93-1.00)	77 (97%)	64	.83 (.73-.91)	
+	+	○	○	3	3	1.00 (.29-1.00) <sup>b</sup>	3	1.00 (.29-1.00) <sup>b</sup>	3 (100%)	2	.67 (.09-.99)	
+	○	+	+	43	23	.53 (.38-.69)	30	.70 (.54-.83)	39 (91%)	26	.67 (.50-.81)	
○	+	+	+	8	6	.75 (.35-.97)	8	1.00 (.63-1.00)	8 (100%)	4	.50 (.16-.84)	
+	+	○	○	17	14	.82 (.57-.96)	14	.82 (.57-.96)	15 (88%)	13	.87 (.60-.98)	
+	○	○	+	39	3	<b>.08 (.02-.21)</b>	<b>3</b>	<b>.08 (.02-.21)</b>	<b>20 (51%)</b>	<b>1</b>	<b>.05 (.001-.25)<sup>c</sup></b>	
○	○	+	+	34	10	.29 (.15-.47)	12	.35 (.20-.54)	27 (79%)	7	.26 (.11-.46)	
+	○	+	○	145	109	.75 (.67-.82)	131	.90 (.84-.95)	142 (98%)	115	.81 (.74-.87)	
○	+	+	○	37	27	.73 (.56-.86)	35	.95 (.82-.99)	37 (100%)	24	.65 (.47-.80)	
○	+	○	+	9	5	.56 (.21-.86)	6	.67 (.30-.93)	9 (100%)	3	.33 (.07-.70)	
+	○	○	○	51	15	.29 (.17-.44)	20	.39 (.26-.54)	44 (86%)	16	.36 (.22-.52)	
○	+	○	○	22	13	.59 (.36-.79)	16	.73 (.50-.89)	21 (95%)	8	.38 (.18-.62)	
○	○	+	○	129	79	.61 (.52-.70)	92	.71 (.63-.79)	124 (96%)	65	.52 (.43-.61)	
○	○	○	+	111	10	<b>.09 (.04-.16)</b>	<b>12</b>	<b>.11 (.06-.18)</b>	<b>69 (62%)</b>	<b>4</b>	<b>.06 (.02-.14)<sup>c</sup></b>	
○	○	○	○	221	19	.09 (.05-.13)	12	.05 (.03-.09)	131 (59%)	5	.04 (.01-.09)	
				<b>973</b>	<b>428</b>		<b>496</b>		<b>791 (81%)</b>	<b>377</b>		
The specific combination of predictor variables present on admission		F	S	B	N	n	Estimate of AHT probability, applying definitional criteria <sup>a</sup>	Patients diagnosed with AHT	Estimate of AHT probability, applying physicians' diagnoses	Patients evaluated with skeletal survey and/or retinal exam	Patients with positive skeletal survey and/or retinal exam	Overall positive yield of completed abuse evaluations <sup>a</sup>
R	B	S	B	N	n	p (95% CI)	n	p (95% CI)	n (%)	n	p (95% CI)	



The specific combination of predictor variables present on admission		F	N	n	Estimate of AHT probability, applying definitional criteria <sup>a</sup>	Patients diagnosed with AHT	Estimate of AHT probability, applying physicians' diagnoses	Patients evaluated with skeletal survey and/or retinal exam	Patients with positive skeletal survey and/or retinal exam	Overall positive yield of completed abuse evaluations <sup>a</sup>
R	B	S			p (95% CI)	n	p (95% CI)	n (%)	n	p (95% CI)
+	+	+	92	.88 (.81-.94)	102	.98 (.93-1.00)	102 (98%)	84	.82 (.74-.89)	
+	+	○	17	.85 (.62-.97)	17	.85 (.62-.97)	18 (90%)	15	.83 (.59-.96)	
+	○	+	132	.70 (.63-.77)	161	.86 (.80-.90)	181 (96%)	141	.78 (.71-.84)	
○	+	+	33	.73 (.58-.85)	43	.96 (.85-.99)	45 (100%)	28	.62 (.47-.76)	
+	○	○	18	.20 (.12-.30)	23	.26 (.17-.36)	64 (71%)	17	.27 (.16-.39)	
○	+	○	18	.58 (.39-.75)	22	.71 (.52-.86)	30 (97%)	11	.37 (.20-.56)	
○	○	+	89	.55 (.47-.62)	104	.64 (.56-.71)	151 (93%)	72	.48 (.40-.56)	
○	○	○	29	.09 (.06-.12)	24	.07 (.04-.11)	200 (60%)	9	.05 (.02-.08)	
			<b>973</b>	<b>428</b>	<b>496</b>		<b>791 (81%)</b>	<b>377</b>		

Abbreviations: AHT=abusive head trauma, B=Bruising of the torso, ear(s), or neck, CI=confidence interval, F=complex skull Fracture(s), P=probability, PediBIRN=pediatric brain injury research network, R=acute Respiratory compromise, S=Subdural hemorrhage(s) or collection(s) that are bilateral or interhemispheric

<sup>a</sup>The data in the columns labelled "Estimate of AHT probability, applying definitional criteria" and "Overall positive yield of completed abuse evaluations" are the data used to create the plots in Figure 1.

<sup>b</sup>One-sided 97.5% confidence interval.

<sup>c</sup>The two rows of data that appear in large bold font present results in the two (PediBIRN-4) "higher risk" subpopulations that were least often evaluated for abuse, and that had the lowest overall positive yields on completed abuse evaluations.