Interrater reliability between surgeons and pediatric emergency providers in the cervical spine assessment of injured children

Fahd A. Ahmad, MD, MSCI, Lorin R. Browne, DO, Nicolaus W. Glomb, MD, MPH, Monica Harding, MS, Lawrence J. Cook, PhD, Rebecca K. Burger, MD, Pradip P. Chaudhari, MD, MS, Alexander J. Rogers, MD, Caleb E. Ward, MB, BChir, MPH, Daniel Rubalcava, MD, MSPH, Kenneth Yen, MD, MS, Nathan Kuppermann, MD, MPH, *and* **Julie C. Leonard, MD, MPH**, *St. Louis, Missouri*

BACKGROUND: Cervical spine injury (CSI) is uncommon in children but an important consideration during trauma evaluation. The Pediatric Emergency Care Applied Research Network (PECARN) derived and validated a CSI prediction rule to guide cervical spine imaging decisions in children after blunt trauma. Our objective was to determine the interrater reliability between EM providers and surgeons for history and physical examination findings used to evaluate children for CSI after blunt trauma.

METHODS: This was a planned secondary analysis of a prospective, observational multicenter study that enrolled children aged 0 year to 17 years evaluated for blunt trauma in 18 PECARN emergency departments (EDs). We collected data on injury mechanisms, history and physical examination findings, imaging ordered, and suspicion of CSI from EM and surgery providers. Kappa, prevalence, and bias-adjusted kappa (PABAK) were used to compare interrater reliability of variables associated with CSI.

RESULTS: Surgeons cared for 8,041 of the 22,430 children enrolled in the parent study. About 18.6% (1494/8041) had data collection forms completed by both EM providers and surgeons and were included in the analysis. Agreement between EM and surgery providers per kappa was moderate (kappa 0.41–0.6) to substantial (kappa 0.61–0.8), while PABAK analyses showed substantial to almost perfect agreement for variables in the PECARN CSI prediction rule. There was agreement between EM and surgery providers in overall clinical suspicion for CSI in 64.2% (959/1494) of patients. Retrospective application of the PECARN Rule indicated that ED and surgical provider assessments would have led to the same imaging decision in 73.7% (1101/1494) of patients.

CONCLUSION: We identified moderate to substantial agreement between EM providers and surgeons for clinical findings that comprise the PECARN Cervical Spine Injury Prediction Rule. Agreement between providers during shared decision-making will strengthen the use of the prediction rule and may lead to decreased cervical spine imaging in EDs. (*J Trauma Acute Care Surg.* 2025;00: 00–00. Copyright © 2025 Wolters Kluwer Health, Inc. All rights reserved.)

LEVEL OF EVIDENCE: Prognostic and Epidemiologic; Level II.

KEY WORDS: Cervical spine injury; pediatric; surgeon; emergency medicine; interrater reliability.

Cervical spine injury (CSI) is uncommon in children, but remains an important cause of morbidity and mortality after blunt trauma.[1] For all children who experience blunt trauma, an evaluation to determine the risk of CSI is essential. When CSI cannot be excluded, implementation of spinal motion restriction (SMR), the process of placing trauma patients in a rigid cervical collar and/or on a rigid board, is used in the prehospital setting and screening for CSI with imaging (plain radiography, computed tomography [CT], and/or magnetic resonance imaging [MRI]) is used in the emergency department (ED).[2,3]

In the ED, decision-making around imaging is significantly influenced by team dynamics, with a recent study emphasizing the importance of effective teamwork and shared decision-making among emergency care providers.[4] This collaborative approach is crucial, especially in high-stakes environments where timely and accurate decision-making can impact

Submitted: December 17, 2024, Revised: April 11, 2025, Accepted: May 5, 2025, Published online: June 19, 2025.

From the Washington University School of Medicine in St. Louis & St. Louis Children's Hospital (F.A.A.), St. Louis, Missouri; Departments of Pediatric and Emergency Medicine (L.R.B.), Medical College of Wisconsin, Milwaukee, Wisconsin; Department of Emergency Medicine and Pediatrics (N.W.G.), University of California, San Francisco, San Francisco, California; Division of Pediatric Emergency Medicine, Department of Emergency Medicine (M.H., L.J.C.), University of Utah School of Medicine, Salt Lake City, Utah; Division of Emergency Medicine, Department of Pediatrics (R.K.B.), Emory University School of Medicine, Children's Healthcare of Atlanta, Atlanta, Georgia; Division of Emergency and Transport Medicine, Children's Hospital Los Angeles, Department of Pediatrics (P.P.C.), Keck School of Medicine of the University of Southern California, Los Angeles, California; Departments of Emergency Medicine and Pediatrics (A.J.R.), C.S. Mott Children's Hospital, University of Michigan Health, Ann Arbor, Michigan; Children's National Hospital & The George Washington University School of Medicine and Health Sciences (C.E.W.), Washington, District Columbia; Division of Pediatric Emergency Medicine, Department of Pediatrics (D.R.), Baylor College of Medicine, Texas Children's Hospital, Houston; Division of Pediatric Emergency Medicine, Department of Pediatrics (K.Y.), UT Southwestern Medical Center, Children's Health, Dallas, Texas; Departments of Pediatrics and Emergency Medicine, GW School of Medicine and Health Sciences and Children's National Hospital (N.K.), Washington D.C.; and Division of Emergency Medicine, Department of Pediatrics (J.C.L.), The Ohio State College of Medicine and Nationwide Children's Hospital, Columbus, Ohio.

This study was presented at the 94th annual meeting of the Pediatric Academic Societies (PAS), April 29, 2023, in Washington, DC.

Supplemental digital content is available for this article. Direct URL citations appear in the printed text, and links to the digital files are provided in the HTML text of this article on the journal's Web site (www.jtrauma.com).

Address for correspondence: Fahd A. Ahmad, MD, MSCI, 660 S. Euclid Ave, Campus Box 8116, St. Louis, MO 63110; email: Fahd.Ahmad@wustl.edu.

DOI: 10.1097/TA.0000000000004695

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

patient outcomes.[5,6] The assessment of historical, mechanistic, and physical examination factors is essential for identifying potential CSI and determining the need for cervical spine imaging.

Despite established guidelines, there is significant variability in the use of cervical spine imaging for pediatric patients across different clinical settings. This variability can lead to overuse of imaging modalities such as CT scans, which expose children to ionizing radiation and increase their lifetime risk for malignancy.[7–12] The ionizing radiation risk is significantly greater with CTs as compared with x-ray. Given the frequent use of CTs in pediatric patients,[13,14] identifying ways to reduce their use without missing injuries is of great importance.[15–18]

Previous clinical prediction rules used to guide cervical spine imaging have had limited pediatric representation, leading to inconsistent application in clinical practice.[19,20] The Pediatric Emergency Care Applied Research Network (PECARN) conducted a multicenter prospective study to derive and validate a clinical prediction rule aimed at identifying children at low risk for CSI after sustaining blunt trauma.[21] Applying this prediction rule to the study's cohort demonstrated that 50% of CTs could have been eliminated from ED care without missing injuries in children being evaluated for CSI. Implementing the PECARN CSI prediction rule could streamline trauma care, reduce exposure to potentially harmful ionizing radiation, and lower costs by decreasing unnecessary cervical spine imaging.

Effective implementation of the PECARN rule requires acceptance and uptake among all providers involved in trauma care. Applying the protocolized and collaborative trauma evaluation is strongly endorsed by Advanced Trauma Life Support (ATLS).[22] When following this approach emergency medicine (EM) physicians, surgeons, nurses, and other clinicians often rely on a single history or physical examination to make diagnostic and treatment decisions. Thus, applying this prediction rule requires that providers of differing training and backgrounds agree on findings reported by team members. Additionally, the presence of SMR independently influences clinicians' decision-making and leads to greater use of radiographic imaging in injured children.[23]

Given the diverse training and backgrounds of EM providers and surgeons, and the myriad of factors influencing their decision-making, it is crucial to determine the level of agreement in their assessments of historical and physical examination findings. High interrater reliability (IRR) would support the generalizability and robustness of the PECARN rule across different clinical settings.

This was a planned secondary analysis of the primary prospective, multicenter study. Our main objective was to determine the IRR of CSI-related history and physical examination findings between EM and surgery providers managing children exposed to blunt trauma. Our secondary objective was to compare the degree of clinical suspicion for CSI estimated for each patient between EM and surgical providers.

METHODS

Study Design

The primary study enrolled children aged <18 years who experienced blunt trauma and were cared for in one of 18 PECARN EDs. The study was approved for all participating sites through a central IRB. The methods are described in the dissemination of the primary study findings.[21] Our study adhered to STROBE reporting guidelines (see checklist in Supplementary Digital Content 1, http://links.lww.com/TA/E574).

Participating Sites

Participating hospitals had pediatric EDs within freestanding children's hospitals or dedicated pediatric EDs within general EDs. Patients were evaluated by pediatric EM physicians (attending or fellow-in-training), general EM physicians, general pediatricians, physician assistants, or pediatric nurse practitioners. All hospitals had either American College of Surgeons Level I Trauma accreditation or their state's equivalent and were staffed by pediatric trauma surgeons and pediatric spine surgeons. The latter maintain either primary surgical accreditation in pediatric neurosurgery or pediatric orthopedic surgery, depending on the site.

Inclusion/Exclusion

Inclusion criteria for this study were 1) patient age younger than 18 years, 2) evaluation in the ED for possible blunt trauma, 3) transport from the scene by EMS, trauma team evaluation, and/or cervical spine imaging during emergency care, and 4) paired study observations by both an EM and surgery provider. We excluded children with only penetrating trauma. This mirrors the criteria from the parent study, with study observations from a convenience sample of surgical providers being the only additional inclusion criteria required for this secondary analysis.

Missed eligible patients were those who had trauma team involvement during ED care but were lacking a completed surgical provider data collection form. Ineligible patients for this cohort were all those who had no trauma team involvement during ED care.

Enrollment occurred during times when research coordinators (RCs) were available, generally between 8:00 AM and midnight, with most sites enrolling 7 days a week.

Data Collection

For enrolled patients, RCs asked treating EM providers and surgeons involved in patient care to complete a tablet-based, branching-logic data collection form stored in a Research Electronic Data Capture® (REDCap) database that collected information regarding mechanism of injury, past medical history, chief complaints, and physical examination findings, imaging ordered (if any), and suspicion of cervical spine injury (CSI). The branching-logic data collection methods and list of variables have been previously published.[21,24]

The study data collection forms were identical for EM providers and surgeons. We did not capture the exact timing of the completion of whether the patient had a shared history and/or physical examination as part of a trauma activation. We did ask EM and surgery providers to complete the forms independently and asynchronously as soon as possible after completing their patient evaluation. Emergency medicine data collection forms were completed by the treating ED clinician (pediatric emergency medicine attending or fellow, general pediatrician, or advanced practice provider) while surgical forms were completed by pediatric surgery attendings or fellows. Our goal was to obtain as many IRR forms collected as possible from surgeons

© 2025 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

that were at the bedside during trauma evaluations. We did not focus on IRR form completion for a specific subset of patients.

If a patient history or physical examination finding was unknown, respondents could indicate this on the data collection form. They could also indicate if an examination finding was not assessed or unable to be assessed (e.g., self-report of pain in a pre-verbal child).

Statistical Methods and Outcomes

We report frequencies and percentages for categorical variables and median and interquartile range (IQR) for continuous variables. We used standard non-parametric methods (Fisher's exact tests with Monte Carlo approximation for categorical variables, Kruskal-Wallis tests for continuous variables) to assess for differences in participant characteristics.

Interrater reliability of key history and examination findings were assessed through prevalence (based on the EM providers and surgeons), percent agreement, and Cohen's kappa with 95% confidence limits,[25] as well as prevalence adjusted bias adjusted kappa (PABAK)[26] to account for variables with low prevalence. We interpreted Cohen's kappa and PABAK as follows: values ≤0 as indicating no agreement and 0.01–0.20 as none to slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1.00 as almost perfect agreement.[27]

We categorized clinical suspicion of cervical spine injury by provider type as <1%, 1–5%, 6–10%, 11–50%, and >50%. These categories were chosen based on their established relevance in pediatric head trauma studies, where they have been demonstrated to be meaningful in assessing patient risk and guiding clinical decision-making. Specifically, studies have shown that such categorizations help in stratifying the risk of injury in the PECARN head trauma prediction rules, which have proven effective in identifying children at low risk for clinically significant traumatic brain injuries.[28,29]

We assessed differences in clinical suspicion using Bhapkar's test for marginal homogeneity.[30] We also assessed the association of variable agreement with cervical spine injury and cervical spine imaging using exact logistic regression. We chose exact logistic regression because of the small number of outcomes. We present odds ratios with their 95% confidence intervals.

For ease of analysis, individual factors were collapsed into assessed (responses: yes, no), not assessed (responses: did not assess, unknown) or missing (unable to assess). We employed McNemar's test to determine marginal homogeneity for the matched pairs of surveys.

RESULTS

Figure 1 provides an overview of patients eligible for analysis. Surgeons were involved in the care of 8041 patients included in the main study, 18.6% (1494/8041) of whom had data collection forms completed by the surgery provider and were eligible for inclusion. The majority of surgical forms were completed by trauma surgeons (85.7, 1280/1494%) with neurosurgeons (26/1494, 1.7%), orthopedic surgeons (16/1494, 1.1%), and surgeons who self-classified as “other” (172/1494, 11.5%) completing the remainder.

Table 1 provides a comparison of the patients eligible and included for analysis versus missed eligible patients where a data collection form was not completed by the treating surgery provider. Compared with missed eligible patients, included patients were slightly younger, less likely to be White or non-Hispanic, more likely to be placed in SMR by EMS, and more likely to be discharged home from the ED. There was no difference in the rate of cervical spine injury between captured eligible patients and those missed for inclusion.

We present percent agreement and kappa results on key history and examination findings from our prediction rule in Table 2. We also include relevant variables from two previously validated adult CSI screening tools, the Canadian C-Spine[19] and NEXUS[20] studies. Kappa analysis demonstrated moderate or substantial agreement between ED and surgical provider observations of all key variables for the PECARN Rule and, when assessed using PABAK, found almost perfect agreement for all investigated variables.

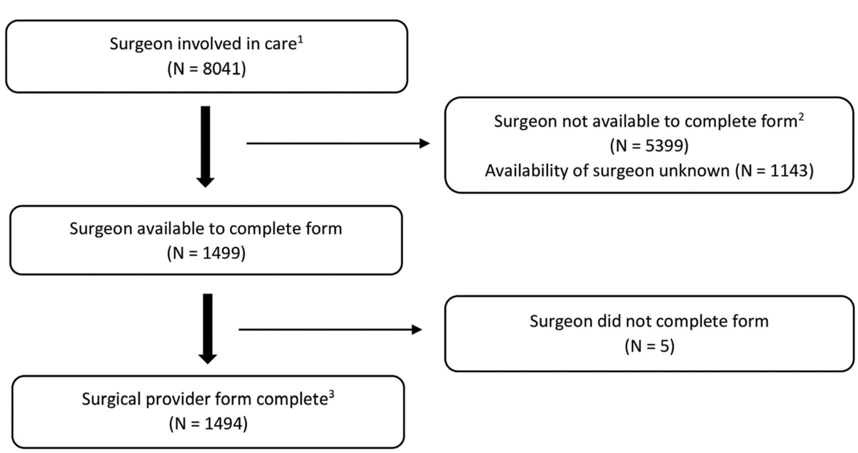

Figure 1. CONSORT diagram. Participants were eligible if a surgeon was involved in care and included in this analysis if the surgical provider form was complete. [1]Eligible for inclusion if participant had a completed ED provider form and surgeon involved in their care. [2]Missed eligible participants. [3]Eligible participants—indicated by involvement of a surgeon, the availability of a surgeon, and a completed surgical provider form.

© 2025 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 1. Demographics and Outcomes of Cohort From Parent Study Who Had Surgeon Involvement With Comparison of Surgeon Form Completion vs. Not Completed

	Cohort		
	Eligible (n = 1494)	Missed Eligible (n = 5,399)	p
Demographics			
Age (y)	7.0 [2.0–12.0]	8.0 [3.0–13.0]	<0.001*
Male	924 (61.8%)	3308 (61.3%)	0.514**
Race			<0.001**
American Indian or Alaskan Native	11 (0.7%)	32 (0.6%)	
Asian	50 (3.3%)	139 (2.6%)	
Black or African American	270 (18.1%)	1107 (20.5%)	
Native Hawaiian or Other Pacific Islander	9 (0.6%)	23 (0.4%)	
White	667 (44.6%)	2877 (53.3%)	
More than one race	10 (0.7%)	68 (1.3%)	
Other	380 (25.4%)	805 (14.9%)	
Unknown	97 (6.5%)	348 (6.4%)	
Ethnicity			<0.001**
Hispanic or Latino	343 (23.0%)	1095 (20.3%)	
Non-Hispanic and Non-Latino	986 (66.0%)	3876 (71.8%)	
Unknown	165 (11.0%)	428 (7.9%)	
Cervical spine imaging obtained	1024 (68.5%)	3934 (72.9%)	0.356**
EMS spinal precautions†			
Cervical collar	309 (20.7%)	824 (15.3%)	0.002**
Rigid long board	178 (11.9%)	419 (7.8%)	0.003**
Vacuum mattress	13 (0.9%)	19 (0.4%)	0.003**
ED provider			
ED Disposition			<0.001**
Home	601 (40.2%)	1974 (36.6%)	
Admit intensive care unit	273 (18.3%)	830 (15.4%)	
Admit floor	499 (33.4%)	2,222 (41.2%)	
Operating room	110 (7.4%)	314 (5.8%)	
Death in the ED	2 (0.1%)	21 (0.4%)	
Other	9 (0.6%)	38 (0.7%)	
Survey respondent			<0.001**
Attending, PEM	944 (63.2%)	3103 (57.5%)	
Attending, EM	126 (8.4%)	476 (8.8%)	
Attending, Pediatrics	11 (0.7%)	158 (2.9%)	
Fellow, PEM	389 (26.0%)	1520 (28.2%)	
APN or PA	20 (1.3%)	137 (2.5%)	
Other	4 (0.3%)	5 (0.1%)	
Outcomes			
Cervical spine injury	67 (4.5%)	230 (4.3%)	0.719**
Congenital anomaly	16 (1.1%)	71 (1.3%)	0.514**
GCS: 3–13, 14, 15			0.181**
3–13	180 (12.0%)	574 (10.6%)	
14	88 (5.9%)	287 (5.3%)	
15	1226 (82.1%)	4538 (84.1%)	

Eligible patients include study participants who had history and examination forms completed by ED providers as well as surgical providers. Missed eligible include study participants who had both ED provider and surgeon involvement but no form completed by surgeon. Frequencies and percentages were reported for categorical variables. Median [Q1, Q3] were reported for continuous variables.

*Kruskal-Wallis test.

**Fisher's exact test with Monte Carlo approximation.

†Presence of cervical collar, rigid long board, or vacuum mattress was determined on the EMS provider form.

Table 3 presents the comparison of clinical suspicion for CSI as assessed by EM and surgery providers. Shaded areas represent when the providers had exact agreement. Estimates for clinical suspicion of CSI were similar between groups. While the percent agreement between providers was substantially higher for those estimated to have <1% chance of CSI compared with other groups, overall, we did not observe a statistical difference in suspicion of injury between the two groups ($p = 0.87$).

We provide the association of EM and surgery provider agreement on key history and examination variables with the use of cervical spine imaging in Table 4. For most variables, we observed that the odds of cervical spine imaging being obtained was greater when both EM providers and trauma surgeons assessed a variable as present compared with both providers assessing a factor as not present. Additionally, when there was disagreement whether a history or physical examination variable was present, we did see a difference in the odds of obtaining imaging for most categories compared with both providers assessing a factor as not present, regardless of whether the surgeon or EM provider assessed the finding as present.

Table 5 shows the imaging recommendations resulting from the PECARN Rule based on EM versus surgeon assessments. There was a small but statistically significant difference in imaging recommendations, with more assessments resulting in plain radiographs among the emergency provider group and more resulting in clinically clearance among the surgery provider group. The shaded numbers on the table represent when provider assessments would have led to the same imaging recommendation, which would have occurred in 73.7% (1101/1494) of patients.

Supplemental Digital Content 2, http://links.lww.com/TA/E575, is a comparison of agreement statistics categorized by whether patients ultimately were diagnosed with a CSI. Overall, we found moderate to substantial agreement using kappa, and substantial to near perfect agreement using PABAK, indicating that EM and surgery providers obtained similar assessments for patients both with and without CSIs. Supplemental Digital Content 3, http://links.lww.com/TA/E576, provides information about when individual risk factors were not assessed by the EM or surgery provider. This excludes patients where either provider indicated they could not assess a patient (i.e., self-reported neck pain in a non-verbal child). Some important risk factors, such as self-reported neck pain and neck tenderness on examination, were not assessed in over 20% of encounters.

IRR completion was variable across sites, with the majority of completed surgeon forms coming from a small subset of sites. Specifically, 70.6% (1055/1494) of completed forms were from the top five sites. Despite this variability, we did not observe major differences in IRR across sites.

DISCUSSION

In this study, we demonstrated moderate to substantial agreement for history and examination findings for CSI risk factors between EM and surgery providers when they care for children who sustain blunt trauma. While evaluations of trauma patients are often protocolized based on ATLS guidelines,[22] in reality patient assessment is fluid and complex. Interactions between providers of different specialties are often collaborative,

© 2025 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 2. Agreement on Key Patient History and Physical Examination Findings

	Prevalence Per EM Provider	Prevalence Per Surgeon	Percent Agreement	Simple Kappa Statistic (95% Confidence Interval)	PABAK Statistic (95% Confidence Interval)
History					
Motor vehicle crash	347 (23.2%)	353 (23.6%)	98.80	0.97 (0.95–0.98)	0.98 (0.96–0.99)
MVC and any one of: intrusion, ejection, death, telemetry	134 (9.0%)	131 (8.8%)	92.70	0.55 (0.47–0.62)	0.85 (0.83–0.88)
Motorcycle ATV motorized scooter crash	114 (7.6%)	123 (8.2%)	98.33	0.89 (0.84–0.93)	0.97 (0.95–0.98)
Hit by car or other motor vehicle (pedestrian, cyclist, other)	145 (9.7%)	140 (9.4%)	98.19	0.90 (0.86–0.93)	0.96 (0.95–0.98)
Fall of height greater than 10 feet	108 (7.2%)	105 (7%)	97.12	0.78 (0.72–0.85)	0.94 (0.93–0.96)
Diving	1 (0.1%)	1 (0.1%)	100.00	1.00 (1.00–1.00)	1.00 (1.00–1.00)
History of loss of consciousness (LOC)	309 (20.7%)	317 (21.2%)	76.64	0.57 (0.53–0.61)	0.65 (0.62–0.68)
Self-reported neck pain*	217 (14.5%)	221 (14.8%)	88.99	0.66 (0.60–0.72)	0.78 (0.74–0.82)
Examination findings					
Abnormal airway breathing or circulation findings*	111 (7.4%)	122 (8.2%)	94.50	0.65 (0.57–0.72)	0.89 (0.87–0.91)
Suspicion for alcohol or drug intoxication	18 (1.2%)	12 (0.8%)	95.98	0.62 (0.39–0.85)	0.92 (0.87–0.97)
Altered mental status*†	291 (19.5%)	278 (18.6%)	91.90	0.74 (0.69–0.78)	0.84 (0.81–0.87)
Signs of substantial‡ head injury*	311 (20.8%)	287 (19.2%)	86.05	0.57 (0.51–0.62)	0.72 (0.69–0.76)
Neck pain upon examination*	182 (12.2%)	180 (12.0%)	91.90	0.70 (0.64–0.76)	0.84 (0.81–0.87)
Substantial torso injury*	150 (10.0%)	119 (8.0%)	92.12	0.52 (0.45–0.60)	0.84 (0.81–0.87)
Thoracic spine tenderness	27 (1.8%)	21 (1.4%)	94.34	0.84 (0.67–1.00)	0.89 (0.76–1.00)
Focal neurologic deficits*	56 (3.7%)	66 (4.4%)	96.51	0.54 (0.43–0.66)	0.93 (0.91–0.95)

Agreement on key patient history and physical examination findings using percent agreement, kappa statistics, and PABAK statistics. Findings were collapsed into three categories: yes, no, and missing. Did not assess, unable to assess, and not applicable findings were all marked as missing.

*Factors from the PECARN cervical spine injury (CSI) prediction rule.

**Prevalence and bias adjusted kappa.

†GCS < 15, AVPU < A, evidence of intoxication, and/or other signs of altered mental status (agitation, somnolence, repetitive questions, slow verbal response).

‡Substantial injuries were defined a priori as those warranting inpatient observation or surgical intervention.

and the potential for differing assessments and care plans exists and occurs naturally. Although many EM physicians and surgeons have some standardized, overlapping training for evaluation of trauma in the ED through ATLS, this does not necessarily lead to a harmonious team model or collaborative decision-making. Our data demonstrate that there is significant overlap in their clinical assessments and potential application of the prediction rule which can support improved team dynamics going forward.[4]

Our findings also highlight the broader applicability of the PECARN CSI prediction rule. To maximize generalizability, variables in a prediction rule should demonstrate adequate interrater agreement to facilitate implementation across different personnel types and clinical settings. The substantial agreement observed between EM and surgical providers suggests that the PECARN rule can be reliably applied by both groups, reinforcing its feasibility for widespread adoption. This consistency is particularly important in trauma care, where multidisciplinary teams rely on shared history and physical examination findings to guide decision-making. Future implementation efforts should focus on integrating the rule into routine clinical workflows through structured educational initiatives and decision-support tools. Prior research has highlighted that decision-making around pediatric cervical spine imaging is inherently collaborative, with shared decision-making and interprofessional interactions playing a crucial role in the imaging process.[4] In addition, prospective implementation studies should evaluate the impact of rule adoption on imaging utilization, patient outcomes, and provider adherence across different trauma centers, including those with varying levels of pediatric expertise. Given the

TABLE 3. Comparison of Suspicion of Cervical Spine Injury Between ED Provider and Surgeon

		Surgical Provider					
		<1% (n = 963)	1–5% (n = 359)	6–10% (n = 97)	<11–50% (n = 45)	>50% (n = 30)	*p*
ED Provider	Clinical suspicion for the presence of CSI						0.87*
	<1% (n = 953)	767 (51.3%)	143 (9.6%)	33 (2.2%)	8 (0.5%)	2 (0.1%)	
	1–5% (n = 375)	156 (10.4%)	155 (10.4%)	40 (2.7%)	16 (1.1%)	8 (0.5%)	
	6–10% (n = 97)	31 (2.1%)	39 (2.6%)	17 (1.1%)	6 (0.4%)	4 (0.3%)	
	11–50% (n = 44)	8 (0.5%)	17 (1.1%)	5 (0.3%)	9 (0.6%)	5 (0.3%)	
	>50% (n = 25)	1 (0.1%)	5 (0.3%)	2 (0.1%)	6 (0.4%)	11 (0.7%)	

*Bhapkar's test for marginal homogeneity.

Agreement of clinical suspicion of cervical spine injury between ED provider and surgical provider. Counts and percentages of agreements represented in the highlighted cells. The simple Kappa statistic for clinical suspicion for the presence of CSI was 0.32 (0.28, 0.35).

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

TABLE 4. Association of Agreement Between ED and Surgical Providers and Outcomes

			Cervical Spine Injury		Cervical Spine Imaging Obtained	
		Overall (N = 1494)	Odds Ratio (95% confidence interval)	*p*	Odds Ratio (95% confidence interval)	*p*
Self-reported neck pain	Agree and not present	1207 (80.8%)	Reference	.	Reference	.
	Agree and present	151 (10.1%)	3.27 (1.71–6.03)	<0.001	29.71 (9.87–146.56)	<0.001
	ED not present, surgeon present	70 (4.7%)	1.16 (0.22–3.75)	0.993	13.45 (4.37–67.25)	<0.001
	ED present, surgeon not present	66 (4.4%)	0.81 (0.09–3.21)	1.000	3.37 (1.68–7.49)	<0.001
Abnormal airway, breathing, or circulation findings	Agree and not present	1340 (89.7%)	Reference	.	Reference	.
	Agree and present	79 (5.3%)	9.62 (4.8–18.69)	<0.001	6.25 (2.71–17.72)	<0.001
	ED not present, surgeon present	43 (2.9%)	5.7 (1.85–14.8)	0.003	2.25 (1.01–5.66)	0.045
	ED present, surgeon not present	32 (2.1%)	9.82 (3.37–25.31)	<0.001	15.92 (2.64–650.96)	<0.001
Altered mental status	Agree and not present	1149 (76.9%)	Reference	.	Reference	.
	Agree and present	224 (15.0%)	8.31 (4.71–14.84)	<0.001	13.46 (7.07–28.79)	<.001
	ED not present, surgeon present	54 (3.6%)	1.73 (0.19–7.25)	0.688	5.03 (2.13–14.51)	<0.001
	ED present, surgeon not present	67 (4.5%)	3.62 (1.05–10.08)	0.043	3.59 (1.79–7.96)	<0.001
Signs of substantial head injury	Agree and not present	1090 (73.0%)	Reference	.	Reference	.
	Agree and present	194 (13.0%)	3.07 (1.57–5.82)	0.001	9.2 (5.16–17.84)	<0.001
	ED not present, surgeon present	93 (6.2%)	3.01 (1.16–6.92)	0.024	4.93 (2.58–10.38)	<0.001
	ED present, surgeon not present	117 (7.8%)	2.67 (1.09–5.89)	0.031	5.78 (3.13–11.69)	<0.001
Examination reported neck tenderness	Agree and not present	1259 (84.3%)	Reference	.	Reference	.
	Agree and present	127 (8.5%)	3.12 (1.54–5.98)	0.002	73.2 (12.8–2923.62)	<0.001
	ED not present, surgeon present	53 (3.5%)	1.51 (0.29–4.96)	0.680	9.68 (3.1–48.81)	<0.001
	ED present, surgeon not present	55 (3.7%)	0.95 (0.11–3.81)	1.000	10.07 (3.23–50.7)	<0.001
Substantial torso injury	Agree and not present	1301 (87.1%)	Reference	.	Reference	.
	Agree and present	76 (5.1%)	5.61 (2.56–11.51)	<0.001	1.56 (0.89–2.85)	0.131
	ED not present, surgeon present	43 (2.9%)	1.46 (0.17–5.97)	0.831	2.49 (1.08–6.67)	0.029
	ED present, surgeon not present	74 (5.0%)	5.22 (2.31–10.95)	<0.001	1.22 (0.71–2.16)	0.537
Focal neurologic deficits	Agree and not present	1401 (93.8%)	Reference	.	Reference	.
	Agree and present	29 (1.9%)	5.09 (1.46–14.31)	0.012	6.53 (1.63–56.91)	0.003
	ED not present, surgeon present	37 (2.5%)	3.82 (1.12–10.41)	0.033	2.5 (1.02–7.38)	0.045
	ED present, surgeon not present	27 (1.8%)	1.96 (0.22–8.19)	0.587	2.13 (0.78–7.24)	0.171

Estimates and *p* values are provided using exact logistic regression. In the cases where the lower confidence limit is 0 or the upper confidence limit is missing (.) this indicates a median unbiased estimate and a one-sided *p* value.

potential for this rule to reduce unnecessary imaging while maintaining patient safety, a key next step is assessing how its application influences real-world clinical practice and workflow efficiency in diverse settings. We are currently planning such a trial which would evaluate implementation strategies in academic pediatric EDs, as well as general EDs who lack pediatric specialists.

To our knowledge, this is the first study to directly examine the IRR of EM and surgery providers when assessing for CSI in pediatric blunt trauma patients. While our study was focused

TABLE 5. Comparison of Imaging Recommendation Between ED Provider and Surgical Provider Using the PECARN Cervical Spine Injury Prediction Rule

		Surgical Provider			
		Clinically Cleared (No Imaging) (n = 753)	CT (n = 215)	Plain Radiograph (n = 526)	*p*
ED Provider	Imaging recommendation per the prediction rule				0.0040*
	Clinically cleared (no imaging) (n = 702)	570 (38.2%)	30 (2%)	102 (6.8%)	
	Computed tomography (n = 205)	20 (1.3%)	146 (9.8%)	39 (2.6%)	
	Plain radiograph (n = 587)	163 (10.9%)	39 (2.6%)	385 (25.8%)	

*Bhapkar's test for marginal homogeneity.

Agreement of imaging recommendation between ED provider and surgical provider. Counts and percentages of agreements represented in the highlighted cells. Participants were grouped in mutually exclusive categories by grouping those receiving both plain radiographs and cervical spine computed tomography with those receiving cervical spine computed tomography only. If neither form of imaging was identified on medical record review, the participant was considered clinically cleared. Projected imaging categories were determined by the PECARN cervical spine injury prediction rule. The simple Kappa statistic for Imaging recommendation per the prediction rule was 0.57 (0.53–0.6).

© 2025 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

exclusively on blunt trauma patients and evaluation of CSI, there is minimal existing literature evaluating agreement in history or examination findings in pediatric trauma populations. Browne et al[31] evaluated interobserver agreement for CSI risk factors between EM and prehospital providers. They found moderate agreement for many CSI risk factors, although they also noted many factors were not assessed and both groups had poor clinical suspicion for CSI identification.

Prior PECARN studies had similarly evaluated interobserver agreement in different conditions. Yen et al[32] evaluated IRR of physical examination findings for pediatric patients with abdominal pain. Pediatric emergency medicine attending physicians and pediatric surgeons had moderate agreement for rebound tenderness, but poor agreement for other examination variables. This contrasts with our results showing high agreement across multiple domains. There are many possibilities for our differing findings. The approach to trauma patients is overall more standardized than for most other patients given the reliance on ATLS training. Additionally, trauma examinations for PEM physicians and surgeons are more likely to occur in close temporal proximity than for isolated evaluations of abdominal pain, where surgery providers are typically evaluating the patients later then PEM physicians. Our study also had a much larger sample size—with nearly 1500 patients in our sample compared with 68 patients[32]—and is more likely to reliably detect differences if they exist.

Limitations

All participating EDs were dedicated pediatric trauma centers with academic affiliations and significant prior experience in PECARN studies. While geographically diverse, these sites represent a subset of institutions with well-established trauma systems and research infrastructure and may not reflect the variability in pediatric trauma care across all settings, such as community hospitals or lower-level trauma centers. Additional research is needed to determine the prediction rule's IRR and implementation feasibility in non-tertiary care settings, where provider experience with pediatric trauma may vary. In addition, we were unable to obtain surgical provider observations on 80% of children included in the primary study. Interrater reliability completion was unevenly distributed across study sites, with the majority of completed surgeon forms originating from a subset of centers. While this may introduce site-specific biases, we did not observe major differences in agreement across sites. However, future studies should evaluate how IRR and implementation feasibility may vary in different clinical settings, particularly in centers with lower research infrastructure or different trauma care models. Finally, while there was high agreement in the history and examination findings, we did not control or record the timing of data collection form completion by the EM and surgery providers relative to obtaining the history and physician examination. In some trauma evaluations, a group examination may have been performed per ATLS protocols, which could bias the responses toward higher agreement.

Our study demonstrated a high interrater agreement in history and examination findings between EM and surgery providers for clinical criteria that comprise the PECARN CSI prediction rule. These findings provide strength to the multidisciplinary use of this tool during the trauma evaluation of injured children and evidence of the generalizability of the rule across trauma specialists. In turn this could lead to decreased radiographic screening for CSI after blunt trauma in children.

AUTHORSHIP

F.A.A., J.C.L., L.R.B., N.W.G., M.H., L.J.C. participated in the conception and Study design. F.A.A., L.R.B., N.W.G., R.K.B., P.P.C., A.J.R., C.E.W., K.Y., D.R. participated in the literature review. All authors participated in the data acquisition. M.H., L.J.C., F.A.A., J.C.L., L.F.B., N.W.G. participated in the data analysis and interpretation. All authors participated in the drafting of the article. F.A.A., L.R.B., N.W.G., J.C.L., M.H., L.J.C. participated in the critical revision.

DISCLOSURE

Funding: This study was funded by the Eunice Kennedy Shriver National Institute of Child Health and Human Development (NICHD; 5R01HD091347 Development and Testing of a Pediatric Cervical Spine Injury Risk Assessment Tool). This information or content and conclusions are those of the authors and should not be construed as the official position or policy of, nor should any endorsements be inferred by NICHD or the US Government. The study was also funded by the Health Resources and Services Administration (HRSA) of the US Department of Health and Human Services (HHS), in the Maternal and Child Health Bureau, under the Emergency Medical Services for Children program through the following cooperative agreements: Data Coordinating Center at the University of Utah (Salt Lake City, UT; UJ5MC30824), Great Lakes Emergency Medical Services for Children Research Node (GLEMSCRN) at Nationwide Children's Hospital (Columbus, OH; U03MC28844), Hospitals of the Midwest Emergency Research Node (HOMERUN) at Cincinnati Children's Hospital Medical Center (Cincinnati, OH; U03MC22684), Pediatric Emergency Medicine Northeast, West and South (PEMNEWS) at Columbia University Medical Center (New York, NY; U03MC00007), Pediatric Research in Injuries and Medical Emergencies (PRIME) at University of California at Davis Medical Center (Sacramento, CA; U03MC00001), Charlotte, Houston, Milwaukee Prehospital EMS Research Node (CHaMP) node at the State University of New York at Buffalo (Buffalo, NY; U03MC33154), West/Southwest Pediatric Emergency Medicine Research (WPEMR) at Seattle Children's Hospital (Seattle, WA,; U03MC33156), and San Francisco-Oakland, Providence, Atlanta Research Collaborative (SPARC) at Rhode Island Hospital and Hasbro Children's Hospital (Providence, RI; U03MC33155). This information or content and conclusions are those of the authors and should not be construed as the official position or policy of, nor should any endorsements be inferred by HRSA, HHS, or the US Government.
Conflict of Interest: The authors have no conflicts of interest. Author Disclosure forms have been supplied and are provided as Supplemental Digital Content (http://links.lww.com/TA/E577).

REFERENCES

1. Kokoska ER, Keller MS, Rallo MC, Weber TR. Characteristics of pediatric cervical spine injuries. *J Pediatr Surg*. 2001;36(1):100–105.
2. National Association of state EMS officials. National Model EMS Clinical Guidelines 2022 [updated March 2022. Available from: https://nasemso.org/wp-content/uploads/National-Model-EMS-Clinical-Guidelines_2022.pdf.
3. American College of Surgeons. Best Practice Guidelines — Spine Injury 2022 [updated March 2022. Available from: https://www.facs.org/media/k45gikqv/spine_injury_guidelines.pdf.
4. Gregory ME, Truelove A, Ahmad F, Corwin D, Tzimenatos L, Oglesbee SJ, et al. Decision-making for pediatric cervical spine imaging after blunt trauma: investigating team dynamics in the emergency department. *JACEP Open*. 2023;4(4):e13024.
5. Ozkaynak M, Wu DT, Hannah K, Dayan PS, Mistry RD. Examining workflow in a pediatric emergency department to develop a clinical decision support for an antimicrobial stewardship program. *Appl Clin Inform*. 2018;9(02):248–260.
6. Madani A, Gips A, Razek T, Deckelbaum DL, Mulder DS, Grushka JR. Defining and measuring decision-making for the management of trauma patients. *J Surg Educ*. 2018;75(2):358–369.
7. Adelgais KM, Grossman DC, Langer SG, Mann FA. Use of helical computed tomography for imaging the pediatric cervical spine. *Acad Emerg Med*. 2004;11(3):228–236.

© 2025 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.

8. Brenner DJ. Estimating cancer risks from pediatric CT: going from the qualitative to the quantitative. *Pediatr Radiol*. 2002;32(4):228–221 discussion 42–4.
9. Brenner DJ, Hall EJ. Computed tomography—an increasing source of radiation exposure. *N Engl J Med*. 2007;357(22):2277–2284.
10. Kleinerman RA. Cancer risks following diagnostic and therapeutic radiation exposure in children. *Pediatr Radiol*. 2006;36(Suppl 2):121–125.
11. Miglioretti DL, Johnson E, Williams A, Greenlee RT, Weinmann S, Solberg LI, et al. The use of computed tomography in pediatrics and the associated radiation exposure and estimated cancer risk. *JAMA Pediatr*. 2013;167(8): 700–707.
12. Pearce MS, Salotti JA, Little MP, McHugh K, Lee C, Kim KP, et al. Radiation exposure from CT scans in childhood and subsequent risk of leukaemia and brain tumours: a retrospective cohort study. *Lancet*. 2012;380(9840):499–505.
13. Broder J, Fordham LA, Warshauer DM. Increasing utilization of computed tomography in the pediatric emergency department, 2000-2006. *Emerg Radiol*. 2007;14(4):227–232.
14. Adelgais KM, Browne L, Holsti M, Metzger RR, Murphy SC, Dudley N. Cervical spine computed tomography utilization in pediatric trauma patients. *J Pediatr Surg*. 2014;49(2):333–337.
15. Brody AS, Frush DP, Huda W, Brent RL, American Academy of Pediatrics Section on R. Radiation risk to children from computed tomography. *Pediatrics*. 2007;120(3):677–682.
16. Connelly CR, Yonge JD, Eastes LE, Bilyeu PE, Kemp Bohan PM, Schreiber MA, et al. Performance improvement and patient safety program-guided quality improvement initiatives can significantly reduce computed tomography imaging in pediatric trauma patients. *J Trauma Acute Care Surg*. 2016;81(2):278–284.
17. Sun R, Skeete D, Wetjen K, Lilienthal M, Liao J, Madsen M, et al. A pediatric cervical spine clearance protocol to reduce radiation exposure in children. *J Surg Res*. 2013;183(1):341–346.
18. Waddell VA, Connelly S. Decreasing radiation exposure in pediatric trauma related to cervical spine clearance: a quality improvement project. *J Trauma Nurs*. 2018;25(1):38–44.
19. Hoffman JR, Mower WR, Wolfson AB, Todd KH, Zucker MI. Validity of a set of clinical criteria to rule out injury to the cervical spine in patients with blunt trauma. *N Engl J Med*. 2000;343(2):94–99.
20. Stiell IG, Wells GA, Vandemheen KL, Clement CM, Lesiuk H, De Maio VJ, et al. The Canadian C-spine rule for radiography in alert and stable trauma patients. *JAMA*. 2001;286(15):1841–1848.
21. Leonard JC, Harding M, Cook LJ, Leonard JR, Adelgais KM, Ahmad FA, et al. PECARN prediction rule for cervical spine imaging of children presenting to the emergency department with blunt trauma: a multicentre prospective observational study. *The Lancet Child & Adolescent Health*. 2024; 8:482–490.
22. Henry S, Brasel K, Stewart R. *Advanced trauma life support*. American College of Surgeons Chicago, IL; 2018.
23. Drain J, Wilson ES, Moore TA, Vallier HA. Does prehospital spinal immobilization influence in hospital decision to obtain imaging after trauma? *Injury*. 2020;51(4):935–941.
24. Ahmad FA, Schwartz H, Browne LR, Lassa-Claxton S, Wallendorf M, Brooke Lerner E, et al. Methods for collecting paired observations from emergency medical services and emergency department providers for pediatric cervical spine injury risk factors. *Acad Emerg Med*. 2017;24(4):432–441.
25. Cohen J. A coefficient of agreement for nominal scales. *Educ Psychol Meas*. 1960;20(1):37–46.
26. Byrt T, Bishop J, Carlin JB. Bias, prevalence and kappa. *J Clin Epidemiol*. 1993;46(5):423–429.
27. Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics*. 1977;33:159–174.
28. Atabaki SM, Hoyle JD Jr., Schunk JE, Monroe DJ, Alpern ER, Quayle KS, et al. Comparison of prediction rules and clinician suspicion for identifying children with clinically important brain injuries after blunt head trauma. *Acad Emerg Med*. 2016;23(5):566–575.
29. Kuppermann N, Holmes JF, Dayan PS, Hoyle JD, Atabaki SM, Holubkov R, et al. Identification of children at very low risk of clinically-important brain injuries after head trauma: a prospective cohort study. *Lancet*. 2009; 374(9696):1160–1170.
30. Bhapkar VP. A note on the equivalence of two test criteria for hypotheses in categorical data. *J Am Stat Assoc*. 1966;61(313):228–235.
31. Browne LR, Schwartz H, Ahmad FA, Wallendorf M, Kuppermann N, Lerner EB, et al. Interobserver agreement in pediatric cervical spine injury assessment between prehospital and emergency department providers. *Acad Emerg Med*. 2017;24(12):1501–1510.
32. Yen K, Karpas A, Pinkerton HJ, Gorelick MH. Interexaminer reliability in physical examination of pediatric patients with abdominal pain. *Arch Pediatr Adolesc Med*. 2005;159(4):373–376.

© 2025 Wolters Kluwer Health, Inc. All rights reserved.

Copyright © 2025 Wolters Kluwer Health, Inc. Unauthorized reproduction of this article is prohibited.