

## CLINICAL PRACTICE

## Availability of Pediatric Emergency Visit Data from Existing Data Sources

Marc H. Gorelick, MD, MSCE, Elizabeth R. Alpern, MD, MSCE, Tasmeen Singh, MPH, NREMT-P, Donald Snowdon, MD, Richard Holubkov, PhD, J. Michael Dean, MD, MBA, Nathan Kuppermann, MD, MPH, for the Pediatric Emergency Care Applied Research Network (PECARN)

### Abstract

**Objectives:** To determine the availability and completeness of selected data elements from administrative and clinical sources for emergency department (ED) visits in a national pediatric research network. **Methods:** This was a retrospective study of 25 EDs in the Pediatric Emergency Care Applied Research Network. Data were obtained from two sources at each ED: 1) extant electronic administrative data for all visits during a 12-month period in 2002 and 2) data abstracted from medical records by trained abstractors for visits during ten randomly selected days over a three-month period in 2003. Epidemiologic data were obtained for all visits and additional clinical data for patients with two target conditions: asthma and fractures. **Results:** A total of 749,036 visits were analyzed from administrative sources and 12,756 medical records abstracted. Data availability varied by element, method of capture, and site. From administrative sources, data on insurance type were the most complete (1.3% overall missing; range, 0%–18.5% for

individual sites), whereas mode of arrival (25.5% missing) and triage time (65.3%) were the least complete. Disposition was missing in only 1.2% of medical records overall (range, 0%–5%) and diagnosis was missing in 3% (range, 0%–16%); these were missing from 14.4% and 10.5%, respectively, of administrative sources. Among visits with injury diagnoses, E-codes were missing in 27% of cases. For patients with asthma ( $n = 861$ ), documentation of specific elements of the clinical examination by nurses and physicians was also variable. **Conclusions:** Data elements important in emergency medical care for children are frequently missing in existing administrative and medical record sources; completeness varies widely across EDs. Researchers must be aware of these limitations in the use of existing data when planning studies. **Key words:** emergency medical services for children; databases; data collection; medical records. ACADEMIC EMERGENCY MEDICINE 2005; 12:1195–1200.

Researchers in emergency medical services for children (EMSC) often rely on existing data sources.<sup>1</sup> Uses of such data in the EMSC literature include reports of the epidemiology of acute childhood illness and injury,<sup>2,3</sup> descriptions of resource utilization in emergency care,<sup>4,5</sup> observational studies of risk factors and outcomes,<sup>6–8</sup> and quality audits.<sup>9</sup> Potentially useful data for such studies in EMSC may be obtained from existing clinical sources (e.g., hospital medical records, emergency department [ED] patient tracking

systems, out-of-hospital “run sheets,” trauma registries) and administrative sources (e.g., hospital billing and registration databases, public or commercial payer databases).

Limitations of such existing sources are well recognized. For example, administrative data sources in general have limited clinical information,<sup>10</sup> while many clinical sources such as paper medical records may require labor-intensive review. Coding errors have also been described in preexisting databases.<sup>11</sup> Finally, information of interest may be missing from a given data source. Although statistical methods are available to address small amounts of missing data, substantial amounts of missing data may lead to inability to answer important clinical questions or may provide biased results.<sup>12,13</sup> While the problem of missing data in studies of pediatric emergency care has been documented,<sup>14–17</sup> to our knowledge, the availability of data from existing data sources relevant to EMSC has not previously been examined on a national scale.

The purpose of this study was to use the infrastructure of a national pediatric emergency care research network to describe the availability and completeness of extant data pertinent to EMSC research across a large sample of hospitals.

From the Medical College of Wisconsin and Children’s Research Institute (MHG), Milwaukee, WI; the Children’s Hospital of Philadelphia (ERA), Philadelphia, PA; the Children’s National Medical Center (TS), Washington, DC; Marquette General Hospital (DS), Marquette, MI; University of Utah (RH, JMD), Salt Lake City, UT; and University of California Davis Medical Center (NK), Sacramento, CA.

The PECARN Study Investigators are listed in Appendix A. Received March 28, 2005; revision received June 15, 2005; accepted June 22, 2005.

Address for correspondence and reprints: Marc H. Gorelick, MD, MSCE, Section of Emergency Medicine, Department of Pediatrics, Medical College of Wisconsin, Children’s Hospital of Wisconsin MS #677, 9000 West Wisconsin Avenue, Milwaukee, WI 53226. Fax: 414-266-2635; e-mail: [mgorelic@mail.mcw.edu](mailto:mgorelic@mail.mcw.edu). doi:10.1197/j.aem.2005.06.028

## METHODS

**Study Design.** We conducted a retrospective study from January 1, 2002, to April 30, 2003, of pediatric ED visits in the Pediatric Emergency Medicine Applied Research Network (PECARN). The PECARN is composed of 25 EDs, organized into four regional nodes, with diverse regional and demographic representation.<sup>18</sup> The institutional review boards of all sites, including the Central Data Monitoring and Coordinating Center, approved the study.

**Study Setting and Population.** All member EDs of PECARN participated in the study, entitled the PECARN Core Data Project. Of member hospitals, 36% are freestanding pediatric hospitals, 68% are designated as pediatric trauma centers, 68% have a separate or freestanding pediatric ED, 20% have a pediatric ED within a general ED, and 12% have a general ED without separate pediatric facilities. All ED patients registered at any of the PECARN sites during the study period were eligible subjects for the PECARN Core Data Project. This article presents data from all patients younger than 19 years.

**Study Protocol.** We conducted the study in two phases.

**Phase I: Comprehensive Retrospective 12-month Electronic Data Capture.** Records for all ED patients younger than 19 years treated in 2002 were included. Data were obtained from existing administrative electronic sources such as registration, billing, and patient tracking systems. Each site determined the most appropriate sources to provide the requested data elements. Data elements requested included date of birth; ED arrival, triage, and discharge date and times; gender; street address; race and/or ethnicity (which we combined into a single race/ethnicity variable); ED disposition; ICD-9 diagnoses (up to five codes for each patient); Current Procedural Terminology codes (up to five for each patient); E-codes (up to three for each patient); arrival mode; and insurance payer type. For compliance with Health Insurance Portability and Accountability Act privacy regulations, all data were encrypted and sent to the Central Data Monitoring and Coordinating Center, which reviewed and summarized the data. We presented descriptive statistics for each variable to each site investigator to determine face validity. One of the authors (ERA) also reviewed all descriptive statistics to confirm face validity.

**Phase II: Comprehensive Concurrent Three-month Electronic Data Capture with Selected Chart Review.** Phase II data included electronic data collection and chart review components. We collected electronic data in an identical manner to phase I for 89 consecutive days during February, March, and April 2003.

For ten randomly selected days within this three-month period, ED medical records of patients younger than 19 years were selected for review, up to a maximum of 60 patients per day per site. This number was chosen by consensus of the investigators as sufficient to obtain representative information without allowing larger centers to dominate the data or overburdening available resources. At sites with  $\leq 60$  visits on a study day, all charts were abstracted. For sites exceeding 60 visits on a study day, a list of 60 random numbers (from one to the census for that day) was generated from the central data management center and accessed via the Internet by the site study personnel. This list of random numbers was used to select charts for review from the daily ED log at the site.

Research assistants abstracted clinical data from all selected charts. Patients with diagnoses of asthma or long bone fracture, selected as representative of common medical and surgical problems, respectively, had additional data fields abstracted. For asthma, variables were selected that are commonly used in clinical severity scores. For long bone fractures, we were particularly interested in variables related to injury epidemiology, such as E-codes and injury circumstances. Pain scoring was not examined.

A manual of operations describing abstraction procedures was developed and disseminated to all sites. Site investigators underwent a live training session in abstraction procedures with sham chart review, and research assistants were subsequently trained on-site by the local investigators. Abstractors recorded data onto paper forms and subsequently entered the data into a specially written data entry program. Quality assurance included double data entry of 10% of charts and chart abstraction by the physician principal investigator for approximately 5% of charts to verify research assistant abstractions.

During the study period, none of the participating institutions had a comprehensive electronic medical record. Sites were asked if they used a generic chart or templated chart with prompts for specific data elements.

**Data Analysis.** For each data element, the primary outcome of interest was the presence of a valid value for that element in either the administrative data or the medical record; results are reported separately for each of these two data sources. Entries that were blank, specified as missing or unknown, or contained invalid values were categorized as missing. We calculated the proportion of missing entries both in aggregate across all sites and by individual site. Results are reported as the overall proportion and the site-specific range.

## RESULTS

Of the 25 participating PECARN hospitals, two were unable to provide the requested electronic data within

**TABLE 1. Frequency of Missing/Unknown Data from Administrative and Clinical Sources**

	Date of Birth	Race/Ethnicity‡	Mode of Arrival	Disposition	Insurance Type	Diagnosis§	Arrival Time	Triage Time	Discharge Time
Electronic administrative data									
No. of sites reporting electronically*	20†	21	18	21	23	23	20	9	19
Overall % missing	0.3†	15.6	25.5	14.4	1.3	10.5	6.6	65.3	13.3
Range of missing data by site (%)*	0–3.4	0–91.9	0–99.5	0–43.1	0–18.5	0–90.6	0–1.6	0–99.7	0–11.4
No. of sites with 20% or more missing data*	0/20	4/21	4/18	1/21	0/23	2/23	0/20	2/9	0/19
Medical record abstraction									
Overall % missing	0	24.7	12.8	1.2	8.7	3.0	21.2	9.0	6.4
Range of missing data by site (%)	—	0–97.7	0.2–91.1	0–5.3	0–16.4	0–16.1	0–71.3¶	0–17.1	0–22.4
No. of sites with 20% or more missing data	0/25	8/25	6/25	0/25	1/25	0/25	4/25	1/25	2/25

$n = 749,036$  for administrative data;  $n = 12,756$  for medical record data.

\*Two sites were unable to provide any electronic administrative data.

†Excludes three sites with combined adult/pediatric EDs in which patients were selected by age derived from date of birth.

‡Only eight of 21 sites reported race and ethnicity separately.

§Records with at least one diagnosis documented were considered complete (up to five diagnoses were allowable).

||Range for 24 sites; one additional site was missing 100%.

¶Range for 21 sites; four additional sites were missing 100%.

the specified time frame of 14–16 months. All 25 sites participated in the chart review. Sites queried a median of two data sources (range, 1–5) to obtain the requested information.

During the 12-month study period of phase I, there were 749,036 visits retrieved from the electronic data sources at the 23 reporting hospitals. During phase II, there were 216,915 visits, with 44,415 occurring during the ten randomly selected study days. Of these, 12,756 charts (28.7%) were abstracted.

The frequency of missing information for each data element from the electronic administrative and clinical data sources is shown in Table 1. Data availability and completeness from both data sources varied by data element and by site. Although race/ethnicity was missing in only 15.6% of cases overall from the administrative sources, of the 21 sites reporting race, only eight reported ethnicity separately.

E-codes were reported by 22 of 23 sites. However, among patients with injury diagnoses (ICD-9 diagnoses 800.0–999.9), E-codes were missing in 27% of the records (range, 1%–99%), with missing rates of 20% or more at five sites.

The documentation of relevant clinical information in the medical records of patients with the two target conditions was also variable. For patients with asthma ( $n = 861$ ), 16.1% were missing information regarding

history of wheezing. Documentation of specific elements of the clinical examination by nurses and physicians was also variable (Table 2). Among patients with fractures ( $n = 252$ ), mechanism of injury was documented 96.4% of the time, but location of the event was recorded for only 54.4%. The use of charts with specific templates did not influence the completeness of documentation. Of the nine sites reported to use a templated asthma chart, 11.5% of records were missing data on all five examination elements, compared with 6.3% of records at the sites not using a templated chart (difference, 5.2%; 95% confidence interval, 0.3% to 10.2%).

Duplicate abstraction of selected information was performed by research assistants and physician investigators for 655 medical records. Of these, research assistants identified a missing diagnosis in 16 (2.4%), and investigators agreed in 13 (81%).

## DISCUSSION

This study provides important information regarding the quality and completeness of existing data sources in a diverse group of EDs providing pediatric care across the United States. Even for basic demographic information, availability of various data elements in both administrative databases and medical charts

**TABLE 2. Availability of Data on Selected Elements of Clinical Examination of Patients with Asthma ( $n = 861$ )**

Examination	Retractions	Wheezing	Grunting	Nasal Flaring	Aeration	All Five Elements Missing	Any Data Element Missing
Physician	42.9%	15.3%	86.3%	79.2%	51.3%	9.5%	91.4%
Nurse	52.4%	33.5%	82.4%	74.8%	60.4%	27.2%	88.5%

Numbers indicate the overall percentage of patients with a given finding not documented on examination by a physician or nurse.

varied widely across sites. Nearly all of the EDs had some missing data, and in many cases the proportion of missing data was substantial. Some data elements seemed to be more readily available from manual record review than from electronic administrative sources, while other variables were more complete from the administrative database. Interestingly, despite extended efforts over several months, two sites were unable to obtain any of the basic data elements requested from their information systems departments.

Although reviews of existing data sources can be an attractive choice for researchers in emergency health services, questions of completeness and accuracy of information pose a challenge to the usefulness and validity of such studies.<sup>19</sup> The degree to which missing data poses a threat to the validity of conclusions is unclear and probably depends on many factors, including the study question, the type of variable, and whether there is any pattern to the missing data.<sup>13,19</sup>

One source recommends excluding variables with more than 10% missing data.<sup>20</sup> Even if a more lenient threshold of 20% missing data is considered unacceptable, basic epidemiologic data from electronic administrative sources would be considered unreliable from 19% of EDs for race, 22% for mode of arrival, and 9% for diagnosis. Detailed clinical information from medical record review for two target conditions (asthma and long bone fractures) was similarly problematic. For example, the information necessary to calculate even the most parsimonious clinical asthma score was missing for most patients. While some clinical findings, such as nasal flaring, may not be assessed or documented routinely unless present, even a widely used finding (i.e., retractions) was missing in a large percentage of cases. Our results are consistent with those of other investigators at single institutions. Teo et al., for example, found high rates of missing information on clinical assessment by nurses and physicians for children with acute asthma in a New Zealand ED.<sup>21</sup> Moll et al. found the location of the accident recorded in only 23% of charts of children involved in bicycle injuries seen at a pediatric ED.<sup>16</sup>

Missing data also have implications beyond the use of existing data for research. Hospitals are increasingly required to report visit data to government agencies, insurers, and other organizations such as the National Center for Quality Assurance Health Plan Employer Data and Information Set. Incomplete data may affect an institution's apparent performance measures and reimbursement.

Given the obvious deficiencies in documentation, both in medical records and electronic databases pertaining to pediatric ED visits, future efforts should focus on improving the completeness of such records. It seems unlikely that guidelines alone will be sufficient. Regarding the data we examined, it should be noted that all of the administrative data elements studied are included in the Data Elements for

Emergency Department Systems minimum recommended data set from the Centers for Disease Control and Prevention,<sup>22</sup> and many are required by the Joint Commission on Accreditation of Healthcare Organizations.<sup>23</sup> Other investigators have demonstrated that use of structured chart templates may improve documentation of selected elements in written ED records.<sup>24,25</sup> However, in our study, documentation of basic data elements for children with asthma was no different at PECARN EDs with or without programmed charts. We cannot determine whether these charts were actually used for the visits included in this study, but it seems that simply making such charts available does not improve documentation.

## LIMITATIONS

This was the first study conducted across the entire PECARN and therefore our first effort at working with information technology and medical records personnel at each site. Data retrieval may improve with greater experience. Manual chart review was conducted only during a three-month period in the winter, which tends to be a period of high ED census for pediatric patients. Documentation may be less complete at such times due to competing time demands on clinical staff. Differences across institutions may reflect variation in training and qualifications of research assistants, although quality assurance activities by site investigators should have ameliorated any such effect. Although the population of EDs represented is diverse, our findings may not be generalizable to other settings. Our goal for this study was limited to an evaluation of the availability and completeness of data available from two disparate sources. However, future studies of the quality of the data and agreement between the two sources are certainly indicated.

## CONCLUSIONS

Data elements important in EMSC are frequently missing from existing administrative databases and clinical records; completeness varies widely across EDs. Researchers must be aware of these limitations in the use of existing data when planning studies and when benchmarking clinical practice.

## References

1. Gilbert EH, Lowenstein SR, Koziol-McLain J, Barta DC, Steiner J. Chart reviews in emergency medicine research: where are the methods? *Ann Emerg Med.* 1996; 27:305-8.
2. Watson WA, Litovitz TL, Rodgers GC Jr, et al. 2002 annual report of the American Association of Poison Control Centers Toxic Exposure Surveillance System. *Am J Emerg Med.* 2003; 21:353-421.
3. Gorelick MH, Baker MD. Epiglottitis in children 1979-1992: effects of *Haemophilus influenzae* type b immunization. *Arch Pediatr Adolesc Med.* 1994; 148:47-50.

4. Alessandrini EA, Lavelle JM, Grenfell SM, Jacobstein CR, Shaw KN. Return visits to a pediatric emergency department. *Pediatr Emerg Care*. 2004; 20:166–71.
5. De Piero AD, Teach SJ, Chamberlain JM. ED evaluation of infants after an apparent life-threatening event. *Am J Emerg Med*. 2004; 22:83–6.
6. Glaser N, Barnett P, McCaslin I, et al., for the Pediatric Emergency Medicine Collaborative Research Committee of the American Academy of Pediatrics. Risk factors for cerebral edema in children with diabetic ketoacidosis. *N Engl J Med*. 2001; 344:264–9.
7. Pitetti R, Glustein JZ, Bhende MS. Prehospital care and outcome of pediatric out-of-hospital cardiac arrest. *Prehosp Emerg Care*. 2002; 6:283–90.
8. Galustyan SG, Walsh-Kelly CM, Szweczuga D, Bergholte J, Hennes H. The short-term outcome of seizure management by prehospital personnel: a comparison of two protocols. *Pediatr Emerg Care*. 2003; 19:221–5.
9. Depiero AD, Ochsenschlager DW, Chamberlain JM. Analysis of pediatric hospitalizations after emergency department release as a quality improvement tool. *Ann Emerg Med*. 2002; 39:159–63.
10. Iezzoni LI. Data sources and implications: administrative databases. In: Iezzoni LI (ed). *Risk Adjustment for Measuring Healthcare Outcomes*, ed 2. Chicago, IL: Health Administration Press; 1997, pp 169–242.
11. Hsia DC, Krushat WM, Fagan AB, Tebbutt JA, Kusserow WM. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. *N Engl J Med*. 1988; 318:352–5.
12. Norris CM, Ghali WA, Knudtson ML, Naylor CD, Saunders LD. Dealing with missing data in observational health care outcome analyses. *J Clin Epidemiol*. 2000; 53:377–83.
13. Dworkin RJ. Hidden bias in the use of archival data. *Eval Health Prof*. 1987; 10:173–85.
14. Miller L, Kent RM, Tennant A. Audit of head injury management in accident and emergency at two hospitals: implications for NICE CT guidelines. Available at: <http://www.biomedcentral.com/1472-6963/4/7>. Accessed Dec 17, 2004.
15. Houry D, Feldhaus KM, Nyquist SR, Abbott J, Pons PT. Emergency department documentation in cases of intentional assault. *Ann Emerg Med*. 1999; 34:715–9.
16. Moll EK, Donoghue AJ, Alpern ER, Kleppel J, Durbin DR, Winston FK. Child bicyclist injuries: are we obtaining enough information in the emergency department chart? *Inj Prev*. 2002; 8:165–9.
17. Schoenfeld PS, Baker MD. Documentation in the pediatric emergency department: a review of resuscitation cases. *Ann Emerg Med*. 1991; 20:641–3.
18. The Pediatric Emergency Care Applied Research Network. The Pediatric Emergency Care Applied Research Network (PECARN): rationale, development, and first steps. *Acad Emerg Med*. 2003; 10:661–8.
19. Worster A, Haines T. Advanced statistics: understanding medical record reviews. *Acad Emerg Med*. 2004; 11:187–92.
20. Wu L, Asthon CM. Chart review: a need for reappraisal. *Eval Health Prof*. 1997; 20:146–63.
21. Teo S, Hanson R, Van Asperen P, et al. Improving asthma documentation in a pediatric emergency department. *J Pediatr Child Health*. 1995; 31:130–3.
22. National Center for Injury Prevention and Control. Data elements for emergency department systems, release 1.0. Atlanta, GA: Centers for Disease Control and Prevention, 1997. Available at: <http://www.cdc.gov/ncipc/pub-res/deedspage.htm>. Accessed Apr 14, 2005.
23. Joint Commission on Accreditation of Healthcare Organizations. *Comprehensive Accreditation Manual for Hospitals: The Official Handbook (CAMH)*. Chicago, IL: Joint Commission on Accreditation of Healthcare Organizations, 2004.
24. Humphreys T, Shofer FS, Jacobson S, et al. Preformatted charts improve documentation in the emergency department. *Ann Emerg Med*. 1992; 21:534–40.
25. Wrenn K, Rodewald L, Lumb E, et al. The use of structured, complaint-specific patient encounter forms in the emergency department. *Ann Emerg Med*. 1993; 22:805–12.

## APPENDIX A

The Pediatric Emergency Care Applied Research Network (PECARN) is supported by cooperative agreements U03MC00001, U03MC00003, U03MC00006, U03MC00007, and U03MC00008 from the Emergency Medical Services for Children program of the Maternal and Child Health Bureau, Health Resources and Services Administration, U.S. Department of Health and Human Services.

Participating centers and site investigators are listed as follows in alphabetical order: Atlantic Health System/Morristown Memorial Hospital (M. Gerardi); Bellevue Hospital Center (M. Tunik); Calvert Memorial Hospital (K. Melville); Children's Hospital of Buffalo (K. Lillis); Children's Hospital of Michigan (P. Mahajan); Children's Hospital of New York–Presbyterian (S. Miller); Children's Hospital of Philadelphia (E. Alpern); Children's National Medical Center (J. Chamberlain, S. Teach); Cincinnati Children's Hospital Medical Center (R. Ruddy); DeVos Children's Hospital (J. Hoyle); Franklin Square Hospital (D. Alexander); Harlem Hospital Center (J. Tsung); Holy Cross Hospital (C. Johns); Howard County Medical Center (D. Monroe); Hurley Medical Center (R. Stanley); Johns Hopkins Medical Center (A. Walker); Marquette General Hospital (D. Snowdon); Medical College of Wisconsin/Children's Hospital of Wisconsin (M. Gorelick); St. Barnabas Health Care System (J. Brennan); University of California Davis Medical Center (N. Kuppermann); University Hospital/SUNY–Upstate Medical University (J. Callahan); University of Michigan (D. Treloar); University of Rochester; University of Utah/Primary Children's Medical Center (H. Corneli); and Washington University/St. Louis Children's Hospital (D. Jaffe).

PECARN Steering Committee: N. Kuppermann (Chair), D. Alexander, E. Alpern, J. Chamberlain, J. M. Dean, M. Gerardi, J. Goepf, M. Gorelick, J. Hoyle, D. Jaffe, C. Johns, N. Levick, P. Mahajan, R. Maio, S. Miller (deceased), D. Monroe, R. Ruddy, R. Stanley, D. Treloar, M. Tunik, and A. Walker. HRSA/MCHB Liaisons: D. Kavanagh, H. Park.

Central Data Management and Coordinating Center: M. Dean, R. Holubkov, S. Knight, and A. Donaldson.

Data Analysis and Management Subcommittee: J. Chamberlain (Chair), M. Brown, H. Corneli, J. Goepf, R. Holubkov, P. Mahajan, K. Melville, E. Stremski, and M. Tunik.

Grants and Publications Subcommittee: M. Gorelick (Chair), E. Alpern, J. M. Dean, G. Foltin, J. Joseph, S. Miller (deceased), F. Moler, R. Stanley, and S. Teach.

Protocol Concept Review and Development Subcommittee: D. Jaffe (Chair), A. Cooper, J. M. Dean, C. Johns, R. Kanter, R. Maio, N. C. Mann, D. Monroe, K. Shaw, and D. Treloar.

Quality Assurance Subcommittee: R. Stanley (Chair), D. Alexander, J. Brown, M. Gerardi, M. Gregor, R. Holubkov, K. Lillis, R. Ruddy, M. Shults, and A. Walker.

Safety and Regulatory Affairs Subcommittee: N. Levick (Chair), J. Brennan, J. Brown, J. M. Dean, J. Hoyle, R. Ruddy, W. Schalick, T. Singh, D. Snowdon, and J. Wright.