Lack of Agreement in Pediatric Emergency Department Discharge Diagnoses from Clinical and Administrative Data Sources

Marc H. Gorelick, MD, MSCE, Stacey Knight, MStat, Evaline A. Alessandrini, MD, MSCE, Rachel M. Stanley, MD, James M. Chamberlain, MD, Nathan Kuppermann, MD, MPH, Elizabeth R. Alpern, MD, MSCE, for the Pediatric Emergency Care Applied Research Network

Abstract

Background: Diagnosis information from existing data sources is used commonly for epidemiologic, administrative, and research purposes. The quality of such data for emergency department (ED) visits is unknown.

Objectives: To determine the agreement on final diagnoses between two sources, electronic administrative sources and manually abstracted medical records, for pediatric ED visits, in a multicenter network.

Methods: This was a cross sectional study at 19 EDs nationwide. The authors obtained data from two sources at each ED during a three-month period in 2003: administrative sources for all visits and abstracted records for randomly selected visits during ten days over the study period. Records were matched using unique identifiers and probabilistic linkage. The authors recorded up to three diagnoses from each abstracted medical record and up to ten for the administrative data source. Diagnoses were grouped into 104 groups using a modification of the Clinical Classification System.

Results: A total of 8,860 abstracted records had at least one valid diagnosis code (with a total of 12,895 diagnoses) and were successfully matched to records in the administrative source. Overall, 67% (95% confidence interval = 66% to 68%) of diagnoses from the administrative and abstracted sources were within the same diagnosis group. Agreement varied by site, ranging from 54% to 77%. Agreement varied substantially by diagnosis group; there was no difference by method of linkage. Clustering clinically similar diagnosis groups improved agreement between administrative and abstracted data sources.

Conclusions: ED diagnoses retrieved from electronic administrative sources and manual chart review frequently disagree, even if similar diagnosis codes are grouped. Agreement varies by institution and by diagnosis. Further work is needed to improve the accuracy of diagnosis coding; development of a grouping system specific to pediatric emergency care may be beneficial.

ACADEMIC EMERGENCY MEDICINE 2007; 14:646–652 © 2007 by the Society for Academic Emergency Medicine

Keywords: International Classification of Diseases, medical records, reproducibility of results, data collection standards

Diagnoses made in the course of pediatric emergency department (ED) visits are routinely recorded in various data sources, including the clinical medical record and administrative sources associated with registration and billing. Information on diagnoses retrieved from these sources is commonly used for

From the Department of Pediatrics, Medical College of Wisconsin (MHG), Milwaukee, WI; Department of Pediatrics, University of Utah (SK), Salt Lake City, UT; Division of Emergency Medicine, Children's Hospital of Philadelphia (EAA, ERA), Philadelphia, PA; Department of Emergency Medicine, University of Michigan (RMS), Ann Arbor, MI; Division of Emergency Medicine, Children's National Medical Center (JMC), Washington, DC; and Departments of Emergency Medicine and Pediatrics, University of California–Davis School of Medicine (NK), Sacramento, CA.

a variety of purposes, including epidemiologic monitoring,[1,2] administration,[3,4] and clinical research.[5–7]

Prior work by other investigators has raised concerns about the accuracy of diagnoses obtained from existing sources of data in other settings[8] and in the ED.[9] Errors may occur in the assignment of a given diagnosis by the clinician, in recording the diagnosis in the primary medical record, in abstracting and transcribing the diagnosis from the record into the administrative data set, or in extracting data from that data source.[10] However, there has been no description of the quality of such diagnosis data for pediatric emergency care (i.e., in the setting of the pediatric ED).

The aim of this study was to determine the agreement on discharge diagnoses for pediatric ED visits obtained from two different data sources: diagnoses extracted from electronically stored administrative sources from hospital information systems, and diagnoses abstracted manually from medical records.

METHODS

Study Design
This was a cross sectional study comparing information obtained for the same patient visits from two data sources: medical paper charts (with abstraction performed by trained research assistants [RAs]), and electronic administrative (usually billing) data. In addition, we compared diagnoses from chart abstraction by the RAs and by physician investigators for a subset of patient visits. The study was approved by the institutional review boards of the participating hospitals and of the Central Data Monitoring and Coordinating Center of the research network in which this study was conducted.

Study Setting and Population
The study was performed in the Pediatric Emergency Care Applied Research Network (PECARN),[11–13] a federally funded network of 25 hospitals with broad geographic representation and a diversity of hospital types, patient populations, and providers.

Study Protocol
Administrative data were obtained from each hospital's information systems department for all ED visits by patients younger than 19 years of age during a three-month period: February to April 2003. Data were then encrypted and sent to the PECARN Central Data Monitoring and Coordinating Center for analysis. Up to ten diagnoses, using the International Classification of Diseases Ninth Revision–Clinical Modification (ICD-9-CM) codes, could be included for each visit. We presented site-specific descriptive statistics for each variable to each site physician investigator to determine face validity of the site statistics (e.g., assuring that reported demographics, admission rate, percentage public insurance, and percentage arriving via ambulance were consistent with the investigator's experience at the site). Additionally, two study investigators (ERA and JMC) also reviewed the summary statistics for all data elements from each site to assess face validity. All outlier data were confirmed by the participating sites or corrected as necessary.

For purposes of medical record selection, during the same three-month period, ten days were selected at random. The same ten days were used at all sites. Medical records for patients younger than 19 years of age seen during those ten study days were reviewed by the trained RA chart abstractors. The background, qualifications, and number of RAs varied by site but primarily consisted of one or several college graduates or nurses employed at each site to assist with PECARN research studies. Supervision was provided by the site physician investigators and by PECARN nodal administrators. The RAs recorded up to three discharge diagnoses for each visit. A maximum of 60 charts per site per day was set; for sites that exceeded this number of eligible visits on a designated study day, a list of 60 visits was randomly selected by the central data management center for abstraction. Data were double entered, and a 5% sample of charts was reabstracted by the physician investigator at each site for quality control and to assess physician–RA agreement. The physician investigator was instructed to follow the same chart abstraction procedures that the RA followed.

Measurements
To attempt to minimize the impact of clinically unimportant variations in coding, we used the Agency for Healthcare Research and Quality Clinical Classification System to group similar diagnoses.[14] This system was originally developed for adult patients and was modified by the investigators to better reflect the diagnoses found in pediatric emergency practice, yielding 104 diagnosis groups.

For the medical record review, RAs and physician investigators were instructed to record the specific diagnoses and then assign each diagnosis to the appropriate group using a standardized manual of operations, developed at the PECARN Central Data Management and Coordinating Center. The manual was pilot tested and then presented to study personnel at several training sessions. During the period of data collection, regular feedback was provided to study personnel through conference calls and circulation of a "frequently asked questions" list. For the administrative data sources, individual ICD-9-CM codes were assigned to the appropriate group using a slight modification of the validated Agency for Healthcare Research and Quality Clinical Classification System mapping software. When referring to agreement, the term "diagnosis" therefore refers to the diagnosis group, not the individual ICD-9-CM code.

Data Analysis
Records for visits from the electronic data were linked to the abstracted data using two methods. For records from institutions that were able to transmit unique identifiers, we performed exact matching on key fields: medical record number, date of birth, and date and time of visit. When unique identifiers permitting exact matching were not available, we used probabilistic record linkage. Probabilistic record linkage is a method in which common variables from the two databases are compared to determine the likelihood that the pair of records refers to the same visit.[15,16] Variables used in the probabilistic linkage were date of birth, date and time of visit, gender, race/ethnicity, mode of arrival, home address, city, zip

code, disposition, discharge date, and discharge time. Based on our file sizes and expected number of matches, only linked pairs attaining a minimum match probability of 90% of being correct were kept.[17] The 90% minimum match probability was chosen based on a desire to minimize false positives as well as past experience with the linkage of similar databases. The physician investigator– and RA-abstracted data were matched using exact matching on medical record number and date of visit.

The outcome of interest was agreement between diagnosis categories in the two data sources. Because visits could have more than one diagnosis, we considered the unit of analysis the diagnosis, not the visit. We began with each diagnosis from the abstracted medical record. If the same diagnosis was found among any of the diagnoses in the corresponding administrative data for that visit, regardless of the order in which the diagnoses were listed, this was considered an agreement. The denominator is therefore the number of unique diagnoses found in the abstracted paper records, and the numerator is the number of matching diagnoses from the linked administrative record. Agreement rates were calculated for all identified diagnoses, as well as separately by hospital and by diagnosis group. For agreement between physician investigator and RA, we chose the site physician investigator as the reference standard and calculated true- and false-positive rates for the RA abstraction compared with the physician investigator. Although we recognize that the physician investigator may be subject to errors as well, we believed that an emergency physician reviewing a medical record retrospectively would be more likely to arrive at the diagnosis intended by another emergency physician compared with a billing coder or an RA.

When appropriate, 95% confidence intervals (CIs) for proportions were calculated using normal approximation methods.

RESULTS

Nineteen of the 25 sites were able to provide data from both electronic administrative sources and RA chart abstraction, and for these 19 sites we compared diagnoses from those two sources. Four additional sites performed chart abstraction but could not provide the electronic administrative data. For analysis of RA and physician investigator agreement, data from all 23 sites are included. Two sites in the PECARN network did not provide any data.

Data for 172,688 visits for patients younger than 19 years old were obtained from 19 PECARN sites during the three-month study period (Figure 1). A total of 9,553 records were manually abstracted during the ten randomly chosen study days. A total of 329 of these records could not be linked to a corresponding record from the administrative source, leaving 9,224 (96.6%) records successfully linked. Of these, 5,643 (61.2%) matched exactly on medical record number, date of birth, and date and time of visit. An additional 1,684 (18.3%) matched exactly on date of birth and on date and time of visit. The remaining 1,897 (20.6%) matched with at least 90% probability on other combinations of variables. No valid diagnosis code was recorded for 364 of the records, mostly

due to the electronic record missing the required fifth digit in an otherwise valid ICD-9-CM code. This left 8,860 records for analysis, with 12,895 diagnoses in the abstracted records and 15,082 diagnoses in the matched electronic source.

Overall, 8,628 diagnoses were reported in both the abstracted record and the corresponding electronic administrative data, resulting in an agreement rate of 66.9% (95% CI = 66.1% to 67.7%). Diagnosis agreement did not differ substantially by the type of variables used to match the electronic and abstracted records. The diagnosis agreement rate for those records that matched exactly on medical record number, date of birth, date of visit, and hour of visit was the lowest, with an agreement of 66.1% (95% CI = 65.0% to 67.1%). For those records that matched exactly on date of birth, date of visit, and hour of visit, the rate of diagnostic agreement was 66.6% (95% CI = 64.7% to 68.5%), and for those records matching on other combinations of variables the diagnostic agreement rate was 69.6% (95% CI = 67.8% to 71.3%). Diagnosis agreement between abstracted records and electronic sources varied by site, with a range from 53.7% to 76.6%. While the differences between sites were statistically significant, there was no pattern of agreement based on site-specific sample size or by the average number of diagnoses per record at a site.

Table 1 shows the agreement by diagnosis groups for the 33 groups with at least 100 observations each; the total number of observations in these 33 groups is 10,989, or 85.2% of the total (an expanded table including all diagnosis groups is included as an online Data Supplement available at http://www.aemj.org/cgi/content/full/j.aem. 2007.03.1357/DC1). There was considerable variation in agreement between groups, from 2.5% to greater than 90% agreement. There was no pattern of agreement by the number of observations. In addition, the ten diagnosis groups with the highest agreement rates and the ten with the lowest rates include a representative combination of injury and noninjury diagnoses.

We examined in detail several of the diagnosis groups with poor agreement. The diagnosis group "strep throat" contains only three ICD-9-CM codes (034, 034.0, 034.1). When strep throat was the diagnosis abstracted from the chart, only 2.5% of administrative data for these patients also had a diagnosis of strep throat, constituting agreement. Of the other diagnoses from the matching administrative source, 76.1% were upper respiratory tract infection, 19.5% unspecified febrile illness, 17.0% pharyngitis, and 12.6% other bacterial infection. If the diagnosis groups for strep throat and pharyngitis are combined into a single category, agreement increases to 59.8%.

When the category "intestinal infection" was abstracted from the medical record, only 8.2% of corresponding electronic data also had a diagnosis of intestinal infection, indicating agreement. The single most common alternative diagnosis group reported in the administrative data for these patients was gastroenteritis (88.5%); other diagnoses commonly found were nausea/vomiting (32.0%), fever (18.9%), and dehydration (fluid/electrolyte disorder) (18.0%). When intestinal infection and gastroenteritis were combined into a single category, the agreement was 90.3%.
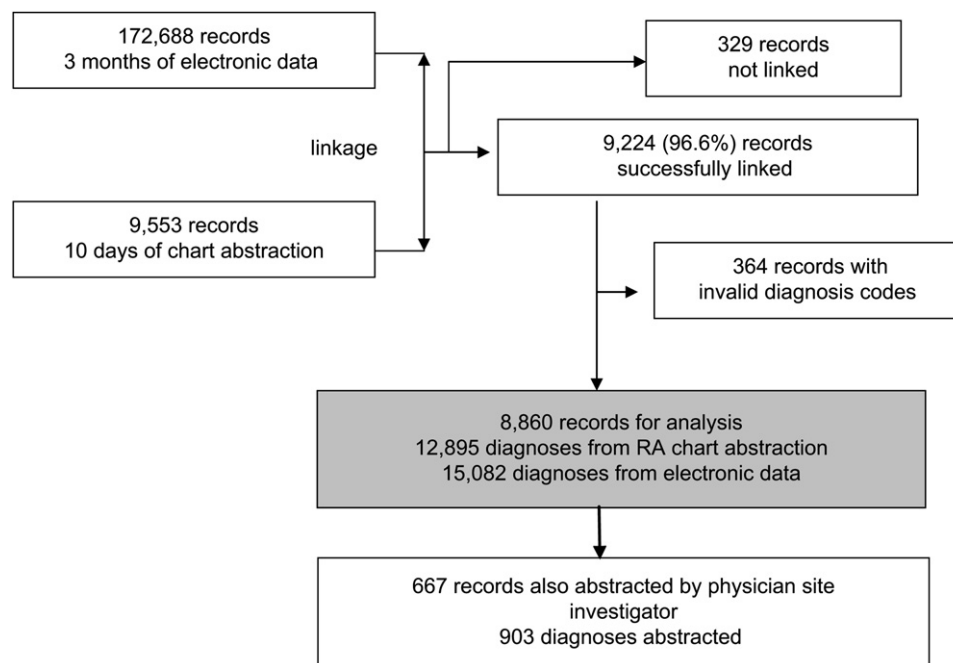
Figure 1. Overview of study data flow. RA = research assistant.

A total of 667 charts (7% of the 9,224 successfully linked) were abstracted by both the site physician investigator and the RA. The physician investigators recorded a total of 903 diagnoses. Using these as the reference standard, the RAs correctly recorded the same diagnosis 689 times, for a true-positive rate (sensitivity) of 76.3% (95% CI = 73.5% to 79.1%). Conversely, of the 956 diagnoses recorded by RAs, 267 were not recorded by the physician investigator, yielding an RA false-positive rate of 27.9% (95% CI = 25.1% to 30.8%).

Of the 903 diagnoses recorded by the physician investigators, 753 (83.4%) came from hospitals that provided electronic administrative data. We linked the diagnoses recorded by the physician investigators to 683 of these records, and 676 of these administrative records reported at least one diagnosis. There was an overall agreement between diagnosis abstracted by the investigator and that obtained in the administrative data of 64.4% (95% CI = 60.6% to 68.0%).

DISCUSSION

In this study, we have demonstrated limited agreement between ED diagnoses abstracted from the medical record and those obtained from administrative billing sources, extending to ED visits a phenomenon previously shown for both inpatient and ambulatory visits. When diagnosis data from existing administrative sources are used for research or other purposes, there may be questions about completeness and the extent to which the administrative data source correlates with the clinical record. We previously reported that the discharge diagnosis was present in the electronic administrative source for 97% of pediatric ED visits in PECARN, but was missing in up to 16% of visits at some individual sites.[18] Studney and Hakstian, in a study of more than 1,200 office

visits to 12 primary care practices in British Columbia, compared diagnoses abstracted from the chart with those listed in the billing record.[19] The overall agreement rate was 60%, similar to that in the current study, with a wide range of agreement across physicians from 34% to 89%. Hsia et al. found a 21% error rate in diagnosis-related group coding of hospital discharges in a sample of U.S. hospitals.[20] Interestingly, although such discrepancies have been widely reported for some time, health professionals may underestimate the degree of the problem. In a national survey of more than 16,000 health information managers at U.S. hospitals, respondents reported an estimated average 5.1% disagreement rate between physicians and coders, with only 7% reporting an estimated disagreement rate of more than 10%.[21]

Other investigators, focusing on adult ED visits, have described limitations of administrative diagnosis data. Nagourney et al. found both underascertainment and overascertainment of acute coronary syndromes using discharge diagnosis compared with medical record review of formal diagnostic criteria.[9] Bazarian et al., studying mild traumatic brain injury, found that ICD-9 codes were reasonably specific but insensitive in identifying patients with traumatic brain injury compared with prospective surveillance.[22] The results of the current study demonstrate similar problems for childhood ED visits across a broad spectrum of diagnosis groups and among a diverse group of hospitals.

Disagreement between two sources of information may arise from numerous factors.[10] First, some of the pairs of records may have been linked in error and not actually refer to the same visit, although this is unlikely because the expected error rate is quite low given the 90% minimum linkage probability and the availability of unique identifiers for the majority of records;

Table 1

Agreement between Diagnosis Abstracted from Chart and Administrative Diagnosis for the Diagnosis Groups with at Least 100 Observations

Diagnosis Group	n	Percent Agreement	95% CI
Otitis media	819	94.0	92.4, 95.6
Laceration: head, neck, trunk	412	93.5	91.1, 95.8
Fracture: upper extremity	197	91.9	88.1, 95.7
Croup	136	90.4	85.5, 95.4
Fluid/electrolyte disorder (including dehydration)	309	89.0	85.5, 92.5
Seizure/epilepsy	179	88.3	83.6, 93.0
Asthma	575	86.8	84.0, 89.6
Pneumonia	266	86.5	82.4, 90.6
Bronchiolitis	271	84.9	80.6, 89.1
Eye problem	175	78.3	72.2, 84.4
Laceration: lower extremity	153	77.1	70.5, 83.8
Abdominal pain	283	74.2	69.1, 79.3
Upper respiratory infection	862	73.9	71.0, 76.8
Sprain/strain	360	73.6	69.1, 78.2
Mental disorder/illness other	104	73.1	64.6, 81.6
Pharyngitis	330	72.7	67.9, 77.5
Gastroenteritis	632	72.2	68.7, 75.7
Superficial injury	555	71.5	67.8, 75.3
Nausea/vomiting	275	71.3	65.9, 76.6
Allergic reaction	166	69.9	62.9, 76.9
Urinary tract infection	123	69.9	61.8, 78.0
Fever	750	64.4	61.0, 67.8
Viral infection	846	62.2	58.9, 65.4
Injury: other	292	58.9	53.3, 64.6
Respiratory disease/cough	166	56.0	48.5, 63.6
Skin condition (including dermatitis and skin infection)	317	45.7	40.3, 51.2
Other	553	30.6	26.7, 34.4
Head injury/concussion	211	21.8	16.2, 27.4
Other bacterial infection	103	19.4	11.8, 27.1
Other lower gastrointestinal disease	160	12.5	7.4, 17.6
Intestinal infection	122	8.2	3.3, 13.1
Respiratory failure, insufficiency, or arrest (includes hypoxia)	128	6.3	2.1, 10.4
Strep throat	159	2.5	0.1, 5.0

moreover, a small linkage error rate would not explain disagreement rates as high as those found here. There could also have been errors in coding by hospital personnel when diagnoses were entered into the administrative data set. Such errors may be relatively random, as might occur with erroneous keystrokes, for example, or may reflect systematic differences in criteria used by coders compared with clinical personnel.[23] Differences in sequences of multiple diagnoses are common,[24] but we examined any occurrence of the same diagnosis, regardless of order; therefore, it is unlikely that this accounted for the disagreement rate. In addition, errors may have occurred in chart abstraction by study personnel. This is supported by our findings of substantial disagreement in abstraction by physician investigators

and RAs. Indeed, we also found substantial interrater disagreement between the physician abstractors and the administrative data.

Finally, disagreement may arise when similar diagnoses are assigned different codes, leading to a distinction without a clinically meaningful difference. This can occur when coders and study abstractors use different levels of specificity in the ICD-9-CM code assignment (e.g., use of three-digit vs. four-digit or five-digit codes) or when less specific diagnoses may be reasonably assigned to one of several codes. An example of the latter would be the use of ICD-9-CM code 009.0, for "infectious enteritis not otherwise specified," in place of one of the codes indicating enteritis caused by a specific infectious agent. Such clinically unimportant variation may arise due to the very large number of individual ICD-9-CM codes (more than 8,000). Other investigators have found high rates of disagreement even when similar diagnoses are grouped together.[25,26] We attempted to minimize such unimportant variation by using a previously developed diagnosis clustering system, the Agency for Healthcare Research and Quality Clinical Classification System. However, even with our modification of this system, the groupings may not be suitable for pediatric ED visits in that some similar diagnoses may be placed into different categories, as appears to be the case for lower gastrointestinal illnesses, for example. Improved agreement, and more meaningful information, may be possible with a diagnosis grouping system developed specifically for pediatric ED visit diagnoses. Although such groupings have been developed on an ad hoc basis for specified purposes,[27] no widely accepted, validated diagnosis grouping system currently exists.

LIMITATIONS

A potential limitation of the current study is the possibility of variation in quality of data and medical record abstraction across sites. In addition, we used diagnosis groups rather than specific diagnoses to minimize unimportant variability in coding. However, this would also tend to overstate the observed agreement. On the other hand, as noted previously, the grouping system we used was not intended specifically for pediatric ED diagnoses and could lead to seemingly related diagnoses being placed in different groups, increasing the apparent disagreement.

Given the recognized problems in obtaining diagnosis data from existing records, how can accuracy be improved? Better, more uniform training of coders, physicians, and research staff may lead to greater agreement. We attempted to standardize data abstraction through training, feedback, and provision of a study manual of operations, but more rigorous training may be beneficial. Agreement may also be improved by having more highly trained clinical personnel performing chart abstraction, although our study was not designed to test this hypothesis. However, we believe our results are applicable to the common practice of using nonclinical research staff to perform chart abstraction. These results suggest that such practices warrant caution. Finally, development of a clinically sensible system of grouping diagnoses in a way that minimizes unimportant variation

and that is specifically applicable to pediatric emergency visits may be useful. Motivated in part by our findings, members of this study team are now working to develop such a grouping system.

CONCLUSIONS

We conclude that there is only moderate agreement between diagnosis codes from electronic administrative data sources and diagnoses abstracted from paper medical records for pediatric ED visits. Those using administrative data for studying disease and injury trends or evaluating outcomes and quality of care must be aware of the potential limitations of these data. Administrative data may not always accurately reflect patient chart information, and coded diagnoses may differ between data sources. When possible, steps should be taken to evaluate the quality of administrative data to minimize errors in diagnosis categorization.

References

1. 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.
2. Gorelick MH, Baker MD. Epiglottitis in children 1979-1992: effects of Haemophilus influenzae type b immunization. Arch Pediatr Adolesc Med. 1994; 148: 47–50.
3. Alessandrini EA, Lavelle JM, Grenfell SM, Jacobstein CR, Shaw KN. Return visits to a pediatric emergency department. Pediatr Emerg Care. 2004; 20:166–71.
4. 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.
5. 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.
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. Demlo LK, Campbell PM, Brown SS. Reliability of information abstracted from patients' medical records. Med Care. 1978; 16:995–1005.
9. Nagourney JT, Brown DFM, Chae C, et al. Disagreement between formal and medical record criteria for the diagnosis of acute coronary syndrome. Acad Emerg Med. 2005; 12:446–52.
10. O'Malley KJ, Cook KF, Price MD, Wildes KR, Hurdle JF, Ashton CM. Measuring diagnoses: ICD code accuracy. Health Serv Res. 2005; 40:1620–39.
11. Pediatric Emergency Care Research Network. The Pediatric Emergency Care Research Network
(PECARN): rationale, development, and first steps. Acad Emerg Med. 2003; 10:661–8.
12. Dayan P, Chamberlain J, Dean JM, Maio R, Kuppermann N. The Pediatric Emergency Care Applied Research Network: progress and update. Clin Pediatr Emerg Med. 2006; 7:128–35.
13. Alpern ER, Stanley RM, Gorelick MH, et al., for the PECARN. Epidemiology of a Pediatric Emergency Medicine Research Network: The Pediatric Emergency Care Applied Research core data project. Pediatr Emerg Care. 2006; 22:689–99.
14. Clinical Classifications Software (ICD-9-CM). Agency for Health Care Policy and Research. Available at: http://www.hcup-us.ahrq.gov/toolssoftware/ccs/ccs.jsp. Accessed Jun 12, 2006.
15. Clark DE. Practical introduction to record linkage for injury research. Inj Prev. 2004; 10:186–91.
16. Jaro MA. Probabilistic linkage of large public health data files. Stat Med. 1995; 14:491–8.
17. Cook LJ, Olson LM, Dean JM. Probabilistic record linkage: relationships between file sizes, identifiers and match weights. Methods Inf Med. 2001; 40: 196–203.
18. Gorelick MH, Alpern ER, Singh T, et al., for the Pediatric Emergency Care Applied Research Network (PECARN). Availability of pediatric emergency visit data from existing data sources. Acad Emerg Med. 2005; 12:1195–200.
19. Studney DR, Hakstian AR. A comparison of medical record with billing diagnostic information associated with ambulatory medical care. Am J Public Health. 1981; 71:145–9.
20. Hsia DC, Krushat WM, Fagan AB, et al. Accuracy of diagnostic coding for Medicare patients under the prospective-payment system. N Engl J Med. 1988; 318:352–5.
21. Lorence D, Ibrahim I. Disparity in coding concordance: do physicians and coders agree? J Health Care Financ. 2003; 29:43–53.
22. Bazarian JJ, Veazie P, Mookerjee S, Lerner EB. Accuracy of mild traumatic brain injury case ascertainment using ICD-9 codes. Acad Emerg Med. 2006; 13:31–8.
23. Romano PS, Chan BK, Schembri ME, Rainwater JA. Can administrative data be used to compare postoperative complication rates across hospitals? Med Care. 2002; 40:847–50.
24. Lloyd SS, Rissing JP. Physician and coding errors in patient records. JAMA. 1985; 254:1330–6.
25. Humphries KH, Rankin JM, Carere RG, et al. Comorbidity data in outcomes research: are clinical data derived from administrative databases a reliable alternative to chart review? J Clin Epidemiol. 2000; 53:343–9.
26. Fisher ES, Whaley FS, Krushat WM, et al. The accuracy of Medicare's hospital claims data: progress has been made, but problems remain. Am J Public Health. 1992; 82:243–8.
27. Sharma V, Simon SD, Bakewell JM, Ellerbeck EF, Fox MH, Wallace DD. Factors influencing infant visits to emergency departments. Pediatrics. 2000; 106:1031–9.

APPENDIX A

Participating centers and site investigators are listed 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 (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 of Michigan (D. Treloar), University of Rochester, University of Utah/Primary Children's Medical Center (H. Corneli), Upstate Medical Center (J. Callahan), 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. Goepp, 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, A. Walker.

Maternal and Child Health Bureau/Emergency Medical Services for Children liaisons: D. Kavanaugh, H. Park.

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

Data Analysis and Management Subcommittee: J. Chamberlain, Chair; M. Brown, H. Corneli, J. Goepp, R. Holubkov, P. Mahajan, K. Melville, E. Stremski, 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, 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, D. Treloar.

Quality Assurance Subcommittee: R. Stanley, Chair; D. Alexander, J, Brown, M. Gerardi, M. Gregor, R. Holubkov, K. Lillis, R. Ruddy, M. Shults, 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, J. Wright.