



Published in final edited form as:

*Acad Emerg Med.* 2009 November ; 16(11): 1103–1109. doi:10.1111/j.1553-2712.2009.00554.x.

## Research Using Emergency Department–related Data Sets: Current Status and Future Directions

**Jon Mark Hirshon, MD, MPH, Margaret Warner, PhD, Charlene B. Irvin, MD, Richard W. Niska, MD, MPH, Daniel A. Andersen, PhD, Gordon S. Smith, MB.ChB., MPH, and Linda F. McCaig, MPH**

Department of Emergency Medicine (JMH), the Department of Epidemiology (JMH, DAA, GSS), and the Charles McC. Mathias Jr. National Study Center for Trauma and EMS (JMH, DAA, GSS), University of Maryland School of Medicine, Baltimore, MD; the Centers for Disease Control and Prevention, National Center for Health Statistics (MW, RWN, LFM), Hyattsville, MD; and the Department of Emergency Medicine, St. John Hospital and Medical Center (CI), Detroit, MI

### Abstract

The 2009 *Academic Emergency Medicine* consensus conference focused on “Public Health in the ED: Surveillance, Screening and Intervention.” One conference breakout session discussed the significant research value of health-related data sets. This article represents the proceedings from that session, primarily focusing on emergency department (ED)-related data sets and includes examples of the use of a data set based on ED visits for research purposes. It discusses types of ED-related data sets available, highlights barriers to research use of ED-related data sets, and notes limitations of these data sets. The paper highlights future directions and challenges to using these important sources of data for research, including identification of five main needs related to enhancing the use of ED-related data sets. These are 1) electronic linkage of initial and follow-up ED visits and linkage of information about ED visits to other outcomes, including costs of care, while maintaining deidentification of the data; 2) timely data access with minimal barriers; 3) complete data collection for clinically relevant and/or historical data elements, such as the external cause-of-injury code; 4) easy access to data that can be parsed into smaller jurisdictions (such as states) for policy and/or research purposes, while maintaining confidentiality; and 5) linkages between health survey data and health claims data. ED-related data sets contain much data collected directly from health care facilities, individual patient records, and multiple other sources that have significant potential impact for studying and improving the health of individuals and the population.

### Keywords

emergency departments; research; data sets; health services

---

© 2009 by the Society for Academic Emergency Medicine

Address for correspondence and reprints: Jon Mark Hirshon, MD, MPH; [jhirs001@umaryland.edu](mailto:jhirs001@umaryland.edu).

This work is the output from a consensus workshop conducted during the May 2009 *Academic Emergency Medicine* consensus conference in New Orleans, LA: “Public Health In the ED: Surveillance, Screening, and Intervention.”

Consensus conference session participants included Srikar Adhikari, Sabina Braithwaite, Kerry Broderick, Chris Buresh, John Finnell, William G. Fernandez, Toni Gross, Michael Handrigan, Nancy Holson, Jeffrey Hom, Yu-Hsiang Hsieh, Chris Kahn, Nancy Kerr, Patrick Chow-In Ko, Damon Kuehl, Rika Maeshiro, Priya Mammen, Mary Pat McKay, Douglas McLachlan, Lawrence Melniker, Ward Myers, Toby Nagurney, David J. Osban, Michael Radeos, Jennifer Setlik, Matthew Scholer, Ashley Sullivan, Carolyn Synovitz, and Adrian Tyndall.

As highlighted by the Institute of Medicine, understanding and strengthening the relationship between public health and medicine is vitally important to better address the health needs of the public.<sup>1,2</sup> Evaluating, measuring, and monitoring the health of populations are primary objectives of public health research, in comparison to the traditional biomedical model that focuses on the evaluation and treatment of diseases in individuals.<sup>3,4</sup> Emergency departments (EDs) play a vital role in our health care system, in part due to their key location at the interface between the populace and health care. They are well positioned to collect data concerning important aspects of public health, medical, and social problems. In many ways, EDs can be viewed as windows on the general health of their communities, allowing for better understanding of the health needs of the public and reasons for health care resource utilization.<sup>5,6</sup> Constantly improving information technology has dramatically increased the availability of health-related information. The data collected from EDs and related systems are vital to the understanding that emergency physicians, researchers, and other health care professionals have concerning patterns of health resource utilization, the magnitude of various public health problems, and the identification of new and emerging health threats. Focused research activities using these data will help to guide more thoughtful and efficacious public health and medical interventions.

The 2009 *Academic Emergency Medicine* consensus conference focused on “Public Health in the ED: Surveillance, Screening and Intervention.” One session of this conference discussed the significant value, from a research perspective, of health-related data sets. This article represents the conference proceedings from that session, primarily focusing on ED-related data sets, and includes an example of the use of a data set, based on ED visits, for research purposes.

## WHAT ARE HEALTH- AND ED-RELATED DATA SETS?

Research using data sets that provide information related to health or the utilization of health services, such as ED visits, is well known to epidemiologists and other public health professionals. This universe of health-related data sets is generally used by researchers to answer scientific and public health questions concerning populations. Health-related data sets are the main tool of many epidemiologists and health services researchers, and the breadth of available data sets is vast and growing. Health statistics and information on many health-related data sets are available at <http://www.fedstats.gov>. The recent National Report Card on the State of Emergency Medicine developed by the American College of Emergency Physicians used a broad range of health-related data sets to develop the 116 objective measures used in the analysis of the current emergency care environment.<sup>7</sup> While there are many health-related data sets that are relevant for public health research and may have implications for emergency medicine, there are a smaller number of data sets that are focused on EDs or include specific questions relevant to the ED, which is the primary focus of this paper.

Health-related data sets that include information on ED use generally include traditional health services encounter data that are collected during the visit. Data sets generated from ED visits are frequently collected for administrative or billing purposes, such as administrative Medicare data collected by the Centers for Medicare and Medicaid Services. Alternatively, data on use of the ED may be collected specifically for research and planning purposes, such as the National Hospital Ambulatory Medical Care Survey (NHAMCS; <http://www.cdc.gov/nchs/ahcd.htm>). In addition, some general purpose health surveys, such as the National Health Interview Survey (<http://www.cdc.gov/nchs/nhis.htm>), collect some information on the ED use of survey respondents.

Many of the large-scale ED-related data sets discussed in this article are sponsored by federal and state governments. However, some private organizations have data useful for emergency care research. For instance, the American College of Surgeons collects trauma registry data from participating trauma centers and produces the National Trauma Data Bank (NTDB; <http://www.facs.org/trauma/ntdb/index.html>). Recently, the American College of Surgeons and the Centers for Disease Control and Prevention (CDC) supported the development of the National Sample Program (<http://www.facs.org/trauma/ntdb/nsp.html>), which is a national probability sample of 100 Level 1 and 2 trauma centers from the NTDB.

Syndromic surveillance systems are another example of data sets collected from ED visits. Syndromic surveillance systems have been developed to provide rapid electronic collection of health data. Most states have syndromic surveillance systems that utilize ED data.<sup>8–10</sup> Syndromic surveillance systems typically consist of data collected on chief complaints of patients, combined with information on other acute health care activities such as medication purchases in a specific population, to offer real-time or near-real-time detection of outbreaks of infectious or bioterrorism-related diseases.<sup>8</sup>

## TYPES OF DATA SETS

ED-related data sets can be categorized by their data collection methods. The primary collection methodologies are 1) complete enumeration, 2) population-based and provider-based sample surveys, 3) non–population-based registries, and 4) linked data. Examples of ED-related data sets based on different collection methodologies are listed in Table 1. Each data set has strengths and limitations, in part based on data collection methodology. The discussion of these details is beyond the scope of this paper.

## IMPORTANCE OF DATA SETS

Health-related data offer researchers a rich opportunity to conduct epidemiologic studies, defined as “how disease is distributed in populations and the factors that influence or determine this distribution.”<sup>11</sup> Data sets generated from ED visits can be used to study the distribution of diseases and injuries treated in the ED. The data have been used to study health threats requiring immediate attention such as injuries from car crashes and to conduct surveillance on emerging public health threats such as influenza. ED data have also been used to investigate disparities in the utilization and provision of services among subpopulations. For instance, ED data have been used to investigate ethnic disparities in the initial management of trauma patients.<sup>12</sup> Additionally, ED data can be linked to data from other sources such as inpatient data to answer important health and health care utilization questions.

Because EDs function as a critical societal health care and social safety net, data sets created from ED clinical and administrative records may hold great promise for understanding treatment patterns for the underserved, including those without insurance and other special subpopulations within the United States. An example of this is recent research based on data from NHAMCS that indicates that African Americans received fewer prescriptions for pain when seen in EDs compared to whites.<sup>13</sup> As another example, studies using trauma registry data indicate that trauma patients are not a cross-section of the general or ED population, but are much more likely to have psychiatric conditions, including substance abuse and dependency disorders.<sup>14</sup> Using this information, researchers and practitioners can provide more effective and efficient interventions by tailoring their programs for the patients in trauma care settings.

Health-related data sets play an important role in public health surveillance. This is especially true for ED-related data sets, because EDs are where the first victims of infectious

disease outbreaks or other medical emergencies (e.g., natural disasters or biologic or chemical attacks) are most likely to present. Utilizing the capabilities of informatics systems to rapidly populate databases to alert front-line medical responders and public health authorities could be crucial in situations requiring immediate action to limit the spread of disease and decrease morbidity and mortality. With the current interest and concern related to potential pandemic influenza, such as novel H1N1, EDs are well positioned to provide near-real-time surveillance of influenza cases. Data collected from multiple data sources, such as ED visits and medication usage patterns from pharmacies, can be used to monitor the health of the population in emergency situations.

## RESEARCH POTENTIAL FOR DATA SETS

Just as the types of ED-related data sets available are multiple and varied, so is their research potential. Research using all types of ED-related data sets can serve both inductive and deductive pursuits through hypothesis generation and hypothesis testing, respectively. The latter use is especially powerful for clinicians or health care workers with research proclivity who notice interesting trends or cases in their clinical practice that may be indicative of a greater public health problem. In particular, data sets related to ED visits can be used to research injury or disease patterns, to identify risk factors in populations, or to tailor interventions for specific health problems. There are numerous additional research applications including, as examples, analysis of costs, quality of care, treatment disparities, and trends in ED utilization or treatment of diseases, injuries, or illnesses.

Many studies have used data to identify populations presenting with particular illnesses or injuries to develop targeted interventions.<sup>15–19</sup> A recent study of ED visits for chronic obstructive pulmonary disease exacerbations used the ED component of the NHAMCS data set and reported that patients in the south United States were less likely to receive systemic corticosteroids.<sup>17</sup> Using NHAMCS data, another study found that sports and recreational activities were the leading external causes of pediatric injury among children visiting EDs.<sup>20</sup>

As noted above, syndromic surveillance systems allow for the timely analysis of ED chief complaint data, frequently on a daily basis, and are currently used for bioterrorism and infectious threat detection. Although these data collection mechanisms may be useful in non-syndromic research, the application has not yet been fully explored.<sup>21</sup>

## COSTS OF ANALYSIS OF DATA SETS FOR RESEARCH

The costs of using ED-related data sets in general varies depending on the source of the data and whether the data are already compiled and collated in a usable format. Establishing health-related data sets based on ED visits can be expensive, but is extremely important considering the approximately 119.2 million ED visits in the United States in 2006.<sup>22</sup> Once established, the cost for ED-related data sets varies depending on the data set utilized. Financial resources required for research based on ED data sets could include the charges associated with obtaining the data, statistical programs for data analysis, and the personnel to analyze the data. Many of the data sets are free (e.g., NHAMCS), while others require a nominal fee (e.g., Healthcare Cost & Utilization Project [HCUP]). However, some may be expensive, such as ED-related data sets purchased from private sector vendors. Funds may be required to analyze confidential data at a research data center.

Without the use of existing ED-related data sets, many investigators would be limited to research pursuits based on primary data collection. Even for investigators who would be able to get funding, the time frame between protocol development and the publication of results can be exceedingly long. Existing health-related data sets can provide preliminary

data for grant development or for initial exploration of health-related questions at a significantly reduced cost relative to the cost of primary data collection.

## BARRIERS TO USE OF DATA SETS

There are numerous reasons for not using currently available ED-related data sets. These include, but are not limited to, concerns regarding the completeness, accuracy, and timeliness of the data, along with the ability to generalize (if the data set is a convenience sample). As noted above, many data sets require statistical analysis packages to analyze the data, which may be a financial and educational barrier for some individuals. However, online data analysis tools are increasingly available, therefore negating the need for sophisticated data analysis tools. The Web-based Injury Statistics Query and Reporting System (<http://webappa.cdc.gov/sasweb/ncipc/mortrate.html>) is an online statistical analysis tool to analyze fatal injury data from the National Vital Statistical System and nonfatal injury data from the National Electronic Injury Surveillance System–All Injury Program. Additionally, HCUPnet (<http://hcupnet.ahrq.gov/>) provides a free online statistical analysis tool for queries of all hospital stays, including mental health hospitalizations, and, for some states, of all ED visits.

Institutional review board (IRB) approval is required for any research project involving human subjects, according to the Code of Federal Regulations (Part 46, Protection of Human Subjects, 46.101).<sup>23</sup> When health-related data sets contain publicly available deidentified data, IRB approval is generally less problematic. However, as the requested data become more specific and the potential for identification of an individual increases, the scrutiny by an IRB will also increase.

## CURRENT RESEARCH WITH AN ED-RELATED DATA SET EXAMPLE: NATIONAL HOSPITAL AMBULATORY MEDICAL CARE SURVEY

NHAMCS, conducted annually by the CDC's National Center for Health Statistics (NCHS), is an example of a nationally representative health services data set. Data are collected on hospital ED and outpatient department visits, as well as hospital and ED characteristics. Yearly reports are published for both types of ambulatory care facilities.<sup>22</sup> NHAMCS uses a complex multistage sampling strategy, beginning with 112 primary sampling units (PSUs), which generally correspond to counties or other similar jurisdictions. Within each PSU, a sample of nonfederal short-stay general hospitals is stratified into four classes based on whether the hospital has an ED only, an outpatient department only, both, or neither. Finally, for the ED component, a sample of visits is selected from each emergency service area over a random 4-week reporting period. Each hospital, ED, and patient visit is weighted according to its inverse probability of having been selected. The somewhat complex weighting process includes an adjustment for nonresponse and is important for the results to be nationally representative. Hospital-, ED-, or patient visit-level research can be conducted using NHAMCS. An example of a hospital-level study is a national report on bioterrorism response planning.<sup>24</sup>

Data are made available to researchers through public use files released yearly through the NCHS Research Data Center (<http://www.cdc.gov/nchs/r&d/rdc.htm>). The survey instrument with variables collected is available online ([http://www.cdc.gov/nchs/data/ahcd/nha-mcs100ed\\_2009.pdf](http://www.cdc.gov/nchs/data/ahcd/nha-mcs100ed_2009.pdf)). To provide oversight, the survey undergoes yearly review and approval at NCHS through its ethics review board, which serves as the IRB for these surveys. Data files undergo rigorous disclosure review and information is not available that could identify either the hospital or an individual patient.

An analysis using NHAMCS provides an example of the research potential of this data set.<sup>25</sup> The authors were interested in studying the epidemiology of blood culture use in EDs in light of published guidelines, associated factors, and time trends. Because a check box for blood cultures on the patient record form was implemented in 2001, the time period chosen was 2001 through 2004. The frequencies of blood cultures over all ED visits and over visit subsets of more relevant interest were generated. Through statistical analyses, the authors were able to show an increasing national trend in blood culture use and that a large proportion of those cultures were for patients without clinical indicators of bacteremia and for patients who were not admitted to the hospital.

## LIMITATIONS OF DATA SETS

Health- and ED-related data sets may be limited by several factors, examples of which include variable data quality, timeliness, and representativeness. Each data set has its own limitations based on a number of factors, such as the method by which the data are collected, the completeness of the data, and the structure and format of the data set. The limitations of the data must be considered prior to analyzing the data, because the limitations will impact both the accuracy and the reliability of the results.

Some ED data sets rely on administrative data using either medical records or claims data. Data from these sources are frequently limited to basic demographic, diagnosis, and health outcome at the time of discharge. The quality of the data even within an institution can vary depending on operational pressures and may have missing and unknown values. For instance, quality of information on race and ethnicity may be poor, as it may be collected using visual inspection rather than asking the patient or may not be recorded in the chart. Problems with classification can be particularly problematic if there is differential misclassification. For instance, if race is missing for younger ages and not for older persons, analysis based on race will be more accurate for older persons than for younger.

Lack of consistency in coding and classification of diseases and injury may also be a limitation of some health-related data sets. In some instances, this is a reflection of the quality of data in the administrative record, but may also be related to the screening and diagnostic technologies available at a particular facility.

For injury research, external cause-of-injury codes (E-codes), which are used to classify injury incidents by intent (e.g., unintentional, homicide/assault, suicide/self-harm, or undetermined) and mechanism (e.g., motor vehicle, fall, struck by/against, firearm, or poisoning), are incomplete in the majority of state inpatient and ED databases.<sup>26</sup> The E-codes are critical to understanding the causes and patterns of injury, as well as for evaluating prevention programs. A recent CDC report recommended strategies to improve data collection, coding, quality assurance practices, analysis, reporting, and dissemination of E-coded data, which are currently being implemented.<sup>26</sup>

Timeliness is frequently a concern for ED-related data sets regardless of data source. For example, many governmental data sets based on surveys or administrative data are made publicly available several years after the data are collected. The time to collect and process the data, particularly for nationally representative samples, can hinder the ability of researchers to see emerging trends. The cost of creating and maintaining ED-related data sets is frequently quite high. Fortunately, the initial costs should be incurred only once and may be outweighed by increases in patient safety and improved outcomes.

Specification of a denominator for calculating rates may be a problem with ED-related data sets, particularly if they are health care provider-based, because clearly defining catchment areas may be difficult due to variability in usage patterns or due to the fact that the

population of a catchment area utilizing a particular facility or group of facilities may vary from the population at large. In addition, ED-related data sets are commonly based on visits, as opposed to individual patients, which may be of concern because a single patient may make multiple visits. Because the data are deidentified, it is often impossible to identify the first visit for a condition. Population-based and household surveys, as well as Medicare data, do not have this limitation. Additionally, some data sets are convenience samples, which will pose a threat to external validity, especially when attempting to extrapolate the data nationally. This includes the National Trauma Data Bank, which is currently the largest trauma registry data set available. However, while these injury data would produce inaccurate national estimates, they may be valuable for planning, implementing, and evaluating the effectiveness of injury prevention programs.<sup>26</sup>

## FUTURE DIRECTIONS AND CHALLENGES

The rapid growth of the electronic medical record opens numerous opportunities for the collection of data from health care providers. First, it may allow for electronic capture of many variables currently manually extracted. This may improve the quality and accuracy of data collection while decreasing the cost. Additionally, the speed of data extraction should improve. However, as the data become more location- and time-specific, there may be additional challenges with maintaining confidentiality.

The move to Internet-based statistical analysis tools, like Web-based Injury Statistics Query & Reporting System (WISQARS) and HCUPnet, may increase the ease of analyzing these databases and therefore increase their use for research. Finally, as electronic data collection evolve, there may be opportunities to link multiple ED visits electronically to see patterns over time, while maintaining data confidentiality.

As an outgrowth of the consensus conference, five main needs related to enhancing the use of ED-related datasets were identified. These were:

1. For applicable data sets, electronic linkage of initial and follow-up ED visits and linkage of information about ED visits to other outcomes, including costs of care, while maintaining deidentification of the data.
2. Timely data access with minimal barriers.
3. Complete data collection for clinically relevant and/or historical data elements, such as the external cause-of-injury codes.
4. Easy access to data that can be parsed into smaller jurisdictions (such as states) for policy and/or research purposes, while maintaining confidentiality.
5. Linkages between health survey data and health claims data.

## RESEARCH QUESTIONS

Research questions related to the above identified needs include:

1. How can the risk of disclosure for individuals included in ED-related data sets be reduced, especially considering that linked data increases the quantity and specificity of information available at the individual level?
2. How will the data be managed and who will be the data stewards once the data are linked, considering that the owners of the data that are to be linked are often different entities, sometimes even from public and private sources?

3. How can the appropriate balance between increased timeliness of data and data quality be determined, since processing raw data and running error checking can be time-consuming, but leads to better data quality?
4. How can the recording of pertinent data elements, such as mechanism of injury, be improved in the clinical record, considering the current pressures that health care providers face?
5. How can methods be developed to parse national data into smaller areas, such as states, that would allow for stable estimates, while maintaining confidentiality?

## SUMMARY

ED-related data sets contain much data collected directly from health care facilities, individual patient records, and multiple other sources that have significant potential impact for studying and improving the health of individuals and the population. These data sources have been used for multiple epidemiologic, health services utilization, and other research studies and have significant potential for use by emergency medicine researchers in the future.

## Acknowledgments

Dr. Hirshon was supported by National Heart, Lung, and Blood Institute Grant 5K08HL073849 and Dr. Andersen was supported by National Institute of General Medical Science Grant T32GM075767.

## References

1. Committee for the Study of the Future of Public Health; Division of Health Care Services. The Future of Public Health. Washington, DC: Institute of Medicine; 1988.
2. Committee on Assuring the Health of the Public in the 21st Century. The Future of the Public's Health in the 21st Century. Washington DC: Institute of Medicine; 2002.
3. Lasker, RD. Medicine and Public Health: the Power of Collaboration. New York, NY: Academy of Medicine; 1997.
4. Rhodes KV, Pollock DA. The future of emergency medicine public health research. *Emerg Med Clin North Am.* 2006; 24:1053–73. [PubMed: 16982352]
5. Hirshon JM. The rationale for developing public health surveillance systems based on emergency department data. *Acad Emerg Med.* 2000; 7:1428–32. [PubMed: 11099436]
6. Varney S, Hirshon JM. Update on public health surveillance in emergency departments. *Emerg Med Clin North Am.* 2006; 24:1035–52. [PubMed: 16982351]
7. Epstein, SK.; Burstein, JL.; Case, RB., et al. The national report card on the state of emergency medicine, 2009 Edition. Dallas, TX: American College of Emergency Physicians; 2008. Available at: <http://www.emreportcard.org/uploadedFiles/ACEP-ReportCard-10-22-08.pdf> [Accessed Aug 9, 2009.]
8. Henning KJ. Overview of syndromic surveillance: what is syndromic surveillance? *MMRW.* 2004; 53(Suppl):5–11.
9. Yuan C, Love S, Wilson M. Syndromic surveillance at hospital emergency departments—southeastern Virginia. *MMWR.* 2004; 53(Suppl):56–58. [PubMed: 15714630]
10. Arboleda N, Fleischauer AT, Sejvar J, et al. An emergency department based syndromic surveillance system for meningitis and encephalitis, Maricopa County, AZ 2004 [abstract]. *Adv Dis Surveil.* 2006; 1:4.
11. Gordis, L. *Epidemiology.* 4. Philadelphia, PA: Elsevier; 2008. p. 3
12. Shafi S, Gentilello LM. Ethnic disparities in initial management of trauma patients in a nationwide sample of emergency department visits. *Arch Surg.* 2008; 143:1057–61. [PubMed: 19015463]



13. Pletcher MJ, Kertesz SG, Kohn MA, Gonzales R. Trends in opioid prescribing by race/ethnicity for patients seeking care in US emergency departments. *JAMA*. 2008; 299:70–8. [PubMed: 18167408]
14. Soderstrom CA, Cole FJ, Porter JM. Injury Control and Violence Prevention Committee. Injury in America: the role of alcohol and other drugs—an EAST position paper prepared by the Injury Control and Violence Prevention Committee. *J Trauma*. 2001; 50:1–12. [PubMed: 11253757]
15. Loder RT. The demographics of equestrian-related injuries in the United States: injury patterns, orthopedic specific injuries, and avenues for injury prevention. *J Trauma*. 2008; 65:447–60. [PubMed: 18695484]
16. Prosser JM, Perrone J, Pines JM. The epidemiology of intentional non-fatal self-harm poisoning in the United States: 2001–2004. *J Med Toxicol*. 2007; 3:20–4. [PubMed: 18072154]
17. Flores AH, Haileyesus T, Greenspan AI. National estimates of outdoor recreational injuries treated in emergency departments, United States, 2004–2005. *Wilderness Environ Med*. 2008; 19:91–8. [PubMed: 18513117]
18. Centers for Disease Control and Prevention. Nonfatal traumatic brain injuries from sports and recreation activities—United States, 2001–2005. *MMWR Morb Mortal Wkly Rep*. 2007; 56:733–7. [PubMed: 17657206]
19. Tsai CL, Sobrino JA, Camargo CA Jr. National study of emergency department visits for acute exacerbation of chronic obstructive pulmonary disease, 1993–2005. *Acad Emerg Med*. 2008; 15:1–9. [PubMed: 18211306]
20. Simon TD, Bublitz C, Hambidge SJ. External causes of pediatric injury-related emergency department visits in the United States. *Acad Emerg Med*. 2004; 11:1042–8. [PubMed: 15466146]
21. Hakenewerth AM, Waller AE, Ising AI, Tintinalli JE. North Carolina Disease Event Tracking and Epidemiologic Collection Tool (NC DETECT) and the National Hospital Ambulatory Medical Care Survey (NHAMCS): comparison of emergency department data. *Acad Emerg Med*. 2009; 16:261–29. [PubMed: 19133850]
22. Pitts, SR.; Niska, RW.; Xu, J.; Burt, CW. National Hospital Ambulatory Medical Care Survey: 2006 Emergency Department Summary. National Health Statistics Reports; No 7. Hyattsville, MD: National Center for Health Statistics; 2008.
23. Code of Federal Regulations. [Accessed Aug 9, 2009.] Part 46, Protection of Human Subjects. Available at: <http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.htm#46.101>
24. Niska, RW.; Burt, CW. Emergency response planning in hospitals, United States: 2003–2004. Advance Data From Vital and Health Statistics; No 391. Hyattsville, MD: National Center for Health Statistics; 2007.
25. McCaig LF, McDonald LC, Cohen AL, Kuehnert MJ. Increasing blood culture use at US hospital emergency department visits, 2001 to 2004. *Ann Emerg Med*. 2007; 50:42–8. [PubMed: 17383772]
26. Centers for Disease Control and Prevention. Strategies to improve external cause-of-injury coding in state-based hospital discharge and emergency department data systems: recommendations of the CDC Workgroup for Improvement of External Cause-of-Injury Coding. *MMWR Recom Rep*. 2008; 57(RR01):1–15.

**Table 1**

## Types of ED-related Research Data Sources

Source Data	Sample/Population	ED and Other Information
<b>Complete enumeration</b>		
Medicare claims and enrollment data <a href="http://www.cms.hhs.gov/home/rsds.asp">http://www.cms.hhs.gov/home/rsds.asp</a>	Medicare claims (bills) for persons with fee-for-service coverage	Medicare eligible 65+ as well as those with end stage renal disease and some disabled
		<ul style="list-style-type: none"> <li>Data on ED visits can be found in the inpatient file (if they result in an admission), outpatient file, or, occasionally, in the (physician) carrier file.</li> </ul>
National Vital Statistics System, Mortality Data <a href="http://www.cdc.gov/nchs/deaths.htm">http://www.cdc.gov/nchs/deaths.htm</a>	State death registration	All deaths in U.S.
		<ul style="list-style-type: none"> <li>Place of death indicates if the death occurred in the ED or outpatient department</li> </ul>
HCUP State Emergency Department Databases <a href="http://www.hcup-us.ahrq.gov/seddooverview.jsp">http://www.hcup-us.ahrq.gov/seddooverview.jsp</a>	ED encounter medical records	All ED encounters in a state
		<ul style="list-style-type: none"> <li>Diagnoses and procedures</li> <li>Charges and expected source of payment</li> </ul>
HCUP State Inpatient Databases <a href="http://www.hcup-us.ahrq.gov/sidoverview.jsp">http://www.hcup-us.ahrq.gov/sidoverview.jsp</a>	Inpatient discharge medical records	All inpatients discharged in a state
		<ul style="list-style-type: none"> <li>Admission through the ED</li> </ul>
<b>Population-based and provider-based sample surveys</b>		
National Health Interview Survey <a href="http://www.cdc.gov/nchs/nhis.htm">http://www.cdc.gov/nchs/nhis.htm</a>	Personal interviews	Persons in civilian households selected from a multistage probability design
		<ul style="list-style-type: none"> <li>Use of medical services including number of ED visits in the past year</li> <li>Demographic characteristics</li> <li>Health insurance status</li> </ul>
Medical Expenditure Panel Survey (MEPS) <a href="http://www.meps.ahrq.gov/mepsweb/">http://www.meps.ahrq.gov/mepsweb/</a>	Five rounds of personal interviews conducted in two years that take place over a two and a half year period	Nationally representative subsample of households that participated in the prior year's National Health Interview Survey
		<ul style="list-style-type: none"> <li>Use of medical services including the ED</li> <li>Costs of care and source of payments</li> <li>Health insurance coverage, income, and employment.</li> </ul>
National Electronic Injury Surveillance System-All Injury Program <a href="http://www.cdc.gov/ncipc/WISQARS/nonfatal/datasources.htm">http://www.cdc.gov/ncipc/WISQARS/nonfatal/datasources.htm</a>	ED encounter medical records	All injury related ED encounters from a national probability sample of hospitals in the U.S.
		<ul style="list-style-type: none"> <li>Cause of injury seen in ED</li> <li>Circumstances of the injury</li> <li>Information on consumer products involved</li> </ul>
National Hospital Ambulatory Medical Care Survey <a href="http://www.cdc.gov/nchs/ahcd.htm">http://www.cdc.gov/nchs/ahcd.htm</a>	Encounter forms completed by physicians or office staff	Sample of visits to EDs and outpatient visits from a national probability sample of hospitals in the U.S.
		<ul style="list-style-type: none"> <li>Characteristics of patient visits</li> <li>Characteristics of treatment</li> <li>Prescribing patterns</li> </ul>
Drug Abuse Warning Network, ED component <a href="http://dawninfo.samhsa.gov/">http://dawninfo.samhsa.gov/</a>	Case report form completed after review of ED medical records	All drug-related visits from a national probability sample of hospital EDs
		<ul style="list-style-type: none"> <li>Information about drug-related ED visits including visits related to illegal drugs of abuse, prescription and over-the-counter medications, dietary supplements, non pharmaceutical inhalants, and alcohol use</li> </ul>
<b>Non-Population Based Registries</b>		

	Source Data	Sample/Population	ED and Other Information
National Trauma Data Bank <a href="http://www.facs.org/trauma/ntdb/index.html">http://www.facs.org/trauma/ntdb/index.html</a>	Encounter data from trauma centers	Convenience sample based upon all visits from 600 U.S. trauma centers submitting data	<ul style="list-style-type: none"> <li>• Characteristics of patients entering trauma care</li> <li>• Abbreviated Injury Scale for trauma patients</li> </ul>
<b>Linked Data</b>			
National Health Interview Survey (NHIS)  Linked Mortality Files <a href="http://www.cdc.gov/nchs/data_access/data_linkage/mortality.htm">http://www.cdc.gov/nchs/data_access/data_linkage/mortality.htm</a>	NHIS survey participants matched with the National Death Index	NHIS survey participants who are eligible for linkage	<ul style="list-style-type: none"> <li>• Data on deaths in the ED combined with detailed demographic characteristics</li> <li>• Data on ED visits in the previous year and follow up on deaths</li> </ul>
NHIS/Medicare Enrollment and Claims Data <a href="http://www.cdc.gov/nchs/data_access/data_linkage/cms.htm">http://www.cdc.gov/nchs/data_access/data_linkage/cms.htm</a>	NHIS survey participants matched with Medicare data	NHIS survey participants who provide information necessary for linkage	Data on ED visits paid for by Medicare can be linked to NHIS patient characteristics