Accuracy of EMS-Recorded Patient Demographic Data

Jane H. Brice, MD, MPH, Kevin D. Friend, MD, Theodore R. Delbridge, MD, MPH

Abstract

Objective. Emergency medical services (EMS) research is frequently dependent on data recorded by prehospital personnel. Linking EMS information with hospital outcome depends on essential identifying data. We sought to determine the accuracy of these data in patients who activated EMS for chest pain and to describe the types of errors committed. **Methods.** We performed a retrospective, consecutive case series study of all prehospital records for patients transported by the City of Pittsburgh Bureau of EMS (annual call volume, 60,000) for chest pain to three area hospitals during a three-month interval. Demographic data, including name, date of birth (DOB), and Social Security number (SSN), for each patient were extracted from the EMS record. These were compared to the definitive information in the hospital records. **Results.** 360 prehospital records were examined, with 341 matches to hospital records. The correct patient name was recorded in 301 records (83.6%), the correct DOB was recorded 284 times (78.9%), and the correct SSN was recorded 120 times (33.3%). The overall error rate of demographic data recorded on EMS records was 73.9% (266/360). If SSN is not included as a demographic variable, then the overall error rate was 25.3% (91/360). **Conclusion.** The use of EMS-generated demographic data demonstrates moderate agreement and linkage with hospital records. Name and DOB are more reliable data elements for matching than SSN. Future research should examine the impact of electronic medical records and EMS identification numbers on data reliability. **Key words:** emergency medical services; demographic data.

PREHOSPITAL EMERGENCY CARE 2008;12:187–191

INTRODUCTION

The EMS Agenda for the Future[1] and the EMS Research Agenda for the Future[2] both cite data linkage as one of the core issues impeding the progress of outcomes research in emergency medical services (EMS) research. Linking the prehospital record to that of the receiving hospital or the medical examiner's office is difficult. While there is a trend toward deployment of electronic prehospital medical records, most EMS systems still use a handwritten paper record.[2] Incomplete information, illegible handwriting, and inaccurate information are often exclusion criteria for EMS research efforts. Of these, electronic records that require EMS personnel to enter all the information will intuitively only eliminate illegible handwriting.

In planning research efforts, it is important to have an *a priori* estimate of the number of prehospital records that will be required to avoid a Type I error in making conclusions. A power calculation will provide the number of completed records necessary to make reliable conclusions. Knowing ahead of time how many records might have to be discarded due to poor linkages between EMS and hospital records would provide additional knowledge to investigators preparing to engage in research.

Many studies have been published regarding the ability to link EMS data to hospital databases. Using probabilistic linkage, these studies have reported linkage rates varying from 14% to 76%, depending on the quality of the source data used.[3–5] As an example, Downing et al. examined data linkages for EMS patients who had been assaulted.[6] However, few of them investigated the reasons that data could not be linked. In our experiences, EMS administrators face similar issues every day when they attempt to generate bills for services provided.

In the present study, we sought to match demographic information gathered in the prehospital phase of care with similar information gathered upon registration in the Emergency Department to 1) determine the rate of linkage of data and 2) examine the accuracy and types of errors made by EMS providers.

MATERIALS AND METHODS

We retrospectively selected the prehospital records for all patients of the City of Pittsburgh Bureau of EMS who activated 911 and were subsequently transported for the chief complaint of chest pain during a three-month period (January 1, 1996 through March 31, 1996). We chose chest pain for the following reasons: 1) it is a common chief complaint, 2) the majority of chest pain patients are transported to a hospital, and 3) chest pain is a symptom paramedics are generally comfortable in

Received May 1, 2007, from the Department of Emergency Medicine, School of Medicine, The University of North Carolina at Chapel Hill, Chapel Hill, North Carolina (JHB), the Department of Emergency Medicine, School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania (KDF), and the Department of Emergency Medicine, School of Medicine, East Carolina University, Greenville, North Carolina (TRD). Revision received September 19, 2007; accepted for publication September 23, 2007.

Address correspondence and reprint requests to: Jane H. Brice, MD, MPH, Department of Emergency Medicine, School of Medicine, The University of North Carolina at Chapel Hill, CB# 7594, Chapel Hill, NC 27599-7594. e-mail: brice@med.unc.edu.

Presented as a poster at the National Association of EMS Physicians 1998 Mid-Year Meeting, July 8–19, 1998, Incline Village, Nevada.

doi: 10.1080/10903120801907687

treating and is less likely to cause errors related to poor organization or speed.

We then grouped the records by receiving hospital and selected the three individual hospitals with the largest frequencies. There were 27 receiving hospitals represented in the data set. Inclusion criteria, therefore, were those patients who called Pittsburgh 911 for chest pain during the three-month study period; were transported by Pittsburgh EMS to University of Pittsburgh Medical Center, Mercy Hospital of Pittsburgh, or West Penn Hospital; and had an EMS record for review. Exclusion criteria consisted of those patients for whom no hospital destination was recorded.

The Pittsburgh Bureau of EMS is municipal third service, and responds to all 911 calls for EMS in the City of Pittsburgh, which has a population of approximately 370,000 contained within 55 square miles. The Bureau staffs 13 ambulances with two-paramedic teams and responds to approximately 60,000 calls for assistance each year. Approximately 8% of calls are for the chief complaint of chest pain.

After approval from the Institutional Review Boards of the University of Pittsburgh Medical Center, Mercy Hospital of Pittsburgh, and West Penn Hospital, demographic data, including name, date of birth (DOB), and Social Security number (SSN), were extracted from the prehospital records of each patient by one of the authors (JB). Paramedics obtained EMS demographic data through patient or family member interview at the scene or enroute to the hospital. EMS-collected data were compared to the gold standard of demographic information contained in the electronic hospital records for that visit by two of the authors (KF and JB). Emergency Department registration personnel obtained hospital demographic data from interview as well as from documents, such as driver licenses and insurance cards.

When searching hospital records, we held date of transport and receiving hospitals as constants. For example, we might have been looking for someone transported by EMS on March 21 to Mercy Hospital. We searched first within 30 minutes on either side of the EMS stated time of hospital arrival, then we searched within 30 minutes of the 12-hour reciprocal of the EMS stated time of hospital arrival. For instance, if the EMS record recorded the time of hospital arrival as 9 o'clock in the morning, we also search around the time of 9 o'clock in the evening. Last, we searched the entire log of patients seen at the hospital for that date. We first searched by last name. If there was no match, we then searched by first name. Next, we searched by middle name if available on the EMS record. Following that, we searched by DOB and then by SSN. Failing to match any of these data elements resulted in the category of "no match."

Once the EMS and hospital records were matched, we then assessed for the accuracy of each data element. For each record, we documented the number and type of errors. The number of documented errors total more than the total number of errors, as some records contained more than one error. We entered data into Microsoft Excel (Microsoft Corporation, Redmond, WA) and accuracy was calculated for each data point collected: name, DOB, and SSN. We calculated the accuracy for each data element as the number of matched EMS records divided by the total number of EMS records. We also calculated the overall error rate as the number of EMS records with any error divided by the total number of EMS records.

RESULTS

During the three-month period, Pittsburgh EMS transported 1140 patients whose chief complaint was chest pain. Of the 1140 records, one did not record a destination hospital and was excluded. Three hundred and sixty patients were transported to the three selected hospitals. One hundred forty-seven were transported to Mercy Hospital of Pittsburgh, 132 were transported to the University of Pittsburgh Medical Center, and 81 were transported to West Penn Hospital. Nineteen records could not be matched for the following reasons: failure to record a date (1), no patient visit found for the recorded date (7), illegible handwriting (6), and no patient by that name known at that hospital (5). Records for which there was no patient by that last name known at that hospital most likely represented a failure of the paramedic to record the correct hospital destination. Three hundred forty-one (94.7% 95%CI 91.9–96.8%) hospital matches for any data element on the day of patient transport were found.

For the 341 records for which a match could be obtained, EMS recorded the correct patient name in 88.2% (301/341) (95%CI 84.4–91.5%) of cases. Errors were use of a diminutive (example Bill instead of William) of the first name (16), misspelled last name (10), use of a middle name in place of a first name (8), misspelled first name (4), failure to record a patient a junior or senior suffix to the last name (4), and failure to record a first name (1). Three records had more than one error.

Correct date of birth was recorded 83.3% (284/341) (95%CI 78.9–87.1%) of the time. Errors included failure to record a date of birth (20), failure to record a complete date of birth (8), illegible handwriting (4) incorrect month (12), incorrect day (6), and incorrect year (12). Five records contained more than one error.

SSN was correct in 35.2% (120/341) (95%CI 30.1–40.5%) of cases. Documented errors were failure to record a SSN (219) and incorrect number (2). In the case of the two records with incorrect numbers, EMS personnel transcribed not one but two digits incorrectly in both cases. Numbers were not transposed. No record contained more than one error.

The overall error rate for the 360 records was 73.9% (266/360) (95%CI 69.0–78.3%); 19 with no match, 219 with no SSN, and 28 records with errors of name and/or DOB not accounted for in the Social Security category. Only 94 records correctly contained every data element of our study. The majority of the error rate is accounted for in the failure to record a SSN. If SSN had not been included as a data element the overall error rate would have been 25.3% (91/360) (95%CI 20.9–30.1%); 19 with no match and 72 records with errors of name and/or DOB.

DISCUSSION

The volume of research dedicated to the prehospital environment has grown steadily,[7,8] and many articles each year are based on data recorded by EMS personnel. Efforts to correlate the EMS information with hospital outcome depend on accurate patient demographic data, such that prehospital and hospital records may be linked. Our study demonstrated moderate accuracy of EMS-recorded demographic data. Nearly 95% of EMS generated records could be matched to a hospital visit. The overall error rate was 73.9%, including SSN as a data element and 25.3% when SSN was not included as a data element. The patient name element proved to be the most reliable followed closely by date of birth. SSN was extremely unreliable as it was infrequently recorded. Fortunately most hospitals and EMS systems are moving away from using SSN as a patient identifier.

Several studies in the literature have examined the usefulness of hospital-recorded demographic data.[9–11] They have looked at the ability to contact patients discharged from Emergency Departments by patient-provided and registration personnel-recorded telephone numbers and most studies have reported error rates consistent with our data. Adams et al. reported an overall error rate of 33% in that 21% of the provided information was for nonworking numbers and another 12% were for incorrect residences.[9] Similarly, 42% of the telephone numbers in a study of asthma follow-up were attributed to disconnected or wrong numbers.[10] In another study, investigators were not able to contact 58% of patients via telephone in follow-up, mostly due to 28% of provided numbers being inaccurate.[11]

In looking specifically at EMS generated reports, Grant et al. demonstrated the variable accuracy of different data sources on motor vehicle accidents.[12] Compared to a Crash Investigation Report gold standard, the ambulance report was 19.3% inaccurate in describing various crash characteristics. The authors speculate that accuracy was poor perhaps, in part, because prehospital providers focus on patient care rather than on crash characteristics. Crashes producing more critical patients were more likely to be inaccurately described by EMS. Another study, by Hunt et al., determined that ambulance records inadequately documented vehicle damage in motor vehicle crashes.[13] Yoon et al. evaluated the data accuracy from various sources responsible for making up the Coverdale Stroke Registries. They found that EMS reports demonstrated an overall incompleteness of stroke data elements of 35.4% and inaccuracy when compared to other data sources of 27.9%.[14] Cone et al. examined the adequacy of EMS documentation for patients refusing EMS care. Of 81 records, they found errors of documentation in 25%.[15]

Downing et al. has conducted a study similar to ours.[6] Taking ambulance call reports for assault in the West Midlands area of the United Kingdom, they linked EMS records probabilistically with hospital records, using DOB, sex, and arrival date/time as the essential date elements on which to match. Of 5384 EMS records, 14.2% (766/5384) were incomplete to the point that they could not be used for matching. Of the 4618 EMS records for which a match was attempted, 84.2% (3889/4618) were eventually linked to hospital records. In their study, Downing et al. were unable to identify those data elements that led to a match failure. These authors recommend a unique identifier that would be used jointly by both EMS and hospital information management systems to link EMS and hospital records.

Probabilistic linkage such as used by Downing et al.[6] has been utilized effectively to link prehospital records with other essential databases, such as death records or hospital admissions.[3–5,17–20] Knight et al. reported a probabilistic linkage of EMS records for refusal of care obtained from the Utah EMS database with death records and hospital records obtained from other Utah sources.[17] They examined the rate of persons refusing EMS transport with subsequent EMS dispatch, Emergency Department visit, hospital admission, or death occurrences. Using 14,109 EMS records as a starting point, they achieved linkage for 465 EMS dispatches, 2790 Emergency Department visits, 174 hospital admissions, and 25 deaths. Had we been examining four large database,s such as managed in the Knight et al. study, it would have been impossible to accomplish this manually.

There are several methods for linking database to one another. According to definition provided by Clark,[21] probabilistic linkage is computer matching based on the probability that one record matches another using a set of common data elements, such as sex, age, or date of service. Deterministic linkage utilizes a common identifying number across records to match one to another and clerical matching uses human judgment about which record matches to another. Clerical matching, such as performed in this study, is considered to be more accurate but is impractical for large databases, as it is being labor as well as time intensive. Probabilistic linkage appears to be an effective method for linking large databases, particularly those that contain only deidentified data.[3–5,16–25]

We believe that several factors contributed to the moderate rate of agreement between our EMS and hospital records. Our study is the only one we are aware of which examines the accuracy of patient-identifying demographic data provided by EMS and describes the reasons for difficulty in matching records. One may expect that patients in distress, such as those with chest pain, may not be able to reliably give basic data. In addition, the paramedics may have been appropriately focused on patient care rather than on collecting demographic data. We did not examine whether the accuracy of data correlated with the severity of the patients' chest pain.

Grant et al.'s study demonstrated that paramedics failed to record essential data and this failure decreased the overall rate of accuracy.[12] We found a similar problem, in that missing prehospital data (especially SSN) frequently decreased the rate of agreement between prehospital and hospital records. In addition to reducing linkage of EMS records to patient records for research purposes, such inaccuracy may have a direct effect on patient care. Emergency Departments have access to patient records from previous admissions and Emergency Department visits and these old records frequently provide valuable information useful in the care of the acutely ill patient. Inability to link to previous hospital records may result in medical errors and reduced capacity to provide quality care for patients.

As EMS systems move toward electronic medical records, it is possible that some of the errors we documented may be eliminated. Those errors of handwriting or failure to record a data element can be remedied by use of an electronic medical record that can force the documentation of data elements. Misspellings and transposition of numbers will not be eliminated by use of an electronic record. Addition of a unique identifier linking EMS and hospital records is being attempted in some systems in an attempt to be able to use deterministic linkage. The success of these linkages has not yet been reported in the literature.

Our study may be limited in its external validity in that we utilized only a single EMS system and limited our data collection to three hospitals. We worked within a large urban system so as to generate enough records for analysis, but other systems may be different or have varying rates of paramedic compliance with recordkeeping. We held the date of service and the hospital destination as constants in our attempt to match records. This may have been a false assumption. It is possible that the date of service or hospital destination were not correctly recorded. This may have resulted in the nearly 6% of records that could not be matched with a hospital visit. Additionally, we held the hospital data as the gold standard, but it is possible that the EMS data was accurate and the hospital data was wrong. We did not attempt to verify information with the true gold standard, the patient. It is also possible that data inaccuracies were attributable to intentionally inaccurate information provided by the patient or to inaccurate data entry on the part of the authors. We did perform duplicate data entry for 10% of our records and found no inconsistencies. This does not, however, mean that there might not have been inaccuracies in the remaining 90% of records.

CONCLUSION

The use of EMS generated demographic data demonstrates moderate agreement and linkage with hospital records. When planning EMS research, investigators should plan to lose approximately 5% of their EMS records due to inability to match with hospital records. Name and DOB are more reliable data elements for matching than SSN. Future research should examine the impact of electronic medical records and EMS identification numbers on data reliability.

References

1. Delbridge TR, Bailey B, Chew JL, Conn AK, Krakeel JJ, Manz D, Miller DR, O'Malley PJ, Ryan SD, Spaite DW, Stewart RD, Suter RE, Wilson EM. EMS agenda for the future: where we are. . . where we want to be. Prehosp Emerg Care 1998;2(1):1–12.
2. Sayre MR, White LJ, Brown LH, McHenry SD; National EMS Agenda Writing Team. National EMS Research Agenda. Prehospital Emergency Care. 6(3 Suppl):S1–43, Jul-Sep 2002.
3. Ferrante AM, Rosman DL, Knuiman MW. The construction of a road injury database. Accid Anal Prev. 1993;25(6):659–665.
4. Dean JM, Vernon DD, Cook L, Nechodom P, Reading J, Suruda A. Probabilistic linkage of computerized ambulance and inpatient hospital discharge records: a potential tool for evaluation of emergency medical services. Ann Emerg Med. 2001;37(6):616–626.
5. Clark DE, Anderson KL, Hahn DR. Evaluating an inclusive trauma system using linked population-based data. J Trauma Sep 2004;57(3):501–509.
6. Downing A, Wilson R, Cooke M. Linkage of ambulance service and accident and Emergency Department data: a study of assault patients in the west midlands region of the UK. Inj Int J Care Inj. 36(6):738–744, 2005.
7. Brice JH, Garrison HG, Evans AT. Study design and outcomes in out-of-hospital emergency medicine research: a ten-year analysis. Prehosp Emerg Care 2000;4(2):144–150.
8. Callaham M. Quantifying the scanty science of prehospital emergency care. Ann Emerg Med. 1997;30(6):785–790.
9. Adams SL, Thompson DA. Inability to follow-up ED patients by telephone: there must be 50 ways to leave your number. Acad Emerg Med. 1996;3(3):271–273.
10. Boudreaux ED, Clark S, Camargo CA. Telephone follow-up after the emergency department visit: experience with acute asthma. Ann Emerg Med. 2000;35(6):555–563.
11. Boudreaux ED, Ary RD, St.John B, Mandry CV. Telephone contact of patients visiting a large, municipal emergency department: can we rely on numbers given during routine registration? J Emerg Med. 2000;18(4):409–415.
12. Grant RJ, Gregor MA, Maio RF, Huang SS. The accuracy of medical records and police reports in determining motor vehicle crash characteristics. Prehosp Emerg Care 1998;2(1):23–28.
13. Hunt RC, Brown RL, Cline KA, et al. Comparison of motor vehicle damage documentation in emergency medical services run reports compared with photographic documentation. Ann Emerg Med. 1993;22(4):651–6.
14. Yoon SS, George MG, Myers S, Lux LJ, Wilson D, Heinrich J, Zheng ZJ. Analysis of data-collection methods for an acute stroke care registry. Am J Prev Med. 2006;31(6S2):S196–S201.

15. Cone DC, Kim DT, Davidson SJ. Patient-initiated refusals of prehsoptail care: ambulance call report documentation, patient outcome, and on-line medical command. Prehosp Disast Med. 1995;10(1):3–9.
16. Dean JM, Vernon DD, Cook L. Probabilistic linkage of computerized ambulance and inpatient hospital discharge records: a potential tool for evaluation of emergency medical services. Ann Emerg Med. 2001;37(6):616–626.
17. Knight S, Olson LM, Cook LJ, Mann NC, Corneli HM, Dean JM. Against all advice: an analysis of out-of-hospital refusals of care. Ann Emerg Med. 2003;42(5):689–696.
18. Newgard CD. Validation of probabilistic linkage to de-identified ambulance records to a state trauma registry. Acad Emerg Med. 2006;13(1):69–75.
19. Gonzalez RP, Cummings G, Mulehar M, Rodning CB. Increased mortality in rural vehicular trauma: identifying contributing factors through data linkage. J Trauma 2006;61(2):404–409.
20. Waien SA. Linking large administrative databases: a method for conducting emergency medical services cohort studies using existing data. Acad Emerg Med. 1997;4(11):1087–1095.
21. Clark DE. Practical introduction to record linkage for injury research. Inj Prev. 2004;10(3):186–191.
22. Cercarelli LR, Rosman DL, Ryan GA. Comparison of accident and emergency with police road injury data. J Trauma 1996;40(5):805–809.
23. Cook LJ, Knight S, Olson LM, Nechodom PJ, Dean JM. Motor vehicle crash characteristics and medical outcomes among older drivers in Utah, 1992–1995. Ann Emerg Med. 2000;35(6):585–591.
24. Smith R, Cook LJ, Olson LM, Reading JC, Dean JM. Trends of behavioral risk factors in motor vehicle crashes in Utah, 1992–1997. Accid Anal Prev. 2004;36(2):249–255.
25. Bull JP, Roberts BJ. Road accident statistics—a comparison of police and hospital information. Accid Anal Prev. 1973;5:45–53.
26. Maas MW, Harris S. Police recording of road accident in-patients. Accid Anal Prev. 1984;16(3):167–184.
27. Rosman DL, Knuiman MW. A comparison of hospital and police road injury data. Accid Anal Prev. 1994;26(2):215–222.