Reprinted From **MEDICAL DECISION MAKING,**
Published by HANLEY & BELFUS, INC., Philadelphia, PA © 1991

Injury Severity:

Better Data through Direct Physician Entry of Anatomic Injuries?

BRAD M. CUSHING, MD, SANDRA D. TEITELBAUM, MLS, WILLIAM BURMAN, MD, DAVID KARGES, MD, WILLIAM BAME, MD

The authors compared the injury diagnoses and Injury Severity Scores (ISSs) generated by three data-collection and -coding methods, and examined the times needed and costs associated with the methods. One method involved direct electronic entry of injury data by a physician in the admitting area. Codes, severity scores, and times and costs varied significantly with the different methods, thus suggesting a need for further study of the derivation of injury severity codes. *Key words:* injury severity codes; data entry. **(Med Decis Making 1991;11 (suppl):S45–S48)**

Outcome from injury is a manifestation of initial injury severity, victim physiologic reserve, and the effectiveness of the clinical care provided. Evaluation of the effectiveness of trauma care and trauma care delivery systems, therefore, requires the accurate measurement of injury severity. The current standard for quantifying injury severity is the Injury Severity Score (ISS).[1] The ISS is used extensively in trauma care research and forms the basis for trauma care quality assurance and evaluation methods.[2,3] The ISS is produced by a process that requires the collection of injury diagnoses and the classification of those diagnoses by the Abbreviated Injury Scale (AIS).[4] The AIS is an anatomic injury classification system that assigns a severity level to each injury category. The ISS is calculated from the AIS according to a formula that uses the AIS severity levels and the body region of the injury. Diagnoses are recorded by the physician either by direct entry to a computer or, more commonly, by notation in a handwritten paper-form medical record. If the paper-form record is used, the diagnoses must be abstracted. Abstraction is usually done by staff who are not physicians and who have variable amounts of medical knowledge. Once the diagnoses are recorded or abstracted, they must be classified by the AIS. AIS coding is done either manually or by computer, directly from the injury description or by conversion from an intermediary ICD-9CM code.[5]

Errors in diagnoses and severity coding can be and are introduced at every step in the process. The extents of inaccuracy introduced by some parts of the coding process, such as inter- and intracoder variation and code conversion, have been quantified,[6,7] but the inaccuracies introduced at other steps have not. The best method to arrive at an accurate severity assessment has not been defined.

As an initial step in the difficult process of defining the optimal method of assessing injury severity, we compared the injury diagnoses and ISSs generated by three methods available at our institution. We also looked at the times and costs associated with the

Received from the National Study Center for Trauma and Emergency Medical Systems, University of Maryland at Baltimore, Baltimore, Maryland 21201.

Address correspondence and reprint requests to Dr. Cushing.

methods. The results of this crude comparison have major ramifications for trauma outcomes research and will be used to guide the design of a better-defined and more comprehensive study.

Methods

The initial injury diagnoses for 24 trauma victims admitted in September 1991 to a level I trauma center were identified and coded by three methods. We collected the AIS severity, ISS, and ICD-9 codes produced by each method as well as the staff times required to generate the codes.

Method 1 ("Med Rec") is the abstraction and coding process of the hospital medical records department. It involves abstraction of the handwritten medical record by medical records personnel after patient discharge, manual ICD-9CM coding, and data entry into a computerized system. For this study, ICD-9 codes were converted to AIS severity and ISS using the conversion program developed by MacKenzie et al.[6] This abstraction, coding, and conversion process is used by many trauma centers and in many investigations that utilize large state data systems. Physician involvement in this process is limited to writing a note in the medical record.

Method 2 ("Dedicated") is employed at The Maryland Institute of Emergency Medical Services Systems (MIEMSS)–Shock Trauma and involves concurrent data abstraction from the handwritten medical record by dedicated trauma data collectors, data verification by the physicians caring for the patients, data entry with automated AIS and ISS severity coding, and manual ICD-9 coding. In this method, the physician is asked to verify the information in his or her note before coding is done.

Method 3 ("MD-Interface") involves the use of a prototype Hypercard data entry and automated coding interface by a single physician present in the admitting area and operating room during the first 24 hours of a victim's admission. The Hypercard interface used for this evaluation was designed and built by one of the authors (Burman). The software employs a point-and-click graphics user interface to depict most injuries. Specific injury information that is difficult to represent graphically is obtained through direct on-screen questioning of the user. Injuries are automatically coded into AIS severity and ICD-9CM codes. An ISS is automatically calculated.

The ISS and ICD-9CM codes generated by each method were compared to assess concordance. Each method produced 18 ISSs, one per patient, but a different number of ICD-9CM codes per patient. We compared the codes from each method with those from each of the other methods. Concordance was calculated as follows. One method was selected as the reference and was compared with each of the other two methods for exact code matches. The number of exact code matches was divided by the number of codes listed by the reference method (the highest number of possible matches). The result was multiplied by 100 to yield the percentage match between the two methods.

Staff time required to generate codes for each method was recorded as follows. The times spent on each task in methods 1 and 2 were recorded on special forms by the persons actually performing the tasks. Some of the method 1 times were estimated by the coding supervisor. Method 2 times were all recorded directly by the coders. The times that physicians spent handwriting the medical record notes used in methods 1 and 2 were not directly recorded but were approximated at 15 minutes per record. This approximation was based on prior observation of the times required for physicians to generate admitting notes. For method 3, the Hypercard interface program automatically recorded interaction time.

Cost was measured for each collection and coding method by multiplying the average time spent generating codes for the study patients by the average wage of the persons involved in each task. The physician costs associated with method 3 are conservative because the interface does not currently complete all the tasks required to generate an admitting note in the medical record. The cost estimates for the nonphysician personnel involved in methods 1 and 2 do not reflect the times required to train and supervise personnel or the indirect costs associated with the employment of staff.

Results

Complete coding data and times by all three methods were available for 18 of the original 24 trauma victims. For six victims, we failed to collect data with either method 1 or method 2. We therefore chose to exclude those patients from the analysis.

The number of injury ICD-9CM codes generated for all 18 patients by the methods were:

Method	*No. of ICD-9CM Codes*
1 (Med Rec)	37
2 (Dedicated)	59
3 (MD-Interface)	43

Minor skin injuries explained most of the variation in the number of diagnoses captured by each method. Both method 1 and method 3 captured nine skin-injury diagnoses, whereas method 2 captured 25. There were 120 unique codes generated in all. Only five diagnostic codes were common to all three methods.

CONCORDANCE

ISS. When the ISS codes generated by method 1 were compared with those of method 2, there were two matches, for a concordance of 11% (2/18). When method 1 codes were compared with method 3 codes, there were also two matches, for a concordance of 11% (2/18). When method 2 codes were compared with method 3 codes, there were three matches, for a concordance of 17% (3/18):

Methods Compared	*ISS Matches—Concordance*	
	No.	%
1 and 2	2	11%
1 and 3	2	11%
2 and 3	3	18%

ICD-9CM. Using the number of method 1 ICD-9CM codes as the reference (37 "total possible matches"), there were 14 method 1–2 matches and seven method 1–3 matches, resulting in concordances of 38% (14/37) and 19% (7/37), respectively. When method 2 was used as reference (59 "total possible matches"), there were 14 method 2–1 matches and nine method 2–3 matches, for concordances of 24% (14/59) and 15% (9/59), respectively. When method 3 was used as reference (43 "total possible matches"), there were six method 3–1 matches and nine method 3–2 matches, producing concordances of 14% (6/43) and 21% (9/43), respectively. Thus, depending on which method is used as the reference, concordance of method 1 to method 2 is between 24% and 38%, concordance of method 1 to method 3 is between 14% and 19%, and for method 2 to method 3, concordance is between 15% and 21%:

Methods Compared	*Concordance*
1 and 2	24%–38%
1 and 3	14%–19%
2 and 3	15%–21%

TIME

Non-physician staff times for methods 1 and 2 included times spent on review of the written patient chart, manual ICD-9CM coding, and entry into a computerized database. For method 2, staff and physician times also included times spent on verification of the injury diagnoses in a weekly team meeting.

When physician time (for writing an admission note on the patient chart) was added to staff times, method 1 required on average a total of 37.1 minutes to produce injury codes per patient, and method 2, with verification, required on average 49.1 minutes per patient. Method 3 required an average of 5.7 minutes per patient.

Average Staff Times Spent per Patient for Injury Coding

Method	*Non-Physician Staff (Min)*	*Physician* (Min)	*Total (Min)*
1	22.1	15	37.1
2	34.2	19.2	49.1
3	—	5.7	5.7

COST

The local average wage for medical records staff responsible for coding and data entry was $12/hour. The figure we used for housestaff physician time was $25/hour.

For method 1, the number of hours spent per patient by non-physician staff was 0.37 hour, for method 2 the time was 0.57 hour. The number of hours of physician time were: method 1, 0.25 hour; method 2, 0.32 hour; and method 3, 0.10 hour.

Average Costs, in Dollars, for Staff Time, per Patient

Method	*Non-Physician Staff*	*Physician*	*Total*
1	$3.08	$6.25	$ 9.33
2	6.84	8.00	14.84
3	—	2.50	2.50

Discussion

The comparison of three data collection and coding methods demonstrated that the injury diagnoses codes and injury severity scores generated by different methods were very different. It also demonstrated wide variations in the times and costs associated with use of different methods.

LACK OF CONCORDANCE

The existence of large method-dependent variations in injury diagnoses and severity codes, if substantiated by further study, has far-reaching ramifications for trauma quality assurance (QA) programs and clinical research. QA programs and research studies routinely stratify study populations by injury severity and physiologic factors and measure the variance in outcome. Variance in outcome is attributed to differences in the therapy rendered. Therapeutic approaches have been accepted as efficacious or rejected on the basis of these analyses. It now seems that much of the variation in outcome could be attributable to the variation in the injury severity of study populations, especially if different methods were used to arrive at severity measures. Approaches to patient care and trauma care delivery systems that were accepted or rejected on the basis of such studies may need to be re-evaluated.

The concordances in this study were very low, in part because the study design demanded that the codes be an exact match. Better concordance would be achieved if near-matches were permitted. It must be understood, however, that the ICD-9 code is already quite nonspecific. For example, 823.9 and 823.8, open vs. closed tibial fractures, vary only by a single digit. A patient with a closed tibial fracture could well have his or her leg casted in the emergency department and be sent home. A patient with an open fracture, however, may well require several operations, a prolonged hospitilization, months of rehabilitation, and possibly a limb amputation. It is clear that small variations in ICD-9CM code can reflect major differences in injury severity and outcome. Thus, the demand of the study for exact code matches in concordance calculations is justified and not a demand for too much precision.

The low concordance of codes generated by the three methods may be explained by specific variations in the processes of data collection and coding employed by the methods. Unfortunately, the study looked only at the end products of the three methods and not at the individual steps in each method. As a result, little can be said about which steps contributed to the variation in end products. It is tempting, and probably correct, to attribute a portion of the variance to the level and type of physician involvement, which does vary substantially between the three methods. A causal relationship, however, cannot be established by this study.

We attempted to compare the accuracies of the diagnoses produced by the methods by making the assumption that the doctor is right and the diagnoses collected by the physician in method 3 represented the "gold standard." In the end, we rejected that assumption. Without a "gold standard" external to the three methods in the study, accuracies could not be compared.

COSTS

The possibility of significant cost savings with direct physician entry and automated coding was clearly demonstrated.

The results have convinced us that further study of methods of deriving injury severity codes is warranted. Among the studies that should be pursued is a more careful evaluation of the effect of direct computer entry of injury diagnoses by physicians. We know it could reduce costs. We need to investigate the effect on the accuracy of injury severity measures by performing a study with an external "gold standard."

The authors thank the people who contributed data for this study: Jan Farrell and the STAPA data coders; Robin Stults, Lori Slencak, and the UMMS Medical Records coders; and Kim Mitchell, MIEMSS trauma registry administrator. Support for the project by the Shock Trauma Associates, P.A. (STAPA) is also gratefully acknowledged.

References

1. Baker SP, O'Neill B, Haddon W, Long WB. The Injury Severity Score: a method for describing patients with multiple injuries and evaluating emergency care. J Trauma. 1974;14:187–96.
2. Committee on Trauma. Resources for optimal care of the injured patient. Chicago: American College of Surgeons, 1990.
3. Porries SE, Gamelli RL, Pilcher DB et al. Practical evaluation of trauma deaths. J Trauma. 1989;29:1607–10.
4. Committee on Injury Scaling. Abbreviated Injury Scale—1990 revision. Des Plaines, IL: Association for the Advancement of Automotive Medicine, 1990.
5. Commission on Professional and Hospital Activities. The international classification of diseases, 9th revision—clinical modification. Ann Arbor, MI, 1980.
6. MacKenzie EJ, Steinwachs DM, Shankar BS. Classifying trauma severity based on hospital discharge diagnoses: validation of an ICD-9CM to AIS-85 conversion table. Med Care. 1989;27:412–22.
7. MacKenzie EJ, Shapiro S, Eastham JN. The Abbreviated Injury Scale and Injury Severity Score: levels of inter- and intrarater reliability. Med Care. 1985;23:823–35.