Quality of external-cause data in the Hospitalization Information System

ABSTRACT

OBJECTIVE: To analyze the coverage and quality of the data on hospitalizations due to external causes in the Hospital Information System of the Brazilian National Health System.

METHODS: Hospitalizations recorded in the Hospital Information System of 11 hospitals in two municipalities (Londrina and Maringá, Southern Brazil), in 2004, were compared with hospitalizations identified as consequences of external causes through investigating medical records, complemented with mortality and pre-hospital data. The crude agreement rate, sensitivity and positive predictive value were calculated. The profile of groups of hospitalization causes recorded in the System was compared with that obtained from investigations.

RESULTS: In Londrina, 3,002 hospitalizations due to external causes were recorded in the System and, in Maringá, 1,403. The investigations found 4,018 and 2,370, respectively. The System presented high positive predictive values for hospitalizations due to external causes in both municipalities: 97.7% in Londrina and 98.6% in Maringá. However, the sensitivity was low: 57.3% in Maringá and 73% in Londrina, thus denoting underrecording of these causes. Comparison between the profiles of the types of causes revealed that there was underestimation of some causes in the System, especially regarding accidents due to exposure to inanimate mechanical forces, assaults and intentional self-harm.

CONCLUSIONS: Underrecording of hospitalizations due to external causes and some distortions regarding the types of causes occurred in the Hospital Information System in both municipalities. Detection of these deficiencies may contribute towards the process of improving the quality of the information in this System.


INTRODUCTION

The use of data routinely gathered by Brazilian information systems, for analyses on the healthcare situation and the impact of interventions, has been increasing over the last few years. The Hospital Information System (HIS) of the Brazilian National Health System (Sistema Único de Saúde, SUS) is the largest national information system and records around 11.5 million hospitalizations per year. Its principal objective is to enable payment for hospitalizations occurring both in public hospitals and in private hospitals that have service...
agreements with SUS. The Authorization for Hospitalization (Autorização de Internação Hospitalar, AIH) is the document that forms each record in the HIS-SUS database.\(^2\)

Despite the limitations inherent to the administrative nature of the HIS-SUS database and the fact that it is not universal (since it only covers hospitalizations paid by SUS), this system presents several advantages. It routinely gathers data in a large number of hospitals; it is made available to the public that has an interest in the system within a short time; it covers approximately 70% of Brazilian hospitalizations;\(^18\) and it provides important epidemiological information that allows large numbers of analyses on the situation of diseases in hospitals and on service evaluation. Moreover, the HIS-SUS database has been used to evaluate other information systems such as the live births system,\(^5\) or as a complementary source of data on diseases and pathological conditions for the purpose of epidemiological surveillance.\(^2\) Thus, the backing that it provides contributes towards planning actions, supporting health surveillance and evaluating interventions.\(^2\) Knowledge of the limits and potential of the information generated by this system is essential for analyses on the profiles that are presented.\(^16\)

Accidents and violence are external causes of morbidity and mortality\(^17\) that account for a large proportion of hospitalizations in Brazil. Although these causes generally present shorter duration of hospitalization, their cost is greater than the cost of hospitalizations due to natural causes,\(^14\) and they have a significant impact on public healthcare resources. In 2005, almost 800,000 hospitalizations due to these causes were recorded in the HIS-SUS database.

Since January 1998, it has been mandatory to assign a code from Chapter XX (External Causes of Morbidity and Mortality) of the International Classification of Diseases, Tenth Revision (ICD-10),\(^2\) to the “secondary diagnosis” field of the AIH, in cases of hospitalization through SUS due to accidents or violent events.\(^a,b\) Thus, in addition to providing knowledge about the consequences of accidents or violence (fractures, burns and wounds, among others), which are coded in the “main diagnosis” field (Chapter XIX of ICD-10),\(^17\) the HIS-SUS database also makes it possible to ascertain the “causes” of these injuries (for example, run over by a vehicle, motorcycle accident, assault or suicide attempt), through the “secondary diagnosis”. In this way, the system contributes towards analyses on the situation and trends regarding such hospitalizations and consequently towards providing backing for the preventive interventions that are needed.\(^1,10,12\)

Despite the importance of the information on the “main” and “secondary” diagnoses of hospitalization, for analyzing healthcare situations, few studies have evaluated the quality of these data in the HIS-SUS database. A review of the literature by Bittencourt et al\(^2\) covering the period 1984-2003 identified only three studies, which were all produced prior to the first ordinance from the Ministry of Health requiring the inclusion of the code for the external cause in the “secondary diagnosis” field. From reviewing the specialized literature subsequent to the issuing of this ordinance, only one study\(^c\) has focused on the quality of the information on hospitalizations due to external causes within the HIS-SUS database. However, that investigation took the starting point of cases described in the HIS-SUS database as due to external causes; thus, hospitalizations due to external causes that were nonetheless coded in the system as resulting from natural causes could not be identified.

The present study had the aim of analyzing the coverage and quality of the information on hospitalizations due to accidents or violent events (external causes of morbidity and mortality) in the HIS-SUS database, taking the medical records of the AIH reports as the gold standard.

**METHODS**

This was a cross-sectional study on hospitalizations through SUS that occurred in the municipalities of Londrina and Maringá, Southern Brazil, in 2004. In that year, the estimated populations of these municipalities were, respectively, 473,741 and 308,206 inhabitants.\(^8\)

Two independent databases were created: one with the hospitalization data appearing in the HIS-SUS database and the other with data from the medical reports of the AIHs due to external causes, which were obtained through manually reviewing these documents. Subsequently, these databases were interlinked into a single file, to compare the causes stated in the HIS-SUS database with those obtained from surveying the medical reports. The latter were taken to be the gold standard.

---


To make up the database from the reports, all of the medical reports of the AIHs issued by all of the hospitals that attended to cases of accidents or violence (six establishments in Londrina and five in Maringá) were reviewed, independent of the municipalities in which the patients lived. Reports that mentioned that the reason for hospitalization was any accident, violence, trauma, injury, poisoning or sequel of an external cause (Chapters XIX and XX of ICD-10), with hospitalization during 2004, were separated out and some of their data were transcribed onto record cards that were devised for this purpose, by healthcare professionals and undergraduate students who had received previous training. Following this, researchers with skills in coding and selecting cause of death and morbidity according to international rules assigned codes to the main injury (main diagnosis) and to the external cause that gave rise to the injury (secondary diagnosis), along with the other study variables. For the main diagnosis, the codes in Chapter XIX of ICD-10 (S00-T98) were used. For the secondary diagnosis, the codes in Chapter XX (V01-Y98) were used. Categories at the level of three characters were used for the coding. With the aim of ensuring higher quality of information, the coding was done in duplicate, with discussion of the divergences until reaching a consensus. The persons responsible for data gathering and those doing the coding were unaware of the causes recorded in the HIS-SUS database.

The electronic processing of the data in the medical reports was done in duplicate through the Epi Info 6.04d software, in order to identify and correct possible typing errors, thus generating the database named “reports”.

In parallel, a database named “HIS-SUS” was created. This was constructed using the data available from the Ministry of Health’s Datasus website. The data were available as compressed files in DBC format, covering all causes for the whole of the State of Paraná. These were decompressed and filtered in electronic spreadsheets, and the hospitalizations relating to the 11 hospitals studied were selected, independent of the municipalities in which the patients lived. Patients at these hospitals with a main or secondary diagnosis classified in Chapters XIX or XX of ICD-10 and with hospitalization during 2004 were selected. These data were then imported into the Epi Info 6.04d software. Considering that the logic followed by the HIS-SUS database is to present the AIHs according to the accounting month, and not according to the date of hospitalization (the AIHs can be presented up to 180 days after discharge), compressed AIH files presented in 2005 but referring to hospitalizations in 2004 were also sought. After June and October 2005, no further hospitalizations that occurred during the previous year were observed in Maringá and Londrina respectively.

The two independent databases were interlinked by means of a “unique identifier” (the number of the AIH), using the Merge (Join) application in the Epi Info 6.04d software, to form a single database with field with different names, in order to compare the two initial databases. In the first analysis, three situations of hospitalization due to external causes were identified: 1- hospitalization identified in both databases; 2- hospitalization identified in the reports but not identified in the HIS-SUS database; 3- hospitalization identified in the HIS-SUS database but not in the reports (Figure).

For the hospitalizations identified in the last two situations, a new search in the files and in the compressed Datasus files was conducted with the aim of incorporating missing records in their respective databases (reports and HIS-SUS). Such cases were treated in the same way as the cases identified earlier, with electronic processing in databases identical to the first ones, and were then incorporated into the single database (Figure).

During this process, it was observed that the type of external cause was not specified in a large proportion of the hospitalizations identified in the medical records. Although the medical reports stated that the case was one of injuries or poisoning, there was no information about the accident or violent event that caused them. Thus, with the aim of obtaining greater specificity of information in the research database, other sources were consulted, in the following order: the mortality information system database, with records identified by name, which was furnished by the municipal health departments of the municipalities studied; and the records from pre-hospitalization care provided through the Integrated Attendance System for Traumas and Emergencies (SIATE). For these comparisons, the names (obtained from the medical reports), the date of the accident or violent event and/or the date of death were used to connect the cases to their respective hospitalizations. For the cases in which it was possible to obtain information on the type of external cause, this was incorporated into the specific field in the “reports” database.

The fields relating to type of cause (external or other) recorded in the HIS-SUS and reports databases were compared with each other. To analyze the coverage and quality of the information on hospitalizations due to external causes in the HIS-SUS database, the following indicators (and their respective 95% confidence intervals) were calculated: crude agreement rate, i.e. the proportion of cases that were concordant with regard to the hospitalization due to external cause in relation to the total number of cases identified in both sources; sensitivity, i.e. the proportion of hospitalizations informed as due to external causes in the HIS-SUS database, taking the researched data as the reference and expressing the coverage of the system in relation to these causes;
and the positive predictive value, i.e. the likelihood that hospitalizations due to external causes informed in the HIS-SUS database were really external causes, taking the researched data as the parameters. The hospitalization profiles were also compared in terms of subtypes of external causes in the HIS-SUS database and those obtained through the search.

The research project was approved by the research ethics committees of the state universities of Londrina and Maringá.

RESULTS

The search through the reports found 4,018 cases of hospitalization due to external causes in Londrina, while the HIS-SUS database recorded 3,002 (74.7%). In Maringá, 1,403 hospitalizations due to external causes were recorded in the HIS-SUS database, thus representing 59.2% of the 2,370 cases identified in the reports (Table 1).

In Londrina, 70 cases informed in the HIS-SUS database as due to external causes were disregarded on the basis of the medical report. Most of those hospitalizations were reclassified as “diseases of the musculoskeletal system and connective tissue” (38 cases), “diseases of the circulatory system” (six cases) and “diseases of the genitourinary system” (six cases). In Maringá, out of 45 hospitalizations in the HIS-SUS database that were disregarded, most were reclassified as “diseases of the circulatory system (11 cases), “diseases of the musculoskeletal system and connective tissue” (seven cases) and “diseases of the nervous system” (seven cases). In both municipalities, the main external cause informed in the HIS-SUS database for these hospitalizations was “fall” (57 in Londrina and 22 in Maringá).

Table 2 presents the concordance indicators for the data on external causes between the HIS-SUS and reports database. In both municipalities, the HIS-SUS database presented a high positive predictive value. However, the sensitivity values showed under-recording of hospitalizations due to accidents and violent events in the HIS-SUS database, although the situation was better in Londrina. Whereas in Londrina around a quarter of the hospitalizations due to external causes failed to be recorded as such, the rate for Maringá was more than 40%, thus indicating that external causes were being incorporated in the HIS-SUS database as natural causes.

Figure. Stages in setting up the single database on hospitalizations due to external causes through the Brazilian National Health System. Municipalities of Londrina and Maringá, Southern Brazil, 2004.

AIH: Authorization for hospitalization
HIS-SUS: Hospital Information System of the Brazilian National Health System (Sistema Único de Saúde, SUS).
The profiles of causes of hospitalization according to the HIS-SUS and reports databases are shown in Table 3. The additional number of hospitalizations detected by searching through the reports consequently gave rise to a higher proportion of cases of events of undetermined intent, i.e. the residual category of Chapter XX of ICD-10 (codes Y10 to Y34), which is attributed when it is not known whether the event was accidental or intentional. These events increased by factors of 16.3 and 403.0 after the search and represented 23.1% and 34.0% of the hospitalizations in Londrina and Maringá, respectively. Nonetheless, some causes were underestimated in the HIS-SUS database in both municipalities, such as accidents due to exposure to inanimate forces, assaults and self-inflicted wounds. There was under-recording of transport accidents in Londrina, which doubled in frequency after reclassification based on the gold standard.

DISCUSSION

The present study is a pioneer with regard to analysis of the coverage of information on hospitalizations due to accidents and violent events recorded in the HIS-SUS database and one of the first to evaluate the quality of this information subsequent to the issuing of the Ministry of Health’s regulations regarding the insertion of a code for the external cause of the injury, in the “secondary diagnosis” field. Moreover, from reviewing the Brazilian literature, only five studies evaluating the quality of the information on hospitalizations were found.6,9,11,20,4 Only three of these allowed some type of evaluation on the quality of the information on accidents or violent events.9,11,4 Specifically in relation to hospitalizations due to external causes, Lebrão9 in 1974 observed that the sensitivity of the information in the State Health Department of São Paulo was 83.3%, in comparison with the same data from hospital medical files. Thus, around 17% of the cases would not have been recognized in that Department’s statistics if no search in the medical files had been undertaken. However, this proportion is smaller than what was seen in the municipalities of the present study.

A study conducted in Maringá in 199211 showed 202 hospitalizations due to injuries and poisoning in the hospital medical files, while the AIHs showed 173 (85.6%). In that investigation, there was also no concordance analysis regarding the circumstances of the event that caused the lesion or poisoning (external cause). The processing of the AIH data and the tabulations at that time (while the ninth revision of the ICD was still in force) only placed value on the consequences of the accident or violent event (present Chapter XIX of ICD-10).17 More recently, Melione4 investigated the concordance between information on external causes recorded in the HIS-SUS database and information obtained from searching through medical files in a hospital in São José dos Campos (SP). However, that study4 did not seek to investigate external causes that might have been informed as due to natural causes in the HIS-SUS database, which impedes comparison with the results from the present study. Nevertheless, the analysis in that study showed that excellent concordance was found for certain groupings of external causes, such as transport accidents and falls.

<table>
<thead>
<tr>
<th>Indicator</th>
<th>Londrina</th>
<th>Maringá</th>
</tr>
</thead>
<tbody>
<tr>
<td>%</td>
<td>71.7</td>
<td>56.2</td>
</tr>
<tr>
<td>95% CI</td>
<td>70.3;73.1</td>
<td>54.3;58.2</td>
</tr>
<tr>
<td>Positive predictive value</td>
<td>97.7</td>
<td>96.8</td>
</tr>
<tr>
<td>95% CI</td>
<td>97.0;98.2</td>
<td>95.7;97.6</td>
</tr>
<tr>
<td>Sensitivity</td>
<td>73.0</td>
<td>57.3</td>
</tr>
<tr>
<td>95% CI</td>
<td>71.6;74.3</td>
<td>55.3;59.3</td>
</tr>
</tbody>
</table>
In the present investigation, in both municipalities, a high rate of underrecording of hospitalizations due to external causes was observed (more than 40% in Maringá and around 25% in Londrina), along with distortions regarding the specific types of these causes in the HIS-SUS database. This situation may have occurred for a variety of reasons, as already pointed out by some researchers: low value placed on information produced for epidemiological and evaluation studies and low but growing use of this information; lack of or insufficient training for the people coding these causes; and the system serves mainly for hospital remuneration; among others. In addition, the lack of material on the importance of healthcare information within medical curricula may be another situation that contributes towards reducing the quality of this information. Doctors who provide attendance in emergency service sectors fail to record the motive (external cause) that gave rise to the injury or poisoning, which makes it difficult or impossible to carry out the subsequent process of coding these causes more specifically. The reasons for these distortions need to be identified and quantified, in both municipalities, so that they can be dealt with appropriately.

Nonetheless, the existence of information on an injury or occurrence of poisoning, even if the circumstances giving rise to it were not detailed on the medical report, would allow the person coding the AIH to assign an external cause code to the “secondary diagnosis” field, albeit generic, i.e. Y10-Y34 – events of indeterminate intent. By adopting this procedure, there would be an increase in the coverage of external causes in the HIS-SUS database, notwithstanding the loss of quality of the specific information regarding whether the event was intentional or not. This could be seen in the present study, in which incorporation of hospitalizations due to injuries or poisoning that were identified through the medical reports increased the numbers of hospitalizations due to external causes of undetermined intent. These causes could not be clarified regarding their intention and/or the type of event, even after searching for complementary information in other data sources (mortality and pre-hospital attendance).

One limitation of the present study results from the decision to compare the external causes and their groupings at the level of only three characters, whereas in the HIS-SUS database the causes are recorded including a fourth digit that is reserved for indicating the place where the event occurred. However, the use of three characters of the codes for external causes was sufficient for the general comparisons intended in the present study, such as the type of hospitalization cause (external or otherwise), groupings according to intention (accidental, intentional or indeterminate) and groupings according to type of external cause within the large groups of intention (falls, transport accidents or assault). In the present study, it was not sought to analyze the specificity of codes assigned (for example, the type of accident or the type of victim in transport accidents, or furthermore, the way in which a fall occurred), although the importance of these factors for surveillance of these events is recognized.

Concern regarding improvements to the quality of information on external causes of hospitalizations in Brazil is not recent, and this field of study is not limited to Brazilian researchers. The Ministry of Health’s decision to make it mandatory to insert a code for the external cause in the HIS-SUS database, for all hospitalizations due to injuries, poisoning or

---

**Table 3. Types of external causes of hospitalization identified from the Hospital Information System of the Brazilian National Health System and from searches. Municipalities of Londrina and Maringá, Southern Brazil, 2004.**

<table>
<thead>
<tr>
<th>Causes (ICD-10 codes)</th>
<th>Londrina</th>
<th>Ratio b/a</th>
<th>Maringá</th>
<th>Ratio b/a</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accidents (V01-X59; Y40-Y84)</td>
<td>2686</td>
<td>0.98</td>
<td>1278</td>
<td>1.13</td>
</tr>
<tr>
<td>Transport accident (V01-V99)</td>
<td>492</td>
<td>2.01</td>
<td>517</td>
<td>0.94</td>
</tr>
<tr>
<td>Falls (W00-W19)</td>
<td>1703</td>
<td>0.48</td>
<td>709</td>
<td>0.87</td>
</tr>
<tr>
<td>Exposure to inanimate mechanical forces (W20-W49)</td>
<td>106</td>
<td>1.88</td>
<td>14</td>
<td>6.50</td>
</tr>
<tr>
<td>Other causes of accidents (W50-X59; Y40-Y84)</td>
<td>385</td>
<td>1.60</td>
<td>38</td>
<td>6.74</td>
</tr>
<tr>
<td>Self-inflicted injuries (X60-X84)</td>
<td>61</td>
<td>1.13</td>
<td>7</td>
<td>1.57</td>
</tr>
<tr>
<td>Assault (X85-Y09; Y35-Y36)</td>
<td>195</td>
<td>1.82</td>
<td>4</td>
<td>19.50</td>
</tr>
<tr>
<td>Undetermined intent (Y10-Y34)</td>
<td>57</td>
<td>16.30</td>
<td>2</td>
<td>403.00</td>
</tr>
<tr>
<td>Sequelae (Y85-Y89)</td>
<td>3</td>
<td>15.33</td>
<td>112</td>
<td>0.26</td>
</tr>
<tr>
<td>Total</td>
<td>3002</td>
<td>1.34</td>
<td>1403</td>
<td>1.69</td>
</tr>
</tbody>
</table>

HIS-SUS: Hospital Information System of the Brazilian National Health System (Sistema Único de Saúde)
ICD: International Classification of Diseases

*There were no causes coded as Y90-Y98 (supplementary factors related to causes of morbidity and mortality classified elsewhere).*
other consequences of accidents or violent events, is in keeping with the proposals from the Centers for Disease Control and Prevention (CDC) for improving the quality of information on these causes, in relation to hospitalizations in the United States.\(^1\) In that country, there is great variability in the quality of the information on hospital morbidity due to external causes between the states. The information is more complete in the states that have adequate regulation and supervision of the compliance with health service practices for reporting external causes.\(^3\) In Brazil, however, from the analysis on the data obtained, it is believed that there is an initial need to invest in increased coverage of the information on external causes in the HIS-SUS database. From this, subsequently or in parallel, the specificity and quality of the information can be improved. The following strategies should be considered: investment in awareness-raising and capacitation of the professionals providing attendance in emergency and admission sectors of SUS hospitals or hospitals under contracts with SUS, so that they can adequately record the external cause on the attendance cards; and adoption of standardized forms for obtaining details on the external cause that gave rise to the hospitalization; and training for the people performing the coding.\(^1,13\)

Since the implementation and regulation of SUS, technicians and policy-makers in Brazilian municipalities have had responsibilities for managing the healthcare information system and for analysis and dissemination of data that did not exist not long ago. Thus, municipalities that, like the one studies here, have services for auditing, evaluating and controlling hospitalizations funded by SUS need to be concerned about analyzing the coverage and quality of the data produced by this system, so that the data may be useful for decision-making within this sector and may provide support for other sectors or organizations involved in facing up to healthcare problems. In the same way that, over recent years, there has been an improvement in the coverage and quality of the information on live births and mortality,\(^15\) there is also room for analysis and improvement of the information on hospital morbidity. Thus, the present study aimed to contribute towards better knowledge of the deficiencies that the system presents and consequently towards the process of qualifying the information within the HIS-SUS database.
REFERENCES


Study funded by the Department of Science and Technology of the Ministry of Health/ National Council for Scientific and Technological Development (Process N. 505.875/2004-7).