Reporting road victims: Assessing and correcting data issues through distinct injury scales

António Couto, Marco Amorim, Sara Ferreira

Faculty of Engineering, University of Porto, Rua Dr Roberto Frias, s/n 4200-465 Porto, Portugal

ABSTRACT

Introduction: The most common measurement for road accidents relies in police reports; however, there is a high portion of underreporting and misclassification, mainly concerning elderly causalities, urban accidents, slightly injured, users of two-wheeled vehicles, and car occupants. *Methods:* In order to assess these issues, road accidents occurring in the Porto Metropolitan Area, Portugal, covering a 6-year period (2006–2011) were analyzed based on police and hospital datasets. By linking hospital data with police data, it is possible to evaluate the misclassification of the victims' severity by the police regarding the maximum abbreviated injury scale (MAIS) classification. Additionally, considering that 29% of the victims recorded by hospitals were not reported by the police, which is in line with the reality of other EU countries, underreporting is further investigated. Thus, we used econometric and statistics tools to measure the correlation between different available data to identify possible causes of underreporting and misclassification. In this sense, factors contributing to the misclassification of casualties by the police are identified using a univariate analysis. On the basis of the linked police–hospital data, and considering those factors and the police classification, a probabilistic model was developed to estimate a MAIS-based classification for all individuals included in the police accident records. *Results:* The results of misclassification indicate a significant over-classification of severe injury by the police. Additionally, a systematic police underreporting phenomenon of around 30% was found. *Conclusions and Practical Applications:* Finally, comparing estimated results and actual data, we were able to produce non-fatality adjustment coefficients to estimate the total casualties taking into account the underreporting and misclassification phenomena and to compare them with the Portuguese and European realities.

© 2016 National Safety Council and Elsevier Ltd. All rights reserved.

1. Introduction

Due to the high number of people killed or seriously injured in road accidents, several studies conducted in different regions have been developed in order to better understand the causes of injury severity. Police reports are often the main source of data for crash investigation and road safety analysis (Tsui, So, Sze, Wong, & Leung, 2009). Police officers typically write their report soon after a crash has occurred. These reports contain information about victims, drivers, and vehicles characteristics, as well as the crash location and conditions. Also, every crash with victims and sometimes without victims but with a high value of property damage is required to be reported by the police. However, previous research has shown that the incompleteness of police crash reports exists in several countries, which may invalidate the conclusions of some studies or, at least, distort the road safety reality (Alsop & Langley, 2001; Amoros, Martin, & Laumon, 2006; Broughton et al., 2010; Farmer, 2003; Guo, Eskridge, Christensen, Qu, & Safranek, 2007; McDonald, Davie, & Langley, 2009; Tsui et al., 2009; Yamamoto, Hashiji, & Shankar, 2008).

In Portugal, as in most countries, accident victims are classified into three categories of injury severity: slight injury (SLI); serious injury (SEI); and fatality (F). In the Portuguese system, the following definition of victim severity is considered: SLI if the person stays in the hospital for no more than 24 h; SEI if the person stays in the hospital for more than 24 h; and F if the person dies within the 30 days following the accident as a result of the suffered injuries. The official classification is provided by the police, who do not have the proper skills to assess the severity of the injuries in a road accident because they do not know, a priori, how long the injured will remain in the hospital. However, the classification by the police is rarely checked with the medical records, except when the injured person dies in the hospital (ANSR, 2012).

Additionally, in the discussion about what the injury classification may report, the severity of the casualties has led the European Commission (EC) to seek an agreement on a new definition based on the maximum abbreviated injury scale (MAIS). The MAIS is derived from the Abbreviated Injury Scale (AIS), which is a severity score of a casualty with several injuries (Gennarelli & Wodzin, 2006). MAIS simply captures the highest severity among the various AIS codes of the victim's injuries. The AIS is the only dictionary specifically designed as a system to define the severity of injuries throughout the body and its first appearance dates back to 1971 (Baker, O'Neill, Haddon, & Long,

1974; Gennarelli & Wodzin, 2006). The scale is an anatomically based, consensus derived, global severity scoring system that classifies each injury by body region according to its relative importance on a 6-point ordinal scale (1 = minor and 6 = maximal; Association for the Advancement of Automotive Medicine, 2014). In contrast, the injury classification based on the length of hospital stay (LS) presents some issues that may lead to biased injury severity interpretation (e.g., if a patient stays longer in the hospital due to other reasons besides the accident injuries, or if administrative reasons lead to a delay on the discharge). In fact, a common agreement seems to exist around the benefits of using the MAIS instead of other common injury severity definitions. For instance, the LS is a classification that is not directly related to the severity of the victims, also depending on the hospital administrative rules and management decisions. In contrast, the MAIS is associated to the medical diagnosis and the probability of dying as a consequence of the injuries suffered.

Although previous studies have been conducted to analyze the level of under and misreporting in several countries, this subject is new in the Portuguese context. The study was developed under the research project LIVE – *Tools to Injury Prevention*, co-funded by the European Commission (EC). Through this project, for the first time in Portugal, police and hospital records were obtained and linked (Amorim, Ferreira, & Couto, 2014).

In this context, the objective of this study is to assess the level of underreporting and misclassification in traffic casualties in the Portuguese context. To analyze the misclassification of the casualties existing in the police records, two different comparisons were assumed. First, considering the police definition for the two levels of injury severity, the MAIS information was used to check the police records. Thus, the EU common definition based on the MAIS was used to classify the victims as SLI and SEI; an SEI is defined whenever the MAIS is equal to or greater than 3, otherwise it is an SLI. Despite the clear advantages of using the MAIS, its full implementation will not be as fast as it would be desirable in some countries like Portugal. Therefore, by comparing the police and the MAIS classifications, adjustment factors may be used as an approximate inference, similarly to the fatality-adjusting coefficient that was used in the past in Portugal; until 2010, the official number of fatalities was set by multiplying the number of fatalities declared at the accident scenes by an adjustment factor of 1.14.

The purpose of the research described here is to develop a model-based method to estimate MAIS classification for all individuals included in the police crash records, using similar conceptual framework of the model-based KABCO-MAIS conversion proposed by Tarko and Azam (2011) applied to the linked medical and crash records. By applying this model to the entire police data crash dataset and comparing the estimated MAIS-based classification with not linked hospital records, results will allow identifying misclassification and underreporting phenomenon. Therefore, considering these results, adjustment factors are provided to estimate, from police crash records, the number of traffic casualties by severity type.

2. Background

The analysis on the reliability of police reports is supported by two different data sources, one from the police records and another from hospitals. To perform such analysis a linkage methodology must be applied in order to correctly connect both police and hospital's databases (Clark, 2004). By linking both data records and comparing the final results, the degree of underreporting and misclassification of the injury severity of the victims can be obtained. In fact, the degree of underreporting found in previous studies is quite significant (Alsop & Langley, 2001; Amoros et al., 2006; Dandona, Kumar, Ameer, Reddy, & Dandona, 2008; Tsui et al., 2009). Official definitions about the injury severity level of casualties differ from one country to another. It can be based on the hospital admission, length of hospital stay, injury

type, and medical treatment required. Regardless of the official injury definition, in almost all of the countries, several differences have been found between the injury severity reported by the police and the hospitals (International Traffic safety Data and Analysis Group, 2012).

In this sense, Broughton et al. (2010) assess the level of underreporting of non-fatal road accident casualty in several European countries, using a three-step methodology: (a) data collection in different countries; (b) definition of a framework for each country to link the road accident database from the police with the medical database from the hospitals; and (c) calculation of the national coefficients to estimate the actual casualties from the police database. A similar methodology was used by McDonald et al. (2009). The Land Transport New Zealand Traffic crash reports between 2000 and 2004 were linked to the New Zealand Health Information Service hospital discharge data. Then, a severity code based on the ICD-10 was assigned and compared to the police classification. The findings demonstrated that the police measurement of non-fatal injuries, in terms of threats to life, has poor validity, suggesting that users of police injury severity data should assess data validity. Using a simple data comparison approach, Bhatti, Razzak, Lagarde, and Salmi (2011) analyzed data from three entities: police, ambulance service, and hospitals. Accident outcomes in a 196 km of a road section in Pakistan were compared in terms of road users. This study showed that the crashes and injuries reported by the police on this road section in a one-year period were several times less than the figures contained in ambulance and hospital records.

Recently, Yannis, Papadimitriou, Chaziris, and Broughton (2014) conducted a disaggregate analysis of road accident injury underreporting in selected European countries. The methodology included a cross-check and match of medical records with police road accident records. Yannis et al. (2014) expressed the level of injury underreporting by underreporting coefficients, estimated as the actual estimated number of road accident injuries of a given severity to the number of related injuries recorded by the police. To assess misclassification the MAIS classification from hospital records was compared to the police assessment, particularly by counting the victims with MAIS ≤2 which were assessed as serious injuries by the police. The same was made for the slight injuries, but now with MAIS >2. Afterward, log-rate models were developed in order to estimate the combined effects of country, road user type, police severity score, and MAIS score on underreporting. Recently, Qiu et al. (2015) made a time trend comparison between the official road accident casualties, obtained through the police reports, and figures from 210 military reports across China. Results showed that the police reported a decrease in injuries/fatalities contrary to the non-decrease trend found in the hospital reports analysis. In the second stage, the study focused on the city of Chongqing. The in-depth city data investigation showed 3 times as many fatalities and 5 times as many injuries as those reported by the police.

In addition to a typical analysis of underreporting and misclassification, some studies have focused on the analysis of specific groups of road users. Tarko and Azam (2011) investigated pedestrian injury severity factors using linked police–hospital data. A bivariate ordered probit model with sample selection was used to check for the presence of the selectivity bias and to account for it in the MAIS estimates. The selectivity bias was considerable in predictions of low injury levels. The pedestrian injury analysis identified and estimated several severity factors, including pedestrian, road, and vehicle characteristics. Male and older pedestrians were found to be particularly exposed to severe injuries. Rural roads and high-speed urban roads appeared to be more dangerous for pedestrians, particularly when crossing such roads. The size and weight of the vehicle involved in a pedestrian crash were also found to have an effect on the pedestrian injury level. Trying to estimate the extent of underreporting of road crash injuries to police among different road user groups, Watson, Watson, and Vallmuur (2015) used databases from Queensland Road Crash Database, Queensland Hospital Admitted Patients Data Collection, Emergency Department