Injuries

Actual incidences of road casualties, and their injury severity, modelled from police and hospital data, France

Emmanuelle Amoros, Jean-Louis Martin, Sylviane Lafont, Bernard Laumon

Background: Nation-wide road casualty figures usually come from police data. In France, as in many developed countries, the reporting of fatalities is almost complete but the reporting of non-fatal casualties is rather low. It is moreover strongly biased. Valid estimates are needed. Methods: Using the capture-recapture method on police data and on a road trauma registry covering a large county of 1.6 million inhabitants, we estimate police under-reporting correction factors that account for unregistered casualties. These correction factors are then applied to the nation-wide police data, with standardization on under-reporting bias factors. Results: In 2004, whereas the police report 108,727 non-fatally injured, the estimation yields 400,200. Over the 1996–2004 study period, the average annual estimated incidence is 871/100,000 for all injured (3.4 times the police incidence), 232/100,000 for hospitalized, 103/100,000 for seriously injured (2.2 times the police incidence) and 12.6/100,000 for casualties with long-term major impairment. The incidence of seriously injured (NISS 9+) is 11.3/100,000 for pedestrians, 9.5/100,000 for cyclists, 36.3/100,000 for motorized two-wheel users and 42.5/100,000 for car users. Conclusions: The estimated incidences are much higher than the police-based ones. This changes the scale of the road injuries issue. The risk of suffering a major impairment from a road crash is equal to the risk of being killed. Motorized two-wheel users experience a large burden of traffic casualties, much larger than that indicated by police data. The approach used can be reproduced in other countries, if an additional medical registration exists.

Keywords: capture-recapture, impairments, incidences, injuries, traffic accidents

Introduction

In most countries, data on road injuries are provided by the police. In the developed countries, the reporting rate of fatalities is usually rather high, but it is rather low for non-fatal casualties. There is under-reporting, both to and from the police. It is strongly biased, mostly on injury severity, mode of transport and number of vehicles involved (single versus multiple). Another source of registration, namely hospital records, can sometimes be used. When both hospital and police data are available, it is possible to obtain a good estimation of the actual number of non-fatal road casualties and of their injury pattern. This is needed to better define prevention programmes.

In France, only the police data are available at the nation-wide level. A national hospital discharge file exists but external causes of injuries are often not recorded. In the Rhône county, a hospital-based road trauma registry provides a second source of data. At this county level, we confront the two registrations, model the probability of registration by accounting for those not reported at all and estimate police under-reporting correction factors. Applying these correction factors to the nation-wide police data, standardizing on the relevant covariates, we obtain a global picture of the health burden of road accidents at a national scale.

Materials and methods

The projection covers all non-fatal road casualties (counts of fatalities are not corrected as they are almost complete). At the nation-wide level, only the police data is available, but at the Rhône county level, both the police data and the road trauma registry are available. This county has 1.6 million inhabitants, with a large city (Lyon), its suburbs and a rural area. Over the 1996–2004 study period, 83,447 non-fatally injured are reported by the registry, 36,329 by the Rhône police file and 1,353,600 by the nation-wide police file.

Police road crash data

The French police have to write a report for every road crash causing at least one casualty. A road crash is officially defined as a crash involving at least one vehicle (motorized ones and bicycles). The police report should describe everyone involved in the crash, and classify them as non-injured, slightly injured, seriously injured or killed. Slightly injured are out-patients and in-patients with a hospital stay of at the most 6 days. Seriously injured are those requiring a longer hospital stay. Killed are those who died within 6 days of the crash.

The Rhône road trauma registry

A road trauma registry has been in operation since 1995 and has been certified by the relevant French authority. It covers all casualties from road crashes in the Rhône county, who seek medical care in health facilities. All health care facilities (from public and private hospitals) in the county and its surrounding area, which may receive crash victims participate: about 220 health units ranging from pre-hospital emergency care, emergency departments, intensive care units, surgery units, etc. to rehabilitation departments. Injury assessment is based on all diagnoses established in the different health services a casualty may have visited. Diagnoses are coded with
the Abbreviated Injury Scale (AIS), which includes a severity score, ranging from 1 (minor) to 6 (beyond treatment).

**Matched casualties between the two registrations**

At the Rhône level, casualties recorded in both the police and the registry were identified thanks to a semi-automated probabilistic record linkage. Linking variables were date and time of the crash, crash location (town/village and possibly street/road), mode of transport, year of birth, month of birth and gender of the casualty.

**Health outcomes**

The outcomes of interest are all injured, hospitalized, seriously injured and injured with major impairment, based on medical classifications (except for the hospitalized). The casualty’s injury severity is measured with the New Injury Severity Score (NISS). It is defined as the sum of squares of the AIS severity scores of a casualty’s three most severe injuries, regardless of body region. The NISS scores are grouped into three categories: NISS 1–3 (minor severity), NISS 4–8 (moderate) and NISS 9+ (serious). Casualties with long-term major impairment(s) are defined with the Injury Impairment Scale (IIS), restricted to those with any injury of IIS 3+. These correspond mostly to serious head injuries (e.g. intraventricular haemorrhage of cerebrum), spine injuries (e.g. contusion, laceration) and lower extremities injuries (e.g. knee dislocation). Information on these health outcomes is available on registry casualties; for police-reported casualties they are estimated by prediction models (Step 1 subsequently).

The whole estimation procedure is presented for the NISS (here and in figure 1); it is repeated for: hospitalized, any IIS 3+ injury and any head, spine or lower extremities IIS 3+ injury separately.

**Predicting the injury severity (NISS) of police-reported casualties (Step 1)**

The police severity classification is not satisfactory but it is correlated with the NISS, and hence useful for its prediction. The prediction model is constructed on non-fatal casualties identified in both the police and registry files, for which both the NISS and police severity are available (22,704 matched casualties). The three-level NISS variable is ordinal, and we model it with a cumulative logit model. The proportional odds assumption being rejected ($P < 0.01$) this corresponds to two binomial models, which are fitted with SAS software, logistic procedure. The variables retained in the predictive model are: interaction between police severity classification and type of police, road user type, interaction between road user type and third party (yes/no), road user type of third party, age, gender, road type, urban/rural and calendar year. For the prediction of the other health outcomes, which are binary (yes/no), we merely use a logistic regression, with the same covariates.

The three-level NISS prediction model is applied to the Rhône non-fatal casualties reported only by the police (figure 1), so that the under-reporting model (which must be based on all observations from the two registrations in the Rhône county) can include the NISS. The three-level NISS prediction model is applied to the nation-wide police data (figure 1), so that in Step 4 the under-reporting correction factors, which are defined according to NISS level (among others), can be applied to these data.

**Improving the record linkage (Step 2)**

Before modelling under-reporting, we need to have a highly reliable picture of how many non-fatal casualties are recorded by the police only, by the registry only and by both. As a result of the record-linkage, there might be false positives: linked
pairs corresponding to two distinct casualties, and false negatives: non-linked pairs corresponding to the same casualty. We estimate their numbers and use them to improve the result of the record-linkage.

**Estimating police under-reporting correction factors, with capture-recapture (Step 3)**

The capture-recapture approach has been used in epidemiology, including in road crash injuries. The major feature of capture-recapture is to quantify those unobserved i.e. not recorded by any registration. The validity of the method is based on four key assumptions; they are discussed later.

A capture-recapture model is fitted to the aggregated Rhône police and registry data (96 238 casualties). A standard result is the total number of road casualties in the Rhône county. Here we aim to estimate under-reporting correction factors, and these can be directly obtained from the fitted model. The model formulates the probabilities of being reported by each source, conditional on the observed data. It enables the estimation of these probabilities, unconditional on the observed data. The inverse of these provides the under-reporting correction factors. The dependent variable is categorical nominal (police only, registry only, both), modelled by a generalized logit, using SAS logistic procedure.

One assumption of capture-recapture is ‘homogeneity of capture’: all subjects should have the same probability of being recorded by a given source. If this is true only within groups (e.g. by injury severity level) one should stratify on or include the corresponding covariate. The probability of being reported by the French police mostly depends on injury severity, road user type, third party (yes/no) and their interaction, road type, urban/rural area and type of police. The probability of being reported by the registry mostly depends on injury severity. In other words, the probability of being reported by a source is homogenous within every category defined by the combination of these aforementioned characteristics. We hence include the corresponding covariates in the model, except for urban/rural, not available in the registry. Year is also included (as a quantitative variable) to allow for a possible time trend in under-reporting. Under-reporting correction factors \( C_j \) are estimated within every category.

**Projection to the nation-wide level (Step 4)**

It is a projection of county-level to the nation-wide level, standardizing on the relevant covariates. It is similar to the indirect standardization of rates, usually over age and sex. The nation-wide estimated number of non-fatal casualties is obtained by \( \sum_{j=1}^{J} C_j \times O_j \), where \( j \) indexes the categories defined by the combination of the under-reporting bias factors, instead of the age and sex categories; \( C_j \) are the \( j \)-specific reference correction factors, instead of the age- and sex-specific reference incidence rates and \( O_j \) are the observed police counts of non-fatal casualties nation-wide, instead of the size or person-years of the study population.

**Results**

Over the study period 1996–2004, for the whole country, the police-based average annual incidence for all injured is 255/100 000; whereas the estimated incidence is 871/100 000 (95% CI: 828–894), i.e. 3.4 times higher. The police-based incidence of seriously injured is 47/100 000 (those requiring more than 6 days of hospital stay) whereas the estimated incidence of NISS 9+ casualties is 103/100 000 (97–106), i.e. 2.2 times higher. As for the hospitalized, the police-based estimated incidence is 104/100 000 (101–107) and the projected one is 232/100 000 (221–239). Table 1 displays annual incidences. The estimated incidences display a lesser decrease than the police-based incidences, especially for seriously injured casualties: a decrease by less than a third instead of by half.

The average annual incidence of all injured is 69/100 000 (66–71) for pedestrians, 94/100 000 (78–112) for cyclists, 203/100 000 (192–210) for motorized two-wheel users and 469/100 000 (450–479) for car occupants. The number of injured cyclists is slightly higher than the number of injured pedestrians. The average annual incidence of seriously injured is 11.3/100 000 (10.5–12.0) for pedestrians, 9.5/100 000 (7.2–11.8) for cyclists, 36.3/100 000 (33.5–38.1) for motorized two-wheel users and 42.5/100 000 (39.8–44.4) for car occupants. The number of seriously injured is in the same order of magnitude for cyclists and pedestrians on the one hand, and for motorized two-wheel users and car occupants on the other hand. Table 2 displays annual estimates. Regarding time trends, motorized two-wheel users experience a much smaller decrease of their number of casualties than the other road users.

Casualties with major impairments (ISS 3+) are quantified in table 3. The ISS 3+ injured body region is principally the head (53.3%), followed by the lower extremities (24.3%) and the spine (15.8%). Cyclists and pedestrians are particularly prone to head ISS 3+ injuries. Motorized two-wheel users have a high frequency of spine and lower extremities ISS 3+ injuries.

**Discussion**

**Limitations and strengths of the method**

Capture-recapture is based on four key assumptions. The ‘homogeneity of capture’ has been dealt with: covariates that are correlated with the registration probabilities are included in the model. A second assumption is the perfect identification of subjects common to the different registrations: the record-linkage is based on many informative linking variables and it has furthermore been corrected for false positives and false negatives. A third assumption is a closed population i.e. no entries or losses between two recordings. Most casualties injured in crashes in the Rhône county live in this county (89.8%) or a county next to it (6.7%) so that even if slightly injured, most casualties will go to a hospital in the Rhône county or its close surroundings i.e. covered by the registry.
## Table 1

Counts and incidences of road casualties per 100,000 inhabitants, police-based and estimated, France

<table>
<thead>
<tr>
<th>Year</th>
<th>All (non-fatally) injured</th>
<th>Seriously injured</th>
<th>Hospitalized</th>
<th>Killed</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Count</td>
<td>Incidence</td>
<td>Count</td>
<td>Incidence</td>
</tr>
<tr>
<td>1996</td>
<td>Police-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>170</td>
<td>117</td>
<td>293.2</td>
<td>554</td>
</tr>
<tr>
<td>1997</td>
<td>Police-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>169</td>
<td>577</td>
<td>291.3</td>
<td>561</td>
</tr>
<tr>
<td>2000</td>
<td>Police-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>162</td>
<td>117</td>
<td>274.7</td>
<td>548</td>
</tr>
<tr>
<td>2001</td>
<td>Police-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>153</td>
<td>945</td>
<td>259.2</td>
<td>524</td>
</tr>
<tr>
<td>2002</td>
<td>Police-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>143</td>
<td>839</td>
<td>236.6</td>
<td>484</td>
</tr>
<tr>
<td>2003</td>
<td>Police-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>133</td>
<td>717</td>
<td>191.5</td>
<td>425</td>
</tr>
<tr>
<td>2004</td>
<td>Police-based</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>123</td>
<td>627</td>
<td>159.4</td>
<td>362</td>
</tr>
<tr>
<td></td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Notes

- a: Based also on police definition (requiring more than 6 days of hospital stay)
- b: Defined as NISS 9+
- c: Estimated. It corresponds to Step 1 in figure 1 applied to the nation-wide police data

### Interpretation of findings

The estimated incidences obtained by projection are much higher than the police-based incidences. This changes the scale of the road injuries issue. These much higher figures are not so surprising since the French police under-reporting rate is 3.7%.26 This rate is well within the average of other countries,2 implying that it is not only a French issue but one that concerns many countries.

The estimated incidences decrease less than the police-based ones, and the difference in decreases is larger for serious casualties. These decreases are coherent with the decrease in the number of road fatalities observed in France, which has been especially large since 2002, when numerous speed control cameras were announced and installed.

Estimated incidences according to age and gender give the same ranking of age and gender groups as those based on police data (data not shown) since these two characteristics are...
not major bias factors in police reporting practices. Results confirm the higher incidence of road injuries among young people and among men. 36

The number of motorized two-wheel casualties hardly decreases; this confirms what is seen in the police-based statistics. 37 Some likely explanations are the ineffectiveness of speed cameras for motorized two-wheel users (because of difficulties in identifying them) and the rather large increase in the fleet of motorbikes. 37

The number of seriously injured (NISS 9+) motorized two-wheel users has reached the same level as seriously injured car users (whereas the number of killed motorized two-wheel users is three times lower than that of car users, and their share in the traffic is highly different). Cyclists roughly undergo the same number of injured (NISS 1+ and NISS 9+) as pedestrians. These two patterns are new. They cannot be seen in the police-based figures because of the biases associated with police reporting. 26

The risk of suffering a major impairment is equal to the risk of being killed. For cyclists and motorized two-wheel users, this risk is in fact higher than the risk of being killed. These criteria should be used for prevention campaigns; the number of fatalities must not be the only criteria.

Motorized two-wheel users experience a large burden of traffic casualties, and much larger than what was known. These casualties are mostly teenagers and young adults. 38 It is crucial that prevention campaigns target these users.

### Acknowledgements

Vallee B, Vallet G, Vancuyck A, Vergnes I, Verney MP, Voiglio EJ, Vourey G, Vuillard J, Westphal M, Willemen L. This study was supported by Ministère des Transports, Direction à la Sécurité et à la Circulation Routière, Programme National de Recherche et d’Innovation dans les Transports Terrestres (PREDIT-DSCR/INRETS n°02/70013 étude n° 6); Ministère de la Santé, Institut de Veille Sanitaire (J06-23).

Conflicts of interest: None declared.

Key points

- We provide the first French nation-wide estimations of non-fatal road casualties that are corrected for under-reporting and its biases (and defined with medical classifications: NISS and IIS).
- The incidence is 871/100 000 for all injured (NISS 1+) and 103/100 000 for seriously injured (NISS 9+), respectively 3.4 and 2.2 times higher than the police-based incidences.
- The number of seriously injured (NISS 9+) motorized two-wheel users is now equal to the number of seriously injured car users (while their share in traffic are highly different).
- The risk of suffering a long-term major impairment (IIS 3+) is equal to the risk of being killed, on average. For cyclists and motorized two-wheel users, the risk of major impairment is higher.
- In terms of morbidity, injury prevention in France should focus on motorized two-wheel users.

References

3 Hvoslef H. Under-reporting of road traffic accidents recorded by the police at the international level. IRTAD special report. Norway: Public Roads Administration, 1994.
38 Martin JL, van Kampen B, Perez K. Deliverable D9 WP3 (Hospital data) data analysis: Pan-European Co-ordinated Accident and Injury Databases (PENDANT); 2006.

Received 13 September 2007, accepted 3 March 2008