Road Traffic Deaths and Injuries Are Under-Reported in Ethiopia: A Capture-Recapture Method

Teferi Abegaz1*, Yemane Berhane2, Alemayehu Worku3, Abebe Assrat4, Abebayehu Assefa5

1 School of Public and Environmental Health, College of Medicine and Health Sciences, Hawassa University, Hawassa, Ethiopia, 2 Addis Continental Institute of Public Health, Addis Ababa, Ethiopia, 3 School of Public Health, College of Health Sciences, Addis Ababa University, Addis Ababa, Ethiopia, 4 Ministry of Transport, Addis Ababa, Ethiopia, 5 World Health Organization (WHO) Ethiopia, Addis Ababa, Ethiopia

Abstract

In low and middle income countries road traffic injuries are commonly under-reported. This problem is significantly higher among those less severely injured road users. The objective of this study was to determine the incidence and the level of ascertainment of road traffic injuries and deaths by traffic police and hospital registry. In this study two-sample capture-recapture method was applied using data from traffic police and hospital injury surveillance, through June 2012 to May 2013. The study was conducted on one of the busiest highways in Ethiopia, the Addis Ababa – Hawassa highway. Primary data were collected by accident investigators and hospital emergency nurses using a structured checklist. Four matching variables; name of the victim, sex, place and time of the accidents was used to get the matched cases. During the study period the police independently reported 224 deaths and 446 injuries/billion vehicle kilometer while hospitals reported 123 deaths and 1,046 injuries/billion vehicle kilometer. Both sources in common captured 73 deaths and 248 injuries/billion vehicle kilometer. Taking the two data sources into consideration, the capture-recapture model estimated the incidence of deaths and injuries ranged 368–390 and 1,869–1,895 per billion vehicle kilometer, respectively. The police source captured 57.4%–60.9% of deaths and 23.5%–23.9% of injuries while the hospital sources captured 31.5%–33.4% of deaths and 55.2%–56% of injuries. Deaths and injuries among females, younger age victims, cyclists/motorcyclists and pedestrians were under-reported by traffic police. In conclusion neither of the two sources independently provided accurate coverage of road traffic incident related deaths and injuries. Strengthening both systems is necessary to obtain accurate information on road accidents and human causalities.

Introduction

Road traffic incidents pose enormous challenges. Globally 1.3 million fatal and up to 50 million non-fatal injuries are reported [1]. It is the second most common cause of death among the most economically active population group (15–44 year-old) [2]. Low income countries with small numbers of vehicles experience a huge burden of fatalities [3–5]. Ethiopia, according to WHO report, is considered one of the worst countries in the world where road transportation kills and injures a large number of road users every year [6].

Estimates of the magnitude of road traffic incident and traffic injuries in low-income countries are primarily obtained from police records and sometimes hospital registry data. However, both sources are affected by under reporting [7–9]. In Ethiopia, traffic police reports are the official source of data for road-related incidents. There is no established trauma care and registry, the health management information system (HMIS) reports traffic incident causalities as a soft tissue injury.

Researchers consistently indicate that high level of under-reporting of injuries and fatalities is observed, particularly in law and middle income countries, when the estimate is based on official records [10]. The problem of under-reporting is substantially higher for the nonfatal injury cases and vulnerable groups of road users (pedestrians, pedal cyclists, and motorcyclists) [7,11].

Capture-recapture approach provides estimates of the outcome of an event based on cases commonly captured by different data sources [12]. Originally the capture-recapture method was developed to estimate wild life populations [13]. However, in recent years, it has been commonly utilized in various epidemiological studies [14–16]; in determining infectious and non-infectious diseases [17,18]; and to estimate the incidence of road traffic incident causalities [19,20].

In low- and middle-income countries knowledge gap jeopardise resource allocation, even though cost effective interventions exist [21]. Few research has been conducted in low- and middle-income countries having poor data registry and high level of under-reporting [10,11]. There was no similar study conducted in Ethiopia. Therefore the objective of this study was to estimate the number of fatal and non-fatal injuries using a capture-recapture approach based on police and hospital data sources.
Method

Study setting
This study was conducted from June 2012 to May 2013, on one of the busiest highways of Ethiopia, the Addis Ababa - Hawassa highway. It is a two-way and two-lane road with an average width of 8 meters covering a distance of 264 kilometers. The first 64 kilometers are part of the main route of the country’s import and export corridor from the port of Djibouti. According to the information obtained from the Ethiopian Road Authority this road network is divided into four segments depending on the average daily vehicle flow. The higher traffic volume (more than 20,000 vehicles) observed near Addis Ababa city and subsequently decreased to 3,000 when it approaches to Hawassa city (Table 1).

Data sources
Hospital data source. The health management and information system (HMIS) report in Ethiopia is incomplete in some of the basic information used to run capture-recapture analysis. Hence, we conducted injury surveillance in five governmental hospitals found across the stretches of the studied highway (Bishoftu referral hospital, Adama referral hospital, Zeway Sher-Ethiopia hospital, Shashemene referral hospital and Hawassa referral teaching hospital). Data were recorded on a structured checklist prepared for the purpose of this study using five trained nurses working in the emergency department. Moreover, two senior emergency nurses and the principal investigator were involved as a supervisor. All injuries and deaths occurred on the studied road network were included. Relevant information of the name of the victim, age, sex, place and time of injury, type of road user and collision were recorded. Fatally injured cases in hospitals were gathered.

Traffic data source. Traffic data were collected by sixteen trained road crash investigators working across the stretches of the highway. Furthermore, two senior police officers and the principal investigator were involved as a supervisor. A structured checklist was prepared and relevant information was collected at the scene of the incidents and followed the victims up to a month to get the final injury outcome. Fatalities include death at the scene of the incident and within one month of the incident. Information on the name of the victim, age, sex, place and time of injury, type of road users and collision were recorded. Fatally injured cases in hospitals included dead on arrival and died prior to discharge or transfer.

Matching criteria
Casualties obtained from police and hospital sources were matched using four key identifier variables; name of the victim, gender, place and date of the crashes. Matched cases were considered if the victims shared at least three of the key identifier variables. For crash location matching, data collectors have sufficient information on the studied road network. Overlapping cases were coded “1” and non-overlapping cases as “0”. Record-linkage was performed manually and after completing the linkage process, we entered the data using Epi-data 3.1 software.

Data analysis
Descriptive analysis was done using SPSS version 16 and by using a manual calculator. Ascertained correctly incidence of deaths and injuries and the completeness of traffic and hospital data sources were determined by capture-recapture method using Chapman estimator [22]. This estimator is less affected by zeros and is said to be less biased than the Lincoln–Petersen estimator [23], and given by:

\[
\hat{N} = \frac{(N_t + 1) \times (N_h + 1)}{(n + 1)} - 1
\]

Where \(N_t\) and \(N_h\) are the numbers of people in traffic and hospital captures respectively, and \(n\) is the number of people identified in both sources (matches). The following formulas were used to calculate the standard error of the estimates (SE), and confidence interval of the estimates (95% CI):

\[
SE(\hat{N}) = \sqrt{\frac{(\hat{N} - N_t)(\hat{N} - N_h)}{n}}
\]

95% CI = \(\hat{N} \pm 1.96SE\)

Stratified Capture-recapture analysis. Homogeneous reporting probability among hospital and police data were checked by using Chi-squared tests for the following factors: gender (male, female); age (<15, 15–44, 45–64 and ≥65); road user type (driver, passenger, pedestrian and bicyclist/motorcyclist); and collision type (with other vehicles, rollover, pedestrian collision and others). A stratified capture-recapture analysis was done for each of these factors on the subgroups, and summing the subgroup results to obtain an overall estimate.

Incidence rate calculation. Incidence of road traffic deaths and injuries were calculated using vehicle kilometer travel as a denominator. Information on the vehicle kilometer travel on the studied road network was obtained from the Ethiopian Road Authority Official website. The incidence rate was estimated using the following formula:

\[
\text{Incidence rate} = \frac{\text{Number of injuries}}{\text{Vehicle Kilometer Travel}}
\]

Table 1. Summary of distance, daily average vehicle flow and vehicle kilometer travel by road segments of the studied road, Addis Ababa–Hawassa, 2012.

| Road segment         | Distance in kilometer | Daily average vehicle flow | Vehicle kilometres travel |
|----------------------|-----------------------|----------------------------|--------------------------|
| Dukem to Debre Zeit  | 13                    | 21,913                     | 284,869                  |
| Debre Zeit to Adama  | 51                    | 17,496                     | 892,296                  |
| Mojo to Shashemene   | 180                   | 3,534                      | 636,120                  |
| Shashemene to Hawassa| 20                    | 3,077                      | 61,540                   |
| Total                |                       |                            | 1,874,825                |
Incidencerate($R$) = \(\frac{C}{V \times 365 \times L} + 1,000,000,000\)

Where: $C$ stands for the total number of deaths or injuries; $V$ is average daily traffic volume on the road segment; 365 stands for the number of days in a year and $L$ is the average length travelled in kilometer. So the result interpreted as total number of death or injuries per billion vehicle kilometer travel on the studied road. According to International Traffic Safety Data and Analysis Group (IRTAD) report expressing deaths and injuries per vehicle kilometer travel is the most objective indicator to describe risk on the road network [24].

Ethical approval

We obtained ethical approval from Hawassa University College of Medicine and Health Science Ethical Review Committee. Written permission was obtained from hospitals and traffic police administrators. In addition, informed verbal consent was obtained from all study participants/caregivers and recorded by the research team on the ethical consent form (with prior approval obtained from our ethics review committee). Written consent was not suggested by Hawassa University College of Medicine and Health Science Ethical Review Committee. After completing the matching process the participants’ names were deleted from the database to ensure confidentiality.

Results

In this study traffic police identified 153 deaths and 305 injuries, while hospital surveillance captured 84 deaths and 716 injuries. Table 2 and 3 illustrates the distribution of road traffic crash related deaths and injuries with regard to various characteristics. In both data sources males and 15–44 year-old age group accounted for the highest percentage of fatal and non-fatal cases. Males were killed and injured 3 times more often than females. Similarly, 71% and 75.6% deaths and 83.9% and 79.6% injuries were in the age range of 15–44 years-old in traffic police and hospitals registry, respectively.

The result showed that in the traffic data there is a statistical significant under-reporting of deaths and injuries among bicycle/motorcycle riders (0.6% and 1.3%, respectively), compared to hospital data (3.5% and 5.5%, respectively). Females, younger age victims and pedestrians were also under-estimated by the traffic data.

Traffic police independently reported 224 deaths and 446 injuries/billion vehicle kilometre, while hospitals reported 123 deaths and 1,046 injuries/billion vehicle kilometer. Both sources in common captured 73 deaths and 248 injuries/billion vehicle kilometer. Taking the two data sources into consideration, stratified capture-recapture analysis provide the incidence of deaths and injuries ranged 368–390 and 1,869–1,895 per billion vehicle kilometer, respectively (Table 2 and 3). Level of under-reporting was higher in both data sources; the police source captured 57.4%–60.9% of deaths and 23.5%–23.9% of injuries while the hospital sources captured 31.5%–33.4% of deaths and 55.2%–56% of injuries.

Discussion

This study illustrates the usefulness of the capture-recapture method to assess the completeness of official data sources. The
Under-Reporting of Traffic Accident Deaths and Injuries

Table 3. Stratified capture-recapture estimates of injuries on the Addis Ababa- Hawassa highway, Ethiopia, June 2012–May 2013.

| Stratified by | Number of unmatched in police data | Number of unmatched in hospital data | Number of matches | Est. Number of injuries (95% CI) | Est. Incidence (95% CI) Per billion vehicle kilometres |
|--------------|-----------------------------------|-------------------------------------|------------------|-------------------------------|-----------------------------------------------------|
| Gender       | 305                               | 716                                 | 170              | 1,282(1,170–1,394)            | 1,873(1,738–2,008)                                   |
| Male         | 226                               | 515                                 | 125              | 929(835–1,023)                |                                                     |
| Female       | 79                                | 201                                 | 45               | 350(291–409)                  |                                                     |
| Total        | 305                               | 716                                 | 170              | 1,279(1,168–1,390)            | 1,869(1,734–2,004)                                   |
| Age <15      | 17                                | 80                                  | 18               | 76(–)                         |                                                     |
| 15–44        | 256                               | 570                                 | 131              | 1,111(995–1,227)              |                                                     |
| 45–64        | 29                                | 54                                  | 18               | 86(76–96)                     |                                                     |
| >65          | 3                                 | 12                                  | 3                | 12(12–12)                     |                                                     |
| Total        | 305                               | 716                                 | 170              | 1,285(1,173–1,397)            | 1,877(1,741–2,013)                                   |
| Road user type |                                  |                                     |                  |                               |                                                     |
| Driver       | 14                                | 85                                  | 14               | 85(85–85)                     |                                                     |
| Occupant     | 214                               | 337                                 | 83               | 864(738–990)                  |                                                     |
| Pedestrian   | 74                                | 255                                 | 72               | 262(254–270)                  |                                                     |
| Bicycle/motorcycle rider | 3   | 39                                  | 1                | 79(–)                         |                                                     |
| Total        | 305                               | 716                                 | 170              | 1,290(1,177–1,403)            | 1,885(1,748–2,022)                                   |
| Collision type |                                  |                                     |                  |                               |                                                     |
| With other vehicle | 116 | 190                                 | 47               | 465(377–553)                  |                                                     |
| Rollover     | 109                               | 204                                 | 36               | 538(415–661)                  |                                                     |
| Pedestrian   | 74                                | 263                                 | 73               | 267(261–273)                  |                                                     |
| Others       | 6                                 | 59                                  | 14               | 27(–)                         |                                                     |
| Total        | 305                               | 716                                 | 170              | 1,297(1,183–1,411)            | 1,895(1,757–2033)                                   |

Analysis showed that the rate of ascertainment by police records and hospital surveillance were incomplete. Road traffic incident deaths and injuries were significantly under-reported by traffic police and hospitals. Police records preferentially reported deaths, while hospital records preferentially reported injuries. The results found in this study are consistent with previous research, in low-income countries traffic police data are more reliable for fatal deaths and injuries were significantly under-reported by traffic police and hospitals. Police records preferentially reported deaths, while hospital records preferentially reported injuries. The results found in this study are consistent with previous research, in low-income countries traffic police data are more reliable for fatal injuries than hospital data, but for non-fatal injuries, hospital data capture more cases than do police records [9,12,25,26].

Low death reporting by hospitals may be explained by poor community awareness on the importance of medical evidence, particularly when the crash occurred in rural areas [7]. On the other hand, the transportation cost of the body to a health facility and back to the village could be also another challenge, especially for serious and fatal cases. Police need medical evidence from hospitals to bring the issue to the court. As a result the ascertainment corrected number will be underestimated. Therefore, the obtained estimates of the total number of casualties must be interpreted as lower bounds [33]. The third assumption is that the capture history of all cases should be independent, (i.e., reporting to one source does not influence the chance of the event being reported to the other source). This assumption is the most difficult to attain in capture-recapture study [12]. Researchers recommended that using three or more sources to control the effect of independency [31,32]. However, we couldn’t do this because of data inaccessibility. We believe some positive dependencies in this regard, since it is a medico-legal issue, especially for serious and fatal cases. Police need medical evidence from hospitals to bring the issue to the court. As a result the ascertainment corrected number will be underestimated. Therefore, the obtained estimates of the total number of casualties must be interpreted as lower bounds [33].

For a capture-recapture estimate to be valid at least the following four assumptions should be respected [29]. The first assumption is that the population should be closed. This assumption may not be a serious concern; since there is very little delay between the police attending the crash scene and the casualty’s visit to the hospital: at most a few days for the most slightly injured casualties [30]. The second assumption is sources are independent, (i.e., reporting to one source does not influence the chance of the event being reported to the other source). This assumption is the most difficult to attain in capture-recapture study [12]. Researchers recommended that using three or more sources to control the effect of independency [31,32]. However, we couldn’t do this because of data inaccessibility. We believe some positive dependencies in this regard, since it is a medico-legal issue, especially for serious and fatal cases. Police need medical evidence from hospitals to bring the issue to the court. As a result the ascertainment corrected number will be underestimated. Therefore, the obtained estimates of the total number of casualties must be interpreted as lower bounds [33]. The third assumption is that the capture history of all cases should be independent, (i.e., reporting to one source does not influence the chance of the event being reported to the other source). This assumption is the most difficult to attain in capture-recapture study [12]. Researchers recommended that using three or more sources to control the effect of independency [31,32]. However, we couldn’t do this because of data inaccessibility. We believe some positive dependencies in this regard, since it is a medico-legal issue, especially for serious and fatal cases. Police need medical evidence from hospitals to bring the issue to the court. As a result the ascertainment corrected number will be underestimated. Therefore, the obtained estimates of the total number of casualties must be interpreted as lower bounds [33].

For a capture-recapture estimate to be valid at least the following four assumptions should be respected [29]. The first assumption is that the population should be closed. This assumption may not be a serious concern; since there is very little delay between the police attending the crash scene and the casualty’s visit to the hospital: at most a few days for the most slightly injured casualties [30]. The second assumption is sources are independent, (i.e., reporting to one source does not influence the chance of the event being reported to the other source). This assumption is the most difficult to attain in capture-recapture study [12]. Researchers recommended that using three or more sources to control the effect of independency [31,32]. However, we couldn’t do this because of data inaccessibility. We believe some positive dependencies in this regard, since it is a medico-legal issue, especially for serious and fatal cases. Police need medical evidence from hospitals to bring the issue to the court. As a result the ascertainment corrected number will be underestimated. Therefore, the obtained estimates of the total number of casualties must be interpreted as lower bounds [33].
as easily calculated by capture recapture method with other official sources.

In conclusion neither of the two sources independently provided accurate coverage of road traffic related deaths and injuries. Strengthening both systems is necessary to obtain accurate information on road incident and human causalities.

Acknowledgments

We would like to thank Hawassa University and Addis Continental Institute of Public health for this research opportunity. The Authors are grateful to management bodies and police officers involved in the data collection process.

Author Contributions

Conceived and designed the experiments: TA YB AW A. Assrat A. Assefa. Performed the experiments: TA. Analyzed the data: TA. Wrote the paper: TA YB AW A. Assrat A. Assefa. Assisting the experiments: YB AW. Principal investigator: TA.

References

1. Sleet D, Baldwin G, Dellinger A, Dinh-Zarr B (2011) The Decade of Action for Global Road Safety. Journal of Safety Research 42: 147–148.
2. World Health Organization (2008) The Global Burden of Disease: 2004 update. World Health Organization Geneva.
3. Mohan D (2002) Road safety in less-motorized environments: Future concerns. Int J Epidemiol 31: 527–532.
4. Nantulya VM, Reich MR (2002) The neglected epidemic: Road traffic injuries in developing countries. BMJ 324: 1139–1141.
5. Ameratunga S, Hijar M, Norrish R (2006) Road-traffic injuries: Confronting disparities to address a global-health problem. Lancet 367: 1533–1540.
6. World Health Organization (2009) Global status report on road safety: time for action. World Health Organization Geneva.
7. Pedes M, Scarfied R, Steet D, Mohan D, Hyder A, et al. (2004) World report on road injury prevention. Geneva: World Health Organization.
8. Sasse S, Varghese M, Killermann A, Lormand J (2005) Prehospital trauma care systems. Geneva: World Health Organization.
9. Tercero F, Anderson R, Petha R, Swanstrom L (1998) Can valid and prevention-oriented information on injury occurrence be obtained from existing data sources in developing countries? An example from Nicaragua. International Journal of Consumer Product Safety 5: 99–105.
10. Samuel JC, Sankhulami E, Ngeshi JS, Babusi P, Thungi C, et al. (2012) Under-Reporting of Road Traffic Mortality in Developing Countries: Application of a Capture-Recapture Statistical Model to Refined Mortality Estimates. PloS ONE 7: e31091.
11. HungTrinhVan, Singhisavanon P, Kaewkungwal J, Suriyasompasian P, Khai LH (2006) Estimation of non-fatal road traffic injuries in Thai Nguyen, Vietnam using capture-recapture method. Southeast Asian J Trop Med Public Health 37: 405–411.
12. Tercero F, Anderson R (2004) Measuring transport injuries in a developing country: an application of the capture-recapture method. Accident Analysis and Prevention 36: 13–20.
13. Pollock K (1991) Modeling capture, recapture and removal statistics for rish and wildlife populations. J Amer Stat Assn 86: 225–239.
14. Laure V, Steibhanie L (2008) Capture-recapture estimates of the local prevalence of problem drug use in six French cities. European Journal of Public Health 19: 32–37.
15. Rongsheng L, Boheng L, Ping Y, Li F, Yong H, et al. (2005) A study on capture-recapture method for estimating the population size of injecting drug users in South West China. Journal of Health Science 5: 405–409.
16. Brugh RF, Swan AV, Hayhurst GK, Fallon MP (1998) A drug misuser prevalence study in a rural English district. European Journal of Public Health 8: 34–36.