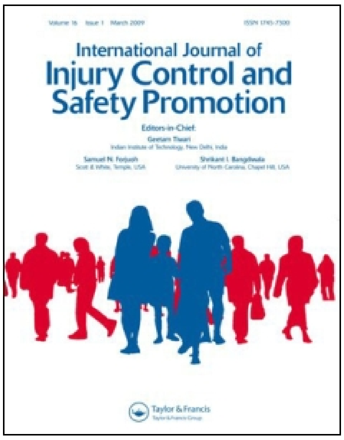

International Journal of Injury Control and Safety Promotion
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713734346

Methods for developing country level estimates of the incidence of deaths and non-fatal injuries from road traffic crashes

Kavi Bhalla [a]; Saeid Shahraz [a]; David Bartels [a];Jerry Abraham [a]
[a] Department of Global Health and Population, Harvard School of Public Health, Boston, MA, USA

Online publication date: 13 November 2009

International Journal of Injury Control and Safety Promotion
Vol. 16, No. 4, December 2009, 239–248

Methods for developing country level estimates of the incidence of deaths and non-fatal injuries from road traffic crashes

Kavi Bhalla*, Saeid Shahraz, David Bartels and Jerry Abraham

Department of Global Health and Population, Harvard School of Public Health, Boston, MA, USA

(*Received 20 March 2009; final version received 9 October 2009*)

The estimates of the incidence of deaths and non-fatal injuries from road traffic crashes are essential inputs for prioritising national health and transport policies. This article sketches a methodology for assembling such estimates at the country level by piecing together data from a wide array of sources that include death registers, hospital records, funeral records, health surveys and police reports. Using examples of the types of data sources available in four developing countries (Iran, Mexico, Ghana and India), methods are proposed for making these consistent and extrapolating to estimates of injury incidence at the country level. This requires filling information gaps, mapping from varying case definitions, deriving population-based incidence estimates from sources that may not track denominator populations, and appropriately reapportioning cases assigned to poorly specified causes. The principles proposed here will form the methodological basis for a series of country reports to be published in the future.

Keywords: road traffic injuries; low-income countries; injury surveillance; health metrics

The need for country estimates

Injuries from road traffic crashes resulted in an estimated 1.2 million deaths worldwide in the year 2002 (Global Burden of Disease 2002). While high-income countries have witnessed decreasing road traffic death rates for several decades now, there are strong indications that with continued economic growth and the concomitant increase in the motor vehicle fleet, the problem will worsen in most low- and middle-income countries (Bhalla, Ezzati, Mahal, Salomon, & Reich, 2007; Kopits & Cropper, 2005; Mathers & Loncar, 2006).

As highlighted by the 2004 World Report on Road Traffic Injury Prevention (Peden et al., 2004) issued jointly by the World Bank and the World Health Organization, controlling the problem would require a concerted effort by national transport and health agencies. Unfortunately, at present, the need to reduce road traffic injuries is not high on the policy agenda of most developing countries. This may be partly due to the underestimation of road traffic injuries in official government statistics relative to other health problems. One reason for the underestimation is that most developing countries rely on police reports for their official road traffic injury statistics even though there is mounting evidence of police underreporting (Aeron-Thomas, 2000; Elvik & Mysen, 1999; London, Mock, Abantanga, Quansah, & Boateng, 2002). To make matters worse, various international agencies are attempting to systematise road traffic injury data by requesting countries to report official statistics (usually from police) to global and regional databases (Global Status Report on Road Safety; Road Safety in Asia and the Pacific). Unless these official statistics are corrected for underreporting, such rechanneling of spurious country-reported data through the international agencies carries the danger of lending increased legitimacy to bad data and underestimating the global burden of road traffic injuries.

We recommend an alternate approach to estimating national road traffic injury statistics by analysing sources other than police reports, including death registers, hospital records, funeral records, injury surveillance studies and health surveys, and triangulating to provide a national snapshot of the incidence of fatal and non-fatal road traffic injuries. These estimates should then be compared with police-based figures to test the validity of official government statistics.

In a recent publication (Bhalla, Naghavi, Shahraz, Bartels, & Murray, 2009), we have demonstrated this

*Corresponding author. Email: kavi_bhalla@harvard.edu

ISSN 1745-7300 print/ISSN 1745-7319 online

DOI: 10.1080/17457300903402184
http://www.informaworld.com

process of building a national estimate of the incidence of road traffic injuries from available data sources in Iran. We now generalise this methodology to a wider array of data sources that may exist in other countries using examples of the types of data sources identified in an ongoing environmental scan of data sources useful for such work (Global Burden of Disease Injury Expert Group, 2009). Because many of these datasets are not intended for such use, deriving incidence rates requires analytical adjustments to make the estimates compatible. The analytical process includes filling information gaps, adjusting for completeness and coverage, and the redistribution and reclassification of cases coded to ill-defined causes. We suggest how such issues should be addressed and describe the process for triangulating to a national estimate from the patchwork of data that may be available in a country.

The methodology for country assessment of road traffic injuries proposed in this article has two key goals – first, to generate reliable assessments of the public health burden of road traffic injuries, and second, to characterise the nature of this burden (i.e. by age, sex, external cause, urban *vs.* rural, etc.) in order to provide the evidence basis for setting national road safety priorities. We start by defining the key concepts of interest. Next, we describe methods for developing estimates for deaths and non-fatal injuries by taking advantage of the strengths of the key data sources, discuss how to fill information gaps and estimate in the absence of data sources. Finally, we close with a discussion that places our proposed methodology in the broader work on injury surveillance.

Definitions – what are we measuring?

The definitions of concepts related with road traffic injuries vary substantially and can lead to difficulties in comparative analysis (Cercarelli, Rosman, Bsc, & Ryan, 1996; Jacobs, Aeron-Thomas, & Astrop, 2000; Ward & Christie, 2000). A detailed review of the definitions literature as it relates to such work has been described elsewhere (Bhalla, 2007). Table 1 summarises the definitions adopted, and this section summarises the main issues.

As shown in Figure 1, injury events are a subset of all crash events and can be further subdivided based on the severity of the injury. The characteristics of road traffic crashes can be further classified depending on the dimension of interest – such as demographic characteristics of the victim, injury severity and type of institutional medical care received. Our definition of a 'road traffic crash' (see Table 1) is derived from the Tenth Revision of the International Statistical Classification of Disease and Related Health Problems (ICD-10) (Manual of the International Statistical Classification of Diseases and Related Health Problems, 1992) because it provides a more inclusive definition for what constitutes a 'road' particularly as related to low- and middle-income countries. The definition does not restrict a 'road' to a path prepared for vehicle use but includes any public path (including, for example a path in a rural field) that is customarily

Table 1. Definitions for describing the national burden of road traffic crashes.

Concept	Concept definition
What is a 'road traffic crash?'	A road traffic crash is an event that produces injury and/or property damage, involves a vehicle in transport, and occurs on a road or while the vehicle is still in motion after running off the road. Note the use of the term 'crash' instead of 'accident'
What is a 'road?'	A road is the entire width between property lines (or other boundary lines) of land open to the public as a matter of right or custom for purposes of moving persons or property from one place to another
Types of road user transport mode	Pedestrian, bicycle, motorised two wheeler, motorised three wheeler, car, van (including pick up trucks), truck, bus, others
What is a road traffic 'fatality?'	Any death for which a road traffic crash is the underlying cause. The 'underlying cause' of a death is the disease or injury which initiated the train of events leading directly to death regardless of how long ago the event occurred. Note that there is no time limit between the crash and the death. There is also no restriction on where the death happens (at crash scene, hospital, home, etc.)
What is a road traffic 'injury?'	A road traffic injury is an injury caused in a road traffic crash. 'Injury' is the reduction in functional health status due to energy exchanges that have relatively sudden discernable effects
Levels of injury severity	Injury severity is defined in terms of the levels of impairment – i.e. reduction in functional health status – e.g. minor/moderate/severe impairments
Types of institutional medical care	A hospital admission is a hospital stay exceeding 24 h. Visits less than 24 h are referred to as outpatient visits
Victim demographics	Sex: male/female Age grouping: 0–1, 1–4, 5–14, 15–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75–84 and 85+ years Urban/rural: national definitions of 'urban' and 'rural' areas as defined by the UN Demographic Yearbook

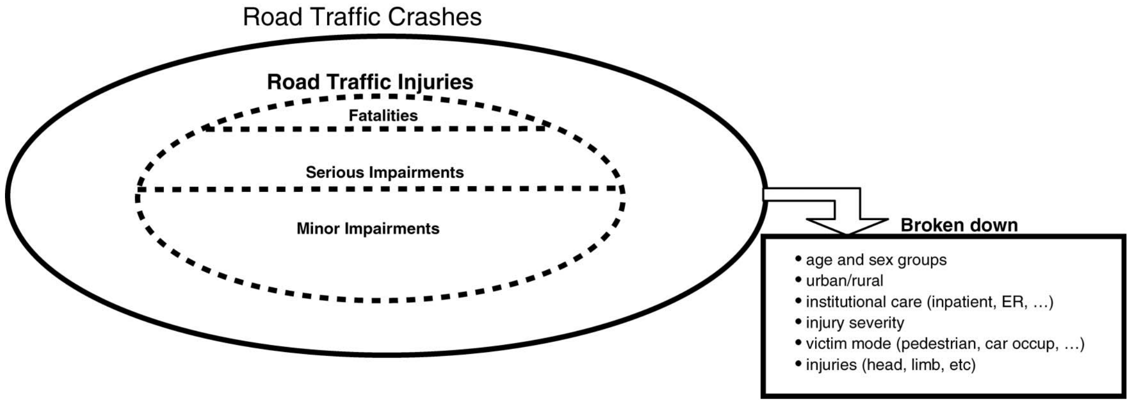

Figure 1. Road traffic crashes and their outcomes.

used for transport in the community. Similarly, the term 'road traffic crashes' encompasses all crashes that occur on the road, regardless of whether they involved motorised or non-motorised vehicles. Note also that the term 'crash' is used instead of 'accident' as has been recommended often in the road safety literature (Davis & Pless, 2001; Neira & Bosque, 2004).

The definition of what constitutes a road traffic death based on how soon after the crash the death occurs has been discussed extensively in the literature (Elvik & Vaa, 2004; Jacobs et al., 2000; Luoma & Sivak, 2007; Odero, Garner, & Zwi, 1997). These time restrictions (death within 1 day, 1 week, 1 month and 1 year) are usually operational definitions that take into account the practicality of reporting for the agency collecting the data. Conceptually, however, there should not be any time restriction on an underlying cause of death. Furthermore, the methods proposed in this article rely extensively on estimating mortality based on death registration data, which typically are coded using ICD rules (Manual of the International Statistical Classification of Diseases and Related Health Problems, 1992) that do not include a time-based restriction on the underlying cause of death. Thus, not adopting a time-based restriction is also operationally easier in such work. Finally, it should be noted that translating between estimates that rely on the various definitions using time-to-death ratios (Jacobs et al., 2000) is relatively straightforward. These time-to-death ratios (Jacobs et al., 2000) also show that only a trivially small fraction ($\sim 3\%$) of road injury deaths happen after 30 days from a crash. The uncertainties introduced in such estimations are substantially smaller than those induced from the other analytical manipulations described in this article.

In comparison to deaths, non-fatal injury outcomes are more difficult to define and classify. Langley and Brenner (2004) propose an energy-based definition for injury – 'the damage produced by energy exchanges that have relatively sudden and discernable effects'. Defining thresholds for levels of injury severity is conceptually difficult and, as a result, has received substantial attention by injury researchers (Baker, O'Neill, Haddon, & Long, 1974; Copes et al., 1988; Gennarelli & Wodzin, 2005, 2006; Osler, Baker, & Long, 1997). The crudest injury classifications are dictated by the operational simplicity of classifying injury severity by hospital admissions and outpatient care. However, especially in comparative work across countries, such a definition is problematic because access to care can vary substantially. Instead, a conceptual definition should be based on medical pathology. The Abbreviated Injury Scale (AIS) (Gennarelli & Wodzin, 2005) and its derivative scales (Copes et al., 1988; Gennarelli & Wodzin, 2006; Osler et al., 1997) are the most commonly used injury severity classifications. However, these existing measures have threat to life as their central focus and do not effectively describe the loss of functional health status that can result from non-fatal injuries. Thus, for instance, an injury requiring an amputation that results in life-long disability, and hence a substantial health burden, can have the same AIS level as an injury with no discernable disability a few days after the event. This is a severe shortcoming of AIS-based injury scales for characterising the public health burden of non-fatal injuries. There is a growing body of work (as reviewed by Segui-Gomez and MacKenzie (2003)) that focuses on measuring the long-term health impairments due to non-fatal injuries. Of the various available metrics for quantifying public health burden, we recommend using the disability adjusted life year (DALY), developed as part of the global burden of disease project. The DALY is a time-based measure that combines years of life lost due to premature death and years of life lost due to life in less than ideal health states. DALYs allow comparing the health burden of injuries with that of other diseases and thus provide a common metric that can be used to estimate the health and economic loss with other health issues.

Finally, Table 1 proposes the use of definitions for victim demographics so as to be consistent with similar work being done for other diseases in the field of health metrics (Institute for Health Metrics and Evaluation at the University of Washington, 2008).

Developing estimates

Types of data sources

Table 2 illustrates the types of data sources (identified as part of an ongoing environmental scan (Global Burden of Disease Injury Expert Group, 2009)) that may be available for such work in a selected set of four low- and middle-income countries spanning four continents: Iran, India, Mexico and Ghana. Table 2 explicitly excludes police-based reporting systems, which exist in all four countries and have traditionally formed the basis of official government statistics.

Data sources for estimating deaths

Data collected by national death registration systems are often the most comprehensive source (in terms of population coverage) for cause of death data at the country level. Such data systems exist in many developing countries although the coverage and completeness of reporting and the quality of cause of death attribution can vary substantially (Lopez, AbouZahr, Shibuya, & Gollogly, 2007; Mahapatra, Shibuya, & Lopez, 2007). Table 2 shows that Iran and Mexico have death registration systems that have high population coverage and are relatively complete. In the absence of national death registration systems, some countries rely on sample registration systems, which rely on a representative sample of deaths. In India, for instance, the Survey of Causes of Death-Rural (Registrar General of India, Vital Statistics Division, 1999) (SCD-Rural) uses verbal autopsy performed by trained paramedics to track causes of deaths in representative rural communities and reports the causes of death (including one category for road traffic injuries) for approximately 50,000 rural deaths annually. For urban areas in India, cause of death information is available from hospitals that report to the national Medical Certification of Cause of Death (Registrar General of India, Vital Statistics Division, 2006) (MCCD) system. Although the sample of deaths included is not representative, it reports approximately 500,000 deaths annually (an estimated 30% of all urban deaths). Finally, some regions of the world, notably Africa (represented by Ghana in Table 2) do not have functional national death registration systems that can provide cause of death statistics. In such cases, community health surveillance projects (such as the International Network of Field Sites with Continuous Demographic Evaluation of Populations and Their Health (INDEPTH) network of disease surveillance system sites (Adjuik et al., 2006)) may be able to provide useful insight into the patterns of causes of death from which road traffic injury death rates can be estimated.

Methods for estimating road traffic deaths

Since most data systems described above rarely provide a full enumeration of all deaths to clearly defined and meaningful causes of death, analytical adjustments are usually necessary for deriving national estimates of road traffic mortality (Figure 2).

Adjusting for completeness and coverage

While most death registration systems aspire to register all deaths that happen in a country, estimates of the true completeness and coverage of these systems are often not readily available along with the datasets. However, these can be estimated by comparing the total number of deaths reported by the registration system with the total national deaths, estimates for which are often available from other, more reliable sources, such as the national population census and surveys processed using standard demographic methods. The United National Population Division (United Nations World Population Prospects, 2004) is a useful source for estimates of all-cause deaths in all countries. Because completeness can vary substantially by age and sex, adjustment for completeness should be done using age–sex specific completeness estimates. Similarly, estimates from a sample registration system, such as India's SCD-Rural (Registrar General of India, Vital Statistics Division, 1999), can be obtained by applying age–sex specific cause of death patterns from the representative sample to the national all-cause death distribution.

Reclassification of deaths attributed to partially specified causes of death

Death registration systems typically code underlying causes of death using the ICD framework, which provides a range of codes to allow for coding deaths for which the full details of causes may not be known. Thus, in the context of estimating national road traffic deaths for the various road users, the death could be fully specified (e.g. identified as 'car occupant') or only partially specified (e.g. identified as 'road traffic death' without information of type of road-user) or completely unspecified (i.e. with no cause specified). In fact, the structure of the ICD cause of death codes has a

Table 2. Selected[a] data sources for estimating the incidence of deaths and injuries from road traffic crashes in four countries.

Country	Deaths	Non-fatal injuries
Iran	*National death registration system* (Bhalla et al., 2009): covers 29 provinces (i.e. all except Tehran); ICD-10 derivative causes of death *National forensic medicine system* (Bhalla et al., 2009): estimates available for all provinces	*Hospital data sample* (Bhalla et al., 2009): Data collected from all hospitals in 12 provinces (outpatient for 4 days, and hospital admissions for 4 weeks), followed back to household post-discharge *Demographic and Health Survey (DHS)* (2000): Approximately 110,000 households, included questions about road traffic injury involvement and care.
India (Global Burden of Disease, 2002)	*National Sample Registration System* (Registrar General of India, Vital Statistics Division, 1999): Nationally representative sample of deaths in India, causes evaluated by verbal autopsy *National Medical Certification of Cause of Death (MCCD) System* (Registrar General of India, Vital Statistics Division, 2006): Cause of death for reporting hospital in urban areas; covers approximately 500,000 deaths from all causes annually.	*World Health Survey (WHS)* (Ustun, Chatterji, Mechbal, Murray, & WHS Collaborating Groups, 2003): representative sample with questions about road traffic injuries and care; conducted in six states *Survey – New Delhi* (Verma & Tewari, 2004): 5412 households, all injury causes *Survey – Bangalore* (Aeron-Thomas, 2000): 20,000 households, stratified by urban/rural and socio-economic status *Survey – near New Delhi* (Mohan & Qadeer, 1990): morbidity patterns in nine villages, 25,000 households, monitored for 1 year *Hospital – Hyderabad* (Global burden of disease injury expert group, 2009): five hospitals, Approximately 800 victims, followed back to household post-discharge
Mexico	*National death registration system:* ICD-10 coded cause of death, estimated to be near complete	*SAEH – Ministry of Health national hospital discharge database*: covers all Ministry of Health hospitals, approximately 115,000 unintentional injury hospital admissions *Instituto Mexicano del Seguro Social (IMSS) national hospital discharge database*: approximately 175,000 injury hospital admissions; external causes not recorded *World Health Survey* (Ustun et al., 2003): representative sample with questions about road traffic injuries and care *Encuesta nacional de Salud y Nutricion (ENSANUT) national health survey* (Avila-Burgos et al., 2008): Sample size 54,068 individuals, included questions on RTI involvement and care
Ghana	*Mortuary data – Kumasi* (London et al., 2002): Data collected from 1996 to 1999 *Demographic Surveillance System (DSS) Sites at Navrongo* (Adjuik et al., 2006): verbal autopsy-based cause of deaths	*World Health Survey* (Ustun et al., 2003): representative sample with questions about road traffic injuries and care *Survey – Kumasi (urban) + Brong Ahafo region (rural)* (Mock, Abantanga, Cummings, & Koepsell, 1999): sample of approximately 21,000 individuals *Hospital records – Accra*: Reporting hospitals

[a]For a full listing see our ongoing environmental scan (Global Burden of Disease Injury Expert Group, 2009) of data sources.

hierarchy of specificity that corresponds to a hierarchy of information content about the causes of death. For instance, in ICD-10, road traffic injuries are coded to the following classes of partially specified causes (listed in the order of decreasing specificity):

- Level 1: road traffic injury – victim's mode of transport not known – excluding pedestrian and bicyclists (ICD-10: V87–V88).
- Level 2: road traffic injury – victim's mode of transport not known (ICD-10: V89).
- Level 3: unspecified transport injury (ICD-10: V99, Y85).
- Level 4: unspecified unintentional injury (ICD-10: X59).
- Level 5: unspecified event, undetermined intent (ICD-10: Y34).
- Level 6: unspecified causes of death (ICD-10: R95–R99).

As a general rule, deaths coded to partially specified causes of death should be reapportioned to

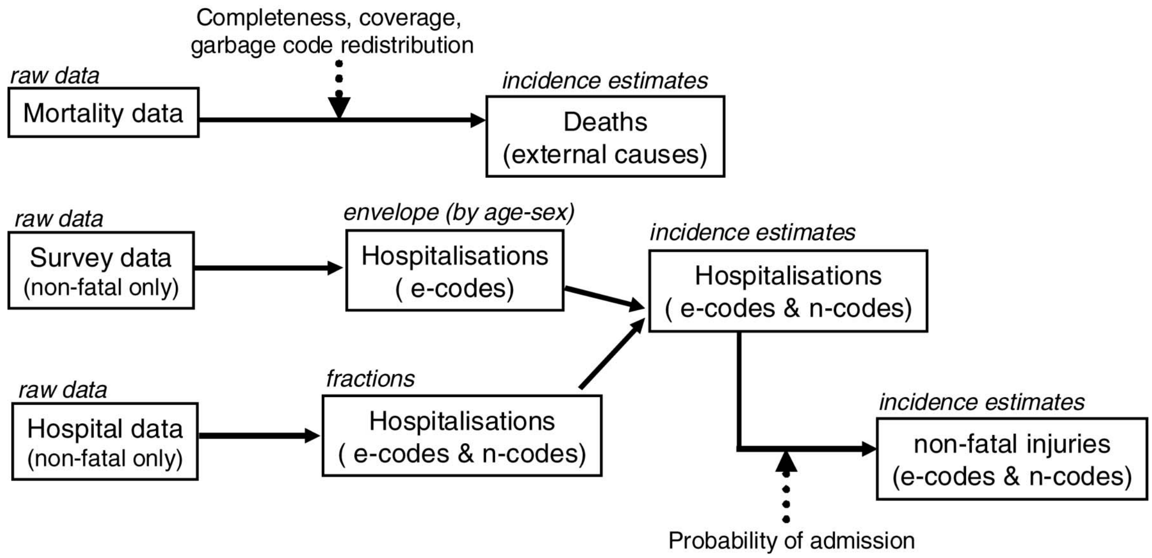

Figure 2. Combining various data sources for estimating national incidence of deaths and injuries from road traffic crashes.

fully specified causes using all available information. Such redistribution should be done within age–sex categories to limit the errors introduced due to the potential biases in the causes of deaths coded to unspecified categories. As an example, in our analysis of death registration data in Iran (Bhalla et al., 2009) we found that among all deaths classified to transport injuries, 6.2% were not further specified. We redistributed these partially specified deaths over all transport categories in proportion (pro-rata) to the distribution of deaths in the transport categories within age–sex groups. Thus, in the listing above, Level 1 unspecified deaths should be redistributed over all sub-categories of road traffic injuries except pedestrian and bicyclist deaths, and Level 3 unspecified deaths should be redistributed over all sub-categories of road traffic injuries and specified non-road traffic transport injuries.

Aggregating and extrapolating from sub-national estimates

Often the data systems for estimating mortality are intended to have sub-national coverage or can only provide reliable statistics for sub-national regions. In such cases, results need to be aggregated to the national level. One example, is using SCD-Rural in India to estimate rural deaths and MCCD data to estimate urban deaths, to come up with a national estimate in India (e.g. Sanghavi, Bhalla, & Das, (2009)). A similar process may be necessary in regions of Africa, where there is limited death registration data available only from selected urban areas but various Demographic Surveillance System (DSS) sites track mortality in selected rural communities. Thus, separate estimates of the pattern of urban and rural deaths combined with all-cause death estimates could be used to estimate plausible estimates of road traffic mortality.

Similarly, in our analysis of road traffic deaths in Iran in 2005 (Bhalla et al., 2009), we found that the Iranian Death Registration System provided reliable estimates for 29 of 30 provinces (excluding Tehran, the most populous province in Iran). However, estimates for Tehran province were available from the national forensic medicine data systems allowing aggregation to a national snapshot.

Data sources for estimating non-fatal injuries

The two primary sources for estimating non-fatal injuries are hospital datasets and health surveys, both of which can be either national or sub-national. As illustrated in Table 2, often countries have some combination of both types of datasets. Thus, this section provides a brief summary of the strengths of each type of data source and then proposes methods for making optimal use of both to estimate the incidence and burden of non-fatal injuries.

Relative pros and cons of hospital data and survey data

Hospital data. Hospital datasets contain detailed medical descriptions of injuries that make them useful for characterising the disability outcomes and, thus, the public health burden of injuries. However, they rarely include complete coverage of a population and, thus, can be difficult to use for estimating population incidence of hospitalised injuries.

Typically, hospital registries involve pooled data from a set of collaborating hospitals, which are unlikely to have a well-defined catchment area. Since victims with more serious injuries may be transported from further away than victim with relatively less serious injuries, the catchment area may be a function of injury severity. Furthermore, road injury cases at hospitals may also include a transit population of motor vehicle users rather than residents. One notable exception to these shortcomings of hospital-based data collection is the Iranian hospital data sample, which was systematically gathered from all hospitals in 12 (of 30) Iranian provinces for a limited time period, thus allowing estimation of population rates.

Another potential shortcoming of hospital data is that often the coding of nature of injuries (e.g. traumatic brain injury, tibia fracture, etc.) is of much higher quality than external causes (e.g. pedestrian, car occupant, etc.). Although this bias in coding in hospital administrative records largely reflects the primary focus of hospitals (i.e. providing medical care for injuries sustained), external causes are more important for policy analysis purposes. In the worst cases, hospitals do not report external causes at all (see for instance, Mexico-IMSS hospitals in Table 2). Advanced analytical methods, such as the use of Bayesian inference (Bhalla, Shahraz, Naghavi, Lozano, & Murray, 2008) or multinomial logistic regression models (Shahraz, Bhalla, Bartels, Naghavi, & Lozano, 2008), that allow harnessing other information about injuries available in hospital records to derive more accurate estimates of external causes are currently in development.

Even when a hospital registry tracks all hospital admissions for a population with high-quality cause coding, such data can at best be used to estimate the incidence of injury hospitalisations. However, the rate of hospitalisations is a poor proxy for incidence of severe injuries especially when comparing injury rates across countries with populations that have vastly differing access to medical care.

Surveys. Surveys that ask questions about injuries resulting from road traffic crashes are ideally suited for estimating the population incidence of non-fatal injuries. Such surveys can be specific to injuries (e.g. in Table 2, India: New Delhi and Bangalore; Ghana: Kumasi and Brong Ahafo) or be broader national health surveys (e.g. in Table 2, Demographic and Health Survey in Iran; World Health Survey in Mexico, India, and Ghana; and ENSANUT in Mexico). Injury-specific surveys typically measure more details about road traffic crashes, but often tend to be sub-national. For our purposes, national surveys, even if they only include a few questions about injuries, have the advantage of providing direct estimates of the national incidence of non-fatal injuries. While estimates from surveys have well-documented problems with potential for recall biases and differential item functioning (King, Murray, Salomon, & Tandon, 2004), our current work suggests that these are less of a problem when used to estimate the incidence of road traffic injury hospitalisations. Surveys that include questions that define injury incidence with reference to a particular impairment threshold (e.g. functional loss for more than 1 day) rather than type of medical care received (e.g. inpatient, outpatient), allow a measurement of injury incidence that is not affected by access to medical care in the community. Given the array of data sources typically available (Global Burden of Disease Injury Expert Group, 2009), surveys have the potential to be the most reliable source of estimating population incidence of non-fatal injuries. Unfortunately, our review of survey instruments suggests that there is little standardisation in injury questions, which makes comparative analysis with survey measurements difficult.

Because surveys collect information from the victim, they can only provide limited descriptions of the nature of injuries, making assessment of injury severity and public health burden difficult. Instead, surveys typically include good information of the external causes of the injury event.

Estimating the incidence of injuries by coupling surveys and hospital data

Given the relative advantages and disadvantages of surveys and hospital registries, one efficient coupling of these two data sources is to estimate the population incidence of road injury hospitalisations from surveys and use the hospital data for estimating the fractional breakdown of the survey-derived injury incidence (Figure 2). Next, the estimates of injury hospitalisations should be adjusted using injury-specific probability of admission (see Figure 2) to get a measure of incidence that is independent of access to medical care. This model harnesses the primary advantage of surveys – i.e. that they are population representative – as well as that of hospitals in that they provide detailed (often ICD coded) descriptions of injury outcomes. Survey-based incidence estimates should be at the level of detail (e.g. age–sex–external cause specific) supported by the survey data collection, which is a function of the sample size of the survey as well as the questions asked in the survey.

Estimating in the absence of data

When a region or country has no data sources for estimating deaths and/or non-fatal injuries from road traffic crashes, it may be necessary to extrapolate from knowledge and insights from other regions where multiple data sources are available. Because statistical extrapolation carries the danger of introducing substantial uncertainty in the estimates, we recommend the use of analytical tools that rely on the structural models of injury causation (e.g. Bhalla et al. (2007)). Such models recognise that the distribution of injuries and the resulting probability of death are a function of the crash characteristics (most importantly, the mode of transports involved) and the biological frailty of the victim (characterised by age and sex). Two such tools are injury pyramids and mappings from the external cause (E) to nature of injuries (N) mappings.

Injury pyramids (ratio of deaths to non-fatal injuries)

When data source availability allows for the prediction of only deaths or non-fatal injuries, injury pyramids can be used to fill in the information gap. Ideally, injury pyramids should be specific to the matrix of combinations described by the victim mode of transport and impacting vehicle (Bhalla et al., 2007), but usually much of this variation can be captured by focusing solely on the victim mode of transport – i.e. separate injury pyramids should be developed for pedestrians, bicyclists, car occupants, etc. Another source of error in the use of injury pyramids is the variation in case fatality rates, which can be addressed partly by the use of age–sex specific injury pyramids.

E → N mappings (distribution of injury outcomes for each external cause)

Estimating the distribution of injury outcomes (e.g. traumatic brain injury, skull fracture, etc.) is an important component of assessing the public health burden of injuries. When hospital data are not available in a region, the incidence of injury outcomes cannot be directly estimated. However, it is reasonable to assume that the pattern of injuries sustained by victims of, for example, pedestrian crashes, does not differ substantially by region. Thus, mappings that relate external causes to nature of injuries developed from medical facility data recorded in other regions can be used for such extrapolation. As with injury pyramids, such mappings should be age–sex specific, as well as specific to the victim's mode of transport.

Discussion

This article outlines a detailed methodology for building a national snapshot of the incidence of deaths and injuries from road traffic crashes from sources other than police reports, which are usually the source for official government statistics. Such estimates are important to derive because of the large body of evidence that police reports underestimate road traffic injuries. For non-fatal injuries, police underreporting is substantial even in high-income countries (Amoros, Martin, & Laumon, 2007) and police-based injury counts are rarely trusted. However, for deaths, it is often assumed that police reporting is complete because in many countries injury deaths are required by law to be investigated by forensic pathologists. While the existence of such a legal framework could improve police recording, the extent to which this happens and the socio-political conditions that lead to high levels of investigation, recording and reporting of injury deaths by the police have never been systematically studied. The methodology we propose of building estimates from alternate sources is one way of critically evaluating the completeness of police-based injury statistics.

There are also other methods for evaluating the completeness of police data. Among these, the most common method is to link records with another independent data source (e.g. hospital, news media reports) and identify the fraction of cases that are missed by both sources (Razzak & Luby, 1998; Tercero & Andersson, 2004). Unfortunately, correct application of the method, which in epidemiology is called 'capture–recapture' or 'mark and recapture', hinges on a set of assumptions that are violated in typical applications to road traffic injuries (Jarvis, Lowe, Avery, Levene, & Cormack, 2000). In the context of developing countries, the most egregious violation is that the data sources used are typically inter-dependent. Victims that are listed in one administrative record (e.g. hospital or death register) have a much higher likelihood of being listed in another administrative record such as a police report. In the worst case scenario, if a hospital has a particularly strong tradition of cross-documenting injury admissions with the police, comparing hospital records with police records will misleadingly suggest high police coverage. Similarly, news reports are unlikely to be an independent data source especially when news agencies rely on police reports for the news stories. Application of capture–recapture using such closely inter-dependent sources produces results that suggest high coverage of police reporting even when this may not be true and risks providing false legitimacy to such record keeping.

Instead, our approach relies on building estimates independently from the police reports using the country-specific patchwork of alternate data sources. Any site with which a road traffic crash victim may come in contact (such as an ambulance system, emergency rooms, hospital wards, civil registration systems, funeral grounds) qualifies as a potential data source. However, appropriate use of these datasets requires careful analytical handling of a range of issues that include estimation of underlying populations for adjustment for completeness and coverage; mapping between various definitions of recorded injury cases; dealing with poor quality data – such as cases coded to poorly specified causes; and extrapolating to regions with missing information. Such analytical manipulations of data can introduce substantial uncertainty in the derived estimates. These uncertainties should be quantified to make cross-country comparisons, or the comparisons of multiple sources within the same country, meaningful. Since most of these issues are subjects of ongoing research and development in the field of health metrics (Alexandrescu, O'Brien, Lyons, & Lecky, 2008; Bhalla et al., 2008; Shahraz et al., 2008), this article provides current best practice recommendations as they apply to estimating road traffic injuries.

The application of this method results in a single snapshot of the incidence of road traffic injuries that can be used to test the validity of official government statistics and provide estimates that can be used for setting national road safety policy priorities. It is important to note that this is fundamentally different from injury 'surveillance', which refers to ongoing data collection (Holder, 2004) and is intended for tracking the growth/decay of a problem. While timeliness of reporting is important for surveillance, it matters less for our goals allowing us to use data sources such as surveys and death certificates which may not be considered timely for injury surveillance (Holder, 2004). On the other hand, in the few developing country settings where injury surveillance infrastructure has been established, the surveillance systems are typically incapable of directly reporting national estimates of the incidence of injuries. Most injury and trauma surveillance systems are established at selected hospital sites and are not intended to be able to generate population rates. Often based in urban settings, they can provide a biased view of the characteristics of the national road traffic injury problem. Thus, to be useful in setting national policies, data from such surveillance systems need to be subject to the same analytical adjustments along with other data sources as has been described in this article.

Although the approach we recommend here is intended to produce evidence to challenge official statistics, national governments could also incorporate this methodology as part of their long-term road traffic injury data strategy. In such a framework, countries would rely on the timeliness of the data provided by the injury surveillance systems to track the evolution of the road safety problem, while periodically performing evaluations from all other data sources to test the accuracy of their official statistics and to push for improvements.

Acknowledgements

This work was funded by a grant from the World Bank Global Road Safety Facility.

References

Adjuik, M., Smith, T., Clark, S., Todd, J., Garrib, A., Kinfu, Y., et al. (2006). Cause-specific mortality rates in sub-Saharan Africa and Bangladesh. *Bulletin of the World Health Organization*, *84*, 181–188.

Aeron-Thomas, A. (2000). *Underreporting of road traffic casualties in low income countries*. PR/INT/199/00. Crowthorne, Berkshire: Transportation Research Laboratory.

Alexandrescu, R., O'Brien, S.J., Lyons, R.A., & Lecky, F.E. (2008). A proposed approach in defining population-based rates of major injury from a trauma registry dataset: delineation of hospital catchment areas (I). *BMC Health Services Research*, *8*(1), 80.

Amoros, E., Martin, J.L., & Laumon, B. (2007). Estimating non-fatal road casualties in a large French county, using the capture–recapture method. *Accident Analysis and Prevention*, *39*(3), 483–490.

Avila-Burgos, L., Medina-Solís, C.E., Pérez-Núñez, R., Híjar-Medina, M., Aracena-Genao, B., Hidalgo-Solórzano, E., et al. (2008). Prevalence of non-fatal road traffic injuries in Mexico: results from ENSA-NUT 2006. *Salud Publica de Mexico*, *50*, 38–47.

Baker, S.P., O'Neill, B., Haddon, W. Jr., & Long, W.B. (1974). The injury severity score: a method for describing patients with multiple injuries and evaluating emergency care. *The Journal of Trauma: Injury, Infection, and Critical Care*, *14*(3), 187.

Bhalla, K. (2007). *Road traffic injury metrics – recommended definitions*. Cambridge, MA: Harvard University Initiative for Global Health Road Traffic Injury Metrics Group.

Bhalla, K., Ezzati, M., Mahal, A., Salomon, J., & Reich, M. (2007). A risk-based method for modeling traffic fatalities. *Risk Analysis*, *27*(1), 125–136.

Bhalla, K., Naghavi, M., Shahraz, S., Bartels, D., & Murray, C.J.L. (2009). Building national estimates of the burden of road traffic injuries in developing countries from all available data sources: Iran. *Injury Prevention*, *15*, 150–156.

Bhalla, K., Shahraz, S., Naghavi, M., Lozano, R., & Murray, C.J.L. (2008). Estimating the distribution of external causes in hospital data from injury diagnosis. *Accident Analysis and Prevention*, *40*(6), 1822–1829.

Cercarelli, L.R., Rosman, D.L., Bsc, D.C., & Ryan, G.A. (1996). Comparison of accident and emergency with police road injury data. *The Journal of Trauma: Injury, Infection, and Critical Care*, *40*(5), 805.

Copes, W.S., Champion, H.R., Sacco, W.J., Lawnick, M.M., Keast, S.L., & Bain, L.W. (1988). The injury severity score revisited. *The Journal of Trauma: Injury, Infection, and Critical Care*, *28*(1), 69.

Davis, R.M., & Pless, B. (2001). BMJ bans 'accidents'. *British Medical Journal (Clinical research ed.)*, *322*(7298), 1320.

Demographic Health Survey (DHS). (2000). Iran: Results from the Demographic and Health Survey 2000. Tehran: Iranian Ministry of Health & Medical Education, Statistical Centre of Iran/UNICEF/UNFPA.

Elvik, R., & Mysen, A.B. (1999). Incomplete accident reporting: meta-analysis of studies made in 13 countries. *Transportation Research Record*, *1665*(1), 133–140.

Elvik, R., & Vaa, T. (2004). *The handbook of road safety measures*. Oxfordshire, UK: Elsevier.

Gennarelli, T.A., & Wodzin, E. (2005). *Abbreviated Injury Scale (AIS) 2005*. Barrington, IL: Association for the Advancement Automotive Medicine (AAAM).

Gennarelli, T.A., & Wodzin, E. (2006). AIS 2005: a contemporary injury scale. *Injury*, *37*(12), 1083–1091.

Global Burden of Disease. (2002). *World Health Organization*. Retrieved, from: http://www.who.int/healthinfo/bodestimates/en/ [26 October 2009].

Global Burden of Disease Injury Expert Group. (2009). *Environmental scan of data sources for estimating the global burden of injuries*. Retrieved, from http://sites.google.com/site/gbdinjuryexpertgroup/ [26 October 2009].

Global Status Report on Road Safety. *World Health Organization 2009*. Retrieved, from http://www.who.int/violence_injury_prevention/road_safety_status/2009 [26 October 2009].

Holder, Y., Peden, M., Krug, E., Lund, J., Gururaj, G., & Kobusingye, O. (2004). Injury Surveillance Guidelines, World Health Organisation, Geneva, Switzerland.

Institute for Health Metrics and Evaluation at the University of Washington. (2008). *The global burden of disease, injuries, and risk factors study – operations manual*.

Jacobs, G., Aeron-Thomas, A., & Astrop, A. (2000). *Estimating global road fatalities*. Crowthorne. TRL Report, No. 445. Crowthorne, Berkshire, UK: Transport Research Laboratory .

Jarvis, S.N., Lowe, P.J., Avery, A., Levene, S., & Cormack, R.M. (2000). Children are not goldfish-mark/recapture techniques and their application to injury data. *British Medical Journal*, *6*(1), 46–50.

King, G., Murray, C.J.L., Salomon, J.A., & Tandon, A. (2004). Enhancing the validity and cross-cultural comparability of measurement in survey research. *American Political Science Review*, *97*(04), 567–583.

Kopits, E., & Cropper, M. (2005). Traffic fatalities and economic growth. *Accident Analysis and Prevention*, *37*(1), 169–178.

Langley, J., & Brenner, R. (2004). What is an injury? *Injury Prevention*, *10*(2), 69–71.

London, J., Mock, C., Abantanga, F.A., Quansah, R.E., & Boateng, K.A. (2002). Using mortuary statistics in the development of an injury surveillance system in Ghana. *Bulletin of the World Health Organization*, *80*, 357–364.

Lopez, A.D., AbouZahr, C., Shibuya, K., & Gollogly, L. (2007). Keeping count: births, deaths, and causes of death. *The Lancet*, *370*(9601), 24 Nov–30 Nov 2007, 1744–1746.

Luoma, J., & Sivak, M. (2007). Characteristics and availability of fatal road-crash databases in 20 countries worldwide. *Journal of Safety Research*, *38*(3), 323–327.

Mahapatra, P., Shibuya, K., & Lopez, A.D. (1992). On behalf of the Monitoring of Vital Events (MoVE) writing group. Civil registration systems and vital statistics: successes and missed opportunities. In C.J.L. Murray & D.B. Evans (Eds.), *The Lancet* (pp. 797–808). Health Systems Performance Assessment, WHO, Geneva, Switzerland.

Manual of the International Statistical Classification of Diseases and Related Health Problems. (1992). *Tenth ICD10 Revision 1*. Geneva: World Health Organization.

Mathers, C.D., & Loncar, D. (2006). Projections of global mortality and burden of disease from 2002 to 2030. *PLoS Medicine*, *3*(11), e442.

Mock, C.N., Abantanga, F., Cummings, P., & Koepsell, T.D. (1999). Incidence and outcome of injury in Ghana: a community-based survey. *Bulletin of the World Health Organization*, *77*(12), 955–964.

Mohan, D., & Qadeer, I. (1990). *Safety of agricultural implements*. New Delhi, India: Center for Biomedical Engineering, Indian Institute of Technology.

Neira, J., & Bosque, L. (2004). The word accident: no chance, no error, no Destiny. *Prehospital and Disaster Medicine*, *19*, 188–190.

Odero, W., Garner, P., & Zwi, A. (1997). Road traffic injuries in developing countries: a comprehensive review of epidemiological studies. *Tropical Medicine & International Health*, *2*(5), 445.

Osler, T., Baker, S.P., & Long, W. (1997). A modification of the injury severity score that both improves accuracy and simplifies scoring. *The Journal of Trauma: Injury, Infection, and Critical Care*, *43*(6), 922.

Peden, M.M., Scurfield, R., Sleet, D., Mohan, D., Hyder, A.A., Jarawan, E., et al. (2004). *World report on road traffic injury prevention*. Geneva: World Health Organization.

Razzak, J.A., & Luby, S.P. (1998). Estimating deaths and injuries due to road traffic accidents in Karachi, Pakistan, through the capture–recapture method. *International Journal of Epidemiology*, *27*(5), 866–870.

Registrar General of India, Vital Statistics Division. (1999). *Sample registration system*. Government of India. www.censusindia.gov.in [26 October 2009].

Registrar General of India, Vital Statistics Division. (2006). *Medical certification of causes of death*. Government of India. www.censusindia.gov.in [26 October 2009].

Road Safety in Asia and the Pacific. *United Nations Economic and Social commission for Asia and the Pacific*. Retrieved, from http://www.unescap.org/ttdw/roadsafety/index.html [26 October 2009].

Sanghavi, P., Bhalla, K., & Das, V. (2009). Fire-related deaths in India in 2001: a retrospective analysis of data. *The Lancet*, *373*, 1282–1288.

Segui-Gomez, M., & MacKenzie, E.J. (2003). Measuring the public health impact of injuries. *Epidemiologic Reviews*, *25*(1), 3–19.

Shahraz, S., Bhalla, K., Bartels, D., Naghavi, M., Lozano, R., & Murray, C.J.L. (2008). Improving the quality of injury statistics by using regression models to redistribute ill-defined events. *Population Health Metrics* (in preparation).

Tercero, F., & Andersson, R. (2004). Measuring transport injuries in a developing country: an application of the capture–recapture method. *Accident Analysis and Prevention*, *36*(1), 13–20.

United Nations World Population Prospects. (2004). Population Division of the Department of Economic and Social Affairs of the UN Secretariat. http://esa.un.org/unpp [26 October 2009].

Ustun, T.B., Chatterji, S., Mechbal, A., Murray, C.J.L, WHS Collaborating Groups. (2003). *The world health surveys*, *370*(9599), 10 Nov–16 Nov 2007, 1653–1663.

Verma, P.K., & Tewari, K.N. (2004). Epidemiology of road traffic injuries in Delhi. Results of a survey. *Regional Health Forum WHO South-East Asia Region*, *8*(1). http://www.searo.who.int/EN/Section1243/Section1310/Section1343/Section1344/ [26 October 2009].

Ward, H., & Christie, N. (2000). *Strategic review of research priorities for accidental injury*. London: Department of Health.