Estimating deaths and injuries due to road traffic accidents in Karachi, Pakistan, through the capture-recapture method

Junaid A Razzak[a,b] and Stephen P Luby[a]

Background   Road traffic accidents (RTA) are an important yet preventable cause of death and disability in developing countries like Pakistan. Yet accurate epidemiological data on injuries in developing country injuries is often difficult to obtain. We applied the capture-recapture method to estimate the death and injury rates due to RTA in Karachi.

Methods   We applied the two-sample capture-recapture method using traffic police records as one source of capture and the logs of a non-government ambulance service as the second capture source for the same 10 months and 20 days for which 1994 data were available. We generated a conservative adjusted estimate of injuries and deaths by considering entries in the two sources as matched if they reported the same date, time, and place, and at least one of the other matching variables, of name, vehicle registration number, vehicle types or patient outcome. We then compared the estimated rates with the police rates.

Results   In 1994 police reported 544 deaths and 793 injuries due to RTA while ambulance records noted 343 deaths and 2048 injuries. The capture-recapture analysis estimated at least 972 (95% CI : 912–1031) deaths and 18 936 (95% CI : 15 507–22 342) injuries attributable to RTA during the study period. Official sources counted only 56% of deaths and 4% of serious injuries. The estimated rates for the year 1994 were 185 injuries and 11.2 deaths per 100 000 population.

Conclusion   Road traffic injuries and deaths in Karachi are a much more substantial health problem than is evident from official statistics.

Keywords   Road traffic accidents, Karachi, Pakistan, capture-recapture method, official data, ambulance data

Accepted   10 December 1997

Road traffic accidents (RTA) are a significant,[1,2] but preventable,[3,4] cause of death, disability and economic loss in developing countries. Among males of the economically active age group, motor vehicle injuries are the third most important cause of death in these countries.[5] It is projected that RTA will be the second most common cause of disability-adjusted life years in developing countries in the year 2020.[6] Despite this, there is little recognition of the health and economic burden of this problem.[7] The World Health Organization (WHO) in its international conference on RTA noted the importance of adequate data on traffic injuries.[8] Indeed, accurate estimates of the public health burden of RTA can establish the priority of

a Department of Community Health Sciences, The Aga Khan University, Karachi, Pakistan.

b Current address: Section of Emergency Medicine, Yale New Haven Hospital, New Haven, CT, USA

Reprint requests to: Junaid A Razzak, 437 Meloy Road, #B1, West Haven, CT 06512, USA.

this public health problem, and provide a rational basis for policy decisions. Yet, accurate epidemiological data from many of the developing countries are difficult to find in the literature.[9]

Injury data from developing countries is primarily based upon either hospital logs[10,11] or is derived from police records.[12–16] It is believed that both these sources underestimate the total burden of the injuries. Hospitals tend to see more severe forms of injuries only.[1] Furthermore, published hospital series are unlikely to reflect all those who seek health care following RTA. In Karachi, Pakistan, for example, it is difficult to find out the total number of hospitals, let alone the RTA related injuries and deaths seen by each of them. Rates based on figures from a few of these hospitals would be expected to underestimate even the severe forms of injuries in the community. Similarly, studies which simply compared the police data with hospital data have consistently found that police data underestimate RTA[15–17] ranging from 60% underestimation in Al Ain City[17] to 80% underestimation in Saudi Arabia.[18] Since none of these

compared police records with the actual rates in the community it is likely that they have underestimated the underreporting by the police. While accurate death and injury rates would best be generated through community-based studies, such studies are expensive and therefore difficult especially for a developing country like Pakistan.

The capture-recapture method has been used extensively in the biological sciences[19] and medicine for estimating difficult to count populations.[20,21] It has been used in many different settings.[22-24] The method is based on matching two independent samples to arrive at an estimate of the total. In this study, the capture-recapture method was applied using combined records from police sources, and a large ambulance service in the city to estimate more accurately the death and injury rates due to RTA in Karachi, Pakistan.

Methods

Karachi is the largest city of Pakistan, and one of the seven largest cities in the developing world with a population of 9.9 million,[25] 859 459 registered vehicles[26] and 245 km of roads.[27]

Police records are the main source of information regarding RTA in the city. Some of the accidents are registered by the police at the time, as is required by official policy when there is loss or damage to life or property. However, most are reported to the police through the medicolegal officer based in five government run hospitals, where presumably most of the trauma patients go for treatment. Although it is a legal requirement to report all motor vehicle collisions resulting in injuries, it has been observed that only a few cases are reported formally. One reason is that people would rather pay the modest bribe to the police rather than get involved in a tedious and often frustrating legal process for which they receive no apparent benefit.[28,29] All cases of officially recorded accidents are aggregated at the traffic police headquarters and a daily report is made, giving details of the traffic accidents reported during the last 24 hours. Information on all the officially reported RTA are sent daily to the Traffic Engineering Bureau, a subsidiary of local government. They then analyse each event with the objective of determining the cause of the RTA and make recommendations to decrease their incidence. A yearly report is then published after an exhaustive analysis of the available data.

Another source of information on RTA related injuries and accidents in Karachi is the Edhi Ambulance Service (EAS). This is a private philanthropic institution that provides, among other services, ambulances to transport the sick. We selected EAS as it is by far the largest in the city; operating 35 (77%) of the 45 ambulance dispatch centres in Karachi. Edhi Ambulance Service covers the entire metropolitan area as its dispatch centres, supported by 75 working ambulances and 150 drivers are spread evenly throughout the city. The other ambulance services have centres only in front of hospitals and none of them have more than two centres each or more than 10% of the total ambulances. These other ambulance services are used primarily by patients of the nearby hospital for transport to home or to another hospital. Typically, just after an accident has occurred, the EAS control tower or the regional centre is informed by one of the witnesses of the accident. The injured or dead are picked up by the ambulance and are transported to the

nearest hospital. Information regarding the name and age of the victim, the time of accident, the place of accident and vehicle type and registration number is noted. The National Identity Card, which is carried by most city dwellers is used in cases where victims are unable to give their name and age. Once the patient has been transported to the hospital, the ambulance driver reports the information to the Edhi information bureau which records all such information in a log.

We used a two sample capture-recapture method. The first capture was all deaths and injuries due to RTA reported by the traffic police in their daily report. The recapture was the EAS logs. Cases were matched for the year 1994 except for 42 days (25 March-5 May) for which the ambulance log was missing.

Each injury case in the police data was matched with EAS data on seven variables: the name of the victim, date, time and place of the accident, the types and the registration number of the vehicle involved, and the outcome (alive or dead). The name of the victim in the ambulance service log was in Urdu, the national language of Pakistan, while police reported it in English. The names were considered matched if the pronunciation was similar in both languages. The street system in Karachi is poorly organized. Street names are hard to find and numerical addresses rarely available. Many major roads of the city have more than one name. Places in Karachi are often identified by a landmark like a well known shop or a mosque or other building. If the study reviewers, who were knowledgeable about the city, judged the place similar in both the capture and recapture, it was considered matched. 'Time' was considered matched if the recorded time from the two sources was within one hour of each other. Date, type of vehicle and registration number of vehicle were considered as matched if they were exactly the same in both the capture and recapture.

The degree of matching was then defined based on four different standards (A-D). Each of the standards required date and time to be the same, and place to be similar. Standard A, the strictest standard required the name of the victim, the date, time and place of the event, the type and registration number of vehicles involved and the patient outcome (survived or died) to be matched. For standards B-D the criteria were progressively slackened, so that each subsequent standard required one less criteria for a match (Table 1).

In some cases one or both data sets did not have the name of the victim and registration number of vehicle. If one of the data sets had no information on one of the variables while the other one had it, the variable was considered unmatched on that variable.

Table 1 The four different standards used to define a match between ambulance and police records of road traffic accidents, Karachi, Pakistan 1994

Standard	Date	Time	Place	Name of victim, registration no., vehicles, outcome
A	Same	Same	Similar	All same
B	Same	Same	Similar	Any three of the above same
C	Same	Same	Similar	Any two of the above same
D	Same	Same	Similar	Any one of the above same

An estimate of the total number of RTA victims was made with the formula

$$\text{Estimation of } n = \hat{n} = \frac{(c1 + 1)(c2 + 1)}{m + 1} - 1$$

Where c1 and c2 are the numbers of people in the first and second capture respectively, and m is the number of people identified in both captures (matches). The following formulae[30] were used to calculate the variance and 95% confidence interval (CI) for the estimate of n:

$$\text{Variance} = \frac{(c_1 + 1)(c_2 + 1)(c_1 - m)(c_2 - m)}{(m + 1)^2\,(m + 2)}$$

95% CI = n±1.96√Variance

Rates and CI for the whole year were calculated by extrapolating the data of 10 months 20 days to the whole year to calculate the annual incidence. This was done by first dividing the estimated number with 323 (i.e. total number of days for which data was available) and then multiplying this per day estimate by 365. Rates were calculated using 9.9 million as the total Karachi's population[26] and 859 459 registered vehicles in Karachi.[27] The length of roads in Karachi was taken from the Sindh Bureau of Statistic report 1994.[28]

Results

Records from traffic police identified 544 deaths and 793 injuries from RTA in Karachi during the 10 month 20 day study period, while the ambulance services data identified 343 deaths and 2048 injuries for the same time period. Most of the officially, as well as unofficially, reported victims were males (93% and 89% for deaths and 88% and 86% for injuries respectively). The most frequent victim in both sources were pedestrians (35% and 30%) and motorcyclists (19% and 20%). The date, time and place of the accident was recorded in all of the cases, while vehicles involved were unknown in 1.4% of cases, name of victim was not mentioned in 10% of cases, and vehicle registration number was not known in 1.5% of cases.

Using the least strict standard for matching, standard D (when the time, date and place matched along with any one of the other four variables i.e. name of the victim, registration number of vehicle, type of vehicle and patient outcome), comparison of

the two groups identified 220 matches, yielding an estimate of 963 deaths/year (95% CI : 912–1031). Using more restrictive standards for matching resulted in progressively larger estimates of injuries and deaths reaching as high as 47 140 injuries and 3116 deaths using the most restrictive matching standard, standard A (Table 2).

The incidence of deaths due to RTA using the capture-recapture method was at least 963 deaths which is equivalent to 9.72 deaths/100 000 population, 11.3 deaths/10 000 vehicles and 3.9 deaths/km of road. Estimated rates for serious injuries due to RTA were 190 injuries/100 000 population, 218 injuries/10 000 vehicles, and 77 injuries/km of road. The case fatality rate was 5.1.

Discussion

Even with the least restrictive (most sensitive) matching criteria, criteria D (requiring a matching date, time and place, with one of the following: name of victim, registration number, vehicle or outcome matched) the serious injury rates from road traffic injuries in Karachi were 21 times higher than what is officially reported by the police. Using more strict (more specific) criteria increased the difference even further. Since most of the data collection reporting process both by the police force and ambulance drivers was by word of mouth, there was a high likelihood of mismatch on recall intensive data like name of the victim, registration number and the type of vehicles. Thus, the least restrictive matching standard, standard D, would likely be most appropriate. These rate estimates would thus be conservative, but would be quite unlikely to falsely overestimate rates. Even these most conservative estimates, however, show that the public health importance of road traffic injuries in Karachi is markedly understated by official statistics. The rates for RTA related injuries estimated by the capture-recapture technique are at least 190/100 000 population compared to 9.0 and 20.9 injuries/100 000 population using the Police or the EAS data separately. Official sources thus counted only 56% of deaths and 4% of serious injuries. Indeed, the total number of injuries for this one city, Karachi, is greater than the official estimate of the road traffic injuries for the whole country, which has over ten times Karachi's population.[23] The death rate of 9.7/100 000 population, on the other hand is 179% higher than the

Table 2 Estimation of deaths and injuries due to road traffic accidents in Karachi for 1994 using the capture-recapture method

Match type	Outcome	Number of matches	Numbers unmatched in police data	Number unmatched in Edhi data	Estimated number of total deaths /injuries	Estimated rate per 100 000 persons per year	95% confidence interval for rate
A	Deaths	68	476	275	3116	31.5	25.5–38
	Injuries	39	754	2009	47 140	476	335–626
B	Deaths	139	405	204	1524	15.4	13.8–17.2
	Injuries	68	725	1980	27 036	273	214–337
C	Deaths	209	335	134	1013	10.3	10.1–11.0
	Injuries	92	701	1956	19 983	202	166–242
D	Deaths	220	324	123	963	9.7	9.2–10.4
	Injuries	98	695	1950	18 760	189	157–226

5.49/100 000 reported by police and 234% higher than the unofficial figure of 3.46 reported by EAS.

Compared to other cities the death rates for Karachi are much higher than cities in developed countries. For example, while 11.3 people die per 10 000 vehicles registered in Karachi, 1.4 people die in Tokyo, and 2.8 people die in Greater Manchester (UK) for a similar number of vehicles. The rates are however, comparable to developing world cities like Bangkok, Thailand with rates of 10 deaths/10 000 vehicles.

Four major assumptions[18,32] must be satisfied for the capture-recapture methodology to produce reliable results. The first assumption is that all the members of the population have the same probability of being captured. All the accidents occurring anywhere in the city were accessible to the extensive networks of both the traffic police and the EAS. Both institutions have good wireless/radio communication systems and both respond to all the RTA which are brought to their notice. As there is no area of the city that was apparently uncovered by either of the systems, all serious RTA in the city have an apparently equal chance of either being taken care of by police or transported by EAS to the hospital or both.

The second assumption is that the capture sources should be independent. This is a difficult assumption to fulfil in most studies which employ capture-recapture methodology.[26] In our study too it will be the more serious injuries which will require ambulance transport and will go to a major government hospital casualty department where there is a medicolegal officer on duty, so there is some degree of positive dependence. Any positive dependence, however, will lead to an overestimate of the number of matches, and so an underestimate of the total number of injuries, thus calculating only the serious forms of injuries. This may be one of the reasons for a higher fatality rate of 5.13% as compared to 1.7% for other developing countries.[33] Indeed, if we assume that the case fatality rate of Karachi is similar to other developing countries the number of serious injuries represented by the number of deaths noted in the analysis would be three times larger.

The third assumption using capture-recapture methods is that the population should be closed. This assumption ensures that there are no major changes in the total population between the time period between the two captures. In our study, since the capture and the recapture was taking place at the same time, there was no chance of any change in the population.

The fourth assumption is that the capture history of all the cases should be accurate. We assume that because of basic legal requirements, police tend to record the date, time, place, vehicle involved and name of the victim much more accurately as opposed to the actual event and burden of mistake. The information on EAS logs are first hand information from the ambulance driver himself who is the actual witness at the site of accident, which in itself reduces the chances of inaccurate reporting. Moreover, to address the difference in the language of the names and inconsistency in the addresses we used loose criteria for matching.

Importantly, for this study, RTA fatality was determined at the time of transport for EAS or the time of emergency room discharge for the police data. Both of these approaches underestimate fatality, and again emphasizes that the mortality estimates provided by this analysis are minimum estimates.

By using the capture-recapture method an estimate of the burden of ill health can be made especially in developing countries where there is no working system to accurately record the number of injuries and deaths. Pakistan, like other developing countries, is likely to have an increase in traffic, which can lead to substantially increased deaths and disability unless efforts are made to understand the problem and steps taken now.

References

[1] Smith GS, Barss PG. Intentional injuries in developing countries: the epidemiology of a neglected problem. *Epidemiol Rev* 1991;**13**:228–66.

[2] Murry CJL, Lopez AD. Mortality by cause for eight regions of the world: Global Burden of Disease Study. *Lancet* 1997;**349**:1269–76.

[3] Viano DC. Cause and control of automotive trauma. *Bull NY Acad Med* 1988;**64**:376–421.

[4] Sosin DM, Sacks JJ, Holmgreen P. Head injury-associated deaths from motor cycle crashes. *JAMA* 1990;**264**:2395–99.

[5] Soderlund N, Zwi AB. Traffic related mortality in industrialized and less developed countries. *Bull World Health Organ* 1995;**73**:175–82.

[6] Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990–2020: Global Burden of Disease Study. *Lancet* 1997;**349**:1498–504.

[7] Zwi AB, Forjouh S, Murugusamphillay S *et al.* Injuries in developing countries: policy response needed now. *Trans R Soc Trop Med Hyg* 1996;**90**:593–95.

[8] *Road Traffic Accidents in Developing Countries.* Report of a WHO Meeting. Mexico City. 9–13 November 1981.

[9] Mohan D. Accidental death and disability in India: A stocktaking. *Accid Anal Prev* 1984;**16**:279–88.

[10] Balogun JA, Abereoje OK. Pattern of road traffic accident cases in a Nigerian University teaching hospital between 1987 and 1990. *J Trop Med Hyg* 1992;**95**:23–29.

[11] Weddell JM, McDougall A. Road traffic injuries in Sharjah. *Int J Epidemiol* 1981;**10**:155–59.

[12] Asogwa SE. Road traffic accidents in Nigeria: A review and a reappraisal. *Accid Anal Prev* 1992;**24**:149–55.

[13] Oluwasanmi AJ. Road accidents trends in Nigeria. *Accid Anal Prev* 1993;**25**:485–87.

[14] Jayasuriya R. Trends in the epidemiology of injuries due to road traffic accidents in Papua New Guinea. *Asia Pac J Public Health* 1991;**5**:41–48.

[15] Bangdiwala SI, Anzola-Perez E. Traffic accidents as a serious health problem in selected developing countries of the Americas. *Bull Pan Am Health Organ* 1972;**1**:38–47.

[16] Mohan D. An analysis of road traffic fatalities in Delhi India. *Accid Anal Prev* 1985;**17**:33–45.

[17] Bener A, Achan NV, Sankaran-Kutty M. Road traffic injuries in Al-Ain City, United Arab Emirates. *J Roy Soc Health* 1992;**112**:273–76.

[18] Ofosu JB, Abouammouh AM, Bener A. A study of road traffic accidents in Saudi Arabia. *Accid Anal Prev* 1988;**20**:95–101.

[19] Cook LM, Brower LP, Croze HJ. The accuracy of population estimation from multiple recapture data. *J Animal Ecol* 1967;**36**:494–95.

[20] Mastro TD, Kitayaporn D, Weniger BG *et al.* Estimating the number of HIV-infected injection drug users in Bangkok: A capture-recapture method. *Am J Public Health* 1994;**84**:1094–99.

[21] Sutter RW, Cochi SL. Pertussis hospitalization and mortality in the United States, 1985–1988: evaluation of the completeness of national reporting. *JAMA* 1992;**267**:386–91.

[22] Hartnoll R, Lewis R, Mitcheson M, Bryer S. Estimating the prevalence of opioid dependence. *Lancet* 1985;**1**:203–05.

[23] van Haastrecht HJA, van den Hoek JAR, Bardoux C, Leentvaar-Kuyprs A, Coutinho RA. The course of the HIV epidemic among intravenous drug users in Amsterdam, the Netherlands. *Am J Public Health* 1991;**81**:59–62.

[24] Frischer M. Estimating prevalence of injecting drug use in Glasgow. *Br J Addict* 1992;**20**:997–1000.

[25] *The Challenge of Urbanization: The World's Large Cities.* United Nations. New York, 1995.

[26] *Number of registered vehicles in Karachi in 1994.* Traffic Engineering Bureau, Karachi Development Authority.

[27] *Pakistan Statistical Yearbook 1995.* Government of Pakistan, Islamabad.

[28] Staff Reporter. Transporters accuse traffic police of harassment. *Daily Dawn*, 9 September 1997.

[29] Azizullah Sharif. Traffic cops expert at harassing motorists. *Daily Dawn*, 22 September 1997

[30] McCarty DJ, Tull ES, Moy CS, Kwho CK, LaPorte RE. Ascertainment corrected rates: Application of capture recapture methods. *Int J Epidemiol* 1993;**22**:559–65.

[31] Jacobs GB, Sayer IA. *Road Accidents in Developing Countries—Urban Problems and Remedial Measures.* Transport and Road Research Laboratory Supplementary Report 1984, p.839.

[32] Wittes JT, Sidel VW. A generalization of the simple capture-recapture model with application to epidemiological research. *J Chron Dis* 1968;**21**:287–301.

[33] Ali J, Naraynsingh V. Potential impact of advanced trauma life support (ATLS) program in a third world country. *Int Surg* 1987;**72**:179–84.