

# ESTIMATION OF NON-FATAL ROAD TRAFFIC INJURIES IN THAI NGUYEN, VIETNAM USING CAPTURE-RECAPTURE METHOD

Hung Trinh Van<sup>1,3</sup>, Pratap Singhasivanon<sup>1</sup>, Jaranit Kaewkungwal<sup>1</sup>, Paibul Suriyawongpaisal<sup>2</sup> and Lap Hoang Khai<sup>3</sup>

<sup>1</sup>Department of Tropical Hygiene, Faculty of Tropical Medicine, Mahidol University, Bangkok;

<sup>2</sup>Community Medicine Center, Faculty of Medicine at Ramathibodi Hospital, Mahidol University, Bangkok, Thailand; <sup>3</sup>Department of Epidemiology, Thai Nguyen Medical College, Thai Nguyen, Vietnam

**Abstract.** Road traffic injuries (RTIs) are increasing in developing countries where accurate routine data are usually not available. Although a capture-recapture technique has increasingly been employed in studies of human populations to provide reliable estimates of the magnitude of problems, it has rarely been used in road traffic injury research. We applied two sample capture-recapture methods using hospital and traffic police records to estimate non-fatal road traffic injuries in Thai Nguyen City during the years 2000-2004. We generated a conservative adjusted estimate of non-fatal RTIs using data from the two sources matched by name, surname, sex of victims and at least one of the other matching variables, of age, address of victim and date of injuries. We then compared the estimated rates with those reported based on police and hospital data. The results show that during years 2000-2004, the police reported 1,373 non-fatal RTIs, while hospital records revealed 6,069 non-fatal RTIs. Most reported victims on both hospital and police reports were males (67.3 % and 74.4%, respectively). More than half the victims on both hospital and police reports were drivers (77.5% and 66.1%, respectively) or pedestrians (10.6% and 7.1%, respectively). Youth and young adults (ages 15-34) constituted the majority of the victims on the hospital and police reports (52.8% and 63.7%, respectively). The capture-recapture analysis estimated that 11,140 (95% CI: 10,626-11,654) subjects were involved in RTIs during the study period. In comparison to the estimated figure, official sources accounted for only 21.9 to 60.1% of total non-fatal RTIs. Estimated rates of non-fatal RTIs were 105.5 injuries/10,000 population per year and 393 injuries/10,000 vehicles. Given the fact that under reporting of RTIs has been a major limitation of routine official data sets in developing countries, we suggested the capture-recapture method be used as a tool to provide affordable and reliable estimates of RTIs in resource-poor countries.

## INTRODUCTION

Road traffic injuries (RTIs) are an important but preventable cause of death, disability and economic loss in developing countries (Viano, 1988; Sosin *et al*, 1990; Smith and Barss, 1991). It has been projected that RTIs will be the second most common cause of disability-adjusted life year loss in developing countries by the year 2020 (Murray and Lopez, 1997). Despite this concern, there has been little recognition of the health and economic burden of RTIs in developing countries.

The World Health Organization (WHO) in its international conference on RTIs noted the importance of adequate data regarding RTIs (WHO, 1981). Accurate estimates of the public health burden of RTIs can assist priority setting and healthy public policy decisions. Accurate epidemiological data regarding RTIs have been lacking in many developing countries (Mohan, 1984).

In most developed countries, it is a legal mandate for police to routinely collect data regarding the nature of RTIs and the characteristics of people and vehicles involved. A comparison of police records and hospital data show that a significant number of hospitalized RTIs were not reported to the police (Hvoslef, 1994). Similarly, RTI data from developing countries is primarily based upon either hospital logs (Weddell and McDougall, 1981; Balogun and

---

Correspondence: Hung Trinh Van, Department of Tropical Hygiene, Faculty of Tropical Medicine, Mahidol University, 420/6 Ratchawithi Road, Bangkok 10400, Thailand.

E-mail: hungtrinhvan@yahoo.com

Abereje, 1992) or derived from police records (Asogwa, 1992; Mohan and Bawa, 1985). It is believed that both sources underestimate the total burden of RTIs. Hospitals tend to see more severe of RTIs (Smith and Barss, 1991). Published hospital data are unlikely to reflect all RTI cases who seek health care. In Thai Nguyen, Vietnam, for example, it was difficult to find out the total number of hospitals, let alone RTIs seen by each of them. Rates based on figures from a few of these hospitals would be expected to underestimate even severe injuries in the community. Studies which compared police data with hospital data have consistently found that police data have underestimated RTIs, ranging from 60% in Al Ain City (Bener *et al*, 1992) to 80% in Saudi Arabia (Ofosu *et al*, 1988). In general, the proportion of missed cases varies from one country to another, from 55 to 83% (Maas and Harris, 1984; Barancik and Fife, 1985; Lindqvist, 1991). Comparisons between police and hospital data may be difficult. First, the origin of the injuries may not be routinely available in the hospital records. Secondly, it may be difficult to link hospital and police records, because people injured in one geographical region might be hospitalized outside that area. More accurate RTI data could be generated through community-based studies, but such studies are expensive and difficult to perform in developing countries.

The capture-recapture method has been used extensively in biological science and medicine for estimating difficult to count populations (Cook *et al*, 1967; Sutter and Cochi, 1992; Mastro *et al*, 1994). It has been used in many different setting (Hartnoll *et al*, 1985; Frischer, 1992). The method is based on matching two independent samples to arrive at an estimation of the total. In this study, the capture-recapture method was applied using the combined records of police sources and hospitals in Thai Nguyen City to estimate more accurately non-fatal RTIs.

## MATERIALS AND METHODS

Thai Nguyen Province is located in the northeastern part of the country. It is bordered by Tuyen Quang, Bac Can, Lang Son, Bac Giang, Vinh Phuc and Ha Noi. Thai Nguyen Province covers 3,769 km<sup>2</sup> with a population of 1,060,136 people (Thai Nguyen Census Report, 2004). The

provincial capital is Thai Nguyen City, which is a center for economic, cultural and social activities for the province and the northeastern part of Vietnam. The total length of the road network in Thai Nguyen is 2,753 km, of which national highways comprise 183 km, provincial roads 105.5 km, district roads 659 km, and inter-commune roads 1,764 km. Most of the national highways and provincial roads are asphalt. Thai Nguyen City has 56,604 registered vehicles with 3.7 people per vehicle (Thai Nguyen Traffic Police Report, 2004).

Police records are the main source of information regarding RTIs in the city. Immediate recording of injuries is undertaken only when there is a loss of life or sizable property damage. Although it is a legal requirement for police to report all motor vehicle collisions resulting in injuries, it has been observed that only a few cases are reported formally. In addition, police are responsible for collecting crash data on the spot. The police standard report form was used to extract data from the activity logbook on a daily basis. The form included variables, such as name, surname, sex, age, address, number of vehicles and road users involved, probable causes, number of victims (killed, severely injured, mildly injured), an estimated cost of damage to property, and environmental and vehicle factors. All police stations throughout Thai Nguyen City send daily traffic crash reports to the Police Information System Center at the Thai Nguyen Traffic Police Department. All officially recorded injuries are collected at traffic police headquarters, where the data are analyzed to determine the cause of RTI in order to guide remedial action. An annual report is available with more extensive analysis. Extraction of data from the daily summaries and the annual reports for 2000-2004 was carried out for our study. Another source of information regarding RTIs in Thai Nguyen is hospital data. Typically, after a crash, the victims are transported to the nearest hospital. Information regarding the name, sex, age, address of the victims, time of injury, place of injury and vehicle type is recorded. When a victim can not give his/her name and age, the National Identity Card or proxy report from relatives or friends was sought for the information used. Hospitalized patients were interviewed by trained personnel using a standard data collection form.

The form included variables, such as individual information, patient history, causes of injury, diagnosis, treatment and prognosis. In this study, the data were collected from hospital records for all the hospitals in the city (3 public hospitals, no private hospitals), which managed all the RTI cases that occurred in Thai Nguyen City.

Based on the two data sets, we applied a two-sample capture-recapture method to provide an estimate of non-fatal RTIs in the city. The first capture was all non-fatal RTIs recorded on the traffic police daily report. The recapture was the RTIs recorded on the hospital data set. Data were compiled and matched case by case for all RTI incidents during the years 2000-2004.

Each non-fatal injury case in the police data was matched with hospital data regarding six important personal characteristics and RTI description variables: name and surname of victim, sex and age, date of the injury and address of the victim. To facilitate computerized matching, the name and surname of each victim in both sources were recoded in Vietnamese and English. Names were considered matched if the characters were exactly the same for both sources. The surnames were considered matched if the characters of last name were exactly the same for both sources and the characters of middle name were the same or similar. The sex variable was considered matched if it was exactly the same for the two sources. Age was considered matched if there was a discrepancy of greater than 2 years. The date of injury was considered matched if the recorded dates from both sources were within 2 days of each other. The address of the victim was considered matched if it was the same in both sources.

The degree of matching was then defined based on four different standards (A-D). Each of the standards required name and sex of a victim to be the same, the last name the same and the middle name similar. Standard A, the strictest standard, required that name, surname, sex, age, date of injury, and address of the victim match. For standards B-D the criteria were progressively slackened, so that each subsequent standard required one less criteria for a match (Table 1). The procedure of matching, criteria used to determine the match, and the criteria used to determine which key variables were defined as the same or similar, were assigned be-

fore data analysis. After that, the estimation of the number of non-fatal injuries and the percentage of under-reporting from both sources was done again. The efficiency of the estimators was considered against the standard error for the estimates (SE), coefficients of variation for the estimates (CV) and confidence interval for the estimates (95% CI).

An estimate of the total number of non-fatal RTIs was made with the formula (Lange, 2003):

$$\text{Estimation of } n = \hat{N} = \frac{(X+1)(Y+1)}{C+1} - 1$$

Where X and Y were the numbers of people in the first and second captures, respectively, and C was the number of people identified in both sources (matches). The following formulas were used to calculate the standard error of the estimates (SE), coefficients of variation of the estimates (CV) and confidence interval of the estimates (95% CI).

$$SE(\hat{N}) = \sqrt{\frac{(N-X)(N-Y)}{C}}$$

$$CV(\hat{N}) = \sqrt{\frac{(1-X)(1-Y)}{C}}$$

$$95\% \text{ CI} = n \pm 1.96 \text{ SE}$$

The lowest values of standard error and coefficients of variation were considered to be the best estimators for these criteria. The range of the confidence interval was also considered to be a best estimator.

Investigation of the percentage of under-reporting from both sources: the expected number of non-fatal injuries due to RTIs, which were not reported in each source compared with the estimated number of non-fatal injuries. Percentages of under-reporting from both sources were calculated using the following expression:

Percentage of under-reporting =

$$\frac{\hat{N} - N_{\text{observed}}}{\hat{N}} * 100$$

## RESULTS

Records from hospital data identified 6,069

non-fatal injuries from RTIs in Thai Nguyen City during the years 2000-2004, while the traffic police data identified 1,373 non-fatal injuries for the same period. Most of the reported victims were males for the hospital and police data (67.3% and 74.4%, respectively). More than half the victims from the hospital and police sources were drivers (77.5% and 66.1%, respectively) or pedestrians (10.6% and 7.1%, respectively). Most victims were young people aged 15-34 who accounted for 52.8% and 63.7% of the total victims according to hospital and police data, respectively (Table 2). The name, surname and sex of the victim were recorded in all of the cases. The time and date of the injury were not recorded for some hospital cases.

Using the least strict standard for matching, standard D (when the name, surname and

sex of a victim matched along with any one of the other variables, such as age, victim address, or date of injuries), comparison of the two groups identified 748 matches, yielding an estimate of 11,140 non-fatal RTIs (95% CI: 10,626-11,654) or 2,228 RTIs per year. The estimated non-fatal RTI rate per 10,000 persons per year was 105.5. Using more strict standards for matching resulted in progressively larger estimates of injuries reaching a high of 32,806 non-fatal RTIs using the most restrictive matching standard, standard A (Table 3).

The point estimate for non-fatal RTIs using the capture-recapture method were at least 105.5 injuries/10,000 population/year. A maximum rate was 310.8 injuries/10,000 population/year, which is equivalent to at least 393 injuries/10,000 vehicles. The percentage of under-re-

Table 1

The four different standards used to define a match between hospital and police records for non-fatal injury due to road traffic injuries in Thai Nguyen City, Vietnam during the years 2000-2004.

Standard	Name	Surname	Sex	Age	Date of injuries	Address of victim
A	Same	Same/similar	Same	Same	Same	Same
B	Same	Same/similar	Same	Same	Same	
C	Same	Same/similar	Same	Same		
D	Same	Same/similar	Same			

Table 2.

Distribution of the non-fatal injuries from road traffic accidents by sex, age group and status of victim comparing hospital and police records in Thai Nguyen City during the years 2000-2004.

Variables	Hospital source		Police source	
	Number	%	Number	%
Sex				
Male	4,086	67.3	1,022	74.4
Female	1,983	32.7	351	25.6
Age group				
<5	129	2.1	8	0.6
5-14	474	7.8	46	3.4
15-24	2,022	33.3	576	42
25-34	1,185	19.5	298	21.7
35-44	1,014	16.7	239	17.4
45-54	642	10.6	127	9.2
≥55	603	9.9	79	5.8
Status of victim				
Driver	4,709	77.5	907	66.1
Passenger	718	11.9	369	26.8
Pedestrian	642	10.6	97	7.1
Total	6,069	100	1,373	100

Table 3

Estimation of non-fatal injuries due to road traffic accidents in Thai Nguyen City during the years 2000-2004 using the capture-recapture method.

Match type	Number of matches	Number of unmatched hospital data	Number of unmatched police data	Estimated number of non-fatal injuries	Estimated non-fatal injury rate per 10,000 persons/year	SE	CV	95% CI	% under-reporting from both sources
A	254	5,815	1,119	32,806	310.8	1,819	1.81	29,168-36,444	78.10
B	441	5,628	932	18,895	179	714	1.37	17,467-20,323	62.95
C	674	5,395	699	12,363	117.1	320	1.11	11,723-13,003	45.26
D	748	5,321	625	11,140	105.5	257	1.06	10,626-11,654	39.91

porting from both sources was at 39.91% in Standard D and 78.1% in standard A.

## DISCUSSION

In our study, the police RTI reporting rate was 22.6% of the hospitalized patient rate, suggesting an important deficit in the police records of collisions leading to hospitalization. This study also illustrates the usefulness of the capture-recapture method in assessing the completeness of information systems regarding road traffic injuries that were based on police and hospital data. The police injury reporting rate was lower than those of other similar studies, which reported a police-reported rate varying from 55 to 83% (Bull and Roberts, 1973; Maas and Harris, 1984; Barancik and Fife, 1985; Rosman and Knuiman, 1994). One possible explanation was that the socio-economic and cultural context in Vietnam causes people involved in a road crash to be less willing to report to police than those in other countries.

Exploring the database revealed the police data could hardly represent a total picture regarding RTIs in Thai Nguyen City. In keeping with other studies, our findings indicate that the reporting rates varied according to type of vehicle involved, time of injury, victim addresses and availability of prehospital care for the victims (Bull and Roberts, 1973; Maas and Harris, 1984; Barancik and Fife, 1985; Harris, 1990; Rosman and Knuiman, 1994).

As far as vulnerable road users are concerned, it should be noted that although 22% of bicycles share in the vehicle mix on Thai Nguyen

roads, bicycle injuries were highly under-represented in the police database. A large deficit of bicycle injuries in the police database is well known in OECD countries (Hvoslef, 1994). Keeping this in mind, this implies that one has to rely on hospital data to shed light on the bicycle injury problem.

Even with the least restrictive (most sensitive) matching criteria (criteria D: requiring a matching name, surname and sex of victim with any one of the following variables: age, address of victim and date of injuries matched) the non-fatal injury rates for RTIs in Thai Nguyen City were 8 times higher than that reported by the police. Using more strict (more specific) criteria increased the difference even further. Since data collection under police and hospital reporting systems relied heavily on verbal reports of victims or bystanders, the quality of data was questionable, especially for variables like age, victim address, and date of injury. Thus, the least restrictive matching standard, standard D, would likely be most appropriate. It provided a conservative estimate of RTIs. Even for the most conservative estimates, the magnitude of the RTI problem in Thai Nguyen City was far greater than that reported by routine police and hospital reporting systems. The annual rate for non-fatal RTIs estimated by the capture-recapture technique was at least 105.5/10,000 population compared to 57.5 and 13.0 non-fatal injuries/10,000 population using the hospital and police data, respectively.

In comparison to reported non-fatal RTIs from other countries, our estimate of 105.5/10,000 population was markedly higher than that

of Karachi (19/10,000 population) (Razzak and Luby, 1998) and Scotland (29.1/10,000) (Morrison and Stone, 2000).

Our study seems to be the first to use the capture-recapture method for counting road traffic injured people in Vietnam. This statistical method, first applied to the study of fish and wildlife, has recently been adapted to epidemiological problems (International Working Group for Disease Monitoring and Forecasting, 1995). To our knowledge, this method has seldom been used to evaluate the number of injured people. Four major assumptions must be satisfied for the capture-recapture methodology to produce reliable results (Ofosu *et al*, 1988). The first assumption is that all the members of the population have the same probability of being captured. The injuries in the city had the same access to the traffic police and hospitals. Both institutions have good communication systems and both respond to all RTIs brought to their notice. As there was no area of the city that was apparently not covered by either of the systems, all serious RTIs in the city had an apparently equal chance of either being taken care of by police or transported to the hospital or both.

The second assumption is that the capture sources should be independent. This is a difficult assumption to fulfill in most studies which employ capture-recapture methodology. In our study, the more serious injuries requiring transportation. Hence, the victims would be more likely to be seen at a major government hospital. A case like this would be more likely, than the less severe cases, to be recorded by the police. So there was some degree of positive dependence between the two sources. Any positive dependence, however, will lead to an overestimate of the number of matches, and underestimate the total number of injuries. Calculating only the serious injuries may be a bias in the estimation. This may be one of the reasons for a higher non-fatal injury rate of 105.5/10000 population in our study compared to 29.1/10,000 for developed countries (Scotland) and 19/10,000 for developing countries (Karachi, Pakistan).

The third assumption using capture-recapture methods was that the population should be closed. This assumption ensures that there are no major changes in the number of population

over time between the two captures. In our study, since the capture and the recapture were taking place at the same time and within a very limited space, there was less chance of any change in the population. According to the population reports of Thai Nguyen City during 2000-2004, the population did not change.

The fourth assumption is that the capture history of all the cases should be accurate. We assume that basic legal requirements effectively encouraged police to record name, surname, sex, age, victim address, time, date, and place of injury much more accurately. Hospital data was first hand information from victims or bystanders accompanying them to hospital. These sources of information helped minimize inaccuracy in reporting. To address the difference in the language of the names, surname and inconsistency in the addresses we used loose criteria for matching.

In conclusion, the capture-recapture method is an affordable alternative to provide an estimate of the burden of RTIs in developing countries where routine reporting systems suffer from under reporting. In both developing and developed countries, police records commonly used in research and policy evaluation of RTIs. Similar to other studies, this study revealed deficit of police data for RTIs. This deficit was particularly noticeable for less severe injuries and for bicycle injuries. We believe that hospital data would be more reliable in assessing the magnitude of serious RTIs. As RTIs are on a rising trend in Vietnam and other developing countries, knowledge of the burden is necessary in policy guidance. With limited resources in establishing and maintaining routine reporting systems, the capture-recapture method may be an affordable alternative to evaluate road traffic injuries.

#### ACKNOWLEDGEMENTS

The authors would like to thank all the health staff of the hospitals, traffic police officers in Thai Nguyen and our colleagues in Thai Nguyen Medical College who helped to complete the data collection process for this study. We are very grateful to Prof Dankmar Boehning and Miss Busaba Suppawattanabodee for their excellent assistance with the capture-recapture analysis.



This study was supported by the Ministry of Education and Training, Vietnam and the Faculty of Tropical Medicine, Mahidol University.

## REFERENCES

- Asogwa SE. Road traffic accident in Nigeria; a review and a reappraisal. *Accid Anal Prev* 1992; 24: 149-55.
- Balogun JA, Abereoje OK. Pattern of road traffic accidents cases in a Nigerian University teaching hospital between 1987 and 1990. *J Trop Med Hyg* 1992; 95: 23-9.
- Barancik JI, Fife D. Discrepancies in vehicular crash injury reporting: Northeastern Ohio Trauma Study IV. *Accid Anal Prev* 1985; 17: 147-54.
- Bener A, Absood GH, Achan NV, Sankaran Kutty M. Road traffic injuries in Al-Ain City, United Arab Emirates. *J R Soc Health* 1992; 112: 273-6.
- Bull JP, Robert BJ. Road injuries statistics – A comparison of police and hospital information. *Accid Anal Prev* 1973; 5: 45-53.
- Cook LM, Brower LP, Croze HJ. The accuracy of population estimation from multiple recapture data. *J Animal Ecol* 1967; 36: 494-5.
- Frischer M. Estimating prevalence of injecting drug use in Glasgow. *Br J Addict* 1992; 20: 997-1000.
- Harris S. The real number of road traffic accident casualties in the Netherlands: a year long survey. *Accid Anal Prev* 1990; 22: 371-8.
- Hartnoll R, Lewis R, Mitcheson M, Bryer S. Estimating the prevalence of opioid dependence. *Lancet* 1985; i: 203-05.
- Hvoslef H. Under reporting of road traffic injuries recorded by the police at the international level. Paris: OECD-International Road Traffic and Injuries Database, 1994.
- International Working Group for Disease Monitoring and Forecasting. Capture-recapture and multiple-record systems estimation I: History and theoretical development. *Am J Epidemiol* 1995; 142: 1047-58.
- Lange JH, LaPorte RE, Chang YF. Use of the capture-recapture methods for in epidemiological studies in determining prevalence. *Acta Neurol Scand* 2004; 109: 79-80.
- Lindqvist KS. Epidemiology of traffic accidents in a Swedish municipality. *Accid Anal Prev* 1991; 23: 509-19.
- Maas MW, Harris S. Police recording of road accident inpatients. Investigation into the completeness, representatively and reliability of police records of hospitalized traffic victims. *Accid Anal Prev* 1984; 16: 167-84.
- 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: 1094-9.
- Mohan D. Accidental death and disability in India: A stock-taking. *Accid Anal Prev J* 1984; 16: 279-88.
- Mohan D, Bawa PS. An analysis of road traffic accident fatalities in Delhi India. *Accid Anal Prev* 1985; 17: 33-45.
- Morrison A, Stone DH. Capture-recapture: A useful methodological tool for counting traffic related injuries? *Inj Prev* 2000; 6: 299-304.
- Murray CJL, Lopez AD. Alternative projections of mortality and disability by cause 1990-2020: global burden of disease study. *Lancet* 1997; 349: 1498-504.
- Ofosu JB, Abouammouh AM, Bener A. A study of road traffic accidents in Saudi Arabia. *Accid Anal Prev* 1988; 20: 95-101.
- Razzak JA, Luby SP. Estimating deaths and injuries due to road traffic accident in Karachi, Pakistan, through the capture-recapture method. *Int J Epidemiol* 1998; 27: 866-70.
- Rosman DL, Knuiman MW. A comparison of hospital and police road injury data. *Accid Anal Prev* 1994; 26: 215-22.
- Smith GS, Barss P. Unintentional injuries in developing countries: The epidemiology of a neglected problem. *Epidemiol Rev* 1991; 13: 228-66.
- Sosin DM, Sacks JJ, Holmgren P. Head injury-associated deaths from motorcycle crashes. *J Am Med Assoc* 1990; 264: 2395-9.
- Sutter RW, Cochi SL. Pertussi hospitalizations and mortality in the United States, 1985-1988. *J Am Med Assoc* 1992; 267: 386-91.
- Viano DC. Cause and control of automotive trauma. *Bull NY Acad Med* 1988; 64: 376-421.
- Weddell JM, McDougall A. Road traffic injuries in Sharjah. *Int J Epidemiol* 1981; 10: 155-9.
- World Health Organization. Road traffic injuries in developing countries. Report of a WHO meeting. Mexico City, 9-13 November 1981.