Exposure to risk on the roads

Hutchinson, T. P.[1], Wundersitz, L. N.[1], Anderson, R. W. G.[1], Kloeden, C. N.[1]
[1]Centre for Automotive Safety Research, University of Adelaide, South Australia 5005
email: paul@casr.adelaide.edu.au

Abstract

The concept of "exposure" to risk is used in the context that number of crashes is the product of exposure to risk and the rate of crashes per unit of exposure. In a practical sense, exposure refers to quantities such as distance travelled, time spent travelling, or number of vehicles passing a point. Comparison of crash rates of different groups of people, different types of vehicle, different roads, different environmental conditions, and so on, may be desired. This paper will examine vehicle registration data, counts of vehicles, surveys of vehicle use, and what is termed induced exposure that attempts to infer risk by distinguishing between crashes in which a party is "innocent" or "responsible". The concept of exposure has not yet achieved all that has been wished for in road safety research, because of problems with both the concept and its practical measurement. However, new technology offers considerable potential.

Keywords

Crash rates, Exposure, Risk, Vehicle registrations, Traffic flows, Travel surveys

Introduction

If different groups of people are found to have different numbers of crashes, should this be attributed to underlying differences in crash risk or differences in exposure to risk? When a question like this is asked, the term exposure is being used in the context of the equation *number of crashes = rate × exposure*. For example, truck drivers may have more crashes per year than car drivers because they drive more.

There are a number of different varieties of exposure, and thus of rates, and similar questions may be asked of intersections having different numbers of crashes, models of car having different numbers of crashes, environmental conditions, and so on. Exposure may refer to a single person or a single intersection, for example, or it may be aggregated to a group of people or intersections. For example, the total crashes to a group of people sharing a common characteristic may be divided by the total exposure of that group of people, resulting in a rate that is intended to be relevant to the group.

While groups having a high *number* of crashes per year can be identified from crash data alone, identification of high crash *risk* groups requires a measure of exposure. There has for decades been rather unsatisfactory treatment of exposure in road safety research, making it difficult to know just what the risks are and how effective road safety countermeasures have been. Appropriate exposure data is often not available or is difficult or expensive to obtain.

This paper is organised in the following sections.

- Vehicle registration data.
- Counts of vehicles.
- Surveys of travel.
- Induced exposure: A method of by-passing measurement problems.
- Discussion.

That is, something is first said about simple data on the numbers of vehicles registered and traffic flows as recorded by automatic equipment, and then more elaborate and expensive surveys of travel are discussed. The "induced exposure" concept stems from the idea that some participants in some crashes are utterly non-responsible (purely passive victims of being in the wrong place at the wrong time), and their numbers reflect exposure. The Discussion section will be split into pessimism and optimism about exposure.

Some of the material in this paper is from a broader discussion of travel exposure measures in Wundersitz and Hutchinson [1].

Vehicle registration data

It is usual for researchers and statisticians to think of exposure to risk as taking place in units of driving events (kilometers driven) or in units of people (per 100,000 population). Vehicle registration data is less commonly used. But there are reasons to consider the use of registration data to represent exposure data for specific purposes, while being mindful of its limitations. Vehicle registration data are technically available in each State of Australia. State registrars of motor vehicles maintain data on current registration and varying amounts of historical data. All states' data are coordinated through the National Exchange of Vehicle and Driver Information System (NEVDIS). For a fee, most States can provide information about current vehicles on register. Perhaps more convenient is the motor vehicle censuses that have been conducted periodically from 1971 to 1999 and annually from 2001 to 2007 by the Australian Bureau of Statistics (ABS). The census includes information on all vehicles registered with a state, territory, or other government motor vehicle registry for unrestricted use on public roads except recreational vehicles (e.g., sand dune buggies, trail bikes), consular vehicles, and vehicles registered by the defence forces. A vehicle registered at the date of the census, or had registration expire less than one month before the date is counted in the census. Consequently, the census provides a snapshot count of registered vehicles. Since 2001, the motor vehicle census has consistently been conducted on 31 March.

Motor vehicle registration statistics are available based on information provided by state and territory motor vehicle registration authorities and reflect information derived from registration documents and, where possible (vehicles manufactured during or after 1990), Vehicle Identification Numbers (VIN). Although records are matched to previous records and additional internal checks are conducted, the ABS acknowledge that there is still some variation in reporting from different jurisdiction registries and care should be taken when making comparisons across jurisdictions. The quality of some detailed data is not known. The vehicle registration data can be sourced from the annual publication "Motor vehicle census Australia" (ABS catalogue No. 9309.0). Data cubes are also available for individual states with information including: make and model of vehicle, postcode of registered owner, vehicle type, and year of manufacture.

The total number of vehicles on register is crude as a measure of exposure, but may be useful when used as an exposure measure on its own. In Australia, the logarithm of the fatality rate per registered vehicle has been declining linearly for at least for 30 years, while registrations have been increasing linearly. Thus the product of an exponential decline in the vehicle based rate and a linear increase in number of vehicles has been predictive of fatality numbers in recent history. However, this observation is not explanatory, nor does it necessarily provide useful information on the fatality risk to an average driver.

Vehicles on register can provide the researcher a measure of exposure, but this is most straightforwardly justified if the researcher can assume that the number of vehicles relates in a predictable way to a direct measure of number of relevant driving events, and therefore to the opportunity for an exposure to risk on the road. One must assume that groups of vehicles have a uniform utilization only cautiously. An example of heterogeneous vehicle utilization is the observation that vehicle use declines with a vehicle's age. Skutenko et al. [2] pooled data from the ABS annual Survey of Motor Vehicle Use from 1998 to 2003 and found that vehicle use declines exponentially with vehicle age. A curve fitted to the ABS data for the whole of Australia suggests an average annual decline in vehicle kilometres driven of 5.1 per cent per year of vehicle age. Thus, registered vehicle data should at least be disaggregated according to vehicle age and weighted appropriately before being used to represent driving exposure. Some vehicle attributes, such as age, are not independent of driver attributes --- for example, younger drivers tend to crash vehicles that are on average more than twice the age of vehicles crashed by drivers of middle age and this may confound risk calculations based only on vehicle age. Where the researcher needs to relate crash risk to some vehicle attribute, registration data can be useful. For example, one may need to calculate crash rates among vehicles that need to be disaggregated by age, make and model, type of ownership, place of residence of owner or age of owner. Some other types of useful disaggregation might include the calculation of crash rates among vehicles of certain size, power-to-weight ratios, front/rear wheel drives and so on; information on the vehicle beyond make and model may be accessed by the interpretation of the Vehicle Identification Number in the registration record.

There are also some situations where one might assume that utilization, though not uniform, at least has a pattern that is stable over time, thus allowing the use of number of registered vehicles to calculate meaningful changes in crash rates.

These comments point to need for care in interpreting registration data in terms of exposure.

One currently useful application of vehicle registration data has been its use to examine the safety profile of the registered fleet. Registration data provides a means of characterising the distribution of safety features and safety related attributes of cars registered in a jurisdiction and provides a means of examining the trends in such characteristics over time. An example of one such analysis is given in Anderson et al. [3].

Counts of vehicles

Individual locations with the highest number of crashes tend to be on high traffic volume roads and particularly where two high traffic volume roads intersect. This is not unexpected, but it does make identifying inherently unsafe locations difficult without some measure of exposure.

Traffic authorities monitor traffic flows at particular sites on the road network with the primary purpose of managing traffic. In the best-case scenario, all individual vehicles passing a given point are monitored with the following variables being recorded --- date, time, direction of travel, speed, wheelbase of vehicle, headway (time between front axle of the previous vehicle passing the site and the front axle of the current vehicle passing the site), gap (time between rear axle of the previous vehicle passing the site and the front axle of the current vehicle passing the site), number of axles, and class of vehicle (Austroads class based on number of axles and wheel bases).

If this data were collected at all sites on a continuous or at least consistent basis, it would provide a good basis for exposure-correcting the crash rates at different sites. However, in the real world there are a number of compromises made.

- Data is not collected continuously.
- Only small proportions of roads and intersections are surveyed.
- The locations tend to be chosen based on sites with traffic congestion problems.
- Surveys at sites are irregular.
- Individual data is often grouped into bins before being stored on the recording equipment.
- Collection of data has been expensive and error prone in that it involves placing tubes across the roadway and manual downloading of the data.

Newer technologies using lasers to detect vehicles and the mobile phone network to deliver data could potentially be more reliable and cheaper to run. Automated count data also suffers from an inability to detect some characteristics of the driver (sex, age, purpose of trip) and the vehicle (commercial, private).

Even given all these limitations it is likely that high volume crash sites will have some data available to enable a gross exposure correction to be made. However, there is also the question of the relative importance of absolute number of crashes versus the rate of crashes. A small improvement at a high volume low risk site could prevent more crashes than a large improvement at a low volume high risk site.

When considering overall crash numbers over time, some measure of the general exposure of vehicles is desirable. Ideally, a selection of sites chosen at random across a range of road types would be continuously monitored to give some indication of how traffic volumes and speeds are changing over time. This data could be used to place the general crash data into a better context.

A limited attempt at such a system has been made in South Australia with 132 sites across 13 road types being surveyed for one week each year [4]. The primary purpose of this on-going study is to track vehicle speeds in South Australia, but the data collected also enables changes in general traffic volumes to be tracked over time. Some initial observations from this study on traffic flows during two successive years are of interest.

- There was considerable variation of traffic flows by time of day and day of week between different sites --- even between sites of the same type that appeared to serve a similar function.
- While, generally, the traffic flows at a given site were consistent from the first year of the survey to the second year, there were some sites where the traffic flow changed substantially either in overall quantity of traffic or in the distribution of traffic over the week. In some cases this could be explained by changes in the road or apparent special events nearby but in other cases there was no apparent reason.

These observations indicate that great caution needs to be taken both when selecting sample sites and when extrapolating from the sample sites to the transport system in general. In addition, traffic engineers seem not to have really embraced the idea of a random sample, and consequently it is difficult to generalise what is found for a limited number of points to a wider geographical area.

Surveys of travel

To obtain information about the characteristics and travel behaviour of different groups or individuals within the population, surveys of people are required. Surveys collecting information about travel behaviour typically gather information by face-to-face interviews, telephone interviews, mail surveys, or a combination of these methods. In addition, the Internet offers a relatively new method of obtaining exposure information.

The most common form of this type of survey involves asking a representative sample of households about their travel patterns. Respondents are typically requested to record details of all trips taken, including walking and cycling, over a specified number of days in a travel diary or logbook. Travel diaries are considered one of the best means of obtaining travel behaviour as they can cover an extended period of time and are not retrospective [5]. Surveys that rely on retrospective recall tend to record typical travel behaviour rather than providing an accurate account [6]. In addition, travel diary-based surveys can provide a comprehensive picture of travel behaviour with information including mode of travel or vehicles used, distance and time taken for travel, characteristics of the traveller such as sex and age, travel costs, and purpose of the journey. In particular, more information about the travel behaviours of cyclists and pedestrians can be obtained from these surveys than observational studies.

Nevertheless, there are some limitations associated with these types of surveys. Firstly, these surveys rely on self-reported data that is dependent on the honesty and thoroughness of the respondent rather than a direct behavioural observation. In particular, travel diaries requiring detailed information tend to be less complete on the last day than on the first [7]. Secondly, the representativeness of the survey is reliant on the population sampled and response rates. Groups with lower response rates, such as young people, are often under-represented in these surveys (statistical weighting can modify this problem to a small extent).

Household surveys based on interviews, compared to observational surveys, are also likely to be somewhat biased towards underestimating the travel of the general population because individuals who travel more are least likely to be home for an interview. Also, household studies are likely to underestimate commercial vehicle travel. Consequently, it is important to compare the results of self-report surveys with the results of observational surveys every few years [5].

To gain a more complete picture of travel behaviour and exposure data, households need to be sampled on different days of the week and at different times of the year. To satisfy this requirement and to provide regular survey data, some travel surveys are carried out on an on-going basis throughout the year (e.g., Sydney Household Travel Survey, New Zealand Travel Survey, and National Travel Survey of Great Britain). The advantage of an on-going survey is that survey data can regularly be combined with crash data to provide up to date estimates of crash risk for different groups of road users. The continuous Sydney survey is designed so that the sample size obtained from three years of data collection is similar to that achieved by a single household travel survey conducted approximately every ten years. However, such continuous travel surveys can be expensive to run. Alternatively, a less costly approach is to conduct frequent (twice yearly) smaller scale surveys that seek to maximise response rates [5]. High response rates can be more important than a large sample size if the characteristics studied are quite variable (e.g., [8]).

In respect of the households that participate in a travel survey, there are at least two types of concern, which might be exemplified as follows. Suppose one wanted to estimate the total distance driven by those aged at least 80.

- Response rates of different ages of people might be different, and the sample would then differ systematically from the population. However, the ages in the population would probably be available from some other source (e.g., the Census), and the ages in the sample could be compared with this, and the total distance driven by the population of those aged at least 80 could be adjusted according to the differential response rates of different ages.
- Response rates of people of different levels of health might be different, and the sample would again differ systematically from the population. But in this case, level of health would be known neither for the sample nor for the population: there would be no way of correcting distance driven for any difference between the sample and the population in respect of level of health.

In the Sydney Household Travel Survey in the period 2004-2007, some 65 per cent of the households approached did respond ([9], p. 47). We fear this is low enough to cause substantial problems if the data were used to calculate crash rates.

Some exposure measures that can be obtained from travel surveys are described in the following subsections.

Aggregate measures

Aggregated measures provide rough overall indicators of exposure, for example, total population, number of registered vehicles, number of licensed drivers, and fuel sales [5]. These measures are popular for estimating crash risk because they are readily available and can be obtained at a low cost. However, such measures often do not allow the estimation of risk for specific groups of road users or in specific situations. In addition, they do not account for differences in risk associated with the amount of time or distance travelled. Generally, the more aggregate the exposure measure, the more intervening variables are introduced [10]. Disaggregate measures of exposure examine individual travel choices, based on individual characteristics and preferences. Generally, disaggregate measures are more complicated and difficult to obtain but they provide a more accurate means for estimating crash risk.

Distance travelled

Distance travelled is a popular measure of exposure that enables disaggregation by travel and demographic groups for meaningful comparisons. However, other aspects of exposure, such as time of day, road type, speed, traffic flow and density of traffic conflicts often vary considerably among drivers. Consequently, the relationship between distance travelled and these other aspects of exposure is not simple [11]. In particular, researchers have observed that distance driven can exaggerate the risk of low mileage groups such as young and elderly drivers [12, 13]. For these low mileage drivers, crashes do not appear to increase in a linear fashion with an increase in the number of kilometres travelled. For example, Forsyth et al. [14] found that crash rates per mile for novice drivers decreased with increased annual mileage. According to Janke [12], individuals travelling shorter distances tend to drive on busier city streets where traffic conflicts are more likely to occur while high mileage drivers spend more time driving on highways where driving is a simpler task and, consequently, the crash risk is lower. (But low congestion leads to higher speeds as well as lower conflicts, and the fatality rate can be high.)

Chu [15] notes that when comparing risk between different modes of transport, distance-based measures of exposure can misrepresent results because of differences in travel speed. For example, one person-mile of walking (relatively slow speed) represents greater exposure to traffic than one person-mile as riding in a vehicle (relatively fast speed).

Nevertheless, problems with distance travelled as a measure of exposure might be overcome by disaggregating risk, that is by separating out the different demographic and environmental factors that are known to vary among drivers (for a discussion, see [16]). Thus, despite the limitations of distance driven as an exposure measure, this measure is popular due to its convenience, particularly when it can be disaggregated by demographic and environmental factors [5].

Number of trips

The number of distinct trips made by an individual, regardless of the distance or time taken for the journey, is another exposure measure. This measure is not particularly useful for assessing risk at specific locations. However, trip data can be used to make comparisons between large jurisdictions and determine changes over time. In addition, trip data can be linked to other survey information obtained at the same time such as trip purpose, mode, time of day etc. Note that pedestrian trips are often under-reported in surveys [17], most likely because they are often only a small part of a multi-modal trip.

Time spent travelling

When explaining crash risk among road users and regions with very different travelling patterns and environments, some researchers argue that time spent travelling is a better measure of exposure than distance travelled [11]. This is because time-based measures take into account different travelling speeds and the environment, factors that are particularly important when comparing risk among different travel modes (e.g. walking, cycling, motorised vehicles). For example, Hakkert and Braimaister [10] suggest normalising exposure by multiplying vehicle kilometres travelled by speed to account for variation in the travel speeds of different modes of transport. While Greene-Roesel et al. [18] argue that time-based exposure measures are the best measures to compare the risks of different travel modes, the European Transport Safety Council [7] suggests using passenger kilometres per year because "…it relates to the decision about how to travel a certain distance rather than how to spend a certain amount of time" (p. 9). Like distance travelled, time spent travelling can be measured on the individual level through surveys or through direct measurements at specific locations. Nevertheless, even supporters of travel time as an exposure measure (e.g., [11]) acknowledge that time-based exposure measures can lead to the conclusion that driving at higher speeds (lower exposure in terms of driving time) leads to fewer crashes, contrary to the known relationship between speed and crash risk [19].

Induced exposure: A method of by-passing measurement problems

Exposure is difficult to define and difficult to measure. Haight [20] identified three general strategies for dealing with the difficulty.

- To accept crude quantities such as distance driven, eliminating confounding factors to as great extent as is practicable with the data available.
- To ignore exposure, concentrate on absolute numbers of crashes, carry out cost-benefit analyses with these, and decide on implementation on this basis.
- To manipulate crash data in such a way as to obtain "exposure-corrected" crash figures, without using any other data such as distances driven.

The third of these is known as the induced exposure approach.

As an example, consider the ratio of the number of a particular category of drivers responsible for crashes to the number of that category innocently involved in crashes. This might be called an over-involvement ratio. The amount of traffic to which this group of drivers is exposed will be reflected both in the number of times they are an innocent party in crashes and in the number of crashes they cause. The inherent danger of the group will only affect the latter, however, and taking the ratio of one to the other results (so it is hoped) in an exposure-corrected crash figure.

However, it is difficult to determine responsibility for each crash, and for work on large datasets of crashes, some alternative is needed. The following assumptions are an example [21].

- Single-vehicle crashes are caused entirely by the driver-vehicle combination concerned. (A driver-vehicle combination means a category such as young drivers in old cars.)
- In each two-vehicle crash, there is a responsible and an innocent party.

- The proportionate involvement of each driver-vehicle combination as the responsible party in two-vehicle crashes is the same as its proportionate involvement among single-vehicle crashes.

Expressed informally, a dangerous category of drivers or vehicles will show greater over-representation in single-vehicle crashes than in two-vehicle crashes, as there is an innocent party in each two-vehicle crash as well as a responsible party.

As already mentioned, ratios may be calculated using either drivers explicitly judged not to be to blame or using drivers presumed not to be to blame because of the circumstances (e.g., their vehicle was stationary). Both methods can both be found in *Research on Road Safety*, at pp. 118-123 [22]. Despite this venerable history, induced exposure methods have achieved only modest popularity: Google Scholar finds 13 articles using the phrase in *Accident Analysis and Prevention* in the period 1998-2007.

One reason for the limited use of these methods may be that the third of the assumptions above is too restrictive: it seems to imply that there is only one kind of fault, that may lead to either a single-vehicle crash or to (responsibility for) a two-vehicle crash. This is surely unrealistic: as simple examples, consider driving fast on an empty road and accepting short gaps in a traffic stream. These are different (even though they might stem from a common source, such as poor judgment or a tolerance of risk) and they will have different effects, increasing single-vehicle crashes in the first case and increasing two-vehicle crashes in the second case. This theoretical objection is reinforced by the empirical finding that in some datasets (e.g., [23]), the model can be seen to fail: for some categories, the ratio of single to two-vehicle crashes is greater than it should theoretically be.

Discussion

> "The utility of exposure data, i.e., road-use data, in road-traffic-accident research is widely acknowledged. This is so in spite of the rather underdeveloped state of exposure research throughout the world."

So wrote Somers and Benjamin [24], and our view is that this is largely true today, also.

There are conceptual and practical problems with exposure that may have been passed over too quickly earlier. The first part of this Discussion will give greater emphasis to them: many studies for which it might be thought that exposure data would be useful would actually be difficult if they were attempted. Road safety practitioners and researchers make only limited use of exposure data --- there is probably good reason for this.

The second part of the Discussion will be more positive and forward-looking. In particular, technology offers solutions to some of the problems.

Pessimism

Whenever something new is found in the crash numbers, exposure is what comes to mind as the likely explanation. However, it is typically difficult to confirm or disconfirm such speculation: exposure as a concept is too indefinite, or the data on exposure is not sufficiently detailed. Some illustrations of this are listed below.

Risks of different people.
(a) Individuals. Some drivers have what seems a lot of crashes, over a period of years. But are they genuinely risky drivers, per kilometre driven, or do they have a high exposure? This is typically not known, because there is no data on how far they drive.
(b) Age/sex groups. On the positive side, it is quite common for results at a macro scale to be presented. But when more detail is required, difficulties very soon become apparent.

- In the case of young novice drivers, the evolution of crash rates on a scale of weeks (after gaining a licence) is likely to be of interest, but sample sizes in surveys of distances driven have not usually been sufficient to study such questions.
- In the case of elderly car drivers, there may be special problems in gaining honest and correct responses, because of the risk of loss of licence on medical grounds.

- Capturing data on non-motorised modes of transport (e.g., walking and cycling) and on being on the road for purposes other than travel (e.g., recreation, working, playing) are perennial problems.

Risk at different places.
(a) Are different crash rates associated with different elements of road design? Examples would include street lighting, pedestrian crossings, a centre median, sealed vs. unsealed surface, shoulder sealing, the maintenance of the road surface, tidying the footpath to remove visual obstructions (and camouflage), and so on. And are different crash rates associated with different elements of intersection design? Examples would include gross design (roundabout vs. traffic signal; crossroads vs. offset junctions), skid-resistant surface treatment, phasing of traffic signals (such as the dangers of nonexclusive right turn), cycle time, road markings, and so on. Unfortunately, it is unlikely to be possible to find exposure data that is sufficiently detailed and sufficiently compatible with crash data.
(b) Some measures are directed at reducing exposure to risk without affecting amount of travel, and are not reflected in the usual measures of exposure. (For example, railings prevent pedestrians venturing on to the carriageway for motor traffic, and traffic lights impose a separation in time on conflicting flows.)
(c) Is a specific intersection more dangerous than would be expected? The traffic flows on its several arms may be known, but it is not clear how these should be combined into a summary measure of exposure. But imagine that there exists a body of high quality research establishing that, say, the sum of flows is a better measure than the product of flows. Even in such a case, the research will probably have involved the comparison (perhaps via regression analysis) of numerous different intersections --- and it is not clear that good evidence about a broad-brush generalisation is any sort of evidence at all about what is appropriate for comparing a specific intersection with other intersections.

Risk of different vehicles.
Are some models of car safer than others? Distances of travel by different models of car are generally not known even at an aggregated level. And it is likely that factors such as nature of ownership (fleet vs. private), age of driver, and circumstances in which they are driven (e.g., urban vs. rural) would have strong effects on crash rates, and these factors might not be independent of model --- with the consequence that differences between models could arise for spurious reasons. Disaggregating distance travelled by different models according to such risk factors is a step or two more difficult than the already difficult task of getting aggregate data for car models. Consequently, crash rates cannot be correlated with such characteristics as size, age of car, front wheel drive vs. rear wheel drive, sedans vs. four wheel drives, ratio of height of centre of gravity to width of track, electronic stability control.

Other contexts for risks.
A lack of appropriate exposure data is likely to be found also when attempting to determine how dangerous are various environmental conditions (such as times of day, daylight vs. dark, and wet vs. dry), conditions of the driver (alcohol, drugs, fatigue, psychological states such as bad temper or feeling stressed), and behaviours (speeding, following distance, presence of passengers, use of mobile phone).

Optimism

The difficulties with exposure should not be over-stated.

- In many contexts, the task that the concept of exposure is being asked to perform is quite a crude one. The differences in numbers of crashes between different groups of drivers seem to be large, and the differences in types of crashes between different models of car seem to be large --- in both cases, it seems likely that differences in exposure are unlikely to be large enough to be responsible (and hence that there is a genuine difference in rate), and quite a rough estimate of exposure should be sufficient to confirm this.
- And in many contexts, practical difficulties can be overcome by survey design targetted at the specific issue and by spending money --- take a sufficiently random sample, take a sufficiently large sample, employ modern technology.

The following paragraphs will discuss some illustrations of this. We will comment on three areas --- technology, the compatibility of crash and exposure data, and improving understanding of exposure and risk without necessarily always calculating rates.

It is possible that modern technology may make collecting data on driving much easier than, say, asking people to keep a diary of their driving.

1. As an example, a sample of young novice drivers could be given a GPS-enabled phone, thereby permitting their pattern of car usage to be continuously tracked. This could continue for, say, the first five years after getting a driving licence. This would provide good data on how much they drive and in what conditions, something that does not exist at present. (Practical difficulties exist --- e.g., it may not be true that teenagers keep their phones with them all the time, or it may be difficult to distinguish between travel as a driver and as a passenger --- but these appear to much less serious than those associated with older technology.) A recent study by Kloeden [25] gives good South Australian data on the crashes and motoring offences of young drivers in their first few years after licensure, but the findings would be much more interpretable and valuable if it were known how distance driven evolves over the months and years following licensure. If a sample truly is a random sample, it need not be very large in order for the mean to be estimated reasonably accurately. Suppose that the population mean and standard deviation are respectively 10000 and 5000. Then a random sample of 25 observations will lead to a standard error of the mean of 1000, which is 10 per cent of the mean. Of course, if it were desired to disaggregate the result by age, sex, experience while learning, district of residence, traffic conditions, and so on, the total size of sample would increase accordingly.
2. Similarly, some electronic device could be attached to a vehicle, and record its movements and thus how much it is used and in what conditions. To validly compare models of car, what is needed in addition is a random sample. Of course, in the case of vehicles, there is already the odometer to give the total usage.
3. To compare lengths of road, the flow of vehicles is likely to be a good enough indicator of exposure for most purposes, and improved technology will make the collection of this cheaper and more accurate, as already discussed. We might add that an assumption that, over the whole range of flows, the number of crashes is proportional to traffic flow may be too bold. Thus rates should only be compared for lengths of road having roughly the same flows.
4. What of intersections, with their several flows, and pedestrian accidents, with pedestrian flows as well as vehicle flows? This is perhaps the point at which the use of a rate does break down and become more trouble than it is worth. After all, even if there were an acceptable method of calculating rate, it would not necessarily be a good guide to action. Presumably, value for money is the criterion that should be used when deciding whether to, for example, improve the layout of an intersection --- and it may be good value for money to improve a busy intersection where the crash rate is quite low when expressed per unit exposure, the number of crashes per year being quite high.

The importance of compatibility of data collected in travel surveys with the crash data was highlighted by Allsop [26]. His argument was not directly that a general-purpose travel survey should obtain data on, say, how many roads of what kinds are crossed and at what types of locations. Instead, it was that there is expertise and experience gained from conducting travel surveys that could be used to design both surveys about the traffic encountered by individuals whether as pedestrians or drivers or other categories of road user, and household-based surveys giving brief information about crash involvement.

It is natural to take the ratio of crashes to exposure in order to get a rate, and then see how this is affected by various independent variables. However, it may sometimes be better to separately present the effect of independent variables on crashes and on exposure, and come to an understanding of the totality of effects by that route. Similarly, with regard to induced exposure methods, there is no need to immediately jump to calculation of some ratio of types of crash. Less formal, exploratory, accounts of how single and two-vehicle crashes (and different types of single and two-vehicle crashes) are similarly affected by some variables and differently affected by others can be influenced by the same ideas that underlie induced exposure, and might be more fruitful.

Summing up the above, we may list six reasons why the future usefulness of exposure (and risk implied by exposure) may be greater than in the past..

1. Availability of technology for tracking people and vehicles.
2. Availability of technology for visual recognition (e.g., of vehicle type or number plate).
3. Increasing practicability of linking different datasets (e.g., the crash, vehicle registration, and driver licence datasets).

4. Random sampling: if a sample is truly random, it does not need to be very big in order to give a good estimate of the population mean. The traffic and transport world has not really embraced the ideas of deciding what exactly is the population of interest and then taking a random sample. The traffic and transport world could choose to change, and use random sampling more widely.
5. Growing awareness of the importance of compatibility between transport and crash datasets.
6. The coordinated exploration of crash and exposure datasets with the ideas behind induced exposure kept in mind, without premature calculation of risk as the ratio of crashes to exposure, might throw up credible interpretations of why certain crash and exposure numbers co-vary while others do not.

Acknowledgements

This paper is based on a study funded by South Australia's Motor Accident Commission (MAC) through a Project Grant to the Centre for Automotive Safety Research. The MAC Project Manager was Ross McColl. The Centre for Automotive Safety Research receives core funding from both MAC and the Department for Transport, Energy and Infrastructure. The views expressed in this report are those of the authors and do not necessarily represent those of the University of Adelaide or the funding organisations.

References

1. Wundersitz, L.N., & Hutchinson, T.P. (2008). Identifying and improving exposure measures. Research Report 053, Centre for Automotive Safety Research, University of Adelaide.
2. Skutenko, D., Cosgrove, D., & Mitchell, D. (2006). Survey of motor vehicle use --- An investigation into coherence. Research Paper, Australian Bureau of Statistics Publication 9208.0.55.005.
3. Anderson R.W.G., Doecke S., & Searson D.J. (2009) Vehicle age-related crashworthiness of the South Australian passenger fleet. Research Report 062, Centre for Automotive Safety Research, University of Adelaide.
4. Kloeden, C.N., & Woolley, J.E. (2009). Vehicle speeds in South Australia 2007. Research Report 051, Centre for Automotive Safety Research, University of Adelaide.
5. Bobevski, I., Clark, B., Lenné, M., Keall, M., Diamantopoulou, K., & Cameron, M. (2007). Development of road safety behaviour, travel and exposure surveys in Victoria. Report No. 269, Monash University Accident Research Centre, Clayton, Victoria.
6. Richardson, A.J., Ampt, E.S., & Meyburg, A.H. (1995). Survey methods for transport planning. Parkville, Victoria: Eucalyptus Press.
7. ETSC (1999). Exposure data for travel risk assessment: Current practice and future needs in the EU. Brussels: European Transport Safety Council.
8. Fogliani, M. (1999). Low response rates and their effects on survey results, methodology. In Advisory committee paper, November 1999 meeting. Canberra: Australian Bureau of Statistics.
9. Transport Data Centre (2008). 2006 Household Travel Survey. Summary Report. Report 2008/01, Transport Data Centre, NSW Ministry of Transport.
10. Hakkert, A.S., & Braimaister, L. (2002). The uses of exposure and risk in road safety studies. Report No. R-2002-12, SWOV Institute for Road Safety Research, Leidschendam, The Netherlands.
11. Chipman, M.L., MacGregor, C.G., Smiley, A.M., & Lee-Gosselin, M. (1993). The role of exposure in comparisons of crash risk among different drivers and driving environments. Accident Analysis and Prevention, 25(2), 207-211.
12. Janke, M.K. (1991). Accidents, mileage, and the exaggeration of risk. Accident Analysis and Prevention, 23(2-3), 183-188.
13. Maycock, G. (1997). Accident liability --- the human perspective. In T. Rothengatter & E. Carbonell Vaya (Eds.), Traffic and transport psychology. Theory and application (pp. 65-76). Amsterdam: Pergamon.
14. Forsyth, E., Maycock, G., & Sexton, B. (1995). Cohort study of learner and novice drivers: Part 3, accidents, offences and driving experience in the first three years of driving. Project Report 111, Transport Research Laboratory, Crowthorne, U.K.

15. Chu, X. (2003). The fatality risk of walking in America: A time-based comparative approach. In Walk21 conference: Health equity and the environment, held in Portland, Oregon.
16. Keall, M. (1995). Pedestrian exposure to risk of a road accident in New Zealand. Accident Analysis and Prevention, 27(5), 729-740.
17. Schwartz, W., & Porter, C. (2000). Bicycle and pedestrian data: Sources, needs, and gaps. Washington, D.C.: Bureau of Transportation Statistics, U.S. Department of Transportation.
18. Greene-Roesel, R., Chagas Diogenes, M., & Ragland, D.R. (2007). Estimating pedestrian accident exposure: Protocol report. http://repositories.cdlib.org/its/tsc/UCB-TSC-RR-2007-5
19. Kloeden, C.N., McLean, A.J., Moore, V.M., & Ponte, G. (1997). Travelling speed and the risk of crash involvement. Report CR172, Federal Office of Road Safety, Canberra.
20. Haight, F.A. (1970). A crude framework for bypassing exposure. Journal of Safety Research, 2(1), 26-29.
21. Thorpe, J.D. (1964). Calculating relative involvement rates in accidents without determining exposure. Australian Road Research, 2, 25-36. Reprinted in Traffic Safety Research Review, 11, 3-8 (1967).
22. Road Research Laboratory (1963). Research on road safety. London: HMSO.
23. Hutchinson, T.P., & Jones, I.S. (1975). The separation of the effects of driver and of vehicle on type of accident. Paper presented at the 5th International Conference of the International Association for Accident and Traffic Medicine, London.
24. Somers, R.L., & Benjamin, T.E.A. (1982). Guest editors' introduction to special issue. Accident Analysis and Prevention, 14(5), 335-336.
25. Kloeden, C. (2008). The crash and offence experience of newly licensed young drivers. Publication No. AP-R331/08, Austroads, Sydney.
26. Allsop, R.E. (2005). Travel survey data required to inform transport safety policy and practice. Transportmetrica, 1, 241-245.