Australian Government
Department of Infrastructure, Regional Development and Cities
Bureau of Infrastructure, Transport and Regional Economics

Location and other risk factors in crashes

At a Glance

This information sheet reports the results of a study into the risk factors of vehicle crashes given that a crash has occurred.

The study includes risk factors normally associated with the consequences of a crash in terms of whether it results in the death of a person involved. These factors include the vehicle type, vehicle age, whether a restraint or helmet was worn, the nature of the crash and the time of day (i.e. split into peak and off-peak periods) the crash occurred.

In addition, the study includes location-specific risk factors such as prevailing environmental conditions (i.e. sunrise, sunset, night-time), the built-up or rural character of an area, and the distance to a high-care emergency medical facility (Principal Referral Hospital).

The rural character of an area significantly increases the risk of being killed in a crash given that it has occurred. The distance to a Principal Referral Hospital is also statistically significant; however, the magnitude of the effect is relatively small when compared with other risk factors.

Other factors identified as increasing the risk of fatality compared with the base case are being male, being over the age of 64, and not wearing a restraint (seatbelt) or helmet, as appropriate to the vehicle type. Motorcyclists and pedal cyclists were found to be at increased risk, while the front and back passenger seats of light vehicles were also found to be associated with an increased risk of fatality.

Crash characteristics that were identified as increasing the risk of fatality compared with the base case are: being involved in a head-on crash, a crash involving a heavy vehicle, a crash where at least one vehicle ran off the road, a single vehicle crash or being involved in a crash where at least one driver failed an alcohol test. Crashes at night or during dusk were also found to significantly increase the risk of death, given that a crash had occurred.

The study was conducted at the national level with the regression analysis including jurisdiction 'dummy' variables. Varying levels of State and Territory reporting of non-fatal outcomes mean that the reported coefficients of the jurisdiction variables reflect differences in data collection and cannot be interpreted as a measure of differences in fatality risk between jurisdictions.

Introduction

In Australia and other countries with remote or rural populations, fatal motor vehicle crashes are a higher proportion of total crashes in regional and remote areas (BITRE 2016). The objective of this study was to investigate the causes of this difference by considering the impact of location-specific risk factors in motor vehicle accidents. In terms of location-specific risk factors, this study specifically investigates differences in access to emergency medical treatment and the difference between built-up or rural areas.

There are two distinct strands of literature that address the issue of location-specific factors in motor vehicle crashes. The first is firmly rooted in health research and focuses on access to emergency healthcare after a crash, and in particular, the length of time before emergency medical treatment is provided. This literature begins with medical studies by authors such as Hoffman (1976), Brodsky and Hakkert (1983) and Bentham (1986) and has grown to include contemporary statistical research, of which a comprehensive overview is provided by Harmsen et al. (2015). The other strand is in road safety research, and focuses on the distinction between crashes that occur in built-up versus rural settings. Comparable studies with an urban/rural distinction include Maio et al. (1992), Siskind (2011) and Lori et al. (2012).

Perhaps with the exception of McAndrews et al. (2017), there is not generally a clear distinction in the road safety literature between the urban/rural character of an area and the correlation to better access to emergency healthcare. Existing studies tend to either include an urban/rural variable or some metric of emergency response, rather than including both. Apart from the defining differences between built-up or rural areas – that is, differences in density of features such as intersections, buildings, vehicles and people – there is evidence of differences in driver behaviour between more and less built-up areas. Previous research by BITRE (2014) has shown increases in injury crashes involving risky/illegal behaviour such as speeding, not wearing a seatbelt, unlicensed driving and driving under the influence of drugs and alcohol is more common in regional and remote areas. It is unclear from existing research whether the built-up form/associated behaviours has an effect on mortality in motor vehicle crashes or if the effect is purely related to the correlation between built-up areas and better access to healthcare.

A secondary consideration of this study is to demonstrate the analytical value and highlight some of the possible areas for improvement of the National Crash Database (NCD). The NCD was developed by BITRE in 2010 for the purpose of monitoring the National Road Safety Strategy 2011–2020 and to support the regular reporting of a core set of agreed performance indicators. Jurisdictions provide NCD data annually for reported casualty crashes (i.e. crashes in which one or more persons have been killed or seriously injured (hospitalised) in a crash on a public road as reported to police and jurisdictional road safety authorities). The NCD contains a sub-set of de-personalised crash data on the crash location/context, vehicle/s, and person/s involved. A person is deemed to have died in a road crash if the person dies within 30 days as a result of injuries sustained in that road crash. This excludes deaths from suicide or natural causes such as a heart attack. Information on deaths is more complete and validated to a higher standard than that of persons injured, while both are more complete than information on non-injured persons.

Strictly the question addressed in this study is:

> *Controlling for other relevant factors, which location-specific factors increase the likelihood that a person involved in a traffic crash will be killed?*

Although this is very narrow, the need to control for and so identify and quantify other relevant factors means that this study looks broadly at the chances of a person surviving a traffic crash, given that they have been involved in one. For this reason the exploration of the data available in the NCD has been fairly broad and has highlighted both its value and some areas for improvement.

Data

A three-year subset of the NCD from 2014 to 2016 inclusive has been used as the basis for this study. This includes records of 398,082 persons, 301,420 vehicles and 178,735 crashes. Once the data was cleaned and records with missing information removed, the study was conducted on 227,566 persons who were in 197,433 vehicles in 133,876 crashes.

Exclusions and missing data

Some categories of road users have been excluded from analysis although some information about them was available. In particular, pedestrians have been excluded as the crash-level factors included in this study do not apply to pedestrians in a way that can usefully be compared with vehicles. Persons in accidents involving 'Other vehicles' have similarly been excluded as this category is too heterogeneous to provide meaningful results. Tables 1 to 3 below show a breakdown of the missing information by injury class. Further information on the data used can be found in Appendix A.

Table 1: Persons with missing information by injury class

	Persons in NCDB	Persons in study	Persons with missing data	Per cent excluded
Fatality	3,169	2,273	896	28%
Injury - hospitalised	57,106	42,957	14,149	25%
Injury - not hospitalised	155,253	108,973	46,280	30%
Not Injured	156,208	65,199	91,009	58%
Unknown	26,357	8,172	18,185	69%
Total	398,093	227,574	170,519	43%

Note: Injury class is not used as a variable in the final model and persons with unknown values are assumed to have survived.

Table 2: Vehicles with missing information by highest level of injury in vehicle

	Vehicles NCDB	Vehicles in study	Vehicles with missing person data	Per cent excluded
Fatality	2,920	2,117	803	28%
Injury - hospitalised	51,081	38,700	12,381	24%
Injury - not hospitalised	132,865	94,464	38,401	29%
Not Injured	113,100	52,495	60,605	54%
Unknown	17,824	5,996	11,828	66%

Note: Injury class is not used as a variable in the final model and persons with unknown values are assumed to have survived.

Table 3: Crashes with missing information by highest level of injury in crash

Injury Class	Crashes in NCDB	Crashes in Study	Crashes with missing person data	Per cent excluded
Fatality	2,877	2,086	791	27%
Injury - hospitalised	48,096	36,328	11,768	24%
Injury - not hospitalised	116,768	82,292	34,476	30%
Not Injured	10,236	4,277	5,959	58%
Unknown	758	183	575	76%
Total	178,735	125,166	53,569	30%

Note: Injury class is not used as a variable in the final model and persons with unknown values are assumed to have survived.

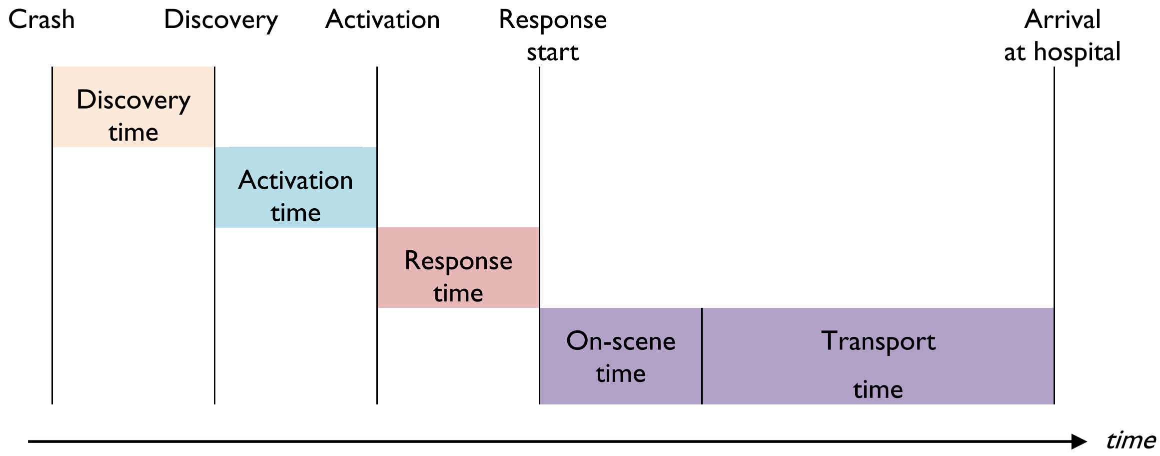
As can be seen in the tables above, the quality of the data varies with the severity of the crash. More data is collected in casualty crashes, especially those involving one or more fatalities. Data is often missing about other persons in the same vehicle if they were not seriously injured. Even in fatal crashes there is very frequently information missing about other vehicles and their occupants if no one in those vehicles was seriously injured or killed. This means that the dataset underrepresents crashes that do not involve a serious injury and overrepresents crashes which involve one or more fatalities. In effect the study has been conducted on a subset of all people involved in vehicle crashes that is largely made up of those who were in a vehicle in which at least one occupant was killed or seriously injured in the crash.

There are also significant differences between jurisdictions in terms of data collection. While all efforts have been made to harmonise data across different jurisdictions this is not always possible, and not all information available for a single jurisdiction is available for all jurisdictions. Some jurisdictions have also not provided some data items, specifically the location of ambulance stations has not been provided by Western Australia and the speed limit of the road where the crash took place has not been provided by the Australian Capital Territory. Significance tests on the distance to ambulance stations was carried out without including Western Australian observations, while the posted speed limit in the ACT has been imputed by matching the latitude and longitude provided to the nearest road within 20 meters.

Access to emergency medical care

Durations have perhaps the most important relationship with the variables of interest in this study and time has a complex relationship with access to emergency medical treatment. For conceptual clarity, the Gantt chart (figure 1) below outlines time as relevant to a crash.

Figure 1: Emergency medical response timeline Gantt chart



Source: BITRE analysis

As can be seen in the figure above, a crash and the path to hospital can be viewed as a sequence of events. This begins with the *crash* and is followed by *discovery*, where the crash is found to have occurred. The next event is *activation*, where a response is mobilised, followed by the *response start*, which is where first responders begin providing first aid etc. and/or transport to emergency medical care if required. The final event is *arrival at hospital*.

This shows that the most accurate model of the effect of time would control for total pre-hospital time by including all of the pre-hospital durations. In many jurisdictions the time between either discovery and response or sometimes activation and response is recorded, but this information is not linked to the NCD. Were this available, they would make useful control variables, although the time between the crash and discovery would remain unknown. The time-on-scene is also a complex consideration as this is a clinical decision and there is no reason to assume that a shorter on-scene time is of greater benefit to crash victims.

In practice this study does not have a control for pre-hospital time. As highlighted by all Ambulance Services contacted for this study, the station location does not provide a good proxy of travel times to incidents due to the majority of responses, particularly in metropolitan areas, not occurring from stations. Although tested, ultimately ambulance station proximity did not have sufficient explanatory power to be included in the final model.

As access to pre-hospital emergency medical care and access to hospital care are highly correlated, it is difficult to separate the effect of each using the data available. The distance to the nearest Principal Referral Hospital has been included in the final model and is effectively a proxy for both types of access. Distance to the nearest emergency department was also tested, but was found to be insignificant, which may reflect the severity of crashes included in the dataset and the corresponding high level of care required.

The built-up character of the crash site was determined by whether the crash occurred within a 2016 Urban Centre or Locality (UCL) or within the remainder of the state. UCLs represent areas of concentrated urban development with populations of 200 people or more. These areas of urban development have been identified by the Australian Bureau of Statistics on the basis of dwelling and population density criteria using data from the 2016 Census (ABS 2018).

Data controlling for other relevant factors

The person-level characteristics of both sex and age have been included, along with four types of vehicle;

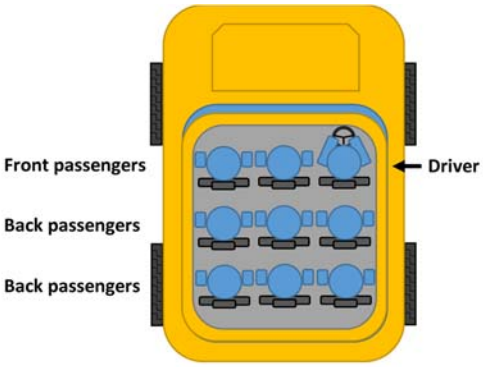
- pedal cycles,
- motorcycles,
- light vehicles (gross vehicle mass less than 4.5 tonnes and not a pedal cycle or motorcycle),
- heavy vehicles (including both trucks with a gross vehicle mass over 4.5 tonnes and buses with 10 or more seats).

The position of people inside vehicles has been grouped into categories depending on the vehicle type;

- pedal cycle and motorcycle riders have been grouped with their respective pillion passengers,
- occupants of heavy vehicles have been grouped into drivers and passengers,
- occupants of light vehicles have been grouped into the categories of driver, front passenger, back passenger and other passenger.[1]

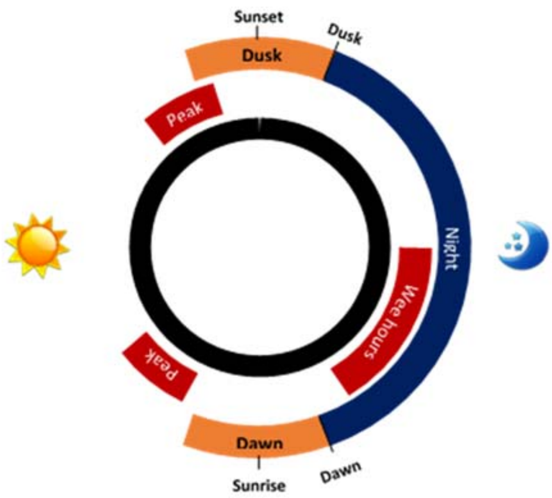
As this differs significantly from the way that the data is reported and coded in the NCD the seat positions for light vehicles have been set out in Figure 2 (below). Each position in a vehicle has been further separated into those wearing a helmet (for pedal cycles and motorcycles) and those wearing a restraint (light and heavy vehicles).

Figure 2: Seat positions for light vehicles



The model also makes use of derived time variables which represent both environmental and human phenomena. These are dawn and dusk windows, night-time, morning and afternoon peak hours (07:30 - 9:30 and 16:30 – 18:00 weekdays), the *wee hours* of the morning (00am – 04am), and weekends (18:00 Friday – 24:00 Sunday). These periods can obviously occur at the same time, for example, evening peak hour may occur during the day, during dusk or even during the night. The periods in environmental time change relative to both each other and to time of day and depend on the exact location of the crash and the time of the year. The relationships between the other (social) times of the day remain fixed with respect to each other. For clarity, the time variables are shown below in Figure 3 on a 24-hour timeline.

Figure 3: Derived environmental and time of day variables



[1] includes not seated

Finally, a number of variables have been included which correspond to the nature of the crash itself. The inclusion of these variables, along with a proxy for speed (the posted speed limit), go some way to controlling for the nature and severity of the crash. These are whether the crash has been identified as:

- occurring at an intersection
- a head-on crash
- a single vehicle crash
- a crash in which at least one vehicle ran off the road
- a crash in which at least one driver failed an alcohol test
- a crash in which a heavy vehicle was involved, where the person was not in a heavy vehicle

An overview of all of the variables used in the final model is included below in table 4.

Table 4: Variable descriptions

Variable	Values
Fatal (Dependent variable)	1 if killed, 0 if survived
Person characteristics	
Sex	1 if male, 0 if female (base case)
Over 64 years of age	1 if over 64, 0 otherwise (base case)
Under 15 years of age	1 if under 15, 0 otherwise (base case)
Position in vehicle	
Light vehicle - Driver (no restraint)	1 if true, 0 otherwise (base case)
Light vehicle - Passenger, back	1 if true, 0 otherwise (base case)
Light vehicle - Passenger, back (no restraint)	1 if true, 0 otherwise (base case)
Light vehicle - Passenger, front	1 if true, 0 otherwise (base case)
Light vehicle - Passenger, front (no restraint)	1 if true, 0 otherwise (base case)
Light vehicle - Passenger, other	1 if true, 0 otherwise (base case)
Light vehicle - Passenger, other (no restraint)	1 if true, 0 otherwise (base case)
Heavy vehicle - Driver	1 if true, 0 otherwise (base case)
Heavy vehicle - Driver (no restraint)	1 if true, 0 otherwise (base case)
Heavy vehicle - Passenger	1 if true, 0 otherwise (base case)
Heavy vehicle - Passenger (no restraint)	1 if true, 0 otherwise (base case)
Motorcycle - Motorcycle rider or pillion	1 if true, 0 otherwise (base case)
Motorcycle - Motorcycle rider or pillion (no helmet)	1 if true, 0 otherwise (base case)
Pedal cycle - Pedal cycle rider or pillion	1 if true, 0 otherwise (base case)
Pedal cycle - Pedal cycle rider or pillion (no helmet)	1 if true, 0 otherwise (base case)
Light vehicle year of manufacture	If light vehicle then years in 10 year increments, centred on the year 2000 (base case), 0 otherwise
Nature of the crash	
Intersection crash	1 if confirmed, 0 otherwise (base case)
Head-on crash	1 if confirmed, 0 otherwise (base case)
Single vehicle crash	1 if confirmed, 0 otherwise (base case)
Run off road	1 if confirmed, 0 otherwise (base case)
Alcohol fail involved	1 if confirmed, 0 otherwise (base case)
Heavy vehicle involved	1 if confirmed and not the driver or passenger of a heavy vehicle, 0 otherwise (base case)
Posted speed limit	Km per hour in 10 km per hour increments, centred around 60km per hour (base case)

Environmental time	
Dawn window	1 if within time window, 0 otherwise (base case)
Dusk window	1 if within time window, 0 otherwise (base case)
Night	1 if within time window, 0 otherwise (base case)
Social time	
Wee hours	1 if local time between 00:00 and 04:00, 0 otherwise (base case)
Weekend	1 if between Friday 06:00 and Sunday 24:00 local time, 0 otherwise (base case)
Morning peak	1 if local time between 07:30 and 09:30, 0 otherwise (base case)
Afternoon peak	1 if local time between 16:30 and 18:00, 0 otherwise (base case)
Spatial factors	
Distance to a Principal Referral Hospital	Km by road network in continuous 100km increments from 0 (base case)
Non-urban	0 if within a UCL (base case), 1 if in the remainder of state
State	Base case = New South Wales
Northern Territory	1 if within the Northern Territory, 0 otherwise
Queensland	1 if within the Queensland, 0 otherwise
South Australia	1 if within South Australia, 0 otherwise
Tasmania	1 if within Tasmania, 0 otherwise
Victoria	1 if within Victoria, 0 otherwise
Western Australia	1 if within Western Australia, 0 otherwise
Australian Capital Territory	1 if within Australian Capital Territory, 0 otherwise

Method

Ideally the model would reflect the data generation process. In the case of vehicle crashes, that process is likely to be hierarchical, with a set of conditions that lead to a person being either injured or not-injured, and a set of conditions that may or may not be the same that result in a serious injury or fatality.

The missing data in the NCD is most concentrated in missing observations from vehicles in which no person was seriously injured or killed. This leaves very poor information on which to model non-injury versus injury outcomes. As a consequences, this analysis has not modelled the data hierarchically, although this would best reflect the data generation process. Instead the data has been modelled using a logistic regression on a binary variable of whether a person involved in a crash and included in the NCD survived (0) or was killed (1).

Technically, people are within vehicles, which are within crashes. This results in two levels of clustering, and would normally make the necessary assumption that each person-level observation is independent problematic. The consequence of clustering in vehicles and crashes is that there is likely to be some correlation between the outcome of persons who are in the same vehicle or crash. In this context, this may result in underestimation of the standard error and, consequently, a higher likelihood of finding significant results when in fact they are not significant (Desai & Begg 2008).

Ideally, the solution would be to use a mixed effects model with random effects for each level of clustering (crashes and vehicles). However, there are insufficient observations to support either level of clustering as there are simply too many crashes and too many vehicles for the number of people involved. With two levels of clustering the combined number of random effects is greater than the number of observations. Each level of clustering was tested individually, however the number of crashes (the smaller level) still contained too many random effects for too few observations and the model was unable to converge.

Consequently, logistic regression has been used without taking into account clustering, bearing in mind that there may be a relationship between observations, and so there may be a bias towards finding results to be significant.

Results

A summary of the model results is included below in Table 5 and is followed by the parameter estimates in Table 6 and the odds ratios in Table 7.

Table 5: Model summary

Observations (persons):	227,566
Number of Fisher Scoring iterations:	9
Convergence status:	Converged
Pseudo R-squared	0.271[2]
Null deviance (227,565 degrees of freedom)	25,462
Residual deviance: on (227,523 degrees of freedom)	18,844

Table 6: Parameter estimates

Variable	Estimate	Std. Error	Pr(<|z|)[3]	Significance
Intercept	-7.16	0.09	< 2e-16	***
Person characteristics				
Sex (male)	0.25	0.05	2.86E-06	***
Over 64 years of age	1.30	0.06	< 2e-16	***
Under 15 years of age	-0.35	0.13	9.51E-03	***
Position in vehicle				
Light vehicle - Driver (no restraint)	2.49	0.10	< 2e-16	***
Light vehicle - Passenger, back	0.55	0.11	9.15E-07	***
Light vehicle - Passenger, back (no restraint)	2.24	0.18	< 2e-16	***
Light vehicle - Passenger, front	0.40	0.08	1.64E-07	***
Light vehicle - Passenger, front (no restraint)	1.94	0.17	< 2e-16	***
Light vehicle - Passenger, other	0.67	0.52	1.97E-01	
Light vehicle - Passenger, other (no restraint)	2.57	0.29	< 2e-16	***
Heavy vehicle - Driver	0.16	0.13	2.21E-01	
Heavy vehicle - Driver (no restraint)	1.85	0.27	7.97E-12	***
Heavy vehicle - Passenger	-0.04	0.31	8.98E-01	
Heavy vehicle - Passenger (no restraint)	1.74	0.40	1.08E-05	***
Motorcycle - Motorcycle rider or pillion	1.74	0.07	< 2e-16	***
Motorcycle - Motorcycle rider or pillion (no helmet)	2.32	0.19	< 2e-16	***
Pedal cycle - Pedal cycle rider or pillion	1.62	0.14	< 2e-16	***
Pedal cycle - Pedal cycle rider or pillion (no helmet)	2.56	0.26	< 2e-16	***
Light vehicle year of manufacture	-0.19	0.03	3.07E-08	***
Nature of the crash				
Intersection crash	-0.11	0.06	8.76E-02	*
Head-on crash	2.05	0.07	< 2e-16	***
Single vehicle crash	0.49	0.08	1.32E-09	***

[2] This is a similar level of fit to Sanchez-Mangas et al. (2010) who also report a (Nagelkerke) Pseudo R-squared ranging from 0.104 to 0.217, depending on the model.

[3] Probability that the value of the coefficient is equal to 0.

Variable	Estimate	Std. Error	Pr(<|z|)[4]	Significance
Run off road	0.79	0.07	< 2e-16	***
Alcohol fail involved	1.07	0.07	< 2e-16	***
Heavy vehicle involved	1.76	0.07	< 2e-16	***
Posted speed limit	0.23	0.02	< 2e-16	***
Environmental time				
Dawn window	0.19	0.14	1.67E-01	
Dusk window	0.32	0.15	2.94E-02	**
Night	0.44	0.07	2.31E-10	***
Social time				
Wee hours	0.11	0.10	2.37E-01	
Weekend	-0.05	0.06	4.00E-01	
Morning peak	-0.31	0.10	1.97E-03	***
Afternoon peak	0.06	0.09	4.91E-01	
Spatial factors				
Distance to a Principle Referral Hospital	0.03	0.01	3.31E-02	**
Non-urban	0.65	0.07	< 2e-16	***
State				
Northern Territory	0.27	0.13	3.49E-02	**
Queensland	0.19	0.06	8.09E-04	***
South Australia	0.06	0.08	4.40E-01	
Tasmania	-0.47	0.12	8.53E-05	***
Victoria	-0.23	0.10	2.08E-02	**
Western Australia	0.96	0.08	< 2e-16	***
Australian Capital Territory	0.63	0.25	1.28E-02	**

Significance level: 0.001 = '***', 0.01 = '**', 0.05 = '*'

To aid the discussion the odds ratios of the parameters of interest are provided below in Table 7. For readers not familiar with logistic regression the following points may assist in interpreting the odds ratio:

- A value below one represents a reduction in the odds of a person being killed in a crash
- A value above one represents an increase in the odds of a person being killed in a crash
- Comparisons of magnitude are possible, however the units of each explanatory variable may not be comparable. For example a 1 unit increase in the posted speed limit (from 60km per hour to 70km per hour) is not in the same unit as a 1 unit increase in the vehicle year of manufacture (from a vehicle built in the year 2000 to a vehicle built in 2010)

When considering the estimates some important features of the base case are:

- The person is: female, between the age of 15 and 64 inclusive, in the driver position, wearing a restraint
- The vehicle is a light vehicle
- The crash is on a 60km per hour road, located in an built up area, 0km from a Principal Referral Hospital
- The time is between 9:30 and 16:30 on a weekday during daylight

[4] Probability that the value of the coefficient is equal to 0.

Table 7: Odds ratios

	Odds Ratio	95 per cent confidence interval Lower	Upper	Significance
Person characteristics				
Sex (male)	1.28	1.16	1.42	***
Over 64 years of age	3.68	3.28	4.13	***
Under 15 years of age	0.71	0.54	0.92	***
Position in vehicle				
Light vehicle - Driver (no restraint)	12.02	9.92	14.56	***
Light vehicle - Passenger, back	1.73	1.39	2.15	***
Light vehicle - Passenger, back (no restraint)	9.39	6.56	13.42	***
Light vehicle - Passenger, front	1.50	1.29	1.74	***
Light vehicle - Passenger, front (no restraint)	6.95	4.98	9.69	***
Light vehicle - Passenger, other	1.95	0.71	5.39	
Light vehicle - Passenger, other (no restraint)	13.09	7.36	23.29	***
Heavy vehicle - Driver	1.18	0.91	1.53	
Heavy vehicle - Driver (no restraint)	6.35	3.74	10.78	***
Heavy vehicle - Passenger	0.96	0.52	1.77	
Heavy vehicle - Passenger (no restraint)	5.70	2.63	12.38	***
Motorcycle - Motorcycle rider or pillion	5.68	4.99	6.48	***
Motorcycle - Motorcycle rider or pillion (no helmet)	10.14	6.93	14.82	***
Pedal cycle - Pedal cycle rider or pillion	5.05	3.87	6.58	***
Pedal cycle - Pedal cycle rider or pillion (no helmet)	12.91	7.76	21.47	***
Light vehicle year of manufacture	0.82	0.77	0.88	***
Nature of the crash				
Intersection crash	0.90	0.79	1.02	*
Head-on crash	7.78	6.75	8.97	***
Single vehicle crash	1.64	1.40	1.92	***
Run off road	2.20	1.91	2.54	***
Alcohol fail involved	2.92	2.55	3.34	***
Heavy vehicle involved	5.80	5.02	6.70	***
Posted speed limit	1.25	1.22	1.29	***
Environmental time				
Dawn window	1.21	0.92	1.57	
Dusk window	1.38	1.03	1.83	**
Night	1.55	1.35	1.77	***
Social time				
Wee hours	1.12	0.93	1.35	
Weekend	0.95	0.85	1.07	
Morning peak	0.73	0.60	0.89	***
Afternoon peak	1.06	0.89	1.27	
Spatial factors				
Distance to a Principle Referral Hospital	1.03	1.00	1.05	**
Non-urban	1.91	1.67	2.18	***

Significance level: 0.001 = '***', 0.01 = '**', 0.05 = '*'

As the odds ratio can be difficult to interpret, probabilities against the base case have been provided for some of the most policy relevant variables in the discussion section. When evaluating the probabilities it is important to bear in mind that that the probabilities change over the values of the other variables – they are only correct with respect to the base case and the event described, not over all cases or all values of the explanatory variable. The magnitude may also be biased if the observations included in the study are not a representative sample of the population of all persons involved in crashes. This is certainly possible due to the high number of missing observations/missing information and their concentration in vehicles in which no persons was injured or killed.

Discussion

In terms of person characteristics, males have higher odds of being killed than females, a result in line with other comparable studies (Maio et al. 1992 and Travis et al. 2012). Persons over the age of 64 have increased odds of being killed, while persons under 15 years have higher odds of survival, though it is worth keeping in mind that survival also depends on seat position and it is not common for a person under 15 to be the driver. Comparable studies have also found that older people have a higher chance of being killed while younger people have a higher chance of survival (for example Maio et al. 1992, Sanchez-Mangas et al. 2010, and Travis et al. 2012)

The results show the importance of wearing a restraint (seatbelt) or helmet, as appropriate to the vehicle type. This is in line with other studies that include the variable, for example Maio et al. (1992) and Travis et al. (2012). Table 8 below compares the probability within the sample of being killed by vehicle type depending on whether the person was wearing a restraint/helmet. Other than the vehicle type and wearing a restraint/helmet, all other variables are held as per the base case. While the results show the effect of restraints and helmets, *given that a crash has occurred*, information on the probability of a crash occurring in the first place is required to properly inform the costs and benefits of policy changes.

Table 8: Benefits of restraints (seatbelts) and helmets

Vehicle type	Base case probability of fatality	Without helmet/restraint probability of fatality
Motorcycle - Motorcycle rider or pillion	0.44 per cent	4.29 per cent
Pedal cycle - Pedal cycle rider or pillion	0.39 per cent	4.83 per cent
Light vehicle - Driver	0.08 per cent	0.93 per cent
Light vehicle - Passenger, front	0.12 per cent	0.80 per cent
Light vehicle - Passenger, back	0.13 per cent	1.25 per cent
Light vehicle - Passenger, other	0.15 per cent	1.95 per cent
Heavy vehicle - Driver	0.09 per cent	0.58 per cent
Heavy vehicle - Passenger	0.07 per cent	0.42 per cent

Note: The base case probability of a fatality for heavy vehicle drivers and passengers are not significantly different from light vehicle drivers.

Table 8 also shows the probabilities of an individual involved in a crash being killed for vehicle types and positions in vehicles (under the conditions that other variables remain as per the base case). Motorcyclists are at the greatest risk if involved in a crash, followed by pedal cyclists, noting that there is known under reporting of serious injury (but non-fatal) crashes for these groups. Light vehicles and heavy vehicles are safest, with the difference between the base case and heavy vehicle drivers and passengers being statistically insignificant.

For light vehicles, holding all other values as per the base case, light vehicle drivers are in the safest position, followed by front passengers, then back passengers. The base case is a vehicle built in the year 2000 and has the base probability of an individual involved in a crash being killed of 0.08 per cent. Holding all other values at the base case, a car built in 1990 is less safe, with the probability increasing to 0.09 per cent, while a car built in 2010 is safer, with the probability decreasing to 0.06 per cent.

The results also show the consequences of various types of crashes. In order of magnitude, head-on crashes, crashes involving heavy vehicles and crashes where a vehicle has run off the road increase the odds of being killed. The increased odds found for vehicles run off the road was also found to be significant in the study by Sanchez-Mangas et al. (2010). In contrast to the study by Al-Ghamdi (2002), intersection crashes appear to increase survival. The model also suggests that there is increased mortality in crashes where at least one driver failed an alcohol test, as has been found in Queensland data by Siskind et al. (2011).

Crashes that occur on roads with higher posted speed limits have a greater chance of being fatal. Holding all other values as per the base case, the in-sample probability of a person involved in a crash of being killed on a 60km per hour road is 0.08 per cent, while at 70km per hour this increases to 0.10 per cent.

The odds of survival are reduced at night and during the dusk window, while crashes during the morning peak correspond to increased odds of survival. Neither the dawn window, wee hours or afternoon peak were significant. The significance of night and the insignificance of the wee hours points to the importance of the environmental conditions of night-time rather than the social conditions and behaviour associated with the early hours of the morning. In terms of comparison to other relevant studies, Travis et al. (2012) have found time variable, specifically the 00:00 – 06:00 window significantly decreases the odds of survival, but did not separate the environmental and social aspects, while Sanchez-Mangas et al. 2010 found night to be a significant factor.

The results show higher mortality in rural areas, controlling for the distance to a Principal Referral Hospital, which is the proxy for access to emergency medical care. Holding all other variables as per the base case the in-sample probability of a person involved in a crash being killed on a road in a built-up area is 0.08 per cent and 0.15 per cent on a rural road. The results also show increased mortality the further a crash is from a Principal Referral Hospital. Again the base-case probability of an individual involved in a crash being killed is 0.08 per cent when 0 km from a Principal Referral Hospital, increasing by around 2.7 per cent per 100km, or to around 0.09 per cent when 500km from a Principal Referral Hospital. This relationship could be related to hospital care, however as the distance to hospital is correlated with access to other emergency health care such as pre-hospital time, it would be an overreach to suggest that proximity to the hospital itself is the primary cause of improved survival.

The model specification also includes state/territory-specific constant variables, which reflect the difference in the proportion of fatalities to persons involved in crashes across jurisdictions, relative to that of New South Wales, and not explained by other factors included in the model. These are highly related to differences in the definitions and scope of data provided to the NCD by each jurisdiction, rather than necessarily being related to unexplained differences in survival.

Conclusion

This study has investigated the consequences of crashes *given that they have occurred*. To calculate the costs and benefits of preventing crashes one not only needs information on the severity of an event, but also information on the likelihood of the event occurring in the first place. For this reason, the contribution of this study to knowledge of the factors influencing survival in a crash is only a contribution to part of the information required for policy making – an important factor to bear in mind when interpreting the results.

Based on this research it appears that increased road-related mortality in rural areas is correlated with both the distance to a Principal Referral Hospital's healthcare and to the nature of rural areas themselves – noting that the analysis controls for differences in the posted speed limit. However, it is not clear from this study how important access to hospital care is relative to care from first responders, as these factors are highly correlated. This area would benefit from further research as it has direct policy implications for areas including the operational decisions of ambulance services and the locations of hospitals.

In order to better understand why people survive, more and more complete data is required for survivors, including those who do not suffer any injury in a crash. This is an area in which current data collection falls short. Extending the data to include all motor vehicle crashes and efforts to reduce the level of missing information would provide much greater insight into why crashes occur and why the consequences vary. This would be of great benefit to policy makers in improving road safety and better directing infrastructure spending.

Beyond an extension of the scope of data to include all persons involved in any crash, the most important extension of the dataset with respect to this type of study is the inclusion of information about ambulance activation, response, on-scene, and transport times.

Other potential extensions that have been found to be significant in comparable studies include: a flag for a crash involving a vehicle travelling in the wrong direction (Al-Ghamdi 2002), a flag for a crash involving a failure to yield (Al-Ghamdi 2002) and a flag for vehicle roll over (Travis et al. 2012). Additional vehicle information might include a deformity index (Maio et al. 1992) or record vehicle damage (Travis et al. 2012).

References

Al-Ghamdi, A.S., 2002. Pedestrian–vehicle crashes and analytical techniques for stratified contingency tables. *Crash Analysis & Prevention*, 34(2), pp.205-214.

Australian Bureau of Statistics (2017), Cat. No. 1270.0.55.004 - Australian Statistical Geography Standard (ASGS): Volume 4 - Significant Urban Areas, Urban Centres and Localities, Section of State, July 2016

Bentham, G., 1986. Proximity to hospital and mortality from motor vehicle traffic crashes. *Social science & medicine*, 23(10), pp.1021-1026.

Bureau of Infrastructure, Transport and Regional Economics (BITRE), 2014. *Impact of road trauma and measures to improve outcomes*, Report 140, December.

Bureau of Infrastructure, Transport and Regional Economics (BITRE), 2016. *Road trauma Australia*, 2016 statistical summary.

Brodsky, H. and Hakkert, A.S., 1983. Highway fatal crashes and accessibility of emergency medical services. *Social science & medicine*, 17(11), pp.731-740.

Desai, M. and Begg, M.D., 2008. A comparison of regression approaches for analyzing clustered data. *American journal of public health*, *98*(8), pp.1425-1429.

Harmsen, A.M.K., Giannakopoulos, G.F., Moerbeek, P.R., Jansma, E.P., Bonjer, H.J. and Bloemers, F.W., 2015. The influence of prehospital time on trauma patients' outcome: a systematic review. *Injury*, 46(4), pp.602-609.

Hoffman, E., 1976. Mortality and morbidity following road crashes. *Annals of The Royal College of Surgeons of England*, 58(3), p.233

Maio, R.F., Green, P.E., Becker, M.P., Burney, R.E. and Compton, C., 1992. Rural motor vehicle crash mortality: the role of crash severity and medical resources. *Crash Analysis & Prevention*, 24(6), pp.631-642.

McAndrews, C., Beyer, K., Guse, C.E. and Layde, P., 2017. Are rural places less safe for motorists? Definitions of urban and rural to understand road safety disparities. *Injury prevention*, 23(6), pp.412-415.

Sánchez-Mangas, R., García-Ferrrer, A., De Juan, A. and Arroyo, A.M., 2010. The probability of death in road traffic crashes. How important is a quick medical response? *Crash Analysis & Prevention*, 42(4), pp.1048-1056.

Siskind, V., Steinhardt, D., Sheehan, M., O'Connor, T. and Hanks, H., 2011. Risk factors for fatal crashes in rural Australia. *Crash Analysis & Pre*vention, 43(3), pp.1082-1088.

Tay, R., 2003. Marginal effects of changing the vehicle mix on fatal crashes. *Journal of Transport Economics and Policy* (JTEP), 37(3), pp.439-450.

Travis, L.L., Clark, D.E., Haskins, A.E. and Kilch, J.A., 2012. Mortality in rural locations after severe injuries from motor vehicle crashes. *Journal of safety research*, 43(5-6), pp.375-380.

Appendix A:

Table A1: Counts of binary variables, 2014 - 2016

Variable	Survived	Killed	Total
Person characteristics			
Sex (male)	127,939	1,659	129,598
Over 64 years of age	21,801	467	22,268
Under 15 years of age	11,722	78	11,800
Position in vehicle			
Light vehicle - Driver (no restraint)	1,265	198	1,463
Light vehicle - Passenger, back	14,697	132	14,829
Light vehicle - Passenger, back (no restraint)	479	45	524
Light vehicle - Passenger, front	27,586	280	27,866
Light vehicle - Passenger, front (no restraint)	648	49	697
Light vehicle - Passenger, other	250	4	254
Light vehicle - Passenger, other (no restraint)	171	15	186
Heavy vehicle - Driver	6,232	69	6,301
Heavy vehicle - Driver (no restraint)	249	18	267
Heavy vehicle - Passenger	1,037	11	1,048
Heavy vehicle - Passenger (no restraint)	307	7	314
Motorcycle - Motorcycle rider or pillion	16,418	489	16,907
Motorcycle - Motorcycle rider or pillion (no helmet)	471	35	506
Pedal cycle - Pedal cycle rider or pillion	8,731	72	8,803
Pedal cycle - Pedal cycle rider or pillion (no helmet)	779	17	796
Nature of the crash			
Intersection crash	112,657	433	113,090
Head-on crash	10,563	531	11,094
Single vehicle crash	41,169	1,139	42,308
Run off road	31,907	933	32,840
Alcohol fail involved	7,988	406	8,394
Heavy vehicle involved	8,760	325	9,085
Environmental time			
Dawn window	6,815	63	6,878
Dusk window	6,196	54	6,250
Night	26,033	559	26,592
Social time			
Wee hours	6,896	217	7,113
Weekend	33,276	454	33,730
Morning peak	25,762	119	25,881
Afternoon peak	21,965	158	22,123
Spatial factors			
Non-urban	42,049	1,496	43,545

Figure A1: Histogram of persons in light vehicles by year of manufacture, 2014 - 2016

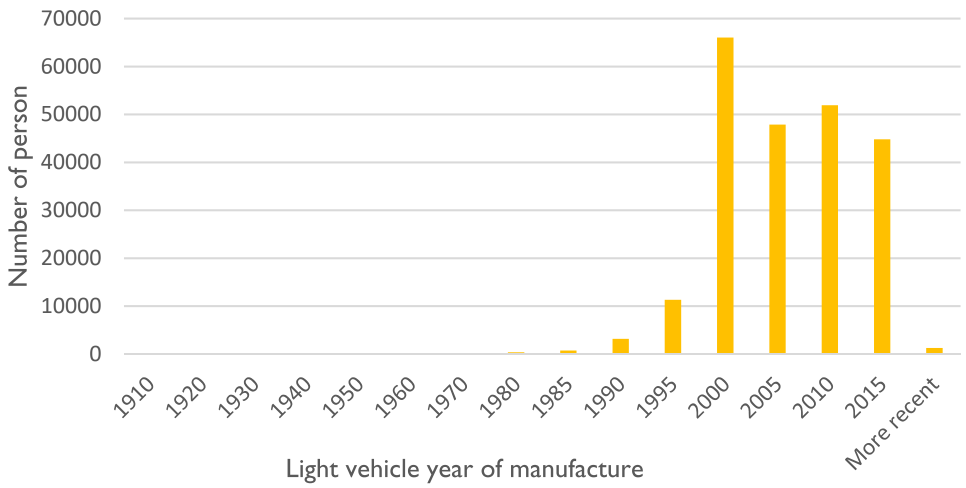


Figure A2: Histogram of posted speed limit, 2014 - 2016

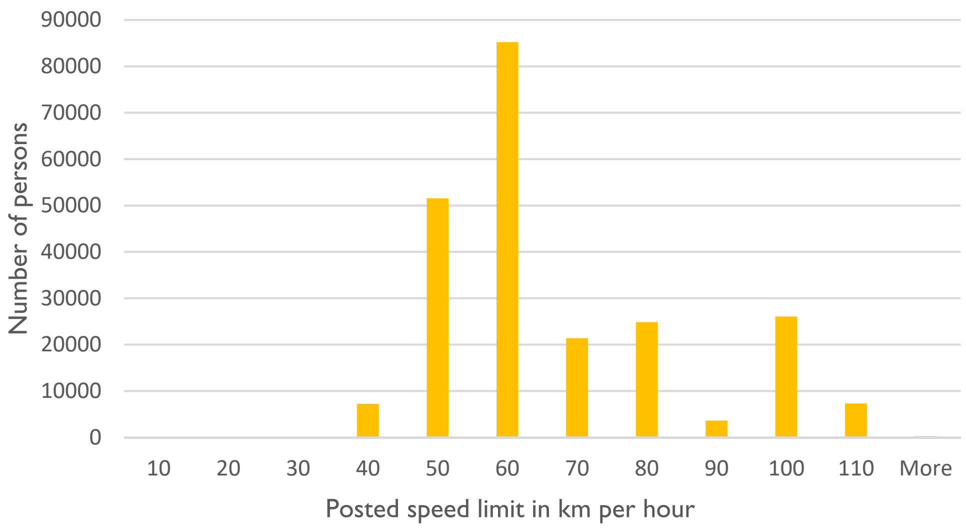
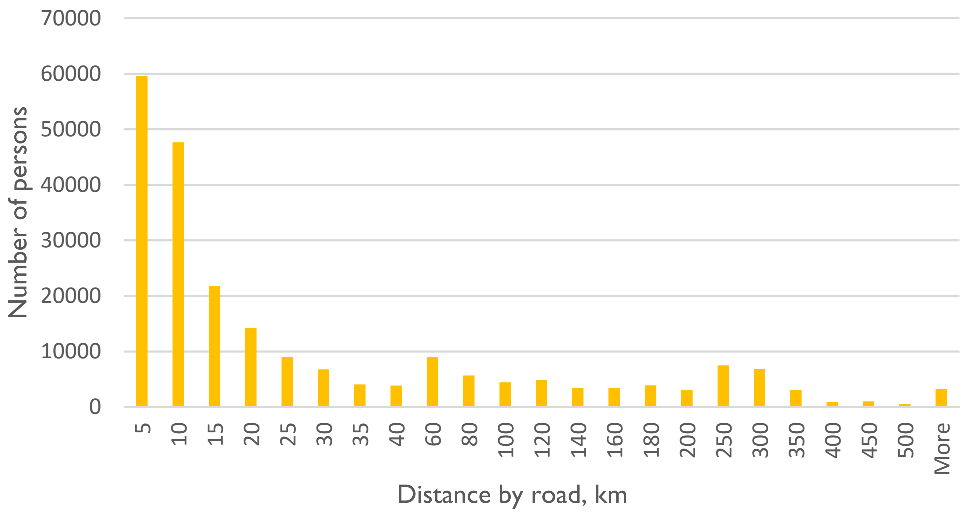


Figure A3: Histogram of distance to a Principal Referral Hospital, 2014 - 2016



ISSN 1440-9593

ISBN 978-1-925701-25-8

INFRA-3556

May 2018

This publication should be attributed in the following way; Bureau of Infrastructure, Transport and Regional Economics (BITRE), 2018. *Location and other risk factors in crashes*, Information sheet 97, BITRE, Canberra.

Acknowledgement

Information sheet authored by Kyle Thomson. The author would like to acknowledge the contribution of Helen Yang, Thomas Rutherford, Peter Johnston, Mark Cregan, Tim Risbey and Mari Adams to this research.

Contact us

Department of Infrastructure, Regional Development and Cities
Bureau of Infrastructure, Transport and Regional Economics (BITRE)
GPO Box 501, Canberra ACT 2601, Australia

Phone: (international) +61 2 6274 7210
Fax: (international) +61 2 6274 6855
Email: bitre@infrastructure.gov.au
Website: www.bitre.gov.au