THE UNIVERSITY OF
NEW SOUTH WALES

NSW Injury Risk Management Research Centre
University of New South Wales

Data Linkage of Hospital and Police Crash Datasets in NSW

Soufiane Boufous, Caroline Finch, Andrew Hayen, Ann Williamson

July 2008

UNSW, SYDNEY NSW 2052, AUSTRALIA
Telephone: +61 (2) 9385 4207, Facsimile: +61 (2) 9385 6040, http://www.irmrc.unsw.edu.au

ISBN 978-0-9804866-1-2

Published by:
NSW Injury Risk Management Research Centre
University of New South Wales,
Sydney NSW 2052, Australia
Telephone: +61 (2) 9385 4207 Facsimile: +61 (2) 9385 6040 http://www.irmrc.unsw.edu.au

Recommended citation:
Boufous S, Finch C, Hayen A, Williamson A. Data Linkage of Hospital and Police Crash Datasets in NSW. Sydney: New South Wales Injury Risk Management Research Centre, July 2008.

Table of Contents

List of Figures

List of Tables

ACKNOWLEDGEMENTS

This project was funded by the IRMRC Core Funding which is provided by the NSW Department of Health, the NSW Roads and Traffic Authority and the NSW Motor Accidents Authority.

Prof Caroline Finch was supported by an NHMRC Principal Research Fellowship and A/Prof Ann Williamson by an NHMRC Senior Research Fellowship.

The Traffic Accidents Data System (TADS) data was obtained from the NSW Roads and Traffic Authority (RTA), Road Safety and Road User Management Directorate (RS & RUM).

The New South Wales (NSW) Inpatient Statistics Collection (ISC) data were accessed via the NSW Department of Health's *Health Outcomes Information Statistical Toolkit* (HOIST), maintained by the Centre for Epidemiology and Research.

Comments on the draft report were received with thanks from:

Ms Liz Develin (Centre for Health Advancement, NSW Department of Health)

Ms Carmel Donnelly (NSW Motor Accidents Authority)

Aarthi Ayaar and Sanja Lujic are thanked for their contributions to this project whilst undertaking their NSW Health Biostatistical Trainee Placement at the IRMRC

Executive Summary

Good quality data on injuries sustained by motor vehicle trauma victims and the causally-related factors leading to the motor vehicle crashes is vital to informing policies designed to reduce the burden of road trauma. It is well recognised that such epidemiological information can rarely be obtained from a single data source and that data linkage of relevant databases has the potential to overcome the limitations of individual data sources, thereby maximising their collective benefit.

The report describes the process of linking hospital separations and traffic crashes datasets to provide a more comprehensive picture on traffic injuries in NSW. Crash records were selected from the Traffic Accident Data System (TADS), which contains information on road crashes on public roads reported to police in which either a person was injured or killed, or at least one vehicle was towed away. Hospital separation records relating to land transport-related accidents, including road and traffic accidents, were selected from the NSW Inpatient Statistics Collection (ISC), a census of all hospital separations from NSW public and private hospitals.

In total, 29,538 records of casualties (injuries and deaths) from police crash records (TADS) were linked with 19,277 hospital separation records for land transport accidents for the period 30th June 2000 and 30th June 2001. Records from the two datasets were linked using probabilistic linkage methods. Records were matched using information that was common to the datasets, such as names, residential addresses and dates of birth. Doubtful links between the two datasets were clerically reviewed.

The adopted linkage process produced comparable matching rates to those found around the world in similar exercises. In total, just under half (44.9%) of the ISC records had a matching crash record. When restricted to ISC records relating to traffic crashes, the linkage rate was 56.2%, and when further restricted to ISC records relating to motor vehicle traffic crashes on public roads the linkage rate was 69.2%.

There were a number of discrepancies between the coding of the type of road users in the two datasets. The agreement (as measured by the proportion of road users in the hospital identified similarly in TADS) for motor vehicle drivers (90%), passengers (87%) and pedestrians (87%) were high; the agreement for pedal cyclists (62%) was low.

The resulting linked dataset contains information on both the causes and consequences of traffic crashes. This report also identifies some of the strengths and weaknesses of the individual datasets.

1. Background

Injuries resulting from motor vehicle crashes are a major international public health problem (World Health Organization, 2004). Unless there is increased global attention given to preventing the problem over the next two decades, road traffic injuries will rank as the third leading global burden of disease and injury problem by 2020. In Australia, while the number of fatalities is decreasing, hospitalisations and the overall burden of road crashes remains significant. In 1996, the total cost of road crashes in Australia was conservatively estimated at approximately $15 billion (Bureau of Transport Economics, 2000).

In order to alleviate the burden of motor vehicle trauma there is a need to fully understand the nature of, and the causal factors underlying, the problem. This can only be achieved with good quality data describing the injuries sustained by motor vehicle trauma victims and the environmental/person/vehicle factors leading to the motor vehicle crash in the first place. It is well recognised that this information can rarely be obtained from a single data source due to limitations of individual datasets. It has therefore become increasingly common to link complementary data sources to enhance the value of motor transport injury data sources. Crash information from police reporting systems has been linked to other injury related datasets, including hospital discharge records, in an attempt to create a useful database for road trauma prevention in various parts of the world – in Australia (Boufous and Williamson, 2006; Ferrante et al., 1993; Rosman, 2001), England (Bull and Roberts, 1973; Cryer et al., 2001), New Zealand (Alsop and Langley, 2001; Langley et al., 2003) and in the United States (Singleton and Qin, 2004; Johnson and Walker, 1996).

In New South Wales, the Roads and Traffic Authority (RTA) Traffic Accident Data System (TADS) and the New South Wales (NSW) Inpatient Statistics Collection (ISC) are the two major sources of information on motor vehicle crashes. While the TADS is rich in information about the circumstances of traffic crashes it contains very little information about injury outcomes. The opposite is true for the ISC, which contains detailed information on injury outcomes but limited information on the circumstances of injury.

A linkage of NSW hospitalisations and police crash records was undertaken in order to minimize the limitations of both datasets and to provide a more complete picture of the circumstances leading to, as well as the nature and outcomes of, motor vehicle injuries. The linkage was carried out using probabilistic record linkage techniques in order to maximize the quality and the validity of the linked dataset compared to traditional deterministic methods.

Record linkage is the joining of information from two or more records that are considered to relate to a common entity, whether that entity is an individual, family, event, business, or address (Newcombe, 1998). When joined, such records are said to be linked. The manner in which record linkage is carried out varies according to the resources available, the type and amount of personal data held in each collection and the level of linkage accuracy which is deemed acceptable.

Probabilistic record linkage attempts to mimic the steps a human would go through mentally when deciding whether two records from two separate datasets belong to the same person. These include steps such as allowing for incomplete and/or error data; evaluating how common a particular name is in the sets of data being compared; assessing how likely it is that a particular pair would match at random; and determining how likely it is that full or partial

agreement on values in a given field is indicative of agreement for the whole record. This form of data linkage was adopted for this project.

The aim of this report is to describe the process of linking hospital and traffic crashes datasets to provide a more comprehensive picture on traffic injuries in NSW. Overall results and matching rates are provided and the implications of the findings on the surveillance and prevention of injury related to road traffic crashes are discussed.

2. Data sources and case selection

Before accessing the relevant datasets for the purpose of data linkage, ethics approval for the project was obtained from the NSW Health Department Ethics Committee and a clearance for access to personal information included in TADS was obtained from the NSW Police. The project was also ratified by the University of New South Wales Human Research Ethics Committee.

Traffic Accident Data System (TADS)

On 1 December 1999, the Traffic NSW Act was replaced by new traffic legislation, including the adoption of the Australian Road Rules. Rule 287 (3) of the Australian Road Rules requires an accident to be reported to police when: any person is killed or injured; when drivers involved in the accident do not exchange particulars; or when a vehicle involved in the accident is towed away (RTA, 2003). Thus, police are only required to attend an accident if:

- a person was killed or injured
- one of the parties failed to stop and exchange particulars
- one or more of the drivers was reported to be under the influence of alcohol or other drugs
- one or more of the vehicles was required to be towed away.

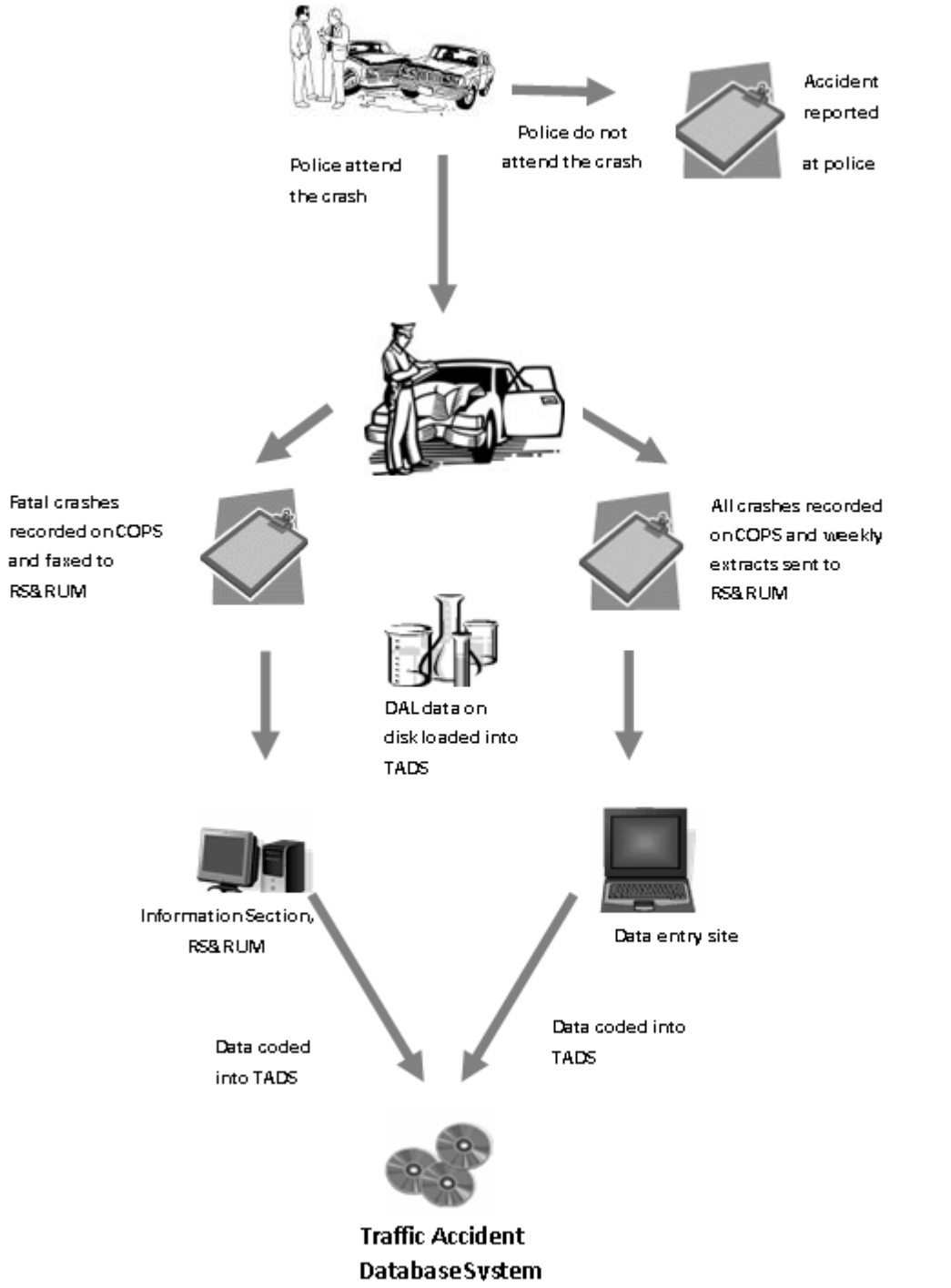
The TADS consists of records of every road traffic accident reported to the police across NSW. The dataset is maintained by, and was obtained from, the NSW Road and Traffic Authority (RTA), Road Safety and Road User Management Directorate (RS & RUM).

A ‘road traffic accident’ is defined to be an apparently unpremeditated event resulting in death, injury or property damage attributable to the movement of a road vehicle on a public road (RTA, 2003). Road vehicles include various means of transport such as bicycles, cars, motorcycles, trains, aircraft as well as vehicles may be propelled by animals,

Figure 1 shows the TADS data collection process. All road accidents attended by the police are recorded by the reporting officer on the Computerised Operational Policing System (COPS). Even if the police do not attend the accident, accidents that fulfil the above criteria may be recorded on COPS if they are later reported to the police. Other minor accidents such as (self reported accidents) can also be recorded on COPS. Other data, relating to alcohol use by controllers, are gathered from the Division of Analytical Laboratories (DAL), Sydney West Area Health Service.

Data are entered into TADS for all accidents in which a person was killed or injured or at least one motor vehicle was towed away (Figure 1). Some accident reports are not reported / received until the calendar year after the accident, which is after the annual accident database has been finalised. These amount to some 2% of casualty and tow-away accidents and are counted in the following year’s statistics.

Figure 1. Data collection procedure for road traffic crashes recorded in TADS



The TADS contains considerable information about the circumstances of traffic crashes reported to NSW police. The collection is subdivided into three components: the accident section, which contains information on the circumstances of the crash (day, time, location of accident, etc); the traffic unit section, which includes information on the vehicle(s) involved in the crash, and the casualty section, which contains information on the characteristics of the casualty. The accident section and the traffic unit sections are merged together using the accident number “unique identifier” and the resulting database is merged with the casualty section using the traffic unit number.

Variables included provide information about the date, time and location of the accident; number and type of vehicle (s) involved; age and sex of persons involved; an indication of the damage caused to the persons and vehicles involved; geographical area of residence of the controller/s of the vehicles; number, age and sex of persons treated and/or hospitalised, and information about the road and weather conditions at the time of the accident. The collection, however, does not include any information as to the exact nature and severity of any of the injuries sustained. For the purpose of the record linkage we also obtained the full name, date of birth and address of people injured in a traffic crash. Appendix I provides a full list of the variables included in TADS.

For the purpose of this data linkage project, 29,538 records of casualties resulting from all traffic crashes reported in TADS between June 30th 2000 and 1st of July 2001 were selected.

Inpatient Statistics Collection (ISC)

The ISC covers all inpatient separations from all public and private hospitals in NSW. The purpose of the data collection is to provide information to assist in the planning of an efficient and equitable distribution of health services, including various indicators of health status and statistical information to monitor the utilisation of NSW hospital services (Centre for Epidemiology and Research, 2003).

The ISC is a financial year collection from 1 July through to 30 June of the following year. Hospitals are required to submit details for every inpatient and for every episode of care. A separate return is processed for each period of inpatient care, irrespective of the time interval between the date of separation and subsequent re-admission. A period of stay in hospital ends with a discharge, transfer or death of a patient (Centre for Epidemiology and Research, 2003). An episode of care ends by either the patient ending a period of stay in hospital (i.e. by discharge, transfer or death) or by the patient becoming a different type of patient within the same period of stay in hospital. Examples of patient service types include acute care, rehabilitation care and palliative care. To illustrate this, consider a patient who has been admitted to a hospital for a multiple fracture of the pelvis as a result of a traffic crash. When first admitted he or she may be classified as an acute patient. After surgery, however, the patient may receive rehabilitation and therefore be classified as a rehabilitation patient. This case would therefore be reported as two separate records, as there are two episodes of care.

The collection contains demographic information, such as age, sex, date and country of birth, as well as clinical information in the form of the International Classification of Disease

Australian modification (ICD-10-AM, 2nd edition) diagnosis codes and an external cause or mechanism of injury. Information on the specific mechanism/circumstances of injuries, including those related specifically to traffic crashes, is very limited. Appendix II provides a detailed list of the subset of variables included in the ISC provided by the NSW Health Department for the purpose of this data linkage project.

The 19,277 cases selected for the purpose of this study included all hospital separations between 1 July 2000 to 30 June 2001 with the first external cause of injury as a land transport accident (i.e an ICD-10-AM code between V01-V89). As only traffic crashes are included in TADS, the subset of land transport accidents coded as traffic crashes were identified from the ISC using the fourth character in the ICD-10-AM external cause code (see Table 1). A traffic accident is defined, according to ICD-10-AM, as any vehicle accident occurring on the public highway. A “public highway” or public road refers to the entire width between property lines (or other boundary lines) of land open to the public as a matter of right of custom for the purposes of moving persons or goods. A small number of cases, when the “traffic” nature of the crash could not be determined according to ICD-10-AM codes, were also included.

Table 1. ICD-10-AM codes identifying traffic crashes

ICD-10-AM code range	4th character
V01 – V06	1, 9
V09, V89	2, 3, 9
V10 – V18, V20 – V28	3, 4, 5, 9
V19,V29,V39, V49, V59, V69, V79	4, 5, 6, 8, 9
V30 – V38, V40 – V48, V50 – V58, V60 – V68, V70 – V78	4, 5, 6, 7, 9
V80, V87	All
V81 – V82	1
V83 – V86	0, 1, 2, 3

3. The data linkage process

Using probabilistic record linkage techniques, hospitalised land transport crash cases were linked to records of casualties resulting from traffic crashes reported in TADS between 1 July 2000 to 30 June 2001.The linkage was carried out using LinkageWiz record linkage software (LinkageWiz, 2002).

Pre-processing

It has been argued that the success of record linkage between two or more datasets is much more dependent on data quality than on the software and record linkage methodology used (Clark, 2004). During the pre-processing phase, both datasets were prepared and any missing data was identified. Standard formats were applied, particularly to variables common to both datasets that were used in the linking process. These variables, also referred to as the

matching variables, included surname, initials, phonetic coding of surname, date of birth, age, gender, postcode, date of the crash (in TADS) and date of hospital admission (in ISC).

Given names and middle names in the hospital data were recoded into initials and grouped in a single variable to conform to the TADS data format for the same variables. For both datasets, the date of birth format was changed to DD/MM/YY. Day, month and year of birth were also separated into different variables. A similar process was applied to the date of the crash and date of admission. Some of the pre-processing tasks were carried out automatically by the LinkageWiz software. Examples include standardising dates, removing hyphens from family names, and so on. Phonetic coding of family names was carried out for both datasets using the linkage software. Phonetic coding, used by LinkageWiz (also referred to as NYSIIS) is a sophisticated phonetic algorithm, developed by the New York State Identification and Intelligence System, that builds a phonetic code of up to 6 letters for each name (LinkageWiz, 2002). The main benefit of using phonetic coding is to take into account spelling errors in the data sources when linking records.

Record linkage between data sources, such as hospitalisation and police crash records for road casualties, where an individual may have more than one crash and be admitted to hospital on several occasions for some or all of those crashes, is termed a 'many-to-many' linkage. Internal linkage of the hospital dataset prior to attempting the between-sources linkage, allowed all records potentially belonging to the same patient to be identified. Sets of multiple records for hospitalised individuals could result from duplicate records being entered incorrectly or from valid records of transfers between hospitals or changes in patients' service type. Additionally, multiple records could result from hospitalisations for more than one traffic accident. In this study, the internal linkage of the NSW Inpatient Statistics Collection resulted

in identifying 8.8% of those hospitalisations coded as due to traffic crashes as being transfers or statistical discharges for the same injury (for which the date of admission was the same as the date of separation of the episode of care).

Assigning linkage weights

Probabilistic data linkage is a process that attempts to replicate the steps a person would follow to manually link records from two sources. It links records between two data sets through the calculation of a linkage likelihood or probability weights, adjusting for data entry errors (such as misspelling of names), as well as incomplete and missing data. In other words, probabilistic data linkage attempts to determine the likelihood that a potential link between records from two separate datasets is in fact a “true” match. In this case, this is the likelihood that the records from the TADS and ISC datasets refer to the same person. Variables used to link datasets were assigned a linkage weight according to their “reliability” and “discriminative power” (Rooss, Wajda & Nicol, 1986). For example, agreement on date of birth is more suggestive of a match than is agreement on sex. Accordingly, matches on date of birth will have a greater weight than those based on sex; this is consistent with the same process that would be used by an individual to manually link records from two data sets.

The likelihood or probability weights are estimated given all observed agreements and disagreements on all data variables used for linking records together. The total weight for a given comparison pair is simply the sum of the agreement/disagreement weights for each matching variable. The probabilistic linkage software, LinkageWiz, initially assigns agreement and disagreement weights for each variable based on their reliability and discriminative power, but also allows the operator to modify the weights in later stages of the linkage.

Weights were therefore modified depending of the outcomes of each pass to fine-tune the linkage process.

In addition to assigning general variable weights, LinkageWiz also assigns value specific weights for some matching variables. This applies particularly to the family name variable where agreements on rarer values (e.g. a surname such as McAlarey) are more suggestive of a match than agreements on more common values (e.g. Smith).

Stratification/ blocking

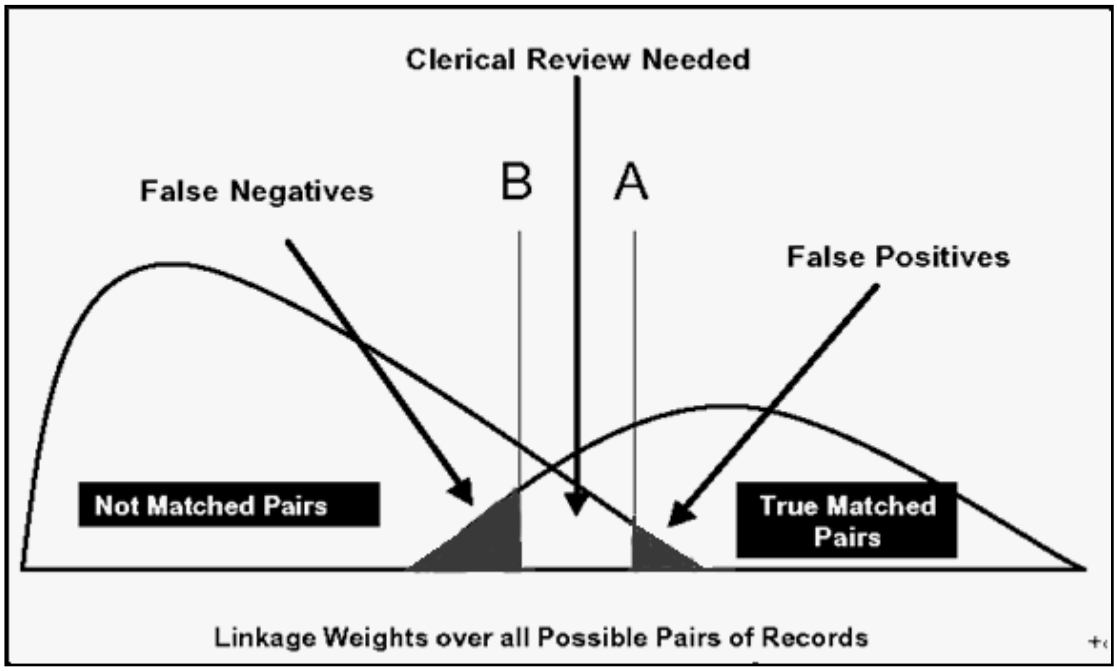
With the probabilistic linkage approach, the number of possible comparisons increases with the file size. This can make it unwieldy when the files are large, such as in this project. Comparisons were therefore restricted to comparisons of "blocks" or "pockets" of cases where one or more variables matched exactly. This process is referred to as “blocking” and simply stratifies the linkage process to minimise the number of comparisons that must be undertaken at a given time. Multiple passes through the data were used for each separate blocking variable. We ran three passes using phonetic name, date of birth and then date of admission/accident. Up to 24 hours between the accident date and that of the hospital admission date was considered a match to allow for any possible lag. These are the variables considered to have the greatest discriminative power and which have been used previously in this type of record linkage. The first pass through the data only compared records where the phonetic representation of the surname field agreed. The second pass only compared records where the date of birth agreed. Finally, the third pass only compared the remaining records which did not match during either the previous passes and where the date of accident and admission agreed. Using multiple passes ensures that any linkages missed by one pass should be picked up by another. For example, a woman who has changed her surname by

marriage would not be picked up on the first pass through the data, but should be picked up on the second pass (assuming that the date of birth has been entered accurately).

Selecting matched records

The ultimate objective of record linkage is to identify matches and combine records for such matches. The task during this stage was to set the 'cut-off' and the “threshold’ weight values in order to identify true links and reject non-links for all comparisons. Figure 2 schematically shows the bimodal distribution of total weight scores for matches and non-matches in record linkage (Blakely & Salmond, 2002).

Figure 2. Distribution of total weight scores for matches and non-matches



Reproduced from Blakely & Salmond, 2002

If the sum of comparison weights for each record pair is below the 'cut-off' value (B), the record pair is rejected as a “non-link”. If the total weight is above a much higher 'threshold' (A), the record pair is defined as a 'definite' or ‘true’ link. Records with values between the 'cut-off' and the 'threshold' are said to be 'possible' or ‘grey area’ links. Records in the grey-area were manually reviewed by two researchers and a decision on linkage status was made based on all the variables included in the file.

One of the major limitations of probabilistic record linkage is its potential to allow multiple “trivial items” of agreement to combine to create an agreement weight which is sufficient to over-ride important indicators of mismatch. For example, agreement on postcode, age, sex and year of the accident may contribute sufficient positive weights to indicate a match, when in fact these are truly random agreements. These links are referred to as false positives and are pairs of records that have been inadvertently identified as “definite links” when they really belong to different individuals. False negatives, on the other hand, are pairs of records that have been rejected as a non-link when they really belong to the same individual (Newcombe, 1998). The extent of false positive and false negatives will depend on the cut-off weight: lowering it (i.e. moving it to the left in Figure 2) will increase the sensitivity, but also increase the number of false positives; increasing it (i.e. moving it to the right in Figure 2) will decrease the sensitivity, but also decrease the number of false positives (Blakely & Salmond, 2002). Trade-offs are always required between the number of false positives and false negatives in a record linkage project. Our strategy was to sacrifice the sensitivity (and incur more false negatives or missed matches) but maintain a high specificity (and incur fewer false positives or incorrect links).

4. Linkage outcomes

At the end of the data linkage process, a total number of 7,891 hospital cases were categorised as definitely linked to TADS records. These pairs largely agreed on all the matching variables. Another 3,215 records were classified as possible links and were checked manually. As a result of the manual checking, a further 764 records were added to the pool of definite links, raising the final number of matches to 8,655 records. After the completion of the matching process, identifiers and potentially identifying items were deleted from the combined dataset. These included variables relating to the person's name, address and date of birth. The date of birth was recoded in the combined dataset to age in years. The date of crash variable in the TADS dataset was retained, as it is important for further analysis of crash risk.

Because it is mandatory to report a traffic crash to the police when a person has been injured in a crash, all cases admitted to hospital should be reported to the police and be classified within police road traffic crash reports as a casualty and should therefore link to a TADS record. However, previous data linkage studies have indicated that this is not necessarily the case with only between 20% and 80% of hospitalised cases for traffic accidents are matched to road crash data (Maas & Harris, 1984; Schelp & Ekman, 1990; Rosman, 2001, Cryer et al, 2001). Table 2 shows the linkage rates achieved in this project compared to those of the only other major Australian hospitalisation and road crash linkage project (Rosman, 2001).

Table 2. Comparison of linkage rates of hospital and police data in NSW and Western Australia

Hospitalisations for land transport accidents	New South Wales, 2000-2001				Western Australia, 1987-1996*			
	Not linked	Linked	Total	Linkage rate	Not linked	Linked	Total	Linkage rate
Non traffic	4,548	880	5,428	16.2%	9432	1,496	10,928	13.7%
Traffic	6,073	7,776	13,849	56.2%	16,838	19,991	36,829	54.3%
Total	10,622	8,655	19,277	44.9%	26,270	21,487	47,757	45.0%

**Rosman, 2001*

While hospitalisation records for traffic accidents are more likely to link to TADS records, as Table 2 shows, 880 cases (16%) of hospitalisations for non-traffic crashes (as classified by the ISC) also linked to TADS. A more detailed analysis is presented in the next section of this report.

Similarly, a comparison of the coding of the type of vehicle-occupant between TADS and ISC indicated various level of discordance betweens the two datasets. For instance, only 62% of cases coded as pedal cyclists in TADS were coded as such in ISC. Similar findings were found in the WA linkage study (Rosman & Knuiman, 1994). Table 3 provides the level of concordance for various vehicle occupant type for both studies.

Table 3. Comparison of vehicle occupant type concordance[+], in NSW and Western Australia data linkage projects

	NSW	WA*
Motor vehicle driver	90%	62%
Motorcycle rider	80%	87%
Pedestrian	97%	94%
Motor vehicle passenger	87%	66%
Motorcycle passenger	80%	69%
Pedal cycle (rider or passenger)	62%	82%

**From Rosman & Knuima ,1994.*
[+]Defined as the level of agreement of vehicle occupant type in linked hospital and police records. The denominator is the total number of hospital cases coded according to ICD as a given road user and the numerator is the number of those which agree with the coding or road user type in TADS.

Record linkage rates varied according to age with the lowest rates observed in younger age groups, particularly those aged 10-14 years. It also varied according to occupant type with vehicle controllers (usually drivers) having the highest rate and the lowest rate being in pedal cycle riders. Linkage rates of traffic crashes appeared to be lower for those residents in areas (postcodes of residence) close to NSW borders, which may be because cases in these areas are more likely to be admitted to hospitals in other states. A detailed analysis of the impact of various factors on linkage rates is described in the next section.

Linkage rates

To examine factors that influenced the probability of linkage, a subset of linked data containing only road traffic crashes for New South Wales residents was used. Because the ISC is based on hospital separations, and not patients, we identified multiple episodes of care for the same injury for the subset used for data linkage. For each patient, we eliminated those episodes of care for which the admission date was the same as the separation date for another episode of care. The remaining 'index' episodes of care are the first episode of care for a patient within a continuous period of hospitalisation.

We then selected ISC records that should, in theory, have a matching TADS record by identifying records for which the place of occurrence was specified as street and highway (ICD-10-AM (second edition) place of occurrence code 'Y92.4') and the incident was classified as a traffic crash. This has reduced the original subset to 17,552 records. Of the ISC records, 22% had either an unspecified or missing place of occurrence code and were therefore not able to be included in the analyses. A number of variables potentially associated with high linkage rates between hospital admissions and police records were identified mainly from previous literature.

Overall, 45.1% (n=7,917) of the ISC subset of 17,552 index traffic transport episodes of care had a matching police record. Of these 17,552 episodes of care, 16,624 were for NSW residents, of which 9,178 were identified as road traffic cases. When restricted to the 9,178 road traffic crashes only, the linkage rate increased to 69.2 % (n = 6351). Our further analysis in this section is restricted to these 9,178 cases of road traffic crashes.

The linkage rates according to specific variables of interest are presented in Table 5 and show varying rates of linkage across categories of each considered variable. There were significant associations between the probability of being linked to TADS and health area of residence, separation mode, payment status category, principal diagnosis of injury, activity when injured, road user type, length of stay and age.

While males had slightly higher linkage rates than females, the difference was not statistically significant. People who live in health service areas bordering other states had a smaller chance of linkage compared to people who did not. Persons with a principal diagnosis of injury had a much higher chance of being linked than those whose principal diagnosis was not injury.

In terms of road user types, the highest linkage rates were observed for motor vehicle controllers (e.g. drivers) and this rate was almost twice as high as that for pedal cyclists. In relation to age, the highest linkage rate was amongst 65-69 year olds, and the lowest was for 10-14 year olds. Patients with hospital payment status indicating insurance compensation had higher linkage rates than non-compensable patients. There was a trend towards higher linkage rates with increasing severity, as measured by length of hospital stay, with the trend stabilising after five days. The lowest linkage rates were for hospitalisations of $\leq$1 day. A complete analysis of matching rates is given in Lujic et al (2008).

Table 4. Linkage rates of traffic crashes related hospitalisations (ISC) to road crash casualty records (TADS), NSW, 1 July 2000- 30 June 2001

	Number of hospital records (ISC)	% Linked to TADS
Age*		
0-4 years	149	69.80%
5-9 years	262	69.08%
10-14 years	361	62.33%
15-19 years	1,123	71.68%
20-24 years	1,319	72.18%
25-29 years	1,005	65.27%
30-34 years	820	69.02%
35-39 years	721	70.87%
40-44 years	673	71.03%
45-49 years	551	70.24%
50-54 years	509	65.23%
55-59 years	328	68.60%
60-64 years	290	72.07%
65-69 years	266	73.31%
70-74 years	260	64.62%
75-79 years	259	62.93%
80-84 years	184	68.48%
85 + years	98	69.39%
Activity when injured		
Sports + Leisure	299	48.49%
Work	637	61.38%
Other/unspecified	8,194	70.60%
Health area of residence		
On the border of another jurisdiction	1,238	61.71%
Not on the border	7,940	70.37%
Length of stay		
≤ 1 day	4,864	65.40%
2 days	882	69.73%
3 days	565	72.74%
4 days	429	74.13%
5 days	331	77.04%
6 days	253	75.89%
7 days	218	74.77%
> 7 days	1,636	74.33%

Occupant type		
Pedestrian	1,376	75.00%
Pedal cycle occupant	545	45.87%
Motor cycle occupant	1,128	62.59%
Motor vehicle controller	3,459	82.51%
Motor vehicle passenger/unknown	2,611	57.22%
Other/unspecified	59	25.42%
Payment status		
Compensable (MVA)	3,394	81.00%
Compensable (other)	713	76.30%
Non compensable	4,732	63.23%
Unknown	339	19.47%
Principal diagnosis		
Non injury	1,380	45.80%
Injury	7,798	73.34%
Separation mode		
Died	126	84.92%
Discharged	8,380	69.01%
Transferred	646	70.90%
Sex		
Male	5,534	69.77%
Female	3,644	68.33%

5. Summary

This probabilistic data linkage of hospital records (ISC) and police crash casualty records (TADS) in NSW has produced comparable matching rates to those found around the world. Record linkage rates varied according to factors, particularly the road user class. Pedal cyclists, in particular, had lower rates of linkage than other types of road users.

The outcomes of the record linkage of hospital records and police casualty crash records suggest that researchers and policymakers should be cautious when examining traffic crashes based on a separate analysis of police crash records and hospital separations as they are individually limited in terms of the scope and the quality of information they contain. For example, using police crash records alone to examine crashes involving cyclists would miss the majority of them and any investigation of this issue would need to be complemented by examining hospitalisation data. On the other hand, hospitalisation data are limited in terms of the availability of information related to the circumstances and the characteristics of traffic crashes and need to be used in combination with the police crash data to examine the relationship between the risk factors and the outcomes of traffic crashes.

The record linkage also indicated a level of disagreement between the TADS and ISC in relation to the coding of road user type (driver, pedestrian, occupant, etc). It is reasonable to assume that, with the exception of cyclists, the coding of the road user type is more likely to be more reliable in TADS as data is collected by police officers on the scene as opposed to health professionals/ clinical coders who might have limited information on the status of the patient in this regard. Unfortunately, there are other possible explanations for this discrepancy, including inaccurate coding of medical records; inaccurate recording by the

police of information by the crash; and inaccurate coding of the information in the police records. It is also possible that some of the inaccuracy could be explained by inaccurate links between the two datasets. However, we were unable to quantify the contributions of each of these factors to the observed discrepancies between the datasets.

Notwithstanding these limitations, the linked dataset has the potential to contribute to the understanding of the causes and consequences of road traffic accidents. It is also able to highlight the some of the strengths and weaknesses of the individual datasets.

References

Alsop, J.C., Langley, J.D. (2001). Under-reporting of motor vehicle traffic crash victims in New Zealand. *Accident Analysis and Prevention*, 33 (3), 353-359.

Blakely T., & Salmond, C. (2002). Probabilistic record linkage and a method to calculate the positive predictive value. *International Journal of Epidemiology*, 31(6), 1246-1252.

Bureau of Transport Economics. (2000). Road Crash Costs in Australia. Report 102. Bureau of Transport Economics: Canberra

Boufous S, Williamson A. (2006). Work-related traffic crashes, a record linkage study. *Accident Analysis and Prevention*, 38(1), 14-21.

Bull, J.P., Roberts, B.J. (1973). Road accident statistics – a comparison of police and hospital information. *Accident Analysis and Prevention*, 5 (1), 45-53.

Centre for Epidemiology and Research, NSW Health Department. (2003). *Health Outcomes and Information Statistical Toolkit (HOIST). Inpatient Statistics Collection.* Sydney: NSW Department of Health.

Clark, D.E. (2004). Practical introduction to record linkage for injury research. *Injury Prevention*, 10(3), 186-191.

Cryer, P.C., Westrup, S., Cook, A.C., Ashwell, V., Bridger, P., Clarke, C. (2001). Investigation of bias after data linkage of hospital admissions data to police road traffic crash reports. *Injury Prevention*, 7(3), 234-241.

Ferrante, A.M., Rosman, D.L., Knuiman, M.W. (1993). The construction of a road injury database. *Accident Analysis and Prevention*, 25 (6), 659-665.

Johnson, S.W., Walker, J., 1996. The Crash Outcome Data Evaluation System (CODES). U.S. Department of Transportation, National Highway Traffic Safety Administration, Washington, DC. Technical report DOT HS 808 338.

Langley, J.D., Dow, N., Stephenson, S., Kypri, K. (2003). Missing cyclists. *Injury Prevention*, 9 (4), 376-379.

LinkageWiz. (2002). Record linkage software, Version 3.04. LinkageWiz Inc, Adelaide.

Lujic, S. Finch C, Boufous S, Hayen A, Dunsmuir W. How comparable are road traffic crash cases in hospital admissions data and police records? An examination of data linkage rates. *Australian and New Zealand Journal of Public Health*, 32 (1), 28-33.

Maas, M.W., & Harris, S. (1984). Police recording of road accident in-patients. Investigation into the completeness, representativity and reliability of police records of hospitalised traffic victims. *Accident Analysis & Prevention,* 16(3), 167–84.

National Centre for Classification in Health. (2000). The International Statistical Classification of Diseases and Related Health Problems, 10th Revision, Australian Modification (ICD-10-AM). Second edition. National Centre for Classification in Health, Sydney.

Newcombe, H.B. (1998). *Handbook of record linkage: methods for health and statistical studies, administration, and business*. London: Oxford University Press.

Road and Traffic Authority. (2003). *Traffic Accident Database System Data Manual.* Sydney: RTA.

Roos, L.J., Wajda, A., & Nicol, J. (1986). The art and science of record linkage: methods that work with few identifiers. *Computers in Biology and Medicine*, 16(1), 45–57.

Rosman, D.L. (2001). The Western Australian Road Injury Database (1987–1996): ten years of linked police, hospital and death records of road crashes and injuries. *Accident Analysis and Prevention,* 33(1), 81–88.

Rosman, D. L., Knuiman, M. W. (1994). A comparison of hospital and police road injury data. *Accident Analysis and Prevention*, 26(2), 215-22.

Schelp, L., & Ekman, R. (1990). Road traffic accidents in a Swedish municipality. *Public Health*, 104 (1), 55-64.

Singleton, M., Qin, H. (2004). Factors associated with higher levels of injury severity in occupants of motor vehicles that were severely damaged in traffic crashes in Kentucky, 2000-2001. *Traffic Injury Prevention*, 5 (2), 144-150.

World Health Organization. (2004). World report on road traffic injury prevention. World Health Organization: Geneva.

Appendix I Data items in the Traffic Accident Data System (TADS)

#	Variable	Name	Type	Description	Values
1	Accident number	ACCNO	Char	Unique 9-digit number assigned to accident.	
2	Degree of accident	ACCDEG3	Num	Severity classification of accident	1 Fatal 2 Injury 3 Non-casualty (towaway)
3	Day	ACCDAY	Num	Day of the week on which accident occurred	1 Sunday 2 Monday 3 Tuesday 4 Wednesday 5 Thursday 6 Friday 7 Saturday
4	Date	DATE	Num	Date of accident in DDMMYY format	dd-mm-yy
5	Time	ACCHOUR1	Num	The 1-hourly time interval in which accident occurred	Every hour: 00 00:01 – 00:59 to 23 23:00 – Midnight 99 Unknown
6	Type of day	ACCDATEG	Num	Type of day on which accident occurred	1 Christmas 2 Easter 3 Other public holidays 4 Other school holidays 5 Other weekends 6 Other weekdays
7	Street name	ACCST	Char	First 14	Street name

#	Variable	Name	Type	Description	Values
				characters of name of street in which accident occurred	Or Unknown/not stated
8	Street type	STTYP	Char	Type of street on which accident occurred	AV Avenue BV Boulevarde BP Bypass CH Chase CI Circuit / Circle / Circlet CL Close CT Court CR Crescent EX Distributor DR Drive ES Esplanade EX Expressway / Freeway GA Gardens GV Grove HY Highway LA Lane MR Main Road MS Miscellaneous PD Parade PW Parkway PL Place / Plaza PR Promenade QY Quay RD Road / Roadway SQ Square ST Street TC Terrace EX Tollway TR Trunk Road WA Way

#	Variable	Name	Type	Description	Values
					UK Unknown
9	Distance	ACCDIST	Num	Distance in metres from identifying feature used to locate accident	00000 On the spot 99000 99 km and over 99999 Unknown / not stated
10	Direction	ACCDIRN	Num	Direction from identifying feature to location of accident	0 On the spot 1 North 2 South 3 East 4 West 9 Unknown / not stated
11	Identifying object	ACCIDOB	Char	First 14 characters of name of identifying feature used to locate accident	Coded as for street name.
12	Identifying object type	IDTYP	Char	Type of identifying feature used to locate accident	If a road, coded as for street type. If not a road, coded as BR Bridge / Causeway CP Caravan Park CB Club CN Corner CK Creek EN Entrance TO Exit FE Ferry BR Floodway OP Flyover BR Ford GT Gate HO Hospital

#	Variable	Name	Type	Description	Values
					PU Hotel HN House number JN Junction LX Level Crossing MO Motel OP Overpass / Overbridge OT Other PK Park / Reserve / Oval PO Post Office GT Property gate TO Ramp PK Reserve RI River SN Station TU Subway / Tunnel TN Town name TO Turnoff / Exit / Ramp OT Other (not listed above) UK Unknown
13	Town	ACCTOWN	Char	First 14 characters of town or place in which accident occurred	
14	Local Government Area	ACCLGAL	Num	LGA including amalgamations	001 City of Sydney 003 Ashfield 004 Auburn to 610 Newcastle City
15	Region	ACCREGN	Num	State Region in which accident occurred	01 Sydney 02 Hunter 03 Illawarra 04 North Coast 05 New England

#	Variable	Name	Type	Description	Values
					06 Orana 07 Central Western 08 South Eastern 09 Riverina 10 Murray 11 Far Western
16	Road Classification	ACCROAD	Num	Regrouped classification of type of road on which accident occurred	1 Freeway/Motorway 2 State Highway 3 Other classified road 4 Unclassified road
17	Urbanisation	ACCURBAN	Num	Area in which accident occurred	1 Sydney metro area 2 Newcastle metro area 3 Wollongong metro area 4 Country urban areas 5 Country non-urban areas 6 Country unknown
18	Location type	ACCLOC	Num	Type of location at which accident occurred	**Intersection Locations** 01 Cross intersection 02 Y-junction 03 T-junction 04 Multiple intersection 05 Roundabout **Non-Intersection Locations** 10 L-junction 11 One-way street 12 Two-way undivided street 13 Divided road (dual carriageway, but not limited access or freeway) 14 Single carriageway limited access road / freeway 15 Dual carriageway limited access road / freeway

#	Variable	Name	Type	Description	Values
					16 Other 99 Unknown / not stated
19	Alignment	ACCALIGN	Num	Alignment of road at location of accident	1 Straight 2 Curved 9 Unknown / not stated
20	Permanent feature	ACCPERM	Num	Permanent feature of location that was a factor in accident	**Construction Features** 01 Narrow roadway 02 Narrow or one-lane bridge 03 Low clearance overhead bridge 04 Other bridge 05 Low clearance structure other than bridge 06 Causeway 07 Floodway or dip 08 Ditch, drain or culvert 09 Embankment or cutting 10 Underpass or tunnel 11 Railway level crossing 12 Steep grade 13 Crest 14 Speed hump, slow point or chicane 15 Footpath, cycle path or nature strip 16 Driveway or entrance 17 Loading Bay 18 Cattle grid, gate or stock crossing **Lane features and road controls** 20 Breakdown lane or road shoulder 21 Climbing or merging lanes 22 Bus or transit lane

#	Variable	Name	Type	Description	Values
					23 Clearway 24 S-lanes or turning bay 25 Bus stop 26 Reserved bus roadway 27 Painted double centre lines 28 Mid-block median opening 29 Channelised intersection with traffic islands 30 Freeway ramp or access road 31 Safety ramp 32 Designated light traffic route 97 Other permanent feature 98 No identifiable permanent feature
21	Hazardous feature	ACCHAZ	Num	Hazardous road surface that was a factor in accident	01 Loose gravel on sealed surface 02 Loose gravel on shoulder 03 Pot holes, corrugations or other rough surface 04 Slippery surface (oily or greasy) 05 Flooded or submerged / water lying on road 97 Other hazardous feature 98 No identifiable hazardous feature
22	Temporary feature	ACCTEMP	Num	Temporary feature at location that was a factor in accident	01 Roadworks / detour / diversion 02 Previous accident 03 Roadblock / Random Breath Testing (RBT) 04 Thick raised dust 97 Other temporary feature 98 No identifiable temporary feature

#	Variable	Name	Type	Description	Values
23	Street lighting	ACCSTRLT	Num	STREET LIGHTING	1 On 2 Off (lights present but off) 3 Nil 9 Unknown / not stated
24	Surface type	ACCSURF	Num	Type of road surface at accident location	1 Sealed 2 Unsealed 9 Unknown / not stated
25	Surface condition	ACCSFCND	Num	Condition of road surface at accident location	3 Wet 4 Dry 5 Snow or ice 9 Unknown / not stated
26	Weather	ACCWTHR	Num	Weather conditions at time of accident	1 Fine 2 Raining 3 Overcast 4 Fog or mist 5 Snowing or sleeting 6 Other (e.g. hail) 9 Unknown / not stated
27	Natural lighting	ACCNATLT	Num	Natural lighting at time of accident	4 Dawn 5 Daylight 6 Dusk 7 Darkness 9 Unknown / not stated
28	Traffic signal operation	ACCSIGOP	Num	Operating status of traffic control signals at accident location	1 On (installed and operating) 2 Off (installed but not operating) 3 Nil (no signals installed) 9 Unknown / not stated
29	Traffic signal number	ACCSIGNO	Char	Traffic control signal identification number at accident location	Traffic signal id number or 9999 Unknown / not stated Null Signal operation is nil

#	Variable	Name	Type	Description	Values
30	Other traffic controls	ACCOTC	Num	Traffic control other than signals that is controlling location of accident	01 Pedestrian crossing 02 Stop sign 03 Give way sign 04 Police 05 No right turn 06 No left turn 07 No U turn 08 No Entry / Wrong Way 09 Rail crossing with flashing signals 10 Rail crossing with stop sign 11 Rail crossing with no signals or stop sign 12 Road / railway worker 97 Other traffic control 98 No traffic controls 99 Unknown / not stated
31	Speed limit	ACCSPEED	Num	Maximum speed limit applicable at accident location	Speed limit or 999 Unknown / not stated
32	Road user movement	ACCRUM	Num	RUM code describing first impact	RUM code or 99 Unknown / not stated
33	First impact (1)	ACCITEM1	Num	Traffic unit (TU) type corresponding to key vehicle in first impact	
34	First impact (2)	ACCITEM2	Num	TU type corresponding to other vehicle or TU involved in first impact	

#	Variable	Name	Type	Description	Values
35	First impact type	ACCIMP1	Num	Type of first impact	01 Vehicle – Vehicle (Head-on) 02 Vehicle – Vehicle (Right angle) 03 Vehicle – Vehicle (Nose-tail) 04 Vehicle – Vehicle (Other angle) 05 Vehicle – Object 06 Vehicle – Pedestrian 07 Vehicle – Animal 08 Vehicle – Train / Aeroplane (trams not included) 09 Vehicle – Rollover 10 Person – Object 99 Other / unknown Vehicle – vehicle (nose-tail)
36	Car accident	ACCCAR	Num	Whether or not the accident involved a car	1 Yes Null No
37	Light truck accident	ACCLGTTK	Num	Whether or not the accident involved a light truck	1 Yes Null No
38	Rigid truck accident	ACCRIGTK	Num	Whether or not the accident involved a heavy rigid truck	1 Yes Null No
39	Articulated truck accident	ACCARTTK	Num	Whether or not the accident involved an articulated truck	1 Yes Null No
40	Bus accident	ACCBUS	Num	Whether or not the accident involved a bus	1 Yes Null No

#	Variable	Name	Type	Description	Values
41	Emergency vehicle accident	ACCEMERG	Num	Whether or not the accident involved an emergency vehicle	1 Yes Null No
42	Motorcycle accident	ACCMC	Num	Whether or not the accident involved a motorcycle	1 Yes Null No
43	Pedal cycle accident	ACCPC	Num	Whether or not the accident involved a pedal cycle	1 Yes Null No
44	Pedestrian accident	ACCPED	Num	Whether or not the accident involved a pedestrian	1 Yes Null No
45	Number of traffic units	ACCTUS	Num	Actual number of TU's involved	
46	Number killed	ACCKILL	Num	Actual number of people killed	
47	Number injured	ACCINJ	Num	Actual number of people injured	
48		XCOORD	Char		
49		YCOORD	Char		
50	Geocoding precision	GEOSTAT	Num	Derived variable on geocoding precision	1 Not estimated 2 Estimated 3 Allocated to LGA
51	Traffic unit number	TUNO	Num	Number assigned to traffic unit	
52	Type of traffic unit	TUTYPE	Num	Classification for type of traffic unit	**Motor vehicles** 01 Car (sedan) or hatchback / liftback 02 Station wagon

#	Variable	Name	Type	Description	Values
					03 Utility based on car design 04 Panel van based on car design 05 Taxi-cab 06 Forward control passenger van 07 4WD vehicle not based on car design 08 Road Train (1998 onwards) 09 B-double (1998 onwards) 10 Light truck / panel van / utility not based on car design 11 Mobile vending vehicle (light truck) 12 Large rigid lorry 13 Rigid tanker 14 Articulated tanker 15 Semi-trailer / low loader 08 Road Train (1998 onwards) 09 B-double (1998 onwards) 16 (pre 1998) Road train or B-double 17 State Transit Authority bus 18 Long distance / tourist coach 19 Other bus 20 Self-propelled plant 21 Ambulance 22 Fire brigade / bushfire brigade vehicle 23 Police patrol car or van 24 Tow truck 25 Other emergency vehicle 26 Motorized wheelchair 27 Tractor 29 Other or unspecified motor vehicle

#	Variable	Name	Type	Description	Values
					Motorcycles 30 Motorcycle (includes motorcycle ambulance) 31 Motorcycle with sidecar 32 Police Motorcycle 33 Motor scooter 34 Mini-bike 35 Moped / Motorized ‘pedal’ cycle 36 Special Mobility Vehicle **Pedal cycles** 40 Pedal cycle (not motor assisted) **Trailers** 50 Small box trailer 51 Small boat trailer 52 Horse float 53 Other small trailer 54 Large trailer 55 Caravan 56 Detached trailer section of semi-trailer 57 Agricultural implement **Other traffic units** 60 Ridden Animal 61 Animal drawn vehicle 62 Train 63 Aeroplane 64 Tram **Pedestrians** 70 Pedestrian

#	Variable	Name	Type	Description	Values
					71 Pedestrian in / operating toy vehicle, pedal car, pram, barrow, billycart or non-motorized wheelchair 99 Other or unknown traffic unit
53	Traffic unit group	TUTYPEG	Num	Group of traffic unit	01 Car / Car derivatives 02 Light trucks 03 Heavy rigid trucks 04 Articulated trucks 05 Buses 17 - 06 Emergency Vehicles 07 Other motor vehicles 08 motorcycles 09 Pedal Cycles 10 Non-motorised vehicles 11 Pedestrians 12 Other traffic units / unknown
54	Street of travel	TUST	Num	Street on which traffic unit was travelling with respect to street of accident	1 Street in which accident occurred 2 Street described as identifying object 9 Unknown / not stated
55	Direction of travel	TUDIRN	Num	Direction of travel of traffic unit	1 North 2 South 3 East 4 West 5 Unknown / not stated
56	Stated speed of vehicle	TUSPEED	Num	Actual speed (km/h) as recorded by Police	Speed in Km/h or 900 Speed not stated but described by police as ‘excessive’ 999 Unknown / not stated

#	Variable	Name	Type	Description	Values
					Null Non-motorised TU
57	Speeding controller	TUEXSPD	Num	Speeding involvement of motor vehicle controller	1 Yes 2 No or unknown Null No controller or non-motorised TU
58	Fatigued controller	TUFATIG	Num	Fatigue involvement of controller	1 Yes (mentioned) 2 No (not mentioned) Null No controller or TU group 10
59	Manoeuvres of traffic unit	TUMAN	Num	Manoeuvre immediately prior to involvement in accident	**Stationary** 01 Stationary in traffic 02 Parked at kerbside / roadside 03 Parked at kerbside / roadside loading or depositing goods / passengers 04 Double parked 05 Broken down in traffic / previous accident 06 Parked or stationary on footpath 07 Parked elsewhere (off road) **Moving along carriageway** 10 Proceeding along lane (on either straight or curved carriageway) 11 Parking (forward) or pulling out from kerb 12 Veering to right to change to a lane moving in the same direction 13 Veering to left to change to a lane moving in the same direction 14 Merging with traffic in same direction 15 Pulling out into opposite stream of traffic 16 Travelling on incorrect side of

#	Variable	Name	Type	Description	Values
					carriageway (including wrong way on one-way street) 17 Cutting back after overtaking **Turning or reversing** 20 Turning right out of own lane 21 Turning left out of own lane 22 Waiting to turn right 23 Waiting to turn left 24 Performing U-turn 25 Entering carriageway from driveway (forward or unspecified) 26 Entering carriageway from driveway (reversing) 27 Moving along footpath 29 Performing other / unspecified forward manoeuvre 30 Reversing in lane (other than parking) 31 Parking (reversing) 39 Performing other / unspecified reversing manoeuvre **Pedestrians** 40 Pedestrian walking across carriageway 41 Pedestrian running across carriageway (see also 54) 42 Pedestrian standing still on carriageway 43 Pedestrian lying / sitting on carriageway 44 Pedestrian working on carriageway

#	Variable	Name	Type	Description	Values
					45 Pedestrian working on vehicle on carriageway 46 Pedestrian playing on carriageway 47 Pedestrian in / on toy vehicle on carriageway 48 Pedestrian moving along edge of c'way with traffic (see also 55) 49 Pedestrian moving along edge of c'way against traffic (see also 56) 50 Pedestrian stepping off / onto kerb 51 Pedestrian stepping off / onto traffic island or median strip 52 Pedestrian on footpath or elsewhere completely off carriageway 53 Pedestrian on skateboard / roller skates or blades 54 Pedestrian jogging (see also 41) 55 Pedestrian moving with traffic but not along edge of carriageway (see also 48) 56 Pedestrian moving against traffic but not along edge of c'way (see also 49) 59 Pedestrian performing other / unspecified manoeuvre 60 Train or aeroplane manoeuvre (tram excluded)
60	Unusual vehicle factor	TUUNUS	Num	Any unusual vehicle factor in	01 This vehicle jack-knifing 02 This vehicle skidding, sliding or aquaplaning

#	Variable	Name	Type	Description	Values
				accident	03 This vehicle swaying 04 Parked / Stationary vehicle slipping 05 Parked / Stationary vehicle door opening 06 The vehicle dangerously parked 97 Other unusual vehicle factor 98 No relevant unusual vehicle factor Null Traffic Unit Group 11
61	Equipment a factor	TUEQUIP	Num	Equipment failure that can be considered factor in accident	20 Brake failure or fault 21 Steering failure or fault 22 Tyre failure or fault (blow out or thrown tread) 23 Tyre tread smooth 24 Wheel, axle or suspension failure or fault 25 Towing or coupling fault or separation 26 Headlamp failure or fault 27 Rear lamp or clearance lamp failure or fault 28 Vehicle with insecure or projecting load 29 Vehicle overloaded 97 Other vehicle equipment failure or fault 98 No relevant equipment factor Null Traffic Unit Group 11
62	Type of first object impacted	TUOBJ1	Num	First object impact	**Fixed objects** 40 Body of water (river etc) 50 Bridge railing or superstructure

#	Variable	Name	Type	Description	Values
					51 Underpass or tunnel (wall or pier) 52 Guide post 53 Guardrail or fence 54 Utility pole 55 Traffic signal pole 56 Signpost or parking meter 57 Traffic island / roundabout / dome median strip / Jersey median / LATM 58 Telephone box / post box / traffic signal box / bus shelter 59 Roadwork materials / temporary signs or barriers 60 Level crossing gates 61 Drain or culvert 62 Embankments / cuttings / rocky outcrops / boulders etc. 63 Trees or bushes 64 Building 65 Vehicle interior 66 Vehicle exterior 69 Any other fixed objects **Falling objects** 70 Object falling from moving vehicle **Other non-fixed objects** 75 Other non-fixed objects **Animals** 80 Straying stock

#	Variable	Name	Type	Description	Values
					81 Stock driven or led 82 Riderless horse 83 Kangaroo or wallaby 84 Emu 85 Other large animals 86 Cat 87 Dog 88 Rabbit 89 Wombat 90 Other small animals 98 No object hit 99 Unknown / not stated
63	Type of second object impacted	TUOBJ2	Num	Second significant object hit by this TU during accident	Coded as for 'type of first object impacted' field.
64	Other traffic unit a factor?	TUOTHTU	Num	Type of TU that was a factor in accident relevant to this TU but not involved in impact of accident	Coded as for 'traffic unit type' field or 98 No such TU involved
65	Vehicle towed away	TUTOWED	Num	Was this vehicle towed away?	1 Yes 2 No 9 Unknown / not stated Null TUTYPEG 9, 10, 11, 12
66	Vehicle catch fire	TUFIRE	Num	Was it mentioned that this TU caught fire	1 Yes 2 No 9 Unknown / not stated Null TUTYPEG 9, 10, 11, 12

#	Variable	Name	Type	Description	Values
67	Stolen mentioned	TUSTOLN	Num	Was it mentioned that this TU was stolen?	1 Yes 2 No 9 Unknown / not stated Null TUTYPEG 9, 10, 11, 12
68	Type of traffic unit towed	TUTOWING	Num	Type of traffic unit towed	Coded as for 'traffic unit type' field
69	Registration number	TUREGST	Num	Registration number of traffic unit	Registration number or 99999998 Not registered 99999999 Unknown / not stated Null TUTYPEG 9, 11, 12
70	Make of vehicle	TUMAKE	Num	The make of this traffic unit	List of makes Or 97 Other make (not listed) 99 Unknown / not stated Null TUTYPEG 9, 10, 11, 12
71	Year of manufacture	TUYEAR	Num	Year of manufacture of this traffic unit	Year of manufacture or 9999 Unknown / not stated Null TUTYPEG 9, 10, 11, 12
72	Vehicle weight (tare)	TUWEIGHT	Num	Weight group to which traffic unit belongs	1 Under 4.5 tonne tare weight 2 Over 4.5 tonne tare weight 9 Unknown / not stated Null TUTYPEG 9, 10, 11, 12
73	Load	TULOAD	Num	Type of load vehicle was carrying	01 NIL (unladen) 02 PET (petrol / oils) 03 LPG (liquid petroleum gas)

#	Variable	Name	Type	Description	Values
					04 FLAM (other flammable loads) 05 COR (corrosive loads) 06 REF (refrigerated loads) 07 TIM (timber) 08 COAL (coal) 09 LIV (livestock) 10 Other hazardous but non-flammable loads 11 Grains or other agricultural produce 12 Gravel, sand or soil 13 Garbage or other refuse / effluent etc. 14 Building materials 15 Glass bottles etc. 16 Canned foods etc. 17 Beer kegs 18 Newsprint / paper rolls etc. 19 Newspapers 20 Furniture 97 OTH (Other) 99 Unknown / not stated Null Traffic Unit Groups 8, 9, 11, 12
74	Vehicle headlights	TULIGHTS	Num	Vehicle headlights	1 On 2 Off 9 Unknown / not stated Null TUTYPEG 9, 10, 11, 12
75	Number of occupants	TUOCCUPS	Num	Actual number of occupants in traffic unit including driver/rider	Number of occupants or 00 Vehicle wa s unoccupied 90 90 or more 99 Unknown / not stated Null TUTYPEG 10, 11

#	Variable	Name	Type	Description	Values
76	Age of controller	TUAGE	Num	Age of controller of this traffic unit	Age of controller or 00 0-11 months to 98 98 + years 99 Unknown / not stated Null No controller / TU group 10
77	Age group	TUAGEG	Num	Age group of TU controller	01 0 - 4 years 02 5 - 16 years 03 17 - 20 years 04 21 - 25 years 05 26 - 29 years 06 30 - 39 years 07 40 - 49 years 08 50 - 59 years 09 60 - 69 years 10 70 - 79 years 11 80+ years 12 Unknown Null No controller or Traffic Unit
78	Sex of controller	TUSEX	Num	Sex of TU controller	1 Male 2 Female 9 Unknown / not stated Null No controller / TU group 10
79	Postcode of controller	TUPCODE	Num	Postcode of controller's address as in Australia Post postcode book	Postcode or 9998 Overseas 9999 Unknown / not stated Null No controller / TU group 10
80	State of license	TULICST	Num	State in which license was	01 A.C.T. 02 N.S.W.

#	Variable	Name	Type	Description	Values
				issued	03 Victoria 04 Queensland 05 S.A. 06 W.A. 07 Tasmania 08 N.T. 09 Overseas 98 Unlicensed 99 Unknown / not stated Null No controller or Traffic Unit Groups 9, 10, 11, 12 Null No controller / TUTYPEG 9, 10, 11, 12
81	Status of license	TULICSS1	Num	Status of license held by TU controller	1 Learner's permit 2 Provisional licence 3 Standard licence 4 Licence expired 5 Unlicensed (includes expired Learner’s permit) 6 Disqualified / Suspended 7 Cancelled 8 Other 9 Unknown / not stated Null No controller or Traffic Unit Groups 9, 10, 11, 12
82	Distraction a factor	TUDISTR	Num	Any distraction of this controller that was a factor in accident	01 Controller with physical infirmity or chronic illness 02 Controller with sudden illness 03 Controller asleep, drowsy or fatigued 04 Controller distracted or vision obscured by passenger (including

#	Variable	Name	Type	Description	Values
					passengers interfering with controls) 05 Controller distracted / vision obscured by something inside vehicle 06 Controller distracted / vision obscured by something outside vehicle 07 Controller being pursued by police 08 Emergency vehicle sounding warning within earshot 09 Controller using hand-held telephone 97 Other distraction a factor 98 No distraction Null No controller or Traffic Unit
83	Error a factor	TUERRFAC	Num	Any error of this controller that was a factor in accident	**Unusual Manoeuvre** 10 Controller error in manipulation of controls 11 Controller swerving to avoid another vehicle 12 Controller swerving to avoid object 13 Controller swerving to avoid animal 14 Controller swerving for any other reason 15 Controller overtaking on left 16 Controller overtaking on right 17 Controller turning right from wrong lane 18 Controller turning left from wrong lane

#	Variable	Name	Type	Description	Values
					19 Controller disobeying traffic control 20 Controller braking hard 21 Described as driving at 'excessive speed' 22 Controller jumping / falling from vehicle 23 Described as 'loss of control' **Other** 29 Pedestrian disobeying traffic controls 30 Passenger jumping / falling from vehicle 31 Controller / passenger protruding from vehicle 40 Pedestrian from behind parked / stationary vehicle 41 Pedestrian from behind other object 42 Pedestrian under influence of alcohol or other drug 43 Pedestrian confused or indecisive 44 Child pedestrian breaking free from supervisor 45 Pedestrian falling / tripping / jumping into path 97 Other error a factor 98 No error Null No controller or Traffic Unit Group 10

#	Variable	Name	Type	Description	Values
84	Decamped	TUNOSTOP	Num	Did the controller leave the scene of the accident?	1 Yes (mentioned) 2 No (not mentioned) Null No controller / TU group 10
85	Seatbelt / Helmet for controller	TUREST	Num	Type of safety device used by TU controller	1 Adult belt worn 2 Belt fitted, but not worn 3 No restraint fitted to this position 4 Open face (jet) helmet worn / bicycle helmet 5 Full face helmet worn 6 No helmet worn 9 Unknown / not stated Null No controller or Traffic Unit Groups 10, 11, 12 (except ridden animal)
86	Surname of controller	TUNAME	Char	First 14 letters of controller's surname	Surname or Unknown Controller's name is unknown Null No controller / TU group 10
87	Initials of controller	TUINITS	Char	Initials of controller's first two given names	Initials or Null Unknown / TU group 10
88	Date of birth of controller	TUDOBB	Char	Date of birth	
89	TU Alcohol group	TUBACGL	Num	Alcohol group of controller	1 Legal 2 .020 - .049 (special range) 3 .050 - .079 4 .080 - .149 5 .150 + 9 Unknown Null No controller or Traffic Unit Groups 9, 10, 11, 12 (except tram drivers)

#	Variable	Name	Type	Description	Values
90	Casualty Number	CASNO	Num	Casualty number as provided by police	
91	Degree of casualty	CASDEG2	Num	Degree of casualty (2 categories)	1 Fatality 2 Injury
92	Class of road user	CASCLASS	Num	Class of road user	1 Motor vehicle driver (TU types: 1-29 / Casualty Position: 1) 2 Motorcycle rider (30-36 / 2) 3 Pedal cycle rider (40 / 2) 4 Pedestrian (70-71 / 20) 5 Motor vehicle passenger (1-29, 50-57 / 3-4, 6-12, 99) 6 Motorcycle passenger (30-36 / 5, 12, 99) 7 Pedal cycle pillion passenger (40 / 5, 99) 8 Other controller (60-63, 64, 99 /1-2) 9 Other passenger (60-63, 64, 99 / 3-8, 10-11, 99)
93	Casualty position	CASPOSN	Num	Casualty position	01 Driver (D) 02 Motorcycle rider / Pedal cycle rider / Animal rider 03 Centre front (CF) 04 Left front (LF) 05 Motorcycle pillion / Pedal cycle pillion / Animal pillion 06 Right rear (RR) 07 Centre rear (CR) 08 Left rear (LR) 09 Other seating position in motor vehicle

#	Variable	Name	Type	Description	Values
					10 Goods area (GC) 11 Elsewhere in / on vehicle (non-seating position) 12 Caravan / trailer / towed vehicle / sidecar 20 Pedestrian 99 Unknown / not stated passenger seating position
94	Sex of casualty	CASSEX	Num	Sex of this casualty	1 Male 2 Female 9 Unknown / not stated
95	Age of casualty	CASAGE	Num	Age of this casualty	Age or 00 Less than 1 year 98 98+ years 99 Unknown / not stated
96	Age group of casualty	CASAGEG	Num	Age group of this casualty	01 0 - 4 years 02 5 - 16 years 03 17 - 20 years 04 21 - 25 years 05 26 - 29 years 06 30 - 39 years 07 40 - 49 years 08 50 - 59 years 09 60 - 69 years 10 70 - 79 years 11 80+ years 12 Unknown
97	Hospital	CASHOSP	Num	Last hospital to which casualty	0401 Auburn district 0208 Balmain

#	Variable	Name	Type	Description	Values
				was transported for treatment or admission	0413 Bankstown to 9997 Not listed 9998 Not applicable (not treated at, admitted to hospital) 9999 Unknown / not stated
98	Seatbelt / helmet of casualty	CASREST	Num	Did the casualty wear a seatbelt or helmet?	1 Adult belt worn 2 Belt fitted, but not worn 3 No restraint fitted to this position 4 Open face (jet) helmet worn / bicycle helmet 5 Full face helmet worn 6 No helmet worn 7 Child restraint (child seat, booster cushion or baby capsule) 9 Unknown / not stated Null Traffic Unit Groups 10, 11 and 12 (except ridden animal)
99	Ejection	CASEJECT	Num	Was it recorded / mentioned that the casualty was ejected from the traffic unit?	1 Yes 2 No or not stated 9 Unknown Null Pedestrian
100	Survival time	CASSURV	Num	Time taken for casualty to die as a result of accident	1 Died instantly 2 > 0 to □1 hour 3 > 1 to □24 hours 4 > 1 to □10 days 5 > 10 to □20 days 6 > 20 to □30 days Null Non-fatal casualty

#	Variable	Name	Type	Description	Values
102	Fatality surname	CASNAME	Char	First 14 letters of fatality's surname	Surname or Unknown Fatality's name is unknown Null Non-fatal casualty
103	Fatality initials	CASINITS	Char	Initials of fatality's first two given names	Initials or Null Initials unknown or non-fatal casualty
104	Casualty surname	CASSNAME	Char	First 14 letters of casualty's surname	Surname or Null Casualty's surname is unknown
105	Casualty initials	CASINIT	Char	Initials of casualty's first two given names	Initials or Null Casualty's initials unknown
106	Casualty's date of birth	CASDOBB	Char	Date of birth of casualty	Date of birth
107	Casualty's postcode	CASPCODE	Num	Postcode of residence	Four character postcode
108	Alcohol group of casualty	CASBACG	Num	Alcohol group of casualty	1 nil 2 .001 - .019 3 .020 - .049 4 .020 - .049 (special range) 5 .050 - .079 6 .080 - .149 7 .150 or more 9 Unknown Null No controller or Traffic Unit Groups 9, 10, 11, 12 (except tram drivers)

Appendix II Data items in the Inpatients Statistics Collection (ISC)

#	Variable	Name	Type	Description	Example Values
1	Year	YEAR	Char	Financial year of admission	00 Financial year 1 July 2000 – 30 June 2001
2	Statistical Local Area	SLARES	Char	SLA of residence	
3	Postcode	PCODE	Char	Postcode of residence	
4	Health area	ADHSRES	Char	Health area / district of residence	000 All NSW 100 Central Syd AHS 105 Northern Syd AHS 120 Western Syd AHS 125 Wentworth AHS 130 S-W Sydney AHS 135 Central Cst AHS 140 Hunter AHS 145 Illawarra AHS To 999 Other
5	Area Health Service	ARHSRES	Char	Area health service of residence	000 All NSW 100 Central Syd AHS 105 Northern Syd AHS 120 Western Syd AHS 125 Wentworth AHS 130 S-W Sydney AHS 135 Central Cst AHS 140 Hunter AHS 145 Illawarra AHS To 999 Other

#	Variable	Name	Type	Description	Example Values
6	Country of birth	COB	Char	Country of birth	Country of bitrth
7	Sex	SEX	Char	Sex	1 Male 2 Female
8	Emergency status	EMERGNCY	Char	Emergency status	1 Emergency 2 Planned 3 Other 4 Maternity / newborn 5 Regular same day planned admissions
9	Date of admission	ADMDATE	Num	Date of admission to hospital	Date
10	Separation mode (M)	CSEPMODE	Char	Separation mode	A Discharged by hospital B Discharged at own risk C Tfrd to nursing home D Tfrd to Psych hospital (same area) E Tfrd to Psych hospital (other area) F Tfrd to Psych hospital (Unknown area) G Tfrd to hospital (same area) H Tfrd to hospital (outside area) I Tfrd to hospital (unknown area) J Died (autopsy) K Died (no autopsy) L Tfrd other accomm. M Type change separation N Discharge on leave O Not known

#	Variable	Name	Type	Description	Example Values
11	Intensive care hours	ICUHOURS	Num	Hours spent in intensive care unit	Number of hours
12	Length of stay	LOS	Num	Length of hospital stay (days)	Number of days
13	Payment status	PAYST_V5	Char	Payment status	20 Public Patient Election – General and Psychiatric 23 Public Patient Election – Overseas Reciprocal 30 Private Patient – General and Psychiatric (Private Facilities Only) 40 Compensable – Workers Compensation 41 Compensable – NSW Motor Vehicle Accident 42 Compensable – Other 45 Unqualified Newborn of Public Patient 46 Unqualified Newborn of Private Patient 50 Department of Veterans' Affairs General 60 Medicare Ineligible – Other
15	Day only length of stay	DOLOS	Num	Length of stay (hours) if day only	Number of hours
16	Principal diagnosis	ICD10D1	Char	ICD10 Principal diagnosis code	ICD code
17-36	Other diagnosis 2 – diagnosis	IDC10D2 – ICD10D21	Char	ICD10 Principal diagnoses codes	ICD code

#	Variable	Name	Type	Description	Example Values
	21				
37-39	ICD10AM External Cause 1 – 3	ICD10EX1 – ICD10EX3	Char	ICD10 codes for external cause	ICD code
40-42	ICD10AM Place of occurrence 1 – 3	ICD10PL1 – ICD10PL3	Char	ICD10 codes for place of occurrence	ICD code
43-45	ICD10AM Activity when injured 1 – 3	ICD10ACT1 - ICD10ACT3	Char	ICD10 codes for activity when injured	ICD code
46	Date of birth	DOB	Num	Date of birth	Date
47	Date of birth doubtful	DOBTAG	Char		
48	Street number	WFARENUM	Char	Street number of patient's address	Street number
49	Street name	WAYFARE	Char	Street name of patient's address	Street name
50	Suburb	LOCALITY	Char	Suburb of residence	Suburb
51	Given name	GNAME	Char	Patient's given name	Name
52	Middle name	MNAME	Char	Patient's middle name	Name
53	Last name	LSTNAME	Char	Patient's last name	Initial