Abstract

Crash databases are an essential tool in analysing crash risk. Databases allow analysis of trends as well as the identification of high crash locations, thereby providing information with which to make strategic safety related decisions, as well the treatment of high risk sites. New Zealand, and each Australian jurisdiction has its own crash database, and great differences exist between these systems. Austroads have funded a project to review Australasian crash databases to assess best practice features. Contact was made with each Australian jurisdiction, and information gathered on available features from each system. A literature review was also undertaken of international practice on this issue, and contact made with key international experts to determine current best practice. Good features from various jurisdictions are presented, and it is hoped that through sharing of this information, all systems in Australia and New Zealand can be improved to provide better analytical tools to those involved in road safety.

Keywords

Crash database, accident database, crash analysis systems

Introduction

Crash databases are an essential tool in analysing crash risk. Databases allow identification of high crash locations, and provide information on crash causation, allowing these crashes to be effectively targeted. The information contained in crash databases also allows analysis of trends in crashes, and so assists greatly in strategic safety related decisions.

Each Australasian jurisdiction has developed its own crash database system, and the analysis tools available for each vary. In order to add value to this data and better inform decision makers, additional tools should be available for use in some or all of these databases. This project aims to identify good practice in database design, including features that may be of high interest to Australasian jurisdictions.

The intention of the project was to collect information on the features of each database system, and not on the design of the system itself. For example, the type of software platform was not relevant to this study. Similarly, details on the data collection process were typically not included. In addition, there are a number of databases that collect differing types of data on road injury. Databases can contain information collected from police crash reports, hospital admissions or other sources (including insurance data or from tow truck companies). This study concentrated on databases managed by the state or national road agency, which may include information from a number of sources, but typically primarily from police reports. It is recognised that other databases may exist that collate information from other sources (such as hospital admissions only).

In addition, as the focus was on useful tools for the assessment of crash risk, including those at specific crash locations, databases that aggregate information from other crash databases were generally not included.

Methods

The method comprised extensive desktop research, including internet, library and database searches to locate all relevant literature; a survey of Australasian database managers; and a survey of overseas database managers.

The literature search was conducted with the aid of the M.G. Lay Library as well as the World Wide Web. The M.G. Lay Library contains the most comprehensive and up-to-date collection of international
literature on land transport issues (particularly roads) in Australia, and is one of the leading technical libraries in its field in the world.

Each of the Australasian databases was assessed. In three cases (Victoria, NSW and New Zealand) a live demonstration was provided to highlight key features of each system. In all cases, a structured interview was conducted (either by phone or in person) to gather information on key features (see Appendix A). The information gathered was then assessed to determine features that were considered to be of interest.

Contact was made with a number of international experts, including those involved in research on crash databases, database managers and commercial providers of databases. A large number of database systems exist internationally, and it was not the intention to gather information on each of these. Instead, information was sought on examples of good databases. Key experts were contacted and asked to provide information on ‘good’ crash database systems in their region. A definition of a good database was not provided, but rather was left to the judgement of those contacted. However, examples of features of interest were given (based on those identified through the literature review) to help guide this judgement.

Information was sought on 17 different database systems, with responses obtained on 11 of these. The database systems used in the following countries were assessed:

- United Kingdom (KeyAccident, AccsMap, MAAP)
- Canada (National Collision Database)
- Denmark (Vegman.dk)
- India (RSMS)
- Japan (Traffic accident data analysis system for road administrators)
- Singapore (TAAM)
- Spain (ARENA)
- Sweden (STRADA)
- US (Traffic Crash Reporting System, Michigan)

Live demonstrations were provided for RSMS and STRADA.

Once these databases were identified, information was sought directly from database managers and from the internet. To facilitate the collection of this information, a questionnaire was developed (see Appendix A. Note that this has been reformatted for inclusion in this paper, with space for responses deleted).

Results

This section provides a summary of some of the useful features available for crash systems based on the review of literature, international contact and contact with Australasian jurisdictions. ‘Bullet point’ features are listed under the following headings (which mirror the structure of the literature review and the questionnaires):

- data entry
- types of data collected
- database navigation
- analysis of crashes – sites of interest
- analysis of crashes – policy level
- mapping
- integration with other data
- quality control issues
- monitoring sites of interest
- dissemination of data.

A full report for this project is currently being prepared and will contain detailed information and examples of these features. It is not to be expected that any one crash system should possess all of these features, but it does indicate the range of options that jurisdictions should be considering when developing or updating their own systems.
Useful features include:

**Data entry**

- in-vehicle computers or hand-held PDA for the reporting of crashes by Police, with direct links to the crash database system
- GIS linkage to allow accurate location of crashes as well as links with external data sources to validate data fields collected by Police as well as provision of additional data fields not already available to Police (e.g. driver and vehicle information)
- increased take-up of GPS in Australia and New Zealand. Accurate use of GPS would lead to more accurate location of crashes, although it was noted that errors can occur when coordinates are recorded manually, and that an electronic link was seen as ideal
- in-built quality control algorithms and logic checks
- tests to ensure completeness of data
- checks against records at the same crash location to compare data (e.g. road features)
- ability to submit details on non-reported crashes either directly to the database or on-line (these can be flagged as non-Police reported)
- higher priority given to data entry of more severe crashes.

**Data items included in the database**

- contributory factors to crashes (level of confidence has been used overseas given the subjective nature of this data field)
- links to hospital data to provide information on injury outcomes, and also information on Police reporting rates. Figure 1 shows an example from the Swedish STRADA system.

![Figure 1: Link to injury data in Swedish system](image)

- information on traffic density at the time of the crash
- information on what pedestrians and motorists were doing at the time of the crash (e.g. trip purpose)
- additional information on heavy vehicles
- the ability to add new data fields without needing to issue a contract to redevelop the database.
Database navigation

- ability to select individual crashes by clicking on crash from a collision map
- pre-defined queries
- user-defined queries
- drop-down menus
- ability to select any combination of variables easily.

Analysis of crashes – Sites of interest

- ranked sites based on crash rates, crash numbers, crash costs or crash commonality
- site summaries
- route assessments. Figure 2 shows an example of a route assessment from Tasmania.

- crash density maps
- comparison of sites against norms for the rest of the database or other similar sites or areas
- alarm reports
- ability to produce reports for corporate applications (including information summaries for local government)
- factor matrix or ‘stick diagrams’
- automatically generated collision diagrams (with correct orientation). An example is shown in Figure 3.
- access to scanned reports (particularly a sketch of the crash scene)
- ability to select crash details by selecting the crash (or group of crashes) from a map
- ability to provide relevant data on one page or in one document (e.g. raw crash data, collision diagram, factor matrix)
- links to traffic volume data
- links to aerial photographs, photos or video of the crash location.

Figure 2: Tasmanian route assessment
Analysis of crashes – Policy level
- pre-defined reports
- user-generated reports
- the ability to easily export data to third party applications (e.g. SAS) for more detailed statistical analysis
- cross tabulations.

Mapping ability
- ability to confirm crash location at data entry stage
- crash selection on a geographic basis
- presentation of crash information, including thematics and pie charts
- ability to link with other GIS sources
- ability to recognise the current map scale, and provide relevant information accordingly.

Integration with other data
- citation and conviction data
- licence information and driver history
- medical records
- coronial data
- motor vehicle registration details
- traffic volumes
- asset data
- population and demographic data (e.g. from census)
- police intelligence data
- vehicle inspection data
- aerial photography, topographic maps, street maps, photos, video
- a simple data export facility to allow easy external integration of data
Quality of data
- consistency or logic checks for data entry
- requirement for a training course before the database is used
- disabling of data fields if conflicting fields are selected
- warnings of possible errors when making queries
- details of search criteria included on all outputs
- ability to determine reporting rates (e.g. through a link with medical data)

Monitoring sites of interest
- some form of site monitoring module allowing assessment of sites or locations where the same treatment has been installed to be assessed before and after treatment installation.

Dissemination
- web-based access
- open access by approved stakeholders
- public access version of the database. Figure 4 shows an example of the VicRoads CrashStats system.

Discussion
Several difficulties were encountered during this study. It was difficult to identify good international examples of crash databases, and then to obtain details on these. In a number of cases, information was sought on databases that had been identified as ‘good’, but no response was provided on these. Even when information was provided, it was difficult to obtain relevant details on good features from crash databases, especially when it was not possible to view the system directly. The questionnaire was designed as a prompt to obtain information on good features, but it is possible that other good features do exist, and were not assessed during this review.

In general, the Australasian databases fared well in comparison to international best practice. There are not many features that are in use overseas that are not already used in at least one of the Australasian databases.
However, there was quite a contrast in the functionality of databases used throughout Australia and New Zealand. The functionality did not seem to correlate well with the size of the jurisdiction or the age of the database. In some cases, smaller jurisdictions have very good databases, and older systems appear to have similar or even better features than more modern systems. This may indicate a lack of knowledge within jurisdictions about possible functionality of database systems, something that this research hopes to address.

Given this apparent lack of awareness of what was happening in other Australasian jurisdictions, there may be advantages in the formation of a crash database user group to help keep track of knowledge and experience in this area.

Every database in Australia and New Zealand could improve the functionality and provision of information based on one or more of the above features. Given the cost of data collection, and the utility of the information provided, an investment in additional features would appear to be highly cost beneficial. None of these features are crucial for an adequate level of database performance (all Australasian databases currently deliver at least basic outputs), but each would help maximise the accuracy, efficiency and/or functionality of data provision.

All of the Australian databases are lacking some of the key features identified above. It would be costly for each jurisdiction to update their systems to provide all of the best practice features identified here. Given broad similarities in the data collected and in the features required from a crash database, there would be large benefits in jurisdictions joining together to create a ‘top of the line’ crash database system. Of greatest practical benefit would be a national database that included all of the features identified here. Ownership of data could be retained at the jurisdiction level, but aggregated nationally to allow much needed data to inform national strategy for road safety. This would require less investment by individual jurisdictions for a greater product.

Although all of the features identified in the results section were seen as being of value, the following features are seen as being particularly useful:

- good linkages with other sources of data (at data entry and analysis stages)
- built-in quality checks
- provision of contributory factors
- linkages to hospital data
- ability to select any combination of variables quickly and easily
- all of the analysis tools listed under ‘sites of interest’
- details of the search criteria included on all outputs for quality reasons
- a monitoring module for sites or lengths where safety improvements have been made.
- web based access
- public access to data.

It is recommended that when updating or renewing crash databases, that these and the other features highlighted be considered. It is also recommended that an informal network of database managers be established to facilitate an exchange in knowledge in this area. The possibility of joint development of databases by jurisdictions should also be assessed.

Acknowledgements

The author would like to thank Austroads for funding this research. Acknowledgement is also made to all those who responded to the questionnaire and provided graphical examples from their database, including those in Australia, New Zealand and overseas. Special thanks to those who provided ‘live’ database demonstrations. Finally, thanks to Will Hore-Lacy who assisted in the literature review, and to Bill Frith who provided comments on draft reports.
Appendix A: Questionnaire
(note that this has been reformatted for inclusion in this paper. Spaces for responses have been removed)

Austroads road safety crash database assessment project

This project aims to identify useful features that are used in road safety crash databases. This checklist has been developed which highlights the areas of interest to this project. For most of these areas of interest, examples are included to help prompt responses. These are given as examples only, and your database may have other features that are of interest.

Background information

What is the name of your database?

What organisation maintains this database?

How long has it been in operation?

What geographic area does it cover?

Data entry

Does your database have any features that makes data entry easier (e.g. direct links to electronic crash forms; use of GPS)?

Are quality checks included in the data entry process to prevent errors, and if so, how?

Data items included in database

Most databases include information on the people involved, severity, type of vehicle, and environment. Do you include information on additional types of factors (e.g. contributory factors, type of objects struck, other)?

Database navigation

Does your database have any special features that make it easy to navigate (e.g. map based selection, drop down menus, predefined queries)?

Analysis of crashes – sites of interest

Does your database have features that make it easy to select sites of interest (e.g. black spots or routes)?

Can you select locations based on radius, lengths, or area?

Can you select crash locations from a map or using street names?

Can you produce crash rates, or select sites based on crash numbers, severity, costs, or crash commonality?

Can you produce ‘alarm reports’ (e.g. statistical tools to detect crash numbers or rates outside normal bounds)?

What types of analysis tools for sites of interest does your database have (e.g. factor matrix, collision diagrams, access to scanned police reports, video or photo of the site, aerial photos of the site)?
Analysis of crashes – policy level

Does your database have tools that make policy level analysis easier (e.g. ability to produce region wide information easily, or to examine change in trends over time)?

Do you have the ability to produce ready made queries (including key indicators for safety) or retrieve data of interest for very specific queries?

Mapping ability

Can you select crashes based on a map location?

Are you able to spatially display information on crash locations? For example, are you able to produce thematic information about crashes (e.g. colour coding by crash severity; produce spatially linked pie charts based on crash features)?

Integration with other data

Are you able to integrate data from your database with other types of relevant information (e.g. traffic volumes, asset information, census data, sanctions and infringement data, registration data etc)?

If you do integrate data, how do you do this (e.g. within the database or externally using common reference numbers or GIS)?

Quality of the data

How up-to-date is the crash data?

How easy is it to interpret data, or to make errors when interpreting? Are there any features which prevent errors in interpreting the data (e.g. reports show details of the query, or flags appear in the system for common interpretation errors)?

Monitoring sites of interest

Is it possible to monitor sites of interest (e.g. sites treated under a black spot program; speed camera sites)?

Does this allow for before and after analysis of treatments? Does it allow selection of control sites (untreated sites)?

Dissemination

Who has access to the database? Can you access the database via the internet? Do the public have access to the data?

Other areas of interest

Are there other features of the database that are useful that have not been covered above? Please also forward any other information that you think might be relevant.