Published paper

Austin, K.P. (1992) *A Linked Police and Hospital Road Accident Database for Humberside.* Institute of Transport Studies, University of Leeds. Working Paper 369

Working Paper 369

July 1992

A LINKED POLICE AND HOSPITAL ROAD ACCIDENT DATABASE FOR HUMBERSIDE

KP Austin

ITS Working Papers are intended to provide information and encourage discussion on a topic in advance of formal publication. They represent only the views of the authors, and do not necessarily reflect the views or approval of the sponsors. This work was sponsered by the Science and Engineering Research Council.

UNIVERSITY OF LEEDS
Institute for Transport Studies

ITS Working Paper 369 *ISSN 0142-8942*

July 1992

A LINKED POLICE AND HOSPITAL ROAD ACCIDENT DATABASE FOR HUMBERSIDE

KP Austin

This work was undertaken on a Research Studentship sponsored by the Science and Engineering Research Council with the assistance of Humberside County Council.

CONTENTS

Page

1.INTRODUCTION

2.PREVIOUS MATCHING METHODS

3.MATCHING METHOD
3.1 Sources of data
3.2 Matching variables
3.3 Matching method

4.REPORTING RATES
4.1 Reporting by user group
4.2 Reporting by gender
4.3 Reporting by age
4.4 Reporting by journey purpose
4.5 Reporting by source of referral
4.6 Reporting by employment status

5.INJURIES TO ROAD TRAFFIC ACCIDENT VICTIMS
5.1 Introduction
5.2 Injury location by user group
5.3 Injury location by passenger seating position
5.4 Injury location by age
5.5 Injury location by vehicle make
5.6 Injury location by manoeuvre
5.6.1 Junctions
5.6.2 Not at junctions

6.CONCLUSION

7.ACKNOWLEDGEMENTS

8.REFERENCES

9.APPENDIX A

ABSTRACT

AUSTIN, K (1992). A linked police and hospital road accident database for Humberside. *ITS Working Paper* **369**, Insitute for Transport Studies, University of Leeds, Leeds.

The current system used to collect road accident statistics provides no detailed information on the injuries sustained, nor does it account for all the road accidents and casualties resulting from them. Other studies have attempted to link hospital casualty data to that contained on the police reports but with limited success. The method outlined here matches 97.3% of casualty records that should be matched with only 1.4% error. This will therefore provide a effective model to test the reporting rates of various groups and the injuries sustained by them.

KEY-WORDS: *author to supply*

Contact: Kevin Austin, Institute for Transport Studies (tel: 0532 335356).

A LINKED POLICE AND HOSPITAL ROAD ACCIDENT DATABASE FOR HUMBERSIDE.

1.INTRODUCTION

The current system used to collect road accident data is via the STATS 19 system which records details of those injury accidents that were reported to the police. However, this provides no information on the injuries sustained to the accident victims, nor does it account for all casualties of road accidents. The Road Safety Code of Good Practice (1989) states that:

> " Local Highway Authorities should consider obtaining supplementary information from local hospital records..."

The aims of this study are to develop an improved technique to link the police casualty data with hospital casualty data by using the name and address of the victim. It investigates the differences in the police reporting of accidents for various groups and the bodily location of injury sustained by the casualty. The data used in this study was covered under the Data Protection Act (1984) and as a consequence all the information was kept within Humberside's accident data system.

2.PREVIOUS MATCHING METHODS

There have been several other studies which have attempted to link hospital casualty data to that contained on the police reports. Nicholl (1980), Sayer and Hitchcock (1984), Stone (1984) and Fife (1989) all use some combination of the variables of gender, age, user class, severity, accident date and location. Barancik and Fife (1985) used similar variables but also had access to the coded name of the casualty.

These studies developed tolerance levels to test if the police and hospital records had a sufficiently close match. That is, some of the variables such as age or accident date were allowed to vary within a predefined range between the police record and the hospital record. On the other hand Fife (1989) created cells based on unique values of the variables and if there was one hospital and one Police record in a cell then it was considered to be a match.

The results of matching vary and this depends on the source of data used as well as the matching procedure. Sayer and Hitchcock (1984) and Barancik and Fife (1985) matched data for all casualties attending hospital and produced matches of 23% and 55% respectively. Nicholl (1980) and Stone (1984) used in-patient data and achieved matches of 41% and 70% respectively. In addition there were 8.9% of erroneous matches in the Nicholl (1980) study and 6.6% of hospital records were involved in multiple matches in the Stone (1984) study. Fife (1989) achieved an 85% match for fatal casualties.

Those hospital records not matched with a corresponding police record can be attributed to either:

- failure to report accidents to the police

- failure in the algorithm to match police and hospital records due to a miscoding in some of the variables used.

Unfortunately the above studies cannot separate these two factors and so the true level of reporting and the accuracy of the matching method cannot be assessed. The only way this can be achieved is by comparing these results with those in questionnaire studies such as Bull and Roberts (1973) or Hobbs at al (1979), where matching based on the casualties name was done manually.

Both Bull and Roberts (1973) and Hobbs et al (1979) found that all fatalities were reported to the police, hence the method devised by Fife (1989) would fail to match about 15% of the fatal records. Bull and Roberts (1973) and Hobbs et al (1979) found about 80% of serious injuries were reported to the police. This category includes injuries that do not require hospital admission, such as some fractures, and so one would expect the reporting rate for in-patients to be higher, at say 90%. Therefore Stone (1984) and Nicholl (1980) failed to identify a quarter and a half of the correct matches respectively. Bull and Roberts (1973) and Hobbs et al (1979) found the reporting rate for all hospital casualties to be between 65% and 72%. So at the very least Sayer and Hitchcock (1983) and Barancik and Fife (1985) failed to identify two-thirds and 15% of correct matches respectively. Sayer and Hitchcock (1984) used data from a less developed country which may have a lower reporting rate, but the use of the persons name may also help to improve the matching process.

3.MATCHING METHOD

3.1SOURCES OF DATA

The hospital records used in the matching process for this study come from the Accident and Emergency Department of Hull Royal Infirmary (HRI). When the casualty arrives at the Accident and Emergency Department a number of basic items are coded into the computer, such as name, address, gender, age, location and date of accident. The medical data are entered later on by the doctor which includes the diagnosis and the treatment given. Between May and December 1991 there were 1593 records recorded on the road traffic accident (RTA) section of the HRI database.

The STATS 19 records contain all the information that is useful for the investigation of accidents, whilst the Police data contains the confidential casualty information to be used in the matching process. The first step was to combine Police and the STATS 19 based on accident reference number and casualty number which could then be used to match with the hospital records.

3.2MATCHING VARIABLES

The variables that were used for the matching process were surname, forename, first line of their address, casualty age, gender and accident date. In an initial study it was found that only 16% of those that should match did so, due primarily to miscoding by either the Police or the hospital.

The length of the variables used were shortened so as to reduce the chance of error, but not so much so that they were no longer unique. Forename was reduced to three digits (and for one part of the matching process to two digits), the surname was reduced to six digits and the address to four digits. Age was not altered.

The date of admission coded on the hospital files could either be the same as the accident date on the Police files or one day after. Unfortunately some correct matches would be missed as a consequence. There were 2 cases where date was one day less in the hospital record, 1 where the hospital record was 2 days less and 1 where it was 2 days more. The widening of the tolerable age range caused a large increase in the number of erroneous matches and so was not considered acceptable. Both date of admission and date of accident are coded onto the hospital database. However there were 43 records on the hospital database where date of accident was omitted, whereas date of admission was coded for all records and so this was considered more reliable.

3.3MATCHING METHOD

The matching algorithm was developed using the Dataease Query Language (DQL) using the Dataease 4.2 database. This algorithm is shown in Table 1.

Table 1:The matching algorithm for linking police and hospital casualty records

	Date	Surname	Forename (3)	Forename (2)	Address	Age	Gender
1	X	X		X			X
2	X				X	X	X
3	X		X		X	X	
4	X	X				X	X
5	X		X			X	X
6	X		X		X		X

The algorithm was devised to act rather like a sieve, in that the police records that are reasonably closely matched with the hospital record would be retained, and those that are not would fall through and not be written to the database. The optimum matching method was developed by varying the combinations of the variables used. There was a trade off between the number of correct matches and the number of erroneous matches. That is, with only a few variables used in the process it would mean that more matches would be obtained since there would be less constraints on the variables used. But as a consequence there would be a higher level of erroneous matches, since an incorrect match may satisfy these few conditions. The converse is also true, in that with a greater number of variables there would be fewer correct matches and fewer errors.

Some variables were considered to be essential to be included in each match. Date was included in all cases to avoid multiple matches when the same person was injured in more than one accident in the study period. In this study three cases corresponded to this factor.

It was essential to distinguish between people of the same family involved in an accident. In this case the surname and address would be identical. Hence, gender and forename or gender and age or age and forename were included in each matching level. In the first matching level the forename was spelt with two initials, so as to allow those matches where the third letter of the forename was incorrect. It was only used in this level since the general use of this increased the level of erroneous matches. There were 19 cases in the matched dataset where the three digit forenames were different but the two digit ones were correct. Of these, three had age different and one had age and address different. Hence by using the two digit forename an extra four matches would be picked up.

The six matching levels were used to take account of the various combinations of errors. If there were more levels it would increase the processing time with only a very small increase in the number of matches found. The system of matching is shown in Figure 1.

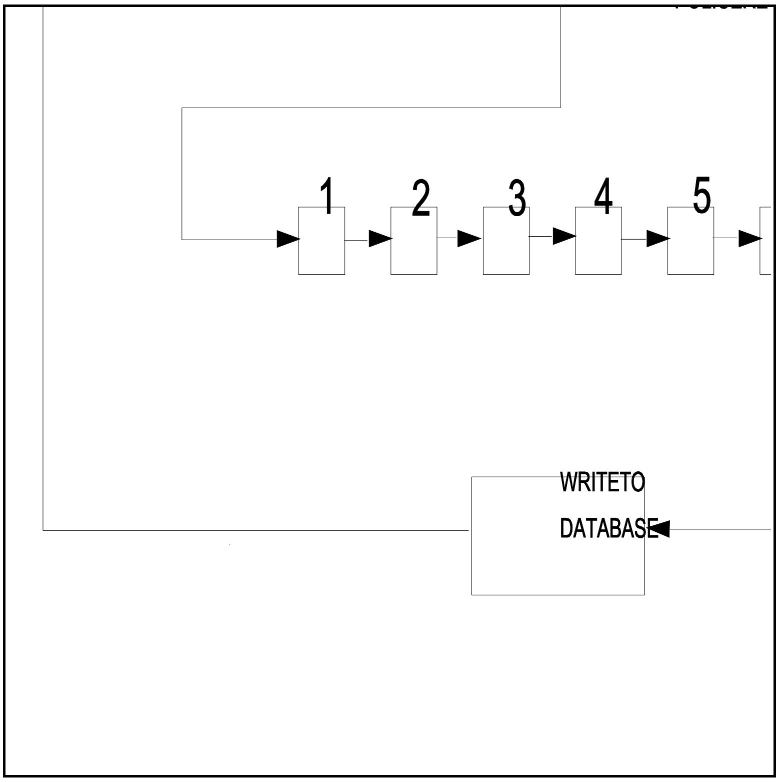


Figure 1: The matching system for police and hospital casualties.

The first hospital record is taken and also the first Police record, the two are compared against the matching algorithm shown above. That is, they are compared firstly using the variables of date, surname, forename and gender. They are then compared using date, address, age and gender, and so on until the sixth stage has been completed. If any of these result in a match then the two records are merged and are written to a file. The next Police record would be taken and the above process is repeated until all the Police records are checked. The next hospital record would be taken and compared with the first Police record, until all comparisons have been made.

A manual match between the two datasets was undertaken and this matched 1067 of the 1593 hospital records, a reporting rate of 67%. The computer algorithm matched 1038 of those records which meant that 97.3% of those that should have matched did so. There were 14 errors in the matching process which amounted to 1.4%.

The studies by Nicholl (1980) and Stone (1984) concerned only in-patients and so to compare with their studies the method was re-run to account for only those who were admitted to hospital. The manual method matched 258 of the 287 in-patients which corresponds to a reporting rate of 89.9%. The computer method matched 249 of those and so there was a success rate of 96.5%. This study recorded only 2 errors which was equal to 0.7%. This is a considerably better method than those highlighted in Section 2. The matching of the records took over a day to link and so the system of matching should be made more efficient. The flow diagram is shown in Figure 2 highlights how this could be achieved.

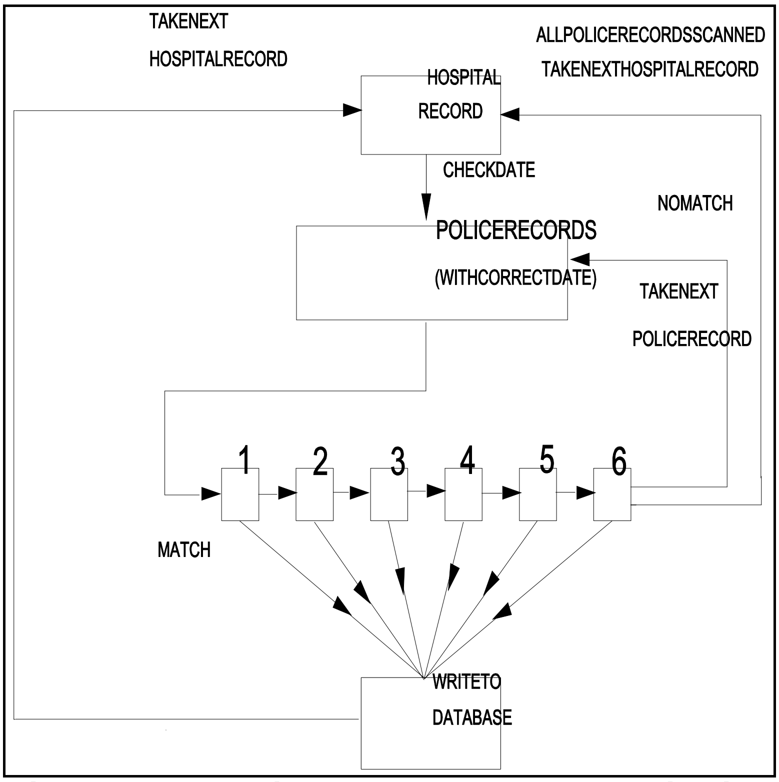


Figure 2: An improved method of linking police and hospital casualty records.

The first hospital record is taken and all the Police records are searched to see if the date is either the same or one day less than that on the hospital record. The ones corresponding to this are put into a temporary file. The police records in this file were then checked against the hospital record, if there was a match then the records would be merged and written to a file and the next hospital record would be taken. If it did not match then the next part of the algorithm would be checked, and so on. If the Police record did not match, then the next Police record would be taken, this would continue until all the Police records in the temporary file are searched. The next hospital record would then be taken and the process repeated.

The results from the matching procedure can therefore be used to assess the levels of Police reporting to those receiving treatment at a hospital for road traffic injuries. It can also be used to investigate injury location by various criteria. In the following sections examples of the results of this data manipulation are shown.

4.REPORTING RATES

4.1REPORTING BY USER GROUP

The levels of reporting by user group are shown in Table 2.

Table 2:Reporting rates by user group

User Group	Hospital File	Matched File	% reported
Pedestrians	302	226	74.8
Drivers	408	250	61.3
Passengers	388	233	60.0
TWMV riders	209	159	76.1
Cyclists	196	131	66.8
Pillion	10	9	90.0
Other	7	6	85.7
Not known	73	24	32.9

The highest reporting rates are for the vulnerable road users, that is, the cyclists, motorcyclists and pedestrians and the lowest reporting rates for the occupants of motor vehicles. This is contrary to other studies that have been undertaken in this field Bull and Roberts (1973), Hobbs et al (1979), Sayer and Hitchcock (1984) and Harris (1990) who found the highest levels of reporting for vehicle occupants and the lowest for cyclists and motorcyclists. This difference may be due to the insurance position, in that the hospital can claim £25 from insurance companies for treatment. Hence, they would wish to collect all those involved in motor vehicle accidents and so the police reporting rate would be low, whereas they may not record all the pedestrian and cyclist casualties on the RTA database since the costs cannot be reclaimed.

4.2REPORTING BY GENDER

It was found that there is little difference in the reporting rates between males and females, 66% of male casualties are reported to the Police compared with 63.8% for females. The chi-squared value with 1 d.f. is 0.82 which is not significant at the 5% level. This result is similar to that of Harris (1984).

4.3REPORTING BY AGE

The reporting rate by age of casualty is shown in Table 3.

Table 3:Reporting rates by age of casualty

Age	Hospital	Matched	% reported
0 - 14	228	161	70.6
15 - 19	260	174	66.9
20 - 34	588	370	62.3
35 - 64	353	229	64.9
65+	98	74	75.5
Blank	66	30	45.4

The oldest (>65) and youngest (0-14) age groups have the highest level of reporting. These were grouped together and compared against those casualties aged 15 to 64. It was found that the 0-14 and 65+ groups had a higher level of reporting. The chi square value with 1 d.f. is 6.82 which is significant at the 1% level. This may be because these groups tend to have a greater proportion of pedestrian accidents which have the highest reporting rates. Pedestrian accidents were removed from the sample and the chi squared test was re-run. It was found that there was no significant difference at the 5% level between the 0-14 and 65+ group and the 15-64 age group with a chi-squared value of 0.0036. In fact the 0-14 group has the lowest level of reporting which is consistent with the results of Maas and Harris (1984).

4.4REPORTING BY JOURNEY PURPOSE

This section relates to the activity that the casualty was undertaking when they were involved in the accident. The reporting rates for these different activities are shown in Table 4.

Table 4:Reporting by journey purpose

Purpose	Hospital	Matched	% reported
Home to work	139	92	68.1
Work to home	103	71	68.9
To/from school	37	18	48.6
Social	814	565	69.4
Working	102	58	56.9
Other	135	91	67.4
Unknown	263	143	54.4

The hypothesis that the reporting rate for children on their way to and from school would be lower than the other categories was tested. The chi-squared value with 1 d.f. is 6.01 which is significant at the 5% level. Those records that had no coding for journey purpose were then re-checked to see if they were coded as children to/from school or not to/from school on the STATS 19 section of the record. 4 were found as to/from school which meant that 53.7% of children on their way to and from school were reported. The difference between children to/from school and the other categories now has a chi-squared value with 1 d.f. of 3.91 which was significant at the 5% level.

The social category has a higher level of reporting than the working category. It has a chi-squared value with 1 d.f. equal to 6.56 which is significant at 5%. To/from work also has a higher level of reporting than working with a chi squared value with 1 d.f. equal to 3.44 which is significant at the 10% level. This may be because those who are working would not want to get involved in legal problems and so would not report the accident.

4.5REPORTING BY SOURCE OF REFERRAL

This section refers to the differences in reporting levels according to who sent them to the hospital. This is shown in Table 5.

Table 5:Reporting by source of referral

Referral	Hospital	Matched	% reported
G.P.	23	7	30.4
Self	444	133	30.0
999	1086	886	81.5
Hospital	7	0	0.0
Work	11	0	0.0
School	1	0	0.0
Deputising	0	0	0.0
Police	17	8	47.1
Other	4	4	100.0

The high level of reporting for 999 calls would be because whenever an ambulance is called the police are also alerted. The difference between 999 calls and the other groups combined has a chi- squared value with 1 d.f. of 405.4 which is significant at 1%. A surprising figure is the low level for those casualties that were referred to hospital by the Police. It is inevitable that for some casualties records may not filled in and this is highlighted in that almost 20% of the 999 calls were also not reported.

4.6REPORTING BY EMPLOYMENT STATUS

The reporting levels by employment status are shown in Table 6.

Table 6:Reporting rates by employment status

Employment	Hospital	Matched	% reported
Employed	856	535	62.5
Retired	120	92	76.7
Unemployed	188	128	68.1
Under 5	57	35	61.4
Student	184	113	61.4
Other	103	78	75.7
Not known	85	57	67.1

The retired casualties have the highest level of reporting, and when compared to the employed group have a chi-squared value with 1 d.f. of 9.19 and so is significant at 1%. This difference may be explained in that the older casualties tend to be more severely injured. In this case 30.4% of retired casualties had either serious or fatal injuries compared to 21.2% for employed persons.

5.INJURIES TO ROAD TRAFFIC ACCIDENT VICTIMS

5.1INTRODUCTION

The use of injury data is twofold, firstly, to investigate how certain groups are affected, and secondly, to look at vehicle design and safety feature effectiveness.

5.2INJURY LOCATION BY USER GROUP

The location of injury by user group is shown in Table 7.

Table 7: Location of injury by user group

	Head	Chest	Abdomen	arms	Legs	Pelvis	Spine
Pedestrian	48.9	3.0	1.5	8.1	33.3	3.0	1.5
Cyclist	42.5	3.5	0.0	23.9	29.2	0.0	0.9
TWMV	10.4	3.9	1.3	27.9	50.0	3.2	3.2
Driver	43.4	13.2	1.5	11.7	11.7	1.9	16.6
Passenger	44.1	15.1	1.7	16.8	12.3	1.1	8.4

All user groups have approximately 40% of injuries to their heads except for motor cyclists who have a considerably lower level and the difference between this group and the others has a chi-squared value with 1 d.f. of 61.6 which is significant at 1%. This is probably due to the compulsory wearing of helmets by motor cyclists which protect their heads.

For chest injuries drivers and passengers have a significantly higher level of injury than other groups, the chi-squared value with 1 d.f. equals 25.8 and this is significant at 1%. This may be because the wearing of seat belts, while reducing the numbers of serious head injuries, actually causes a number of chest injuries.

The motor cyclists and pedal cyclists have a higher proportion of injuries to their arms presumably because this would often be the first part of the body to hit the ground as the rider breaks his/her fall. The difference between these two groups summed together and the other groups has a chi-squared value with 1 d.f. equal to 23.2 and is significant at 1%.

The drivers of vehicles have a higher level of spine injuries, this may be because of their seating position at the time of impact. The difference between this group and the other groups summed together has a chi-squared value with 1 d.f. is 35.9 and is statistically significant at 1%.

Leg injuries occur to a much greater proportion to pedestrians, motor cyclists and pedal cyclists. The difference between these groups summed together and the other groups has a chi-squared value with 1 d.f. of 72.7 and this is significant at 1%.

5.3INJURY LOCATION BY PASSENGER SEATING POSITION

The injury location by the passenger seating position is shown in Table 8 and is compared to the Nicholl (1980) study.

Table 8:Injury location by passengers seating position

	Front				Rear			
	Nicholl		Austin		Nicholl		Austin	
	no	%	no	%	no	%	no	%
Head/neck	268	75.6	33	33.7	98	75.3	42	58.3
Chest/ shoulders	24	6.9	31	31.6	9	6.9	7	9.7
Arms	13	3.7	11	11.2	6	4.6	6	8.3
Spine	3	0.9	11	11.2	3	2.3	4	5.6
Abdomen/hips Pelvis	6	1.7	2	2.0	4	3.1	3	4.2
Legs	35	10.0	10	10.2	10	7.7	10	13.9

The results of Nicholl (1980) show that front and rear passengers have the same proportion of injuries to the head and neck. In the present study head injuries in general form a lower proportion of total injuries than in Nicholl (1980), this may be because head/neck injuries form a larger proportion of injuries to in-patients, since they may be kept in for observation. In this study the proportion of head/neck injuries to rear seat passengers are larger than to front seat passengers and the chi-squared value with 1 d.f. equals 10.2 and is significant at 1%. Legislation for the use of rear seat belts came into force on July 1st 1991, but there was no decline in the proportion of head/neck injuries to this group after it's introduction as one would expect. Chest and shoulder injuries form a larger part of injury to front seat passengers than rear seat passengers and this may be because of seat belt use, the chi-squared value with 1 d.f. equals 11.5 and so is significant at 1%.

An investigation of the written diagnosis revealed that front seat passengers still have a higher proportion of neck strains and whiplash than rear seat passengers, the chi-squared value with 1 d.f. equals 6.06 and is significant at the 5% level.

5.4INJURY LOCATION BY AGE

Injury location by age can be investigated and this is shown in Table 9.

Table 9:Injury location for car occupants by age

Age	Head/neck		Trunk		Limbs	
	Nicholl	Austin	Nicholl	Austin	Nicholl	Austin
0 - 14	77.8	71.0	5.6	12.9	16.7	16.1
15 - 25	74.3	43.4	8.8	23.2	16.9	33.2
26 - 55	69.9	43.7	14.3	34.0	16.0	22.2
56+	70.8	30.0	18.4	47.5	10.8	22.5
All	72.2	44.5	12.0	29.6	15.7	25.9

The results show that as the casualty becomes older there is a lower level of head injury and a higher level of trunk and limb injury. The results follow a similar trend to that of Nicholl (1980) although possibly more pronounced. The same study could be done for the other user groups.

5.5INJURY LOCATION BY VEHICLE MAKE

Injuries can be linked to the vehicle that the casualty was associated with or the vehicle that the pedestrian was hit by. The matched hospital record which contained the casualty details was linked to the STATS 19 vehicle record to obtain the vehicle type involved and was then linked to the police vehicle record so as to include the make and type of vehicle. This achieved 908 matches in total. Injuries can therefore be investigated by the model of car, this example uses just Fords and the results can be seen in Table 10.

Table 10:Injury location by vehicle make

	Head	Arms	Legs	Chest	Abdomen	Pelvis	Spine
Sierra	13	2	6	1	0	1	1
Escort	23	9	13	3	2	0	5
Fiesta	16	8	11	4	0	2	3
Orion	3	1	7	1	0	0	2
Cortina	9	2	3	3	2	0	1
Transit	6	5	1	0	0	0	1
Others	4	4	1	1	0	0	1
Total	77	31	42	13	4	3	17

This can be further subdivided into user groups, for example, for injuries to pedestrians involved in contact with Ford Sierras, 3 had injuries to their heads, 1 to arms and 3 to legs. Although the sample is small and significant results cannot be achieved, it demonstrates the feasibility of the system for analyzing car types and injury patterns if a larger amount of data were to be obtained. With the appropriate data manipulation it would be possible to look at cycle and motorcycle injuries when they are involved in collisions with vehicles of a specific make and model.

5.6INJURY LOCATION BY MANOEUVRE

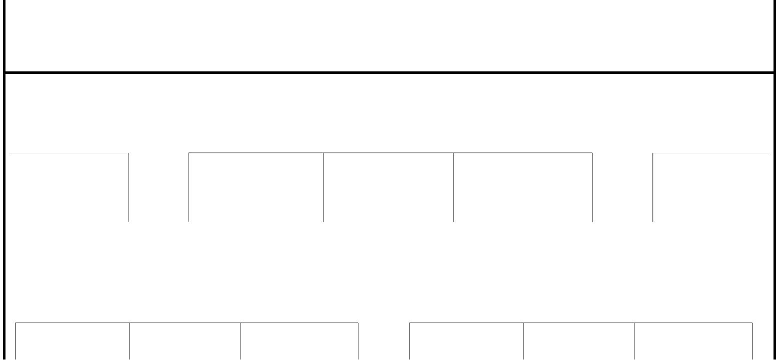
The location of injury by manoeuvre can be subdivided by user group and whether the accident occurred at a junction or not. In this case car occupants are the only category analyzed.

5.6.1Junctions.

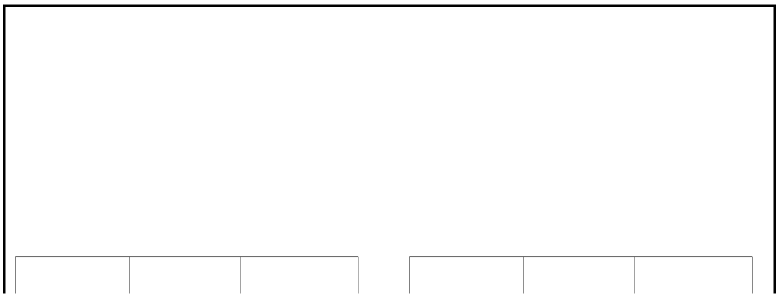
The manoeuvres were subdivided into:-

- Nose to tail: This was when one vehicle turning left or right and was being hit from behind by a vehicle going straight on, or two vehicles collided that were travelling in the same direction.

•Cross over: This would include one vehicle turning and being hit by a vehicle going straight on.



•Head on: This would be vehicles travelling in the opposite direction.



•Single vehicle: This was where only one vehicle was involved.

For car occupants the results are shown in Table 11.

Table 11:Injuries to car occupants by manoeuvre at junctions

	Head	Arms	Legs	Chest	Abdomen	Spine	Pelvis
Nose-tail	24	4	3	7	0	12	2
Cross over	43	7	15	31	1	15	2
Head on	2	2	0	1	1	1	0
Single vehicle	11	4	2	3	0	2	1

Cross over accidents produce a higher proportion of chest injuries than nose to tail accidents with a chi-squared value with 1 d.f. of 3.81 which is significant at the 10% level.

5.6.2Not at junctions.

This was divided into:

- Head on
- Nose to tail
- Overtaking
- Single vehicle
- Others

The results are shown in Table 12.

Table 12:Injury location for car occupants by manoeuvres not at junctions

	Head	Arms	Legs	Chest	Abdomen	Spine	Pelvis
Head on	17	6	7	6	1	2	0
Nose-tail	25	4	7	1	0	5	0
Overtaking	2	3	1	0	0	1	1
Single Vehicle	33	11	9	4	1	6	0
Other	2	2	0	0	0	0	0

The level of chest injuries for nose to nose is higher than for nose to tail. The chi-squared value with 1 d.f. is 4.33 and is significant at 5%. With a larger source of data a much clearer investigation of this concept would be possible

6.CONCLUSION

The use of confidential information such as name and address in a matching process, such as that highlighted above, provides a much better system than any used before. 97.3% of those that should have matched did so by this method with only 1.4% error.

Because of the very low rates of error and omission it was possible to investigate the reporting by various groups. It was also possible to investigate the location of injury for various groups and this produced several significant results. With a larger amount of data it will be possible to provide reliable answers to such questions as "what are the likely injury locations to a cyclist turning right across the path of a G registered Vauxhall Cavalier."

The ability has been allowed for to link in data for in-patients in short-stay wards. The in-patient and the merged data will be linked using the unique accident and emergency reference number. The data included would be admission and discharge date which can be used to calculate the length of stay, and the discharge remarks. The link between these two databases should be almost 100% since a unique number coded at the same time is used. This can then be used to assess the costs of certain types of accident and should aid the Highway Authority in reducing the bed days in hospital.

7.ACKNOWLEDGEMENTS

The work undertaken in this project would not have been possible without the help of the Director of Technical Services, Humberside County Council, the Chief Constable of Humberside Police and the Consultant of the Accident and Emergency Department of Hull Royal Infirmary who allowed the data to be used. It was undertaken as part of a research studentship sponsored by the Science and Engineering Research Council.

The author would like to thank Philip Redfern and Peter Shepherd of Humberside County Council, Chief Inspector Steven Madson of Humberside Police and Dr Peter Haly of Hull Royal infirmary for their help and encouragement and would also like to thank Howard Kirby and Dr Miles Tight of Leeds University for their helpful suggestions which have contributed to the work reported here.

8.REFERENCES

BARANCIK, JI and FIFE, D (1985). Discrepancies in vehicular crash injury reporting. Northeastern Trauma Study IV. *Accident Analysis and Prevention,* **17**(2), pp 147-154.

BULL, JP and ROBERTS, BJ (1973). Road accident statistics. A comparison of Police and hospital information. *Accident Analysis and Prevention,* **5**(1) pp 45 - 53.

DEPARTMENT OF TRANSPORT (1991). Car and driver injury accident rates. Great Britain 1989. *Transport Statistics Report.* HMSO (London).

FIFE, D (1989). Matching Fatal Accident Reporting System cases with National Centre for Health Statistics motor vehicle deaths. *Accident Analysis and Prevention,* **21**(1) pp 79-83.

HOBBS, CA, GRATTAN, E and HOBBS, JA (1979). Classification of severity by length of stay in hospital. *LR871,* Transport Road Research Laboratory (Crowthorne, Berkshire).

LOCAL AUTHORITIES ASSOCIATION (1989). *Road Safety Code of Good Practice.* Association of County Councils (London).

MAAS, MW and HARRIS, S (1984). Police recording of road accident in-patients. *Accident Analysis and Prevention,* **19**(2) pp 91 - 104.

NICHOLL, JP (1980). The use of hospital in-patient data in the analysis of injuries sustained by road accident casualties. *SR628,* Transport Road Research Laboratory (Crowthorne, Berkshire).

SAYER, I and HITCHCOCK, R (1984). An analysis of police and medical road accident data: Sri Lanka 1977-1981. *SR 834,* Transport Road Research Laboratory (Crowthorne, Berkshire).

STONE, RD (1984). Computer linkage of transport and health data, *LR 1130,* Transport Road Research Laboratory (Crowthorne, Berkshire).

TUNBRIDGE, RJ (1987). The use of linked transport-health road casualty data. *RR 96,* Transport Road Research laboratory (Crowthorne, Berkshire).

9. APPENDIX A

DATA FIELDS ON THE MATCHED DATABASE

From the Police/STATS 19 database:

Accident date
Accident reference number
Vehicle reference number
Casualty reference number
Casualty class
Casualty gender
Casualty age
Casualty severity
Pedestrian location
Pedestrian movement
Pedestrian direction
School pupil
Seat belt use
Car passenger
PSV passenger
School number
Casualty surname
Casualty forename
Casualty address

From the hospital the fields are:

Accident and Emergency number
Date of accident
Accident location
User group
Vehicle type
Seat belt/helmet use
Journey purpose
Police informed
Hospital number
Date arrival
Time arrival
Casualty surname
Casualty forename
Date of birth
Casualty age
Casualty gender
Casualty address
Casualty postcode
Casualty telephone number
Complaint
religion
GP code

GP name
GP address
Employment status
Marital status
Source of referral
Mode of arrival
Place of accident
Attend
Examination room/cubicle
At risk number
Receptionist code
Time see doctor
Time discharged
Doctor
Admit
Ward
Consultant
Clinic code
Clinic date
Clinic time
Discharge
Mode discharge
Diagnosis 1
Attendant type
Coded diagnosis 1
Coded anatomical site 1
Diagnosis 2
Coded diagnosis 2
Coded anatomical site 2
X-ray
Treatment
Drugs
Tetanus
Self certificate
Med3
Private certificate
no certificate

wptnlist\wp369.kpa