Reporting of Road Traffic Accidents in London: Matching Police STATS19 Data with Hospital Accident and Emergency Department Data

by

Heather Ward and Sandy Robertson (UCL)

Tiffany Lester and Annette Pedler (TRL Limited)

PR/T/069/2002

TRL Limited

PROJECT REPORT PR/T/069/2002

REPORTING OF ROAD TRAFFIC ACCIDENTS IN LONDON: MATCHING POLICE STATS19 DATA WITH HOSPITAL ACCIDENT AND EMERGENCY DEPARTMENT DATA

by **H Ward and Sandy Robertson (UCL)**
T Lester and A Pedler (TRL Limited)

Prepared for: **Project Record:** **Matching of Road Traffic Accidents**
Client: **London Accident Analysis Unit (LAAU)**
Transport for London

	Approvals
Project Manager	
Quality Reviewed	

Contents

Executive Summary

Objectives

The objective of the project is to estimate the level of reporting to the Police of accidents in London in order to provide:

- Better understanding of the extent of the accident and casualty problem.
- Better and more robust information on the reporting of the severity of injuries.
- A base-line against which the effect of policy initiatives can be properly assessed, given that the level of reporting might be influenced and change over time. Such initiatives might include publicity campaigns, road safety engineering programmes, and congestion charging; and
- To engender better working relationships and co-operation with Health Authorities and Hospitals.

The reporting rate

If all road traffic injuries were reported to the police the STATS19 records, (the official national accident report STATS19 of personal injury accidents that occur on the public highway), would be an accurate estimate of risk. Not all accidents are reported because there are people who either do not know they should report injury accidents or, for other reasons, decide not to do so. However, there are certain circumstances in which the accident does not need to be reported (e.g. if correct documents are produced to an authorised person at the time of the accident, or details exchanged between involved parties, even if there is personal injury involved).

One underlying hypothesis of the work to match police and hospital records is that if a person is injured in a road traffic accident, it is unlikely that their decision to attend a hospital for treatment would be affected by knowledge or experience of road safety or other traffic management initiatives.

However, such initiatives may influence their decision to report their accident to the police, and earlier work, (Gloucester Safety City Project), is indicating that the reporting rates can increase relative to "control" areas. The consequence of this is that expected safety improvements can be apparently lost due to higher reporting rates and the benefit of safety programmes under-estimated.

The alternatives to attending hospital after a road accident are:

- people can be treated at the roadside
- people can treat their injuries at home
- a visit to a GP, who may treat the injury at the surgery or send to a minor accident unit

There may, therefore, be injured people who appear in the police records who do not appear in the hospital records.

We also do not have information on people who are injured and do not go to hospital nor report their accident to the police. This is a group whose size and composition cannot be ascertained in this study.

In this project we used the records of the A&E Departments of three representative hospitals (outer, inner and central London) to supplement those of the police to allow us to estimate of the

proportion of people injured who report their injury accident to the police. From these two sets of records we can estimate the reporting rate to the police and at the same time improve our estimate of the level of injury in London.

The reporting rate used in this study can be described as **all casualties known to the police divided** by **all known casualties**. Previous work has indicated that the levels of reporting in free-standing English towns is of the order of 52-60%.

How the reporting rate was estimated

Electronic records were requested from three London hospitals out of a total of 31 hospitals with full time A&E Departments. They were chosen to represent Central (St Mary's Hospital, Paddington), Inner (King's College Hospital, Camberwell)and Outer London (Barnet General Hospital, Barnet).

A&E records do not have Ordnance Survey Grid References as do the Police STATS19 records and the English language description of the location is not precise, e.g. 'Oxford Street', 'Edgware Road'. To enable us to define sets of police and hospital data from comparable areas around each hospital, London Ambulance Service information was used to guide the study team as to where to draw an artificial boundary around each hospital. The procedure for this is described in Section 2.3 of the report. This allowed TfL to provide police STATS19 data for a defined polygon around each hospital and for the study team to manually assign each hospital record to one of four categories:

1. Definitely in the area
2. Definitely not in the area
3. On a road that passes through the area and continues outside it, or on a road that starts (or finishes) in the area and crosses the boundary
4. Unknown location (no location information)

This is shown schematically in Figure 9 on p17 `

Because the hospital location information is imprecise, is was not possible to locate where casualties were injured on roads that entered the area. This is category 3 above. A procedure for assigning such casualties to within or outside of the area was developed and is described in detail in Section 2.5 and Appendix 2. Because of uncertainties about location three estimates of the reporting rate were produced according to a matching procedure described in Section 3.

1) Hospital data for inside the area only.
2) All hospital data for inside and cross border casualties
3) Hospital data for inside and adjusted unmatched cross border casualties

The first, using hospital data for inside the area only will give the highest reporting rate and is obviously an overestimation because only a proportion of the 'true' hospital dataset is being used. This is shown in the tables in the column headed **Upper Estimate**.

The second set of estimates gives the lowest reporting rate and is an underestimation because all the hospital data that lie on boundary roads is included in the 'true' dataset, and in reality only some of these will be in the area. This is shown in the tables as the **Lower Estimate.**

The **Best Estimate** of reporting rate lies between these two figures and is given in the final column of the following table. Using our adjustment described in Section 2 we have attempted to guide the reader approximately as to where this reporting rate might lie within the range. Until further work is done on refining this process we consider it to be our best estimate at the moment.

The table below gives these lower, upper and best estimates for each hospital and for the major road user groups

The overall headline rate

The overall rates shown are considerably higher than those observed in previous studies in free-standing towns, which range between 50% and 60% . The rates estimated for the areas surrounding King's College Hospital (Inner London) and Barnet General Hospital (Outer London) are similar at 71% and 70% respectively of all known casualties being reported to the police. St Mary's Hospital (Central London) is considerably higher at 87%. The study team are not confident that this represents the reporting rate in Central London and would like to reserve judgement at the moment until further investigations are carried out. One issue is that there are, unusually, more casualties known to the police than to the hospital, bringing us towards the idea that the St Mary's data may not be complete or that casualties may have been taken to other central London hospitals.

Different road users groups

The level of reporting of pedestrian injuries is in line with previous studies, with a best estimate of about 70%. The rate for pedal cycles at about 66-70% is also in line with other studies. The reporting rate for two wheeled motor vehicles (TWMVs) is unusually high, but in London there are many couriers and others who use their motor cycles and scooters for work purposes. The rate for car occupants is also higher than reported elsewhere and a similar hypothesis may be put forward, i.e. that it is the high proportion of business users, coupled with a strong police presence in London that contribute to the higher reporting rates. A high motivation to report injuries for insurance and organisational reasons coupled with fast response times for police in London, may explain the higher reporting rates measured in this study.

Age Groups

Not shown in the table below, but calculated and shown in Section 4 of the report were rates for different age groups. These were close to the average reporting rate of 70% but with the reporting rate for children aged 0-15 years in Barnet being considerably below this at 58%. In none of the hospital areas was there a difference in the reporting rates between males and females, (all ages taken together).

Estimates of the headline reporting rates to the police for each hospital area together with percentage of records matched.

Rate=All casualties known to police/All casualties.

Road User Class	Hospital	Area in London	Lower estimate [1] (%)	Upper estimate [2] (%)	Current best estimate [3] (%)
All casualties	King's	Inner	66	80	71
All casualties	Barnet	Outer	61	75	70
All casualties	St Mary's	Central	77	91	87
Pedestrian casualties	King's	Inner	62	78	67
Pedestrian casualties	Barnet	Outer	72	78	77
Pedestrian casualties	St Mary's	Central	66	84	78
Pedal cyclist casualties	King's	Inner	61	75	66
Pedal cyclist casualties	Barnet	Outer	60	75	69
Pedal cyclist casualties	St Mary's	Central	89	98	96
TWMV [4] casualties	King's	Inner	74	87	78
TWMV [4] casualties	Barnet	Outer	65	80	73
TWMV [4] casualties	St Mary's	Central	74	92	85
Car occupant casualties	King's	Inner	67	78	72
Car occupant casualties	Barnet	Outer	58	73	68
Car occupant casualties	St Mary's	Central	82	92	90

[1] Casualties identified as being in the area only for police data and all those that are in the area and on cross border roads for the hospital

[2] Casualties identified as being in the area only for hospital and police data

[3] Casualties identified as being in the area only for police data and all those that are in the area plus adjusted number for those on cross border roads for the hospital

[4] TWMV - Two Wheeled Motor Vehicle

Severity of injury
In each area the reporting rate for serious injury was lower than for slight injury. Only about two thirds of serious injuries are recorded by the police as such. This could be because police officers being untrained medically systematically underestimate severity of injury especially where internal or head injuries are not immediately obvious to the untrained eye.

To conclude
If we take as a very rough estimate the 70% reporting rate and apply it to the number of casualties reported to the police in 2001 (44,500) we estimate that there may have been about 63,000 injured on the roads of London.

More work is need to refine this figure. More work is also needed to refine the estimates of the number of hospital casualties injured on roads in the cordon area given the imprecise nature of the location descriptions at the hospitals.

This study has estimated the number of people injured on the roads around three representative hospitals in London. It has provided an estimate of reporting in London and gives TfL a series of baseline reporting rates for different classes of road user or severity against which to measure the effect of future initiatives on casualty occurrence.

The reporting rate is not a precise figure and the values for Central London need to be treated with caution until a better estimate is available.

1. Background and objectives.

1.1 Objectives

The objective of the project is to estimate the level of reporting to the police of accidents in London in order to provide:

- Better understanding of the extent of the accident and casualty problem.
- Better and more robust information on the reporting of the severity of injuries.
- A base-line against which the effect of policy initiatives can be properly assessed, given that the level of reporting might be influenced and change over time. Such initiatives might include publicity campaigns, road safety engineering programmes, and congestion charging; and
- To engender better working relationships and co-operation with Health Authorities and Hospitals.

If all road traffic injuries were reported to the police the STATS19 records (the official national accident report STATS19 of personal injury accidents that occur on the public highway) would be an accurate estimate of risk. Not all accidents are reported because there are people who either do not know they should report injury accidents or, for other reasons, decide not to do so. However, there are certain circumstances in which the accident does not need to be reported (e.g. if correct documents are produced to an authorised person at the time of the accident or details exchanged between involved parties, even if there is personal injury accident involved).

If a road user is injured in a road traffic accident they often seek medical attention for their injury. If it severe, or suspected to be severe, then the casualty is most often taken by ambulance to the nearest hospital Accident and Emergency Department (A&E). If the injury is less severe the casualty may visit the hospital A&E Department themselves or their GP who may refer them to the hospital. Minor injuries can also be treated at the roadside.

One underlying hypothesis of the work to match police and hospital records is that if a person is injured in a road traffic accident it is unlikely that their decision to attend a hospital for treatment would be affected by knowledge or experience of road safety or other traffic management initiatives. However, such initiatives may influence their decision to report their accident to the police.

There are people who may treat their injuries at home or visit the GP, who may treat the injury at the surgery or send casualty to a minor accident unit. Therefore, there may be injured people who appear in the police record who do not appear in the hospital record. These people cannot be included in this exercise, nor can we estimate their numbers, because there is no way at present of gathering this information.

There is another group about whom we do not have information. These are people who are injured and do not go to hospital nor report their accident to the police. Similarly, this is a group whose size and composition cannot be ascertained in this study.

We can use the records of the A&E Department to supplement those of the police to allow us to estimate of the proportion of people injured who report their injury accident to the police. From these two sets of records we can estimate the reporting rate to the police and at the same time

improve our estimate of the level of injury in London. The reporting rate is explained in more detail in section 1.2.

Because the STATS19 data only includes personal injury accidents that occur on the public highway. These exclude injuries that occur in other places such as car parks. Casualties occurring in such places had, therefore to be excluded from the hospital data.

Consideration was given to trying to use London Ambulance Service (LAS) data to assist the matching process. The LAS collect geo-coded information on where a casualty was collected and to which hospital they were delivered plus a field for injury type. Whilst using LAS data would have given more accurate accident location information it was another layer of matching on an already complex task. For this reason its use at this stage was precluded

As a consequence of estimating the reporting rate we may be able to use this information to assess the effectiveness of new road safety or other policy initiatives. When road safety programmes, or other policy initiatives are being implemented, awareness of road safety issues can be increased and with it may come the tendency to report more accidents to the police than before. If only police records were used and the public did report a higher proportion of their accidents, then reporting differences would be apparent in the data and conventional before and after comparisons could underestimate the casualty reduction effect of the project or policy. The matching process allows an estimate to be made of this reporting difference.

1.2 The reporting rate

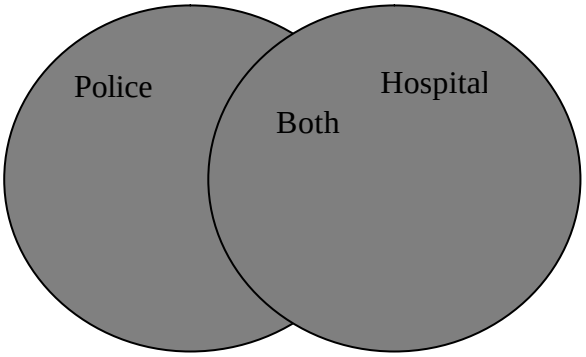
The method of calculation of reporting rate may vary between different studies. Here the reporting rate to the police is calculated as follows:-

(p+b)/(p+h+b)

Where
p= number of casualties reported to the police only
b= number of casualties reported to both police and hospital
h= number of casualties reported to the hospital only

In other words **all casualties known to the police** (only known to the police plus known to both) **divided by all known casualties**. This is shown schematically below in Figure 1. The reporting rate is the shaded areas divided by the total area.

Figure 1. The overlap between police and hospital casualty records.



1.3 Previous studies

A number of studies have looked at reporting rates but in many cases (e.g. Simpson 1996), the focus has been from the perspective of hospital data and the proportion of casualties (or even admitted casualties) that were known to the police together with comparisons of recorded severity. With studies that focus on the hospital's perspective a direct comparison to the data in this report is not appropriate. In some cases the definition of under reporting is not made explicit so that there is a degree of uncertainty about what the reported level of under reporting actually means. James (1991) noted overall reporting rates from a range of studies, (the basis upon which the reporting rate was estimated was not made explicit), these ranged from 50 to 72 per cent.

In a study, comparable with the one reported here, Ward, Cave, Morrison, Allsop and Evans (1994) found that for pedestrians in Northampton, under reporting was 24% (i.e. reporting rate was 76%).

Recorded numbers of road traffic casualties in the Gloucester Safer City project area were analysed for the years 1996 to 2000. Casualties recorded by the police and casualties recorded by the Gloucester Royal Infirmary were matched to identify those known only to the police, known only to the hospital and known to both. Casualty numbers and reporting rates were analysed. A similar analysis was undertaken for the control town of Cheltenham for the years 1996, 1998 and 2000. This was to establish whether underlying reporting rates had changed across the region associated with the implementation of an area wide safety initiative. Ward and Robertson (2002)

In Gloucester, the indications are that there was a statistically significant increase in the number of people reporting their injuries to the police over the five-year period 1996-2000. However, the number of people attending the hospital and admitting to being injured in a road traffic accident appears to be unchanged. This leads towards the conclusion that the number reporting their injuries to the police increased substantially. The indications are that were no changes to the reporting rate for casualties in the control town of Cheltenham.

In general, there were no significant changes for the different road user types with the exception of pedal cycles where reporting to the police increased and the number of casualties reported to the hospital decreased. It may be tentatively concluded that pedal cycle casualties have decreased (given that there is not a plausible mechanism to induce pedal cycle riders to under-report at the hospital).

There were no significant differences in reported casualties over the period for different age bands with the exception of the 25-34 and the 35-44 year age groups. It was concluded that there may be an increase in reporting to the police for the former age group.

A comparison of the reported severity of casualties and hospital disposal codes indicated that there were substantial numbers of casualties that while recorded as slight by the police had hospital disposal codes that were consistent with the police definition of a severe injury.

This study of reporting rates in London will estimate the overall rate, those for each of the major road user and groups, males and females, and finally will consider differences in estimates of severity for the police and hospitals.

2. Selection of hospitals for inclusion in matching exercise

2.1 How the hospitals in the study were selected.

Ideally we would like to have a complete coverage of hospitals in London but in practice this is not possible given that there are 31 hospitals with full time A&E Departments each with between 1000 and 2000 road traffic accident casualty records annually. To make the job tractable, it was decided to sample three hospitals as being broadly representative. One each from :

- Central London
- Inner London
- Outer London

Hospital Trusts do not have a unified system for data recording and several systems operate within London. We visited Ruth Brown, the chair of the London A&E Consultants Group and her advice was to contact hospitals that used the Footman-Walker database for A&E departments. This is a system developed by the software firm Footman-Walker Associates which records a minimum dataset that usually includes:

- date of attendance,
- age,
- sex,
- geographic location of accident, and
- disposal which can give us an idea of severity.

Disposal can include, admitted, referred to fracture clinic, died on ward, transferred to other hospital, did not wait, no further/GP follow up. Representative lists of disposal codes are given in Appendix 3.

There are about eight hospitals in London using this system. Ruth Brown sounded out the consultants at a meeting in January and provided details of the following who said they might be prepared to help. The following were contacted:

St Mary's Hospital, Paddington
Royal London Hospital, Whitechapel (from 1 April 2002)
Mayday Hospital, Croydon
King's College Hospital, Camberwell
Barnet General Hospital
Bromley Hospital

The Mayday Hospital does not record information about the location of the accident and the Bromley Hospital didn't respond to our request. The Royal London is a busy hospital located in the heart of Tower Hamlets and was about to move over to a fully automated Footman-Walker system in April 2002 and has no historical data in a format we could easily use.

The three hospitals we chose were

- St Mary's Hospital, Paddington – Central London
- King's College Hospital, Camberwell – Inner London
- Barnet General Hospital – Outer London

2.2 Composition of data files from hospitals

This section briefly describes the data received from the three hospitals. The numbers received in each dataset are shown in Table 1.

Table 1: Number of casualties in datasets in 2001

Hospital	Barnet	St. Mary's	King's
Casualties	1739	929	1832

The format of hospital datasets are shown below.

Barnet	St Mary's	King's
Episode No.		
Disposal / Ward	Disposal / Ward	Disposal / Ward
Date & Time Reg	Date & Time Reg	Date & Time Reg
Triage Category		
Date of Birth		
Date & Time of Incident	Date & Time of Incident	Date & Time of Incident
Complaint		
Road User Category	Road User Category	Road User Category
Place of Incident	Place of Incident	Place of Incident
Referral Source		
Time since Incident		
Sex	Sex	Sex
	Age	Age

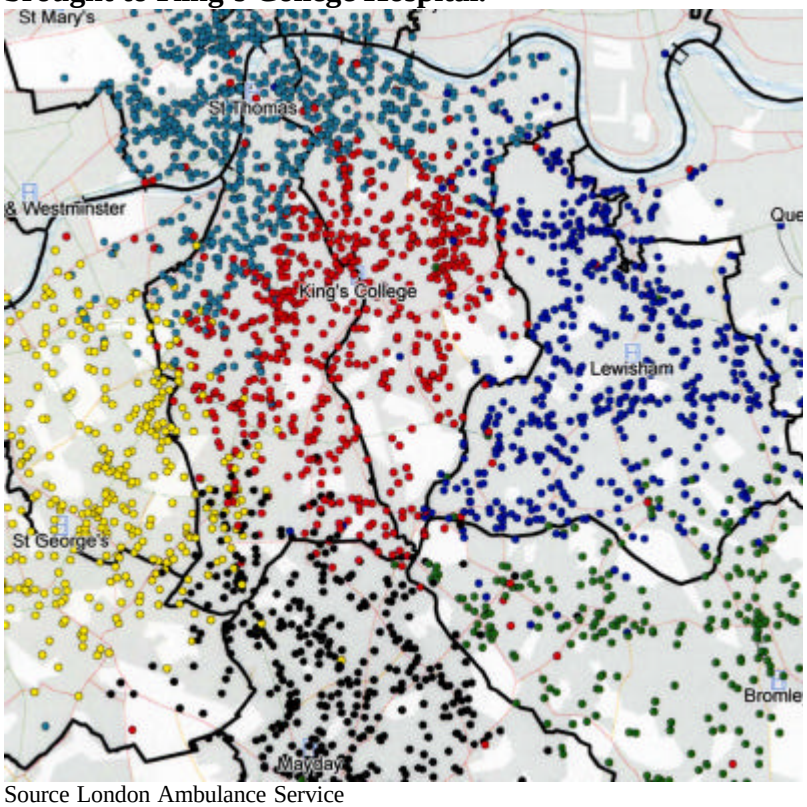
2.3 Use of London Ambulance Service data to select cordon around hospitals

One issue regarding the hospital data relates to which records to match with the police STATS19 data . The police data includes 10 digit grid references for the location of every accident but the hospital data only records a road name, e.g. "Oxford Street", "Marylebone Road", "Edgware Road". Whereas we can define an area using a grid-referenced polygon for the police data we cannot do this for the hospital data. Also there is no absolute boundary for each hospital where a person might be expected to be taken by ambulance, or go to on their own, in the event of an injury. The norm is for the London Ambulance Service (LAS) to take the casualty to the nearest hospital but in times of severe congestion sometimes they are taken to the nearest by time and not by distance. Also, if the casualty has suffered a head injury and they are within a reasonable distance of a hospital with a specialist head injuries unit, they may be taken there and not to the nearest hospital. King's College Hospital is one such hospital with a head injuries unit.

The task was to draw a cordon around each hospital in order to create an area within which we could identify casualty records for linkage and matching.

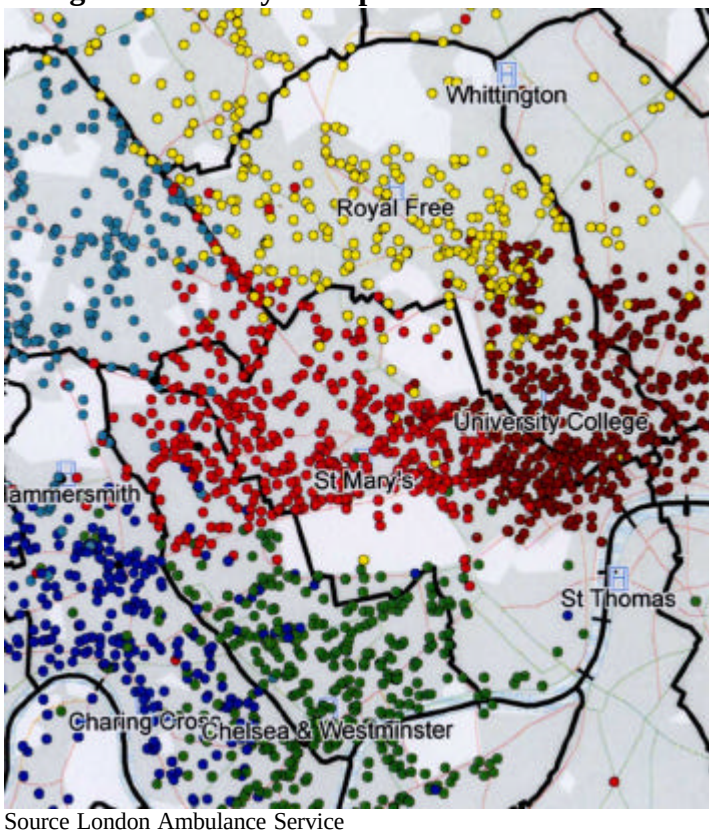
As can be seen from Figures 2,3 and 4, provided by LAS, there is a considerable overlap in where casualties are taken. To try to minimise overlaps with other hospitals a cordon was drawn that excluded those casualties in areas where there were casualties taken to other nearby hospitals.

Figure 2: Locations of road traffic accidents occurring in 2001 from which casualties were brought to King's College Hospital.



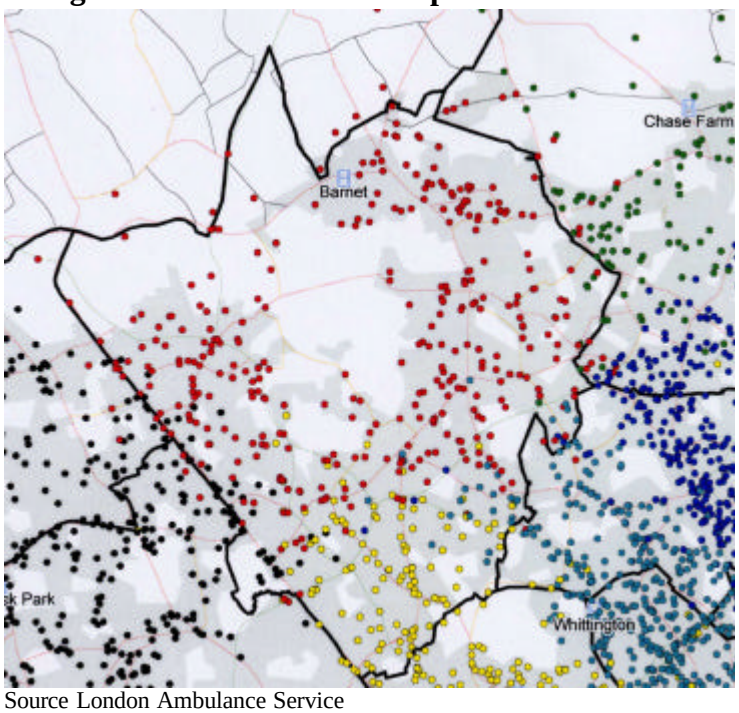
Source London Ambulance Service

Figure 3: Locations of road traffic accidents occurring in 2001 from which casualties were brought to St. Mary's Hospital.



Source London Ambulance Service

Figure 4: Locations of road traffic accidents occurring in 2001 from which casualties were brought to Barnet General Hospital.



Source London Ambulance Service

The boundaries agreed between UCL and TfL are shown in Figures 5, 6, 7 and 8 (the grid references are given in Appendix 1). The approximate area bounded by each cordon is as follows:

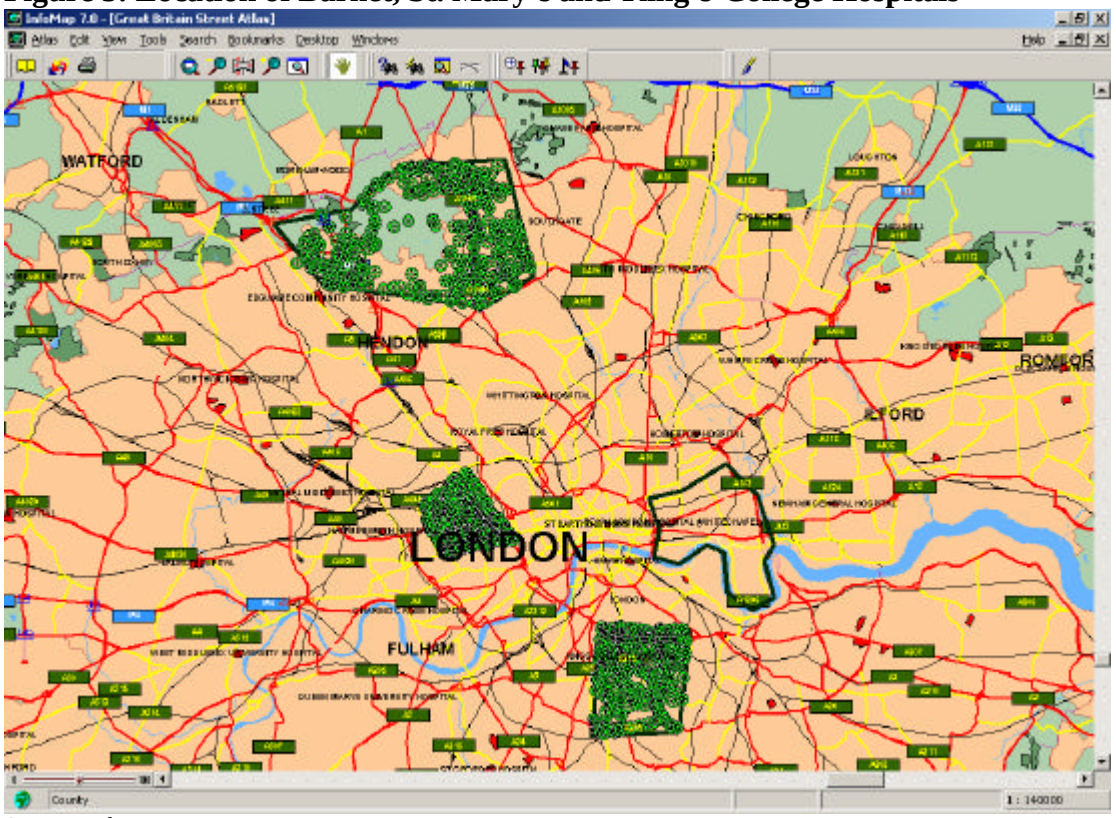
- St. Mary's Hospital 7.9 Km sq
- King's College Hospital 16.2 Km sq
- Barnet General Hospital 44.3Km sq

A software package called InfoMap was used and the following Figures come from this system. Figure 5 shows all the hospital areas.

The sample areas are outlined in black and are shown below for London. Each area is shown separately and the associated grid references are given in Appendix 1.

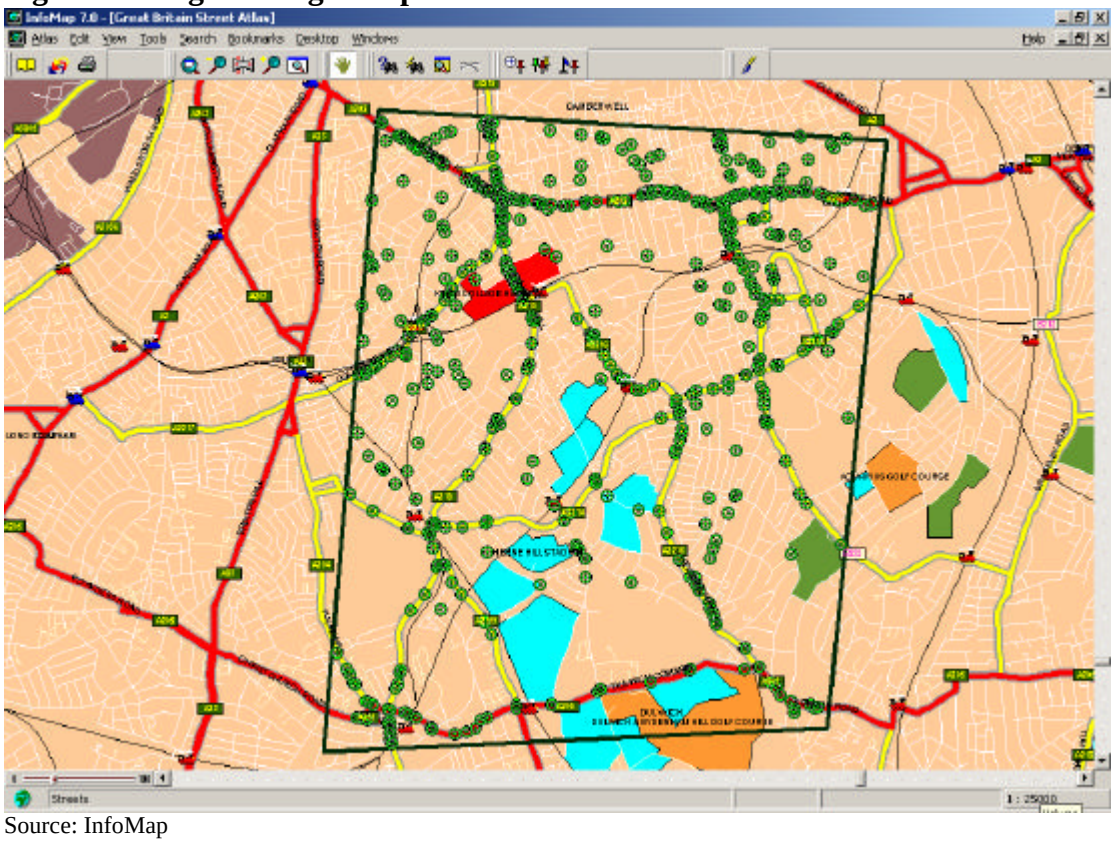
For the purposes of drawing the boundaries on the map the interface with the software only permits an approximate grid reference to be obtained so there may be a slight mismatch between the given grid references and the plotted area on the map.

Figure 5: Location of Barnet, St. Mary's and King's College Hospitals



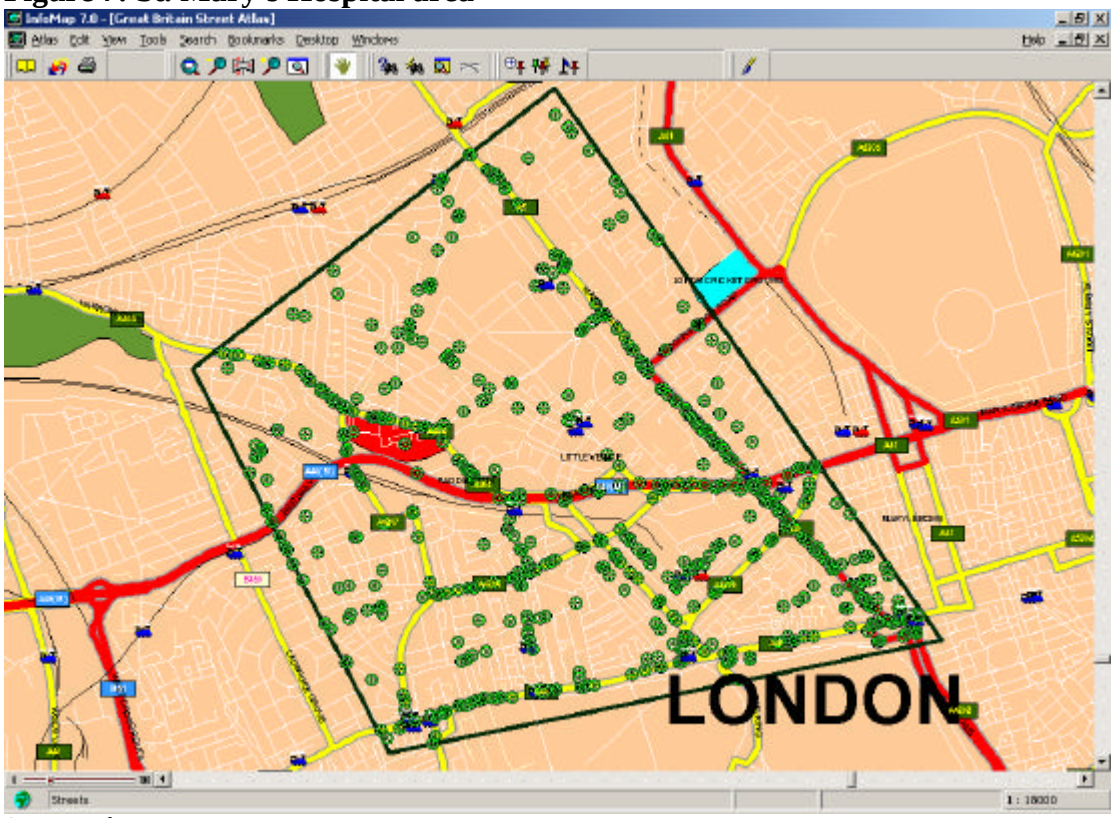
Source: InfoMap

Figure 6: King's College Hospital area



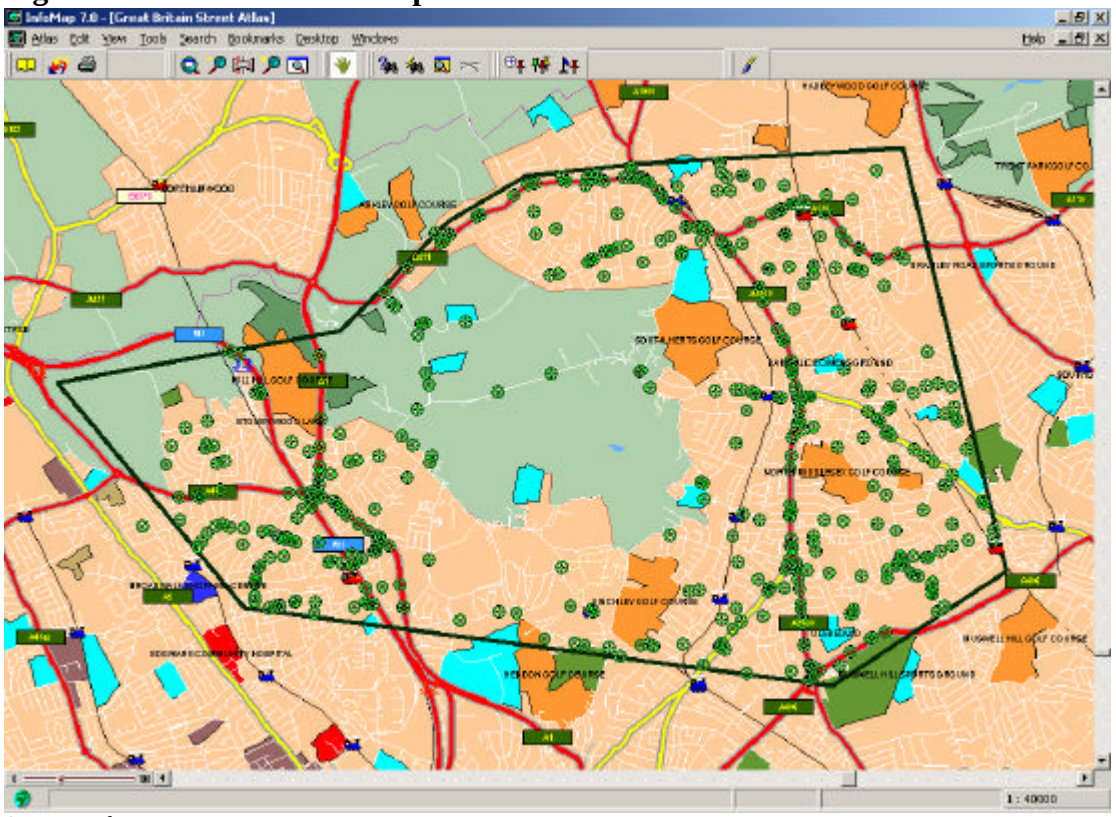
Source: InfoMap

Figure 7: St. Mary's Hospital area



Source: InfoMap

Figure 8: Barnet General Hospital area



Source: InfoMap

2.4 Police STATS19 data for each hospital area

Police STATS19 data were requested from TfL for the polygons represented by the grid references given in Appendix 1 for each Figure above.

Table 2: Number of Police casualty records in 2001 for each hospital area.

Hospital	Barnet	St. Mary's	King's College
Police data	665	862	837

2.5 Sorting hospital records according to location in or outside the cordon

Sorting the hospital data is made difficult because only vague descriptions of location are given. Using InfoMap a list was extracted of all roads lying within the cordon and each of the hospital records was checked manually using the location information. This resulted in the records being partitioned into four sets:

1 Definitely in the area
2 Definitely not in the area
3 On a road that passes through the area and continues outside it, or on a road that starts (or finishes) in the area and crosses the boundary
4 Unknown location (no location information)

2.5.1 What do we do about casualties on roads running through the area and out the other side?

Figure 9 illustrates the issues associated with set 3 in the list. Here we do not know whether the casualty was injured on the part of the road that is in the area or the part of the road that is outside the area. The size of these 'boundary' datasets will differ according to the road characteristics of the cordon area. This has implications for the size of the hospital dataset and hence the denominator used in the calculation of reporting rate.

By definition, all of the police records used are within the boundary, (since we are assuming that the grid reference is correct), so these issues do not arise here.

If we use only the data that are known to be in the area for matching with Police data we will be 'discarding' hospital casualties that lie on the roads passing through the area even though we cannot identify these records individually. We would therefore overestimate the reporting rate. See Figure 9.

However, if we include all the boundary data as being within the hospital set we will be overestimating the size of the hospital dataset and thereby underestimate the reporting rate.

The actual number of casualties known only to the hospital will lie somewhere between the values described above. In this report we have included the high and the low estimates of reporting rate together with our best estimate of the actual value.

For the purposes of this report, we have made an estimate of the proportion of casualties that are known to the hospital only (unmatched hospital data in the boundary area) on roads which cross the boundary of the study area. For details of the estimation process, see Appendix 2.

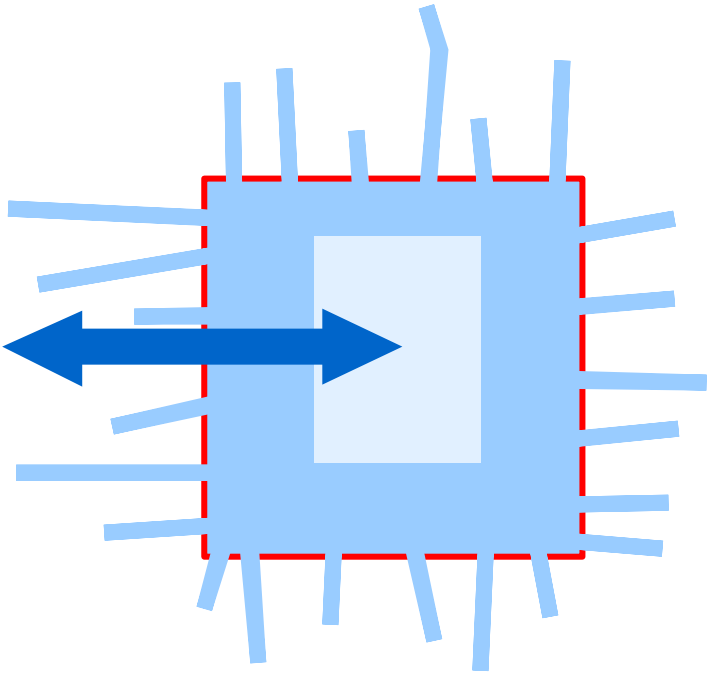
Figure 9. Schematic diagram of roads within and roads crossing the boundary to the study area.

Hospital data for casualties that occurred on roads with their entire length inside the catchment

Hospital data for casualties that occurred on roads with some length inside the catchment and some length outside the catchment

Hospital data for casualties that occurred at some unknown or unspecified location

Boundary of the catchment that all the police data lies within



This estimate is calculated separately for each hospital area and is based on the number of matched casualties known to be in the study area divided by the number of casualties in the hospital dataset that are known to be wholly within the study area. Then the number of casualties in the cross border dataset are divided by this proportion to give the total number of casualties we might expect to have occurred within the study area on the roads that cross the borders of the study area. To get to our final estimated number we then have to remove the casualties that are deemed to be outside the boundary, but as we don't know which these are we exclude, and put to one side, the ones that already match then remove the others according to a random selection process.

This estimate is the best that can be provided within the scope of this report and in terms of the time available for the analysis. We have made the assumption that the reporting and other casualty characteristics for casualties occurring within the study area are the same as those occurring on the roads crossing the boundaries of the study areas. This assumption is not very robust as the characteristics of roads that cross the study area boundaries can be markedly different from those within the study area in terms of size, traffic flow, composition and numbers of casualties. An example of this would be the M1 in the Barnet study area. On the other hand,

there are some roads which cross the study area that would be more typical of the roads found entirely within the study area.

2.5.2 Hospital data where accident location was not recorded.

A similar approach could be applied to casualties known to the hospital only where there is no record of location at all. Time constraints did not permit a similar treatment at this stage. There were not many of these and as there were few matches for casualties where the hospital location was unknown it was concluded that the difference in the resulting estimate of numbers would be small.

3. The data-matching process

3.1 The matching process

The matching process was undertaken in two stages with an automated and a manual component. The police data and hospital data (having been pre-processed to a consistent format) were placed on the same page of a spreadsheet which allowed automated and visual comparison of the two datasets. This section will describe the datasets used, and the matching process.

3.2 The combined dataset

The data from the hospitals and from the police were converted to a compatible format for the matching exercise. This included parsing of date and time data to a form that could be used in a spreadsheet. These are shown below where 'p' denotes police and 'h' hospital.

Police data was converted to:

Refnum	Sex(p)	Age(p)	Cas Class(p)	Severity(p)	Date	Police =1	Time	East	North	Location	Postcode	Match	Date

Hospital data was converted to:

Date	Age(h)	Sex(h)	Class(h)	Severity(h)	Hospital=2	RTA Date	RTA Time	Location of Accident	In area	Ref number

There are two date fields in each of the datasets. The date field at the end of the police and the one at the beginning of the hospital data share the same spreadsheet column and for matched data become a common field. A further set of fields converted the police and hospital codes to a common set of codes that were used in the analyses. These fields were:

P/H/B	Age(hp)	Severity(ph)	Class(ph)	Year code	sex(hp)

The field P/H/B identifies whether the line of data is known to police only, to hospital only or to both. The year code is to allow year by year comparison in the future.

3.3 Data coding

Data coding was key to the success of the work. The data codings are shown in Appendix 4. Some of the key data coding issues are described below.

3.3.1 Severity

The severity coding used was the STATS19 levels of fatal, seriously injured and slightly injured. The hospital descriptions of severity were in the form of disposal codes or outcomes and these were mapped onto the STATS19 severity codes. Each hospital had a different set of disposal codes (or discharge outcomes) which were individually mapped on to the severity codes mapped. These can be found in Appendix 3.

3.3.2 Codes for vehicle class

The codings used for the class of a road user casualty by the police and by the hospital differed, so a set of categories that encompassed both coding systems was required. The following categories were used:-

Pedestrian
Pedal cycle
Two Wheel Motor Vehicle (TWMV)
Car, bus, Heavy Goods Vehicle (HGV), Light GoodsVehicle (LGV), etc
Other (horse etc)
Unknown

Police user class data were taken as definitive for tables. It was assumed that the vehicle type was more likely to be correctly identified by police at the scene of the crash than by hospital A&E staff. The hospital user codes are shown in Appendix 4.

3.3.3 Age band codes

Hospital age data were taken as definitive for tables. Where these were not available the police age data were used. The following age bands were used

0 -15
16-24
25-59
60+
Unknown

3.3.4 Codes for area

The locations of the hospital casualties were coded in terms of whether they were on roads that were within the study area. These data were used to allow selection for data for different analyses.
The categories were:-

On roads wholly within the study area.
On roads that were partially within the study area.
On roads whose location was unknown.
On roads outside the study area.

3.3.5 Codes for date and time

For the hospital data, both the time of registration at the hospital and the time of the incident were recorded. Where the incident date was missing it was replaced by the date of registry at the hospital. Where the incident time was missing in the hospital data this was manually set to 00:12 . This allowed unknown times to be identified in the manual matching process.

3.4 The matching algorithm

Experiments were undertaken using 1 month of data (121 entries) to determine the effectiveness of various algorithms. 5 fields were used in the algorithm:

Date of incident
Date and time of incident (allowed incidents spanning midnight to be included)
Age
Gender
User class

A number of runs using different combinations of mandatory matches were undertaken and the total number of matches recorded as shown in the following table.

Number of matching variables	**Mandatory matching variables**	**Matches**	**Of which false positives**
5	All	12	0
4	Date	24	5 False 2 uncertain
4	Date & age	20	2 uncertain
4	Date & gender	20	2 uncertain
4	Date age & class	20	2 uncertain
4	Date & time	12	0
4	Date or time, and age & gender	20	2 uncertain

Given the high proportion of false positives obtained where 4/5 variables with same date were used it was decided not to investigate the effect of 3 matches. The algorithm finally selected for the auto matching was at least 4 of the fields matching with the following fields mandatory.

- age
- gender
- date or time

3.5 Automated matching

The data sheet consisted of two sets of columns of data one for the police and one for the hospital data with one common column for the date field. When two lines of data matched they were placed in a single line. The matching process moved lines of hospital data to the same line as the matching police data. Codes for age, gender and whether a given line contained data from police, hospital or both etc were generated in further fields from which pivot tables were used to produce results tables.

The automated matching process used an algorithm written in visual basic for an application running within an Excel spreadsheet. The automated matching process was run using all police and all hospital data including those hospital data that had been provisionally identified as being outside the study area. This allowed a double check on the exclusion of hospital data. The automated matching process did not take account of the plain English description of locations.

Codes for matches were inserted in the 'matches' field on the combined data page. A value was inserted to indicate a match. A blank was left to indicate no match. In the auto sort routine, a

number was also inserted where additional lines also matched the specific police data. These were used to help inform the manual matching process.

- auto match where 4 variables matched: - (date or time) and (age and gender)
- auto match where 5 variables matched: - (date, time, age, gender, class)

The matching process was run on all police and all hospital data supplied (this included data which had been provisionally identified as outside the area so as to provide a double check on location data).

3.6 Manual matching

After the automated matching process, a manual matching process was undertaken on the output of the automated matching process.

For the manual matching process there was an element of qualitative judgement applied to the matches based on the experience of the people undertaking the matching process. Possible additional matches and false matches were identified by one individual and the decisions on whether to include these were made by two or, in difficult cases, by three individuals.

The fields used in the manual matching process were:-

Date
Time
Age
Gender
User class
Location (plain English)

In some cases the following codes were included in the matched spreadsheet

Code 6 = Manual match where the matchers had high confidence.
Code 7 = Manual match where the matchers had less than high confidence.

This process was undertaken on the matched datasheet. The datasheet used earlier to identify the location was not updated. Where matches were made on data that had been previously been identified as being outside the area, a careful examination of location information was undertaken. In some cases additional data that were within the area were identified. Matching lines of data were moved manually within the spreadsheet.

The manual matching required some qualitative assessment. In a number of cases the original data had to be revisited especially where there were inconsistencies of date. An example would be where the date of incident was recorded by the hospital as being on a particular day and the date of registration at the hospital was recorded as being on the day before, as was the police record of the accident. Such a combination of factors would imply a match. Other types of inconsistency included misuse of the 12/24 hour clock system so time of incident could sometimes appear to be 12 hours apart.

3.7 Selecting subsets of the data

The original matching exercise was undertaken on the full dataset, but for a number of the analyses a sub set of the data was required, (for example hospital casualties on roads that were entirely from within the study area). Where subsets of the data were required, the unwanted

hospital data was identified by sorting on the "in area" field then removing the unwanted hospital data and saving the spreadsheet under a new name.

4. Estimation of levels of reporting to the police

4.1 Summary of results

For each area the following datasets were matched with the police data as described in Section 3.

1) Hospital data for inside the area only.
2) All hospital data for inside and cross border casualties
3) Hospital data for inside and adjusted unmatched cross border casualties

This gives us three estimates of the reporting rate. The first, using hospital data for inside the area only will give the highest reporting rate and is obviously an overestimation because only a proportion of the 'true' hospital dataset is being used. This is shown in the tables in the column headed **Upper Estimate**. The second set of estimates gives us the lowest reporting rate and is an underestimation because all the hospital data that lie on boundary roads is included in the 'true' dataset, and in reality only some of these will be in the area. This is shown in the tables as the **Lower Estimate.**

The **Best Estimate** of reporting rate lies between these two figures and is given in the final column of the tables. Using our adjustment described in Section 2 we have attempted to guide the reader approximately as to where this reporting rate might lie within the range. Until further work is done on refining this process we consider it to be our best estimate at the moment. Tables 3,4, and 5 summarise the results for each hospital and Table 6 gives the estimates of reporting rates to the police for each hospital area together with percentage of records matched.

Where time permitted we have extracted the data in the 'unknown location', (location data missing) category because we do know from the police data that some of the casualties were injured within the area. We have not presented this data as we are unable to estimate the proportion which might have been so.

Table 3: King's College Hospital -Lower and upper estimates of percentages of casualties reported to the police with our current best estimate of reporting rate for each category where:

Rate= (casualties known to only to police +casualties known to both)/All casualties.

Casualty breakdown	Lower estimate	Upper estimate	Current best estimate
All casualties	66	80	71
Pedestrians	62	78	67
Pedal cyclist	61	75	66
TWMV*	74	87	78
Car occupant	67	78	72
0-15 years	68	79	72
16-24 years	63	77	67
25-59 years	66	79	71
Over 60 years	69	80	75
Males	67	80	71
Females	66	78	71
Fatal	86	86	86
Serious	59	77	65
Slight	68	80	72

Table 4:Barnet General Hospital - Lower and upper estimates of percentages of casualties reported to the police with our current best estimate of reporting rate for each category.

Casualty breakdown	Lower estimate	Upper estimate	Current best estimate
All casualties	61	75	70
Pedestrians	72	78	70
Pedal cyclist	60	75	69
TWMV*	65	80	73
Car occupant	58	73	68
0-15 years	47	65	58
16-24 years	57	71	68
25-59 years	61	75	70
Over 60 years	80	86	83
Males	61	75	71
Females	61	76	70
Fatal	80	80	80
Serious	54	70	63
Slight	62	76	72

*TWMV = Two Wheeled Motor Vehicle

Table 5: St. St. Mary's Hospital -Lower and upper estimates of percentages of casualties reported to the police with our current best estimate of reporting rate for each category.

Casualty breakdown	Lower estimate (%)	Upper estimate (%)	Current best estimate (%)
All casualties	77	91	87
Pedestrians	66	84	78
Pedal cyclist	89	98	96
TWMV*	74	92	85
Car occupant	82	92	90
0-15 years	69	83	77
16-24 years	74	91	86
25-59 years	78	91	87
Over 60 years	78	91	86
Males	78	92	88
Females	75	89	85
Fatal	86	100	86
Serious	65	80	74
Slight	79	93	89

Our first analysis indicated that there may be a problem with the St Mary's data. The data were checked and the hospital contacted to ascertain whether some records had been left out. There was no evidence that we had received anything other than the full dataset. The issues surrounding the St. Mary's data can be summarised as follows:

- The number of casualty records provided by the hospital is about half the number at the other two hospitals, whilst the number of police recorded casualties is about the same as King's College Hospital.

Casualty Source	Barnet	St. Mary's	King's
Hospital	1739	929	1832
Police	665	862	837

- The reporting rates were higher than at the other two hospitals which is consistent with having fewer St Mary's hospital records to match with the police set. There are alternative explanations for these observations which are discussed in Section 5.

However, the percentage of records in the datasets which match is broadly similar across all three hospitals.

Table 6: Estimates of the headline reporting rates to the police for each hospital area together with percentage of records matched.

Casualty match	Lower estimate	Upper estimate	Current best estimate
Estimates of percentages of all casualties reported to the police (reporting rate)			
All casualties King's	66	80	71
All casualties Barnet	61	75	70
All casualties St Mary's	77	91	87
Estimates of percentages of all casualties known to both/all known casualties (matching rate)			
	Roads in Study Area + whole of cross-border roads	**Roads Wholly in Study Area**	**Factored Best Estimate**
All casualties King's	29	22	31
All casualties Barnet	31	30	36
All casualties St Mary's	27	21	31

4.2 King's College Hospital

4.2.1 Hospital data for inside the King's area only (Upper Estimate).

Table 7: Overall numbers and reporting rate

All casualties: King's, 2001

Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
605	214	232	1051	80%

Hospital data for those casualties that occurred on roads that are completely within the study only.

Table 8: By casualty class

Road user Class	Casualties: King's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	120	52	65	237	78%
Pedal cycle user	48	22	17	87	75%
TWMV user	129	26	52	207	87%
Car,van,goods, bus user	307	111	98	516	78%
Other	1	3	0	4	25%
Unknown	0	0	0	0	
All	605	214	232	1051	80%
Hospital data for those casualties that occurred on roads that are completely within the study only.					

Table 9: By Age

Age Band	Casualties by age: King's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	66	26	32	124	79%
16 to 24	95	42	47	184	77%
25 to 59	370	133	138	641	79%
Over 60	37	13	15	65	80%
Unknown	37	0	0	37	100%
All	605	214	232	1051	80%
Hospital data for those casualties that occurred on roads that are completely within the study only.					

Table 10: By Gender

Gender	Casualties by gender: King's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Male	377	125	138	640	81%
Female	228	89	94	411	78%
All	605	214	232	1051	80%
Hospital data for those casualties that occurred on roads that are completely within the study only.					

Table 11: By Severity

Severity	Casualties by severity: King's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Fatal	4	1	2	7	86%
Serious	65	30	38	133	77%
Slight	536	181	192	909	80%
Unknown	0	2	0	2	0%
All	605	214	232	1051	80%
Hospital data for those casualties that occurred on roads that are completely within the study only.					

4.2.2 All King's hospital data for inside and cross border casualties (Lower Estimate).

Table 12: Overall numbers and reporting rate

All casualties: King's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
470	425	367	1262	66%
Hospital data for those casualties that occurred within the area and on all roads that cross the boundaries to the study area				

Table 13: By Class

Road user Class	Casualties: King's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	81	113	104	298	62%
Pedal cycle user	38	42	27	107	61%
TWMV user	101	64	80	245	74%
Car,van, goods, bus user	249	200	156	605	67%
Other	1	4	0	5	20%
Unknown	0	2	0	2	0%
All	470	425	367	1262	66%
Hospital data for those casualties that occurred within the area and on all roads that cross the boundaries to the study area.					

Table 14: By Age

Age Band	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
	Casualties by age: King's, 2001				
0 to 15	41	46	57	144	68%
16 to 24	72	84	69	225	63%
25 to 59	292	269	220	781	66%
Over 60	32	23	20	75	69%
Unknown	33	3	1	37	92%
All	470	425	367	1262	66%
Hospital data for those casualties that occurred within the area and on all roads that cross the boundaries to the study area					

Table 15: By Gender

Gender	Casualties known only To the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
	Casualties by gender: King's, 2001				
Male	294	255	219	768	67%
Female	176	170	148	494	66%
All	470	425	367	1262	66%
Hospital data for those casualties that occurred within the area and on all roads that cross the boundaries to the study area					

Table 16: By Severity

	Casualties known only To the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
	Casualties by severity: King's, 2001				
Fatal	4	1	2	7	86%
Serious	41	72	62	175	59%
Slight	425	347	303	1075	68%
Unknown	0	5	0	5	0%
All	470	425	367	1262	66%
Hospital data for those casualties that occurred within the area and on all roads that cross the boundaries to the study area					

4.2.3 King's Hospital data for inside and estimated unmatched cross border casualties (Best Estimate)

Table 17: Overall numbers and reporting rate

All casualties: King's , 2001

Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
470	339	367	1176	71%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 18: By Class

Casualties: King's, 2001

Road user Class	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	81	92	104	277	67%
Pedal cycle user	38	34	27	99	66%
TWMV user	101	50	80	231	78%
Car, van, goods, bus user	249	158	156	563	72%
Other	1	3	0	4	25%
Unknown	0	2	0	2	0%
All	470	339	367	1176	71%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 19: By Age

Casualties by age: King's, 2001

Age Band	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	41	39	57	137	72%
16 to 24	72	70	69	211	67%
25 to 59	292	211	220	723	71%
Over 60	32	17	20	69	75%
Unknown	33	2	1	36	94%
All	470	339	367	1176	71%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 20: By Gender

Gender	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
	Casualties by gender: King's, 2001				
Male	294	206	219	719	71%
Female	176	133	148	457	71%
All	470	339	367	1176	71%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 21: By Severity

	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
	Casualties by severity: King's, 2001				
Fatal	4	1	2	7	86%
Serious	41	55	62	158	65%
Slight	425	279	303	1007	72%
Unknown	0	4	0	4	0%
All	470	339	367	1176	71%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

4.3 Barnet General Hospital

4.3.1 Hospital data for inside the Barnet area only (Upper Estimate).

Table 22: Overall numbers and reporting rate

All casualties: Barnet, 2001				
Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
397	220	268	885	75%
Hospital data are for those casualties that occurred on roads wholly within the study area only				

Table 23: By Class

	Casualties: Barnet, 2001				
Road user Class	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	53	34	68	155	78%
Pedal cycle user	19	9	8	36	75%
TWMV user	50	20	29	99	80%
Car, van, goods, bus user	272	157	163	592	73%
Other	3	0	0	3	100%
All	397	220	268	885	75%
Hospital data are for those casualties that occurred on roads within the study area only					

Table 24: By Age

	Casualties by age: Barnet, 2001				
Age Band	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	29	34	35	98	65%
16 to 24	90	63	65	218	71%
25 to 59	199	108	130	437	75%
Over 60	53	15	37	105	86%
Unknown	26	0	1	27	100%
All	397	220	268	885	75%
Hospital data are for those casualties that occurred on roads within the study area only					

Table 25: By Gender

Gender	Casualties by gender: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Male	229	128	152	509	75%
Female	168	92	116	376	76%
All	397	220	268	885	75%
Hospital data are for those casualties that occurred on roads within the study area only					

Table 26: By Severity

Severity	Casualties by severity: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Fatal	3	1	1	5	80%
Serious	35	42	61	138	70%
Slight	359	175	206	740	76%
Unknown	0	2	0	2	0%
All	397	220	268	885	75%
Hospital data are for those casualties that occurred on roads within the study area only					

4.3.2 All Barnet hospital data for inside and cross border casualties (Lower Estimate)

Table 27: Overall numbers and reporting rate

All casualties: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
324	429	341	1094	61%
Hospital data are for those casualties that occurred on roads within the study area and all casualties on roads that cross the boundaries to the study area.				

Table 28: By Class

Road user Class	Casualties: by class of road user Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	44	47	77	168	72%
Pedal cycle user	15	18	12	45	60%
TWMV user	28	43	51	122	65%
Car,van,goods, bus user	234	321	201	756	58%
Other	3	0	0	3	100%
All	324	429	341	1094	61%

Hospital data are for those casualties that occurred on roads within the study area and all casualties on roads that cross the boundaries to the study area.

Table 29: By Age

Age Band	Casualties by age: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	24	72	40	136	47%
16 to 24	73	120	83	276	57%
25 to 59	161	213	171	545	61%
Over 60	44	23	46	113	80%
Unknown	22	1	1	24	96%
All	324	429	341	1094	61%

Hospital data are for those casualties that occurred on roads within the study area and all casualties on roads that cross the boundaries to the study area.

Table 30: By Gender

Gender	Casualties by gender: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Male	181	247	200	628	61%
Female	143	182	141	466	61%
All	324	429	341	1094	61%

Hospital data are for those casualties that occurred on roads within the study area and all casualties on roads that cross the boundaries to the study area.

Table 31: By Severity

Severity	Casualties by severity: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Fatal	0	1	4	5	80%
Serious	21	83	75	179	54%
Slight	303	340	262	905	62%
Unknown	0	5	0	5	0%
All	324	429	341	1094	61%

Hospital data are for those casualties that occurred on roads within the study area and all casualties on roads that cross the boundaries to the study area.

4.3.3 Barnet Hospital data for inside and estimated unmatched cross border casualties (Best Estimate)

Table 32: Overall numbers and reporting rate

All casualties: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
324	280	341	945	70%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 33: By Class

Road user Class	Casualties: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	44	36	77	157	77%
Pedal cycle user	15	12	12	39	69%
TWMV user	28	29	51	108	73%
Car,van,goods, bus user	234	203	201	638	68%
Other	3	0	0	3	100%
All	324	280	341	945	70%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 34: By Age

Age Band	Casualties by age: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	24	46	40	110	58%
16 to 24	73	73	83	229	68%
25 to 59	161	143	171	475	70%
Over 60	44	18	46	108	83%
Unknown	22	0	1	23	100%
All	324	280	341	945	70%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 35: By Gender

Gender	Casualties by gender: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Male	181	159	200	540	71%
Female	143	121	141	405	70%
All	324	280	341	945	70%

Hospital data for those casualties that occurred on roads that cross the boundaries to the stud area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 36: By Severity

Severity	Casualties by severity: Barnet, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Fatal	0	1	4	5	80%
Serious	21	56	75	152	63%
Slight	303	219	262	784	72%
Unknown	0	4	0	4	0%
All	324	280	341	945	70%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

4.4 St. Mary's Hospital

4.4.1 Hospital data for inside the St. Mary's area only (Upper Estimate).

Table 37: Overall numbers and reporting rate

All casualties: St. Mary's, 2001				
Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
665	86	197	948	91%
Hospital data are for those casualties that occurred on roads which are entirely within the study area.				

Table 38: By Class

	Casualties by user class: St. Mary's, 2001				
Road user Class	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	131	33	47	211	84%
Pedal cycle user	78	2	21	101	98%
TWMV user	126	15	48	189	92%
Car, van, goods, bus user	328	34	80	442	92%
Other Unknown	2	2	1	5	60%
All	665	86	197	948	91%
Hospital data are for those casualties that occurred on roads which are entirely within the study area.					

Table 39: By Age

	Casualties by age: St. Mary's, 2001				
Age Band	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	41	11	13	65	83%
16 to 24	107	14	42	163	91%
25 to 59	411	53	126	590	91%
Over 60	61	8	16	85	91%
Unknown	45	0	0	45	100%
All	665	86	197	948	91%
Hospital data are for those casualties that occurred on roads which are entirely within the study area.					

Table 40: By Gender

Gender	Casualties by gender: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Male	425	47	129	601	92%
Female	240	39	68	347	89%
All	665	86	197	948	91%

Hospital data are for those casualties that occurred on roads that are entirely within the study area.

Table 41: By Severity

	Casualties by severity: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Fatal	4	0	2	6	100%
Serious	77	28	37	142	80%
Slight	584	58	158	800	93%
Unknown				0	
All	665	86	197	948	91%

Hospital data are for those casualties that occurred on roads that are entirely within the study area.

4.4.2 All St. Mary's hospital data for inside and cross border casualties (Lower estimate)

Table 42: Overall numbers and reporting rate

All casualties: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
555	256	307	1118	77%

Hospital data are for those casualties that occurred on roads within the study area and on roads that cross the boundaries to the study area.

Table 43: By Class

Road user Class	Casualties by user class: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	105	90	73	268	66%
Pedal cycle user	63	12	36	111	89%
TWMV user	97	60	77	234	74%
Car, van, goods, bus user	288	88	120	496	82%
Other Unknown	2	6	1	9	33%
All	555	256	307	1118	77%

Hospital data are for those casualties that occurred on roads within the study area and on roads that cross the boundaries to the study area.

Table 44: By Age

Age Band	Casualties by age: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	35	24	19	78	69%
16 to 24	86	54	64	204	74%
25 to 59	338	156	201	695	78%
Over 60	54	22	23	99	78%
Unknown	42	0	0	42	100%
All	555	256	307	1118	77%

Hospital data are for those casualties that occurred on roads within the study area and on roads that cross the boundaries to the study area.

Table 45: By Gender

Gender	Casualties by gender: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Male	356	156	198	710	78%
Female	199	100	109	408	75%
All	555	256	307	1118	77%

Hospital data are for those casualties that occurred on roads within the study area and on roads that cross the boundaries to the study area.

Table 46: By Severity

	Casualties by severity: St. Mary's, 2001				
	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Fatal	4	1	2	7	86%
Serious	53	61	61	175	65%
Slight	498	194	244	936	79%
Unknown				0	
All	555	256	307	1118	77%
Hospital data are for those casualties that occurred on roads within the study area and on roads that cross the boundaries to the study area.					

4.4.3 St. Mary's Hospital data for inside and estimated unmatched cross border casualties (Best Estimate)

Table 47: Overall numbers and reporting rate

All casualties: St. Mary's, 2001				
Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
555	131	307	993	87%
Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.				

Table 48: By Class

	Casualties by user class: St. Mary's, 2001				
Road user Class	Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Pedestrian	105	51	73	229	78%
Pedal cycle user	63	4	36	103	96%
TWMV user	97	30	77	204	85%
Car, van, goods, bus user	288	44	120	452	90%
Other Unknown	2	2	1	5	60%
All	555	131	307	993	87%
Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.					

Table 49: By Age

Age Band	Casualties by age: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
0 to 15	35	16	19	70	77%
16 to 24	86	24	64	174	86%
25 to 59	338	78	201	617	87%
Over 60	54	13	23	90	86%
Unknown	42	0	0	42	100%
All	555	131	307	993	87%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 50: By Gender

Gender	Casualties by gender: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Male	356	76	198	630	88%
Female	199	55	109	363	85%
All	555	131	307	993	87%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

Table 51: By Severity

	Casualties by severity: St. Mary's, 2001 Casualties known only to the Police	Casualties known only to the hospital	Casualties known to both	All casualties	Reporting rate
Fatal	4	1	2	7	86%
Serious	53	41	61	155	74%
Slight	498	89	244	831	89%
Unknown				0	
All	555	131	307	993	87%

Hospital data for those casualties that occurred on roads that cross the boundaries to the study area have been factored to give an estimate of the number of those casualties which actually occurred within the study area.

5 Interpretation of results

The reporting rates obtained in this study were considerably higher than had been expected. High reporting rates occur when;

- the proportion of casualties known to the police is high in relation to the total number of casualties, or,
- the proportion of unmatched casualties in the hospital dataset is small.

5.1 Why might our reporting rates be higher than we expected at the outset?

We propose three explanations for why the high reporting rate was observed:

1) Characteristics of London hospitals
2) The data reporting and collection process.
3) There is higher reporting to the police in London.

5.1.1 Characteristics of the use of Hospitals by casualties.

London has 31 hospitals with full time A&E departments, and as far as we can tell from the London Ambulance Service data (presented in Figures 2-4) there is some considerable overlap in their 'catchment areas'. If ambulances take people to hospitals other than our target hospitals then people who 'walk-in' also have a choice of hospital. This is not generally the case in towns and cities with only one hospital.

We speculated that with many commuters and visitors to London there may be a tendency by commuters who had been injured in London to report their accident to the police but to attend a hospital near where they live, which might be elsewhere in, or outside, London. In the case of RTA injuries this would result in a reduced number of casualties known to the hospital for a given number known to the police. This would only be the case where injuries were not of sufficient severity to prevent the casualty returning home by themselves.

If this was occurring, we would expect a lower proportion of slight injuries reported by the hospital. This might be coupled with a higher proportion of road users whose injuries tend to be more severe (Pedestrian, pedal cyclists and two wheel motor vehicle).

A key test would be that there is a higher proportion of people living locally to the hospital appearing in the hospital dataset than in the police dataset. We cannot know this as we do not have address information. It might also be expected that those hospitals serving areas where the proportion of commuters was lower would be less affected. (e.g. Barnet General least affected and St. Mary's most), but our results do not support this hypothesis.

5.1.2 The data recording and collection process.

The following could affect the levels of recording in either the police or hospital datasets:

5.1.2.1 Accuracy of recording by police or by hospital.

The hospital data is a combination of ambulances reporting to the hospital they are bringing in an RTA casualty, and the self-reporting of people walking-in to the A&E department. There is an element of self-reporting in the police data as well since not all accidents are attended by the police. We have no estimates of the accuracy and veracity of reporting by members of the public. If there was a systematic organisational problem with the accuracy this might affect one particular dataset but not others.

5.1.3 Higher levels of reporting to the police in London

In Central London there are a high number of police and their response rate to incidents is often very rapid. It could be hypothesised that this would make reporting accidents easier. Given the nature of traffic in London, an RTA may well be more likely to cause congestion and hence attract the attention of the police. Therefore, a greater proportion of all RTA casualties would be known to the police. If this were the case one would expect the proportions of the different user groups and age groups to be similar for hospital data and for police data.

5.1.3.1 Economic incentives to report to the police**.**

It may be hypothesised that there are more business vehicles active in London than in other places, hence there would be organisational pressure on drivers to report RTAs in the correct manner.

5.2 How do the police and hospital datasets vary?

This section compares the police and the hospital datasets in terms of the user characteristics to help inform the interpretation of the data. Figures 10-15 show the proportion of casualties of various categories in terms of the total numbers of police and of hospital casualties. The dataset used for this estimate was the 'best estimate dataset' described earlier.

From examination of the comparison charts for road user class, age group and severity, we can determine if any of the datasets have greatly different characteristics from each other which helps us judge whether there are anomolies in one particular dataset. This task is made more complex because each study area has different characteristics.

Figure 10. Composition of police and hospital datasets by road user class. (Hospital data includes estimated casualties from roads crossing the boundary of the study area).

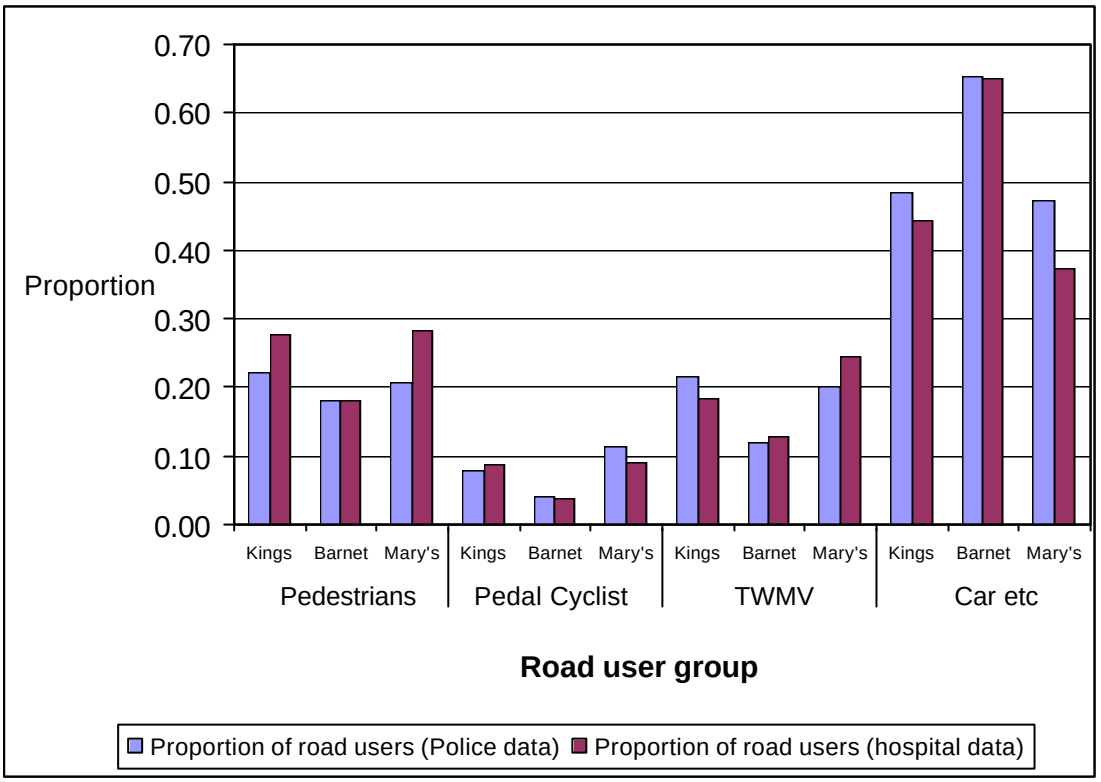


Figure 11. Composition of police, hospital and unmatched hospital datasets by road user class. (Hospital data includes estimated casualties from roads crossing the boundary of the study area).

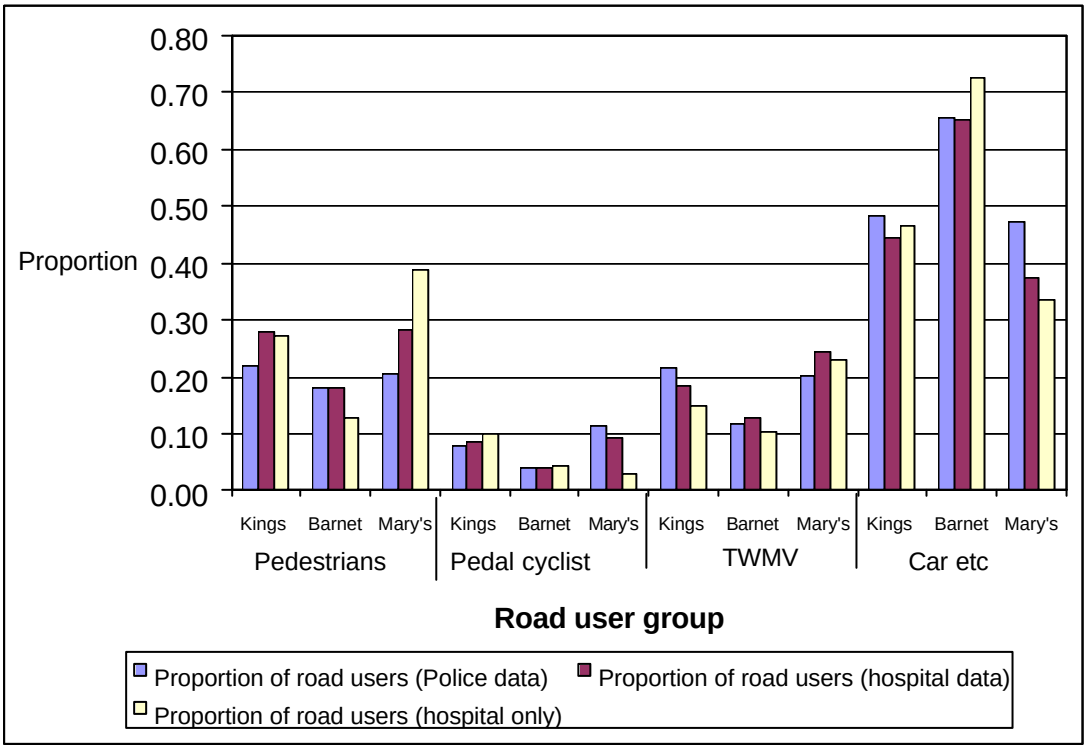


Figure 12. Composition of police and hospital datasets by age group. (Hospital data includes estimated casualties from roads crossing the boundary of the study area).

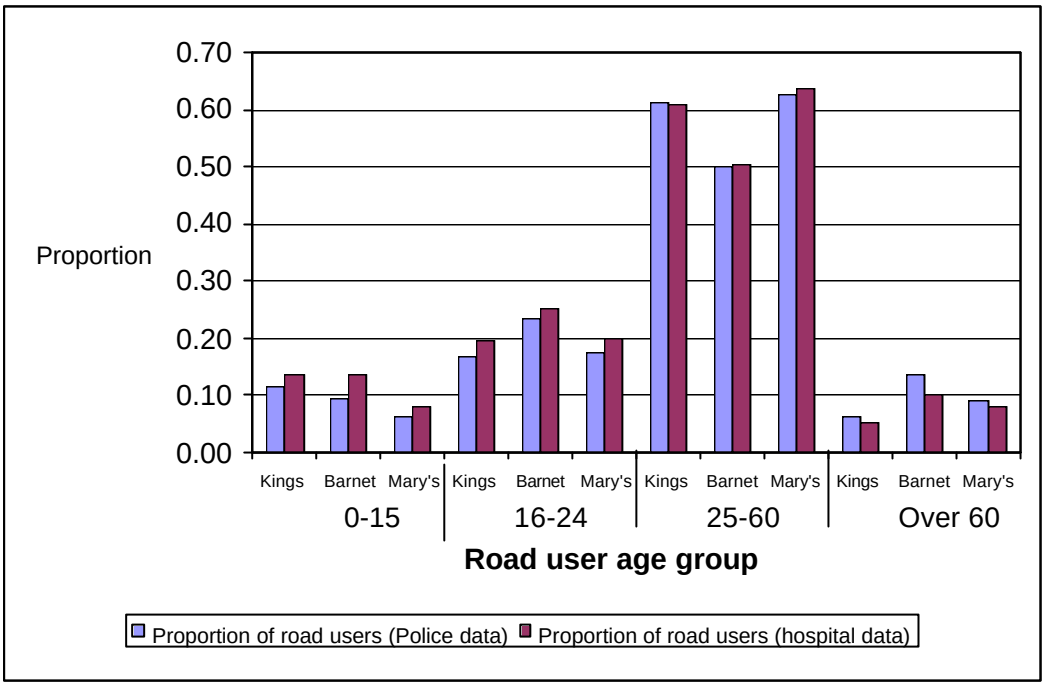


Figure 13. Composition of police, hospital and unmatched hospital datasets by age group. (Hospital data includes estimated casualties from roads crossing the boundary of the study area).

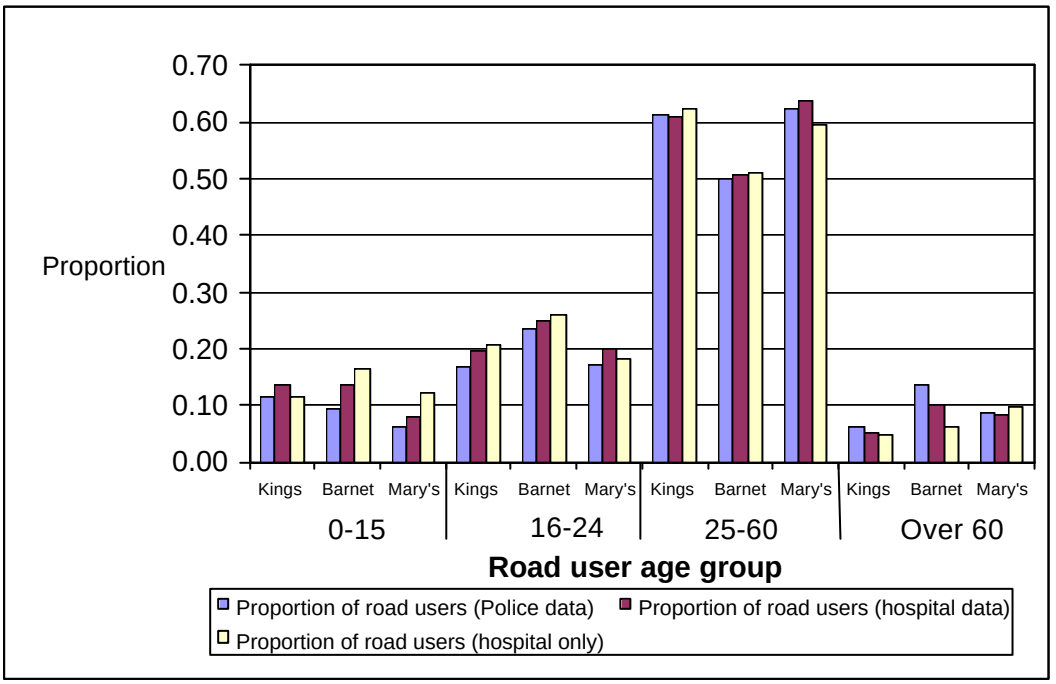


Figure 14. Composition of police and hospital datasets by severity. (Hospital data includes estimated casualties from roads crossing the boundary of the study area).

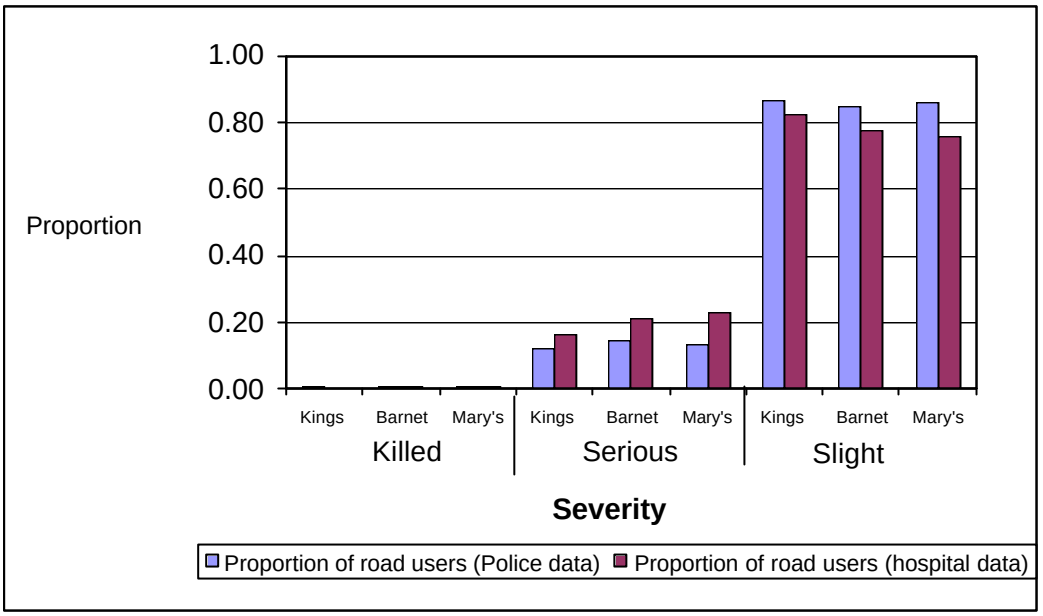
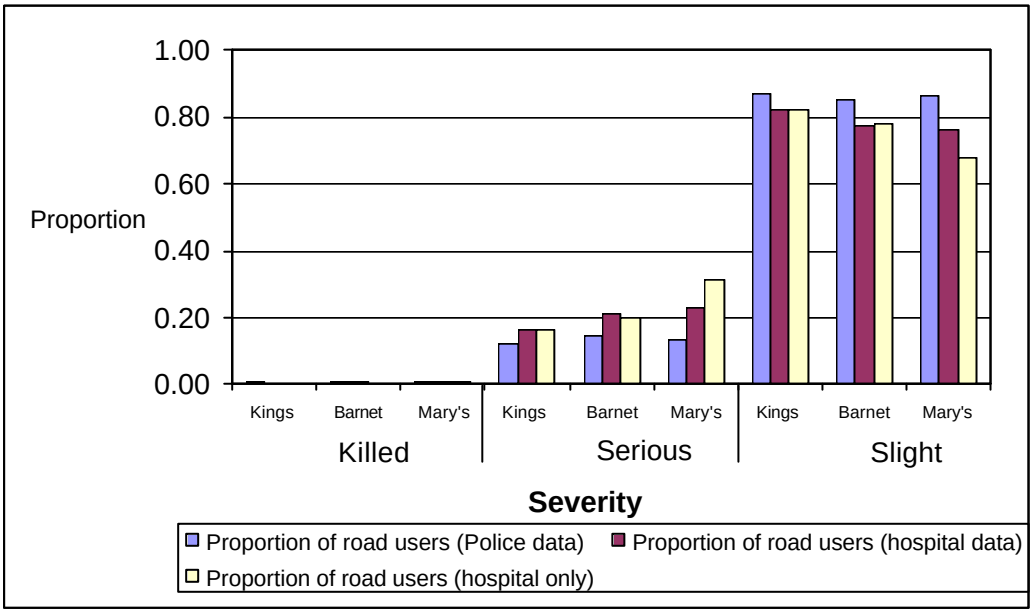


Figure 15. Composition of police, hospital and unmatched hospital datasets by severity. (Hospital data includes estimated casualties from roads crossing the boundary of the study area).



The charts show broad agreement between the hospitals. In the cases where there are large differences (e.g. proportion of 25-60 year olds in the Barnet area, Figure 12), both police and hospital datasets show similar types of differences. The implication is that none of the datasets have a systematic error.

At this stage no statistical analysis has been undertaken, but there appear to be some systematic differences. There is a higher proportion of serious injuries recorded by the hospital compared to the police data (Figure 14). This would be expected as the lower levels of slight injury recorded by the police include minor injuries such as cuts and bruising which are not referred to the hospital Accident and Emergency Unit. A second factor operating here is that the police are not

medically trained and do not accurately estimate severity of injury. There could be a tendency for the police to underestimate severity especially where there is internal, or head injury without obvious outward signs (personal communication).

In the cases where the fatal records are not matched we may assume that as the hospital record is an A&E record and not a general hospital record – if a person dies after being admitted to a ward this information will not be on the A&E disposal code. We have no explanation for the three (from a total of 19) deaths known to the A&E and not to the police.

There is a lower proportion of car occupants and a higher proportion of pedestrian casualties recorded at King's and St. Mary's compared to the corresponding police data (Figure 10). Possible explanation is that speeds in central London are generally low and pedestrian activity high. This might lead to slight car occupant injuries being reported to the police but the casualty not attending hospital. The reverse seems to be the case for pedestrian injuries. This could possibly be due to 'slight' casualties who would otherwise take themselves to an A & E department, not knowing where to find a hospital or not being able to get themselves to it.

As described at 5.1.3, the greater presence of police and awareness of incidents on roads in Inner and Central London possibly makes the reporting of injury easier in these areas. It follows that some people who would have chosen not to report their own injury are recorded by the police as a casualty but subsequently do not report themselves to A & E.

5.3 Reporting rates

Table 52: Estimates of the headline reporting rates to the police for each hospital area together with percentage of records matched.

Rate=All casualties known to police/All casualties.

Class	Hospital	Lower estimate [1]	Upper estimate [2]	Current best estimate [3]
All casualties	King's	66	80	71
All casualties	Barnet	61	75	70
All casualties	St Mary's	77	91	87
Pedestrians	King's	62	78	67
Pedestrians	Barnet	72	78	70
Pedestrians	St Mary's	66	84	78
Pedal Cyclist	King's	61	75	66
Pedal Cyclist	Barnet	60	75	69
Pedal Cyclist	St Mary's	89	98	96
TWMV	King's	74	87	78
TWMV	Barnet	65	80	73
TWMV	St Mary's	74	92	85
Car occupant	King's	67	78	72
Car occupant	Barnet	58	73	68
Car occupant	St Mary's	82	92	90

The overall results tell us that the reporting rate in London was considerably higher than in previous studies, such as the Safer City Project in Gloucester. Cheltenham and Gloucester are freestanding towns each having a population in the order of 100,000. For Cheltenham and Gloucester, there is a single hospital serving each area. It is known that in a small number of cases a casualty is taken to another hospital for example for specialist treatment.

Our study areas do have fundamentally different characters from those in Gloucester or Cheltenham. A number of hypotheses have been suggested above which look at reasons why this study has measured higher levels of reporting.

Given the similarity in the composition of the hospital and police datasets, we come towards the conclusion that the higher levels of reporting in London compared to Gloucester, Cheltenham or Northampton is a genuine effect. The overall rate is similar at 70% for King's and Barnet (Inner

[1] Casualties identified as being in the area only for police data and all those that are in the area and on cross border roads for the hospital

[2] Casualties identified as being in the area only for hospital and police data

[3] Casualties identified as being in the area only for police data and all those that are in the area plus adjusted number for those on cross border roads for the hospital

and Outer London) and considerably higher at 87% for St Mary's (Central London). These similarities between inner and outer London may be seen in the rates for pedal cycles and pedestrians, again with Central London being higher in both cases. For car occupants and TWMVs, St Mary's is still the highest but in both cases King's is higher than Barnet.

One of the objectives of the study is to try to gross up these results to the whole of London. Possible ways of doing this are described in Section 5.3.

5.4 Estimation of the number of people injured in road traffic accidents for the whole of London

There were about 44,500 casualties recorded by the police in London during 2001. The best estimates of reporting rates for King's, Barnet and St Mary's hospitals have been calculated as 71%, 70% and 87% respectively. If we use the lower figure of a 70% reporting rate we might assume that the total is nearer to 63,000.

Clearly more work needs to be done in this area, especially in trying to get a better estimate of the Central London reporting rate and to refine the techniques for estimating the number of cross-border casualties.

However, there are questions outstanding regarding the St Mary's example which is so much higher than the other two hospitals. At this stage there is little evidence to suggest a data error but with a sample of one there is insufficient evidence to conclude that this is typical of Central London. It is recommended that a second Central London hospital's data be matched with police data.

The evidence is more in line with expectations for King's and Barnet and they may be more typical of Inner and Outer London. Further work could be undertaken to try to use the ratio of RTA records to total A&E records for a larger sample of hospitals.

Finally the higher reporting rates found for London may go some way to explaining the 'London Effect' found so often in accident data studies. This is an observation where the accident and casualty rates for London are considerably higher than the rest of the country and are thus modelled or displayed separately in reports.

5.5 Further work

It is recommended that further work is carried out in this area particularly in relation to:

- Investigating whether there is a better way of estimating the number of casualties occurring on cross-border roads, given they make up a substantial proportion of the total number of casualties in the hospital dataset.
- Investigating St. Mary's data more fully as there may be a shortfall in the number of casualties reported as RTA's. The way to do this would be to take St. Thomas' or UCL Hospitals as central London hospitals and calculate their reporting rates.
- Calculate ways to gross up our findings to the whole of London. In the context of this study we have not had time to do this other than provide a very rough estimate.

6 Acknowledgements

The help and interest of the A&E consultants at the following hospitals is acknowledged and our thanks go to them for providing data so quickly.

Ruth Brown, Consultant in Emergency Medicine St Mary's Hospital Trust
Brijendra Shravat, A&E Consultant Barnet Chase Hospital Trust
Edward Glucksman, A&E Consultant King's College Hospital

The London Ambulance Service who provided data in the form of maps:
Michael Damiani, Information Analyst Management Information, London Ambulance Service
Tim Coates, A&E Consultant, Royal London Hospital, Whitechapel for his help and insights to the workings of a busy A&E Department.

Finally, to John Devenport and Jon Fray for their guidance and helpful comments at every stage of this study.

7. References

DETR (2000) Road Accidents Great Britain 2000. HMSO

Geographers A-Z Map Co. (2001) London A-Z, Geographers A-Z Map Co. Sevenoaks.

James H.F. (1991) Under-reporting of road traffic accidents. Traffic Engineering and Control 32 p574-583

Pjsoft (1999) InfoMap ver.7.0,

Simpson H.F. (1996) Comparison of hospital and police casualty data: A national study. TRL Report 173. TRL. Crowthorne.

Ward H, Cave J, Morrison A, Allsop R, Evans A, Kuiper C and Willumsen L (1994) Pedestrian activity and accident risk. AA Foundation for Road Safety Research., Basingstoke.

Ward H and Robertson S (2002) Do public perceptions of safety affect reporting rates? Paper presented at 6th World conference: Injury prevention and control. May 12-15 Montreal. Le Presses de l'Universite de Montreal P55-56

Appendix 1: Grid references for hospital boundaries

For Figure 6: Grid references for King's from NW corner going clockwise

531640	177377
535329	177160
534887	172901
531237	172743

For Figure 7: Grid references for St Mary's from NE corner going clockwise

525987	183813
528053	180875
525107	180272
524053	182331

For Figure 8: Grid references for Barnet General from NE corner going clockwise

517941	194103
521158	194709
522401	195952
523287	196465
527608	196791
528805	191942
526800	190652
520070	191538

Appendix 2: Estimation of number of casualties in the cross- border condition

Reporting rate:
(p+b)/(p+h+b)
or

P/(P+h)

Where
H all casualties known to hospital (i.e. h+b)
h casualties only known to hospital
b casualties known to both
P all casualties known to police (i.e. p+b)
p casualties known to police only
subscript i for inside the study area
subscript b cross-border data
The police data supplied are known to be all inside the study area.

The proportion (K) of hospital casualties that are known to the police were calculated. i.e. $K= b_i/H_i$
for the hospital data that are known to be inside the study area

An estimate of the number of casualties on roads that cross the boundary to the study area and known by the hospital alone ($h_{estimated}$) can now be determined. This estimate is based on the number of casualties in the cross boundary roads known to both (b_b) and the total number of casualties know to the hospital (H_b).

The total number of casualties that would be expected to be known to the hospital in the cross boundary area ($H_{estimated}$) is calculated as follows
$H_{estimated}= b_b/K$

To obtain the expected number of casualties known to the hospital alone subtract the number known to both.
$h_{estimated} = H_{estimated} - b_b$

$h_b - h_{estimated}$ excess unmatched casualties in the cross boundary area were then stripped out randomly from the unmatched hospital casualties

Selection of the casualty data to be stripped out was undertaken using generated random numbers from a spreadsheet function. A random number was assigned to each casualty then the casualties were sorted by that random number. The top $h_b - h_{estimated}$ casualties in the list were stripped out leaving $h_{estimated}$ casualties.

Example calculation for King's 2001.

Within study area	p	= 605
Within study area	h_i	= 214
Within study area	b_i	= 232
Within study area	H_i	= 446
Proportion of hospital data known to police	$K = b_i/H_i$	= 0.52017
On cross boundary roads	h_b	= 211
On cross boundary roads	b_b	= 135
On cross boundary roads	H_b	= 346
Expected total number on cross boundary roads	$H_{estimated} = b_b /K$	= 259.53
Number to be removed from cross boundary roads	$H_b - H_{estimated}$	= 86.47

Appendix 3: Hospital Disposal Codes

Each hospital uses its own set of disposal codes. These are records of where the patient was sent after A&E. We have assigned our own codes which broadly correspond with the STATS19 severity definitions to allow matching of severity to be undertaken.

Severity Code	Casualty Severity
1	Fatal
2	Serious
3	Slight

A.3.1 St. Mary's disposal codes 2001

Disposal	Severity code
A&E Alcohol Declined	3
A&E Dressing Clinic Allcroft	3
A&E Ref OPD	3
A&E Review Clinic	2
A&E Review Clinic A Bourne I – Labour	2
A&E Review Clinic Albert	2
A&E Review Clinic Aleck Bourne II	2
A&E Review Clinic Allcroft	2
A&E Review Clinic Almroth Wright	2
A&E Review Clinic CCU	2
A&E Review Clinic Great Western/PICU	2
Admitted Grand Union	2
Admitted Great Western	2
Admitted Other	2
Admitted PICU	2
Admitted SMH Trust A Bourne I – Labour	2
Admitted SMH Trust A&E	2
Admitted SMH Trust Almroth Wright	2
Admitted SMH Trust Charles Pannett	2
Admitted SMH Trust Dickson Wright	2
Admitted SMH Trust Douglas A&E	2
Admitted SMH Trust Handfield Jones	2
Admitted SMH Trust ITU (Milne)	2
Admitted SMH Trust Joseph Toynbee	2
Admitted SMH Trust Manvers HDU	2
Admitted SMH Trust Other	2
Admitted SMH Trust Samaritan Ward	2
Admitted SMH Trust Samuel Lane	2
Admitted SMH Trust Valentine Ellis	2
Admitted SMH Trust William Coulson	2
Admitted SMH Trust Zachary Cope	2
AutoDischarge	3
Did Not Wait	3
Died in Dept	1

Discharged - no FU	3
Discharged - no FU None	3
Fracture Clinic	2
GP Followup Appt	3
GP Ref OPD	3
GP ROS	3
GP to Register	3
Home	3
OPD	3
Orthopaedic Clinic	2
Removed by security	3
Review Clinic	3
Took own Discharge	3
Transfer other Trust	2
Transfer other Trust A&E	2
Transfer other Trust Beverley	2
Transfer other Trust Other	2
Transferred	2

A.3.2 Barnet hospital disposal codes 2001

Unknown=999

Discharge Code	Severity Code
A&E Obs Unit Aeobs	2
A&E Obs Unit Holly	2
Admit Aeobs	2
Admit Aspen Ward	2
Admit Beech Ward	2
Admit Cedar Ward	2
Admit Damson Ward	2
Admit Day Services Unit	2
Admit H.D.U.	2
Admit Holly	2
Admit I.T.U.	2
Admit Juniper/Wd 2	2
Admit Olive/Wd 8	2
Admit Theatre	2
Admit Treasure Island	2
Admit Unity (Gynae)	2
Admit Victoria (Maternity)	2
Dead On Arrival	1
Did Not Wait	3
Died In Department	1
Disch Fracture Cln	2
Disch Home G.P. F/Up	3
Disch Home No F/Up	3
Disch Psychy Dept	3
Disch. O.P.D.- B.G.H	3
Disch.A&E Review	2
Dressing/Rso	3
Maxillo-Facial Team	3
Medical Team	2
No Record	999
Obs/Gynae Team	999
Orthopaedic Team	2
Paed Ward Review	2
Paediatrics	2
Refused Treatment	3
Self Discharged	3
Surgical Team	2
Trans To Other Hosp Mount Vernon	2
Trans To Other Hosp Royal Free Hospital	2
Trans To Other Hosp Trans Other Hosp	2
Trans To Other Hosp Trans Other Hosp	2

A.3.3 King's College Disposal codes 2001

Discharge code	Severity Code
A&E Physio clinic	3
Admit King's/Dul ACW	2
Admit King's/Dul Butlin	2
Admit King's/Dul Cotton	2
Admit King's/Dul David Ferrier	2
Admit King's/Dul Dawson	2
Admit King's/Dul Donne/SRU	2
Admit King's/Dul ICU	2
Admit King's/Dul Lion	2
Admit King's/Dul Lister	2
Admit King's/Dul Lonsdale	2
Admit King's/Dul Mary Ray	2
Admit King's/Dul Matthew Whiting	2
Admit King's/Dul Observation Unit A&E	2
Admit King's/Dul Oliver	2
Admit King's/Dul PICU	2
Admit King's/Dul Princess Elizabeth	2
Admit King's/Dul Sambrooke	2
Admit King's/Dul Sylvia Henley	2
Admit King's/Dul Trundle & Wadd	2
Admit King's/Dul Twining	2
Admit King's/Dul Victoria & Albert	2
Admit Other Hospital	2
Admit Other Hospital Guy's	2
Admit Other Hospital Observation Unit A&E	2
Admit Other Hospital Other Hospital	2
Admit Other Hospital St Thomas'	2
Computer Discharged	3
Computer Discharged Observation Unit A&E	3
Dental Hospital	3
Did not wait (DNA)	3
DIED	1
Discharged	3
Discharged Observation Unit A&E	2
Dressing Clinic A&E	3
Eye Clinic	3
Fetal assessment unt	9
Fracture Clinic King	2
Gynae Scanning	9
OPD King's/Dul	3
Own GP	3
Own GP Observation Unit A&E	2
REF ONCALL SPECIALTY	2
Removed from dept	3
Review Clinic A&E	3
Self discharge	3
(blank)	9

Appendix 4: Codes used for data matching

Codes for Severity

The severity coding used was STATS19 i.e. killed, seriously injured and slightly injured as shown below. Police levels of severity were taken as definitive.

Severity	Code
Killed	1
Serious	2
Slight	3
Unknown	999 or 9

Hospital Severity Cases

The hospital descriptions of severity were in the form of disposal codes and these were mapped onto the STATS19 severity codes. Each hospital had a different set of disposal codes (or discharge outcomes) which were individually mapped on to the severity codes mapped. These can be found in Appendix 3.

Codes for Gender

Male 1
Female 2
Unknown 3

Codes for vehicle class

The coding used for the class of a road user casualty by the police and by the hospital were different so a set of categories that encompassed both coding systems was required. The following table shows the coding scheme used.

Class	Code
Pedestrian	1
Pedal cycle	2
TWMV	3
Car, bus, HGV,LGV etc	4
other (horse etc)	5
Unknown	9

Police user class data were taken as definitive for tables. It was assumed that the vehicle type was more likely to be correctly identified by police at the scene of the crash than by hospital A&E staff.

Police user class data

Vehicle type: Police description	Code
Agricultural Vehicle	5
Bus or Coach	4
Car	4
Goods	4
Goods 3.5 to 7.5T MGW	4
M/C	3
M/C >125cc	3
Moped	3
Other	5
Other Motor Vehicle	5
Other Non Motor	5
Pedal Cycle	2
Pedestrian	1
Ridden Horse	5
Taxi	4
Unknown	9

Vehicle type hospital codes were broadly similar between hospitals, though some differences in capitalisation required separate lookup tables for coding each hospital's user class data.

Hospital user classes, Barnet	Code
Cyclist	2
Driver	4
Front Seat Passenger	4
Motor Cyclist	3
Not Entered	9
Pedestrian	1
Rear Seat Passenger	4

Age band codes

Take hospital as definitive for tables

Age band	Code
Unknown	-1
0-15	1
16-24	2
25-59	4
60+	5
Unknown	9

Codes for area

0 out
1 definitely in
2 spelling error (used in initial process of identifying location of casualties
3 Roads cross boundary
4 Unknown