A Methodology to Investigate and Visualise the Geographical Provenance of Road Traffic Casualties in Deprived Areas

Sanjay Rana and Heather Ward

Centre for Transport Studies, University College London

Gower Street, London WC1E 6BT, UK

Abstract

This paper demonstrates a methodology based on spatial overlay techniques, casualty data and national deprivation scores, to quantify and visualise the geographical provenance of road traffic casualties in deprived areas. A case study of four districts with varying deprivation in Greater Manchester, UK is presented. It is shown that most injuries to pedestrians and car occupants occur in areas of similar levels of affluence/deprivation to that of where the casualties live. Thus, it is proposed that the phenomenon underlying the cause of road traffic injuries are probably universal despite the differences in factors such as demography and level of deprivation.

Acknowledgements

This study was funded by the Department for Transport as part of the evaluation of the Neighbourhood Road Safety Initiative. Greater Manchester Transportation Unit provided the STATS19 data and the digital maps were obtained from the EDINA archive of Ordnance Survey Maps © Crown Copyright under the CHEST agreement.

Running Headline

Geographical Provenance of Road Traffic Casualties in Deprived Areas

1. Introduction

The relationship between area deprivation and the occurrence of road traffic accidents (RTAs) has been known for sometime (Preston, 1972; Christie, 1995; White et al., 2000; Graham et al., 2005). Christie (1995) has shown that residents of deprived areas tend to have relatively higher number of accidents while Graham et al. (2005) presents a strong relationship between deprivation at an accident site and the occurrence of child and adult pedestrian casualties. Thus, the level of deprivation at the accident site and casualty residence appears to have a relationship with the risk of injury in a road traffic accident. However, it remains unknown what it is exactly about a deprived area that leads to this higher accident risk. With the use of binomial regression between various "area type" factors (e.g. traffic flow, population etc.), deprivation and the child and adult pedestrian casualties, Graham et al (1995) proposed that only deprivation had a statistically significant influence on the RTAs. Christie (1995) presented several risk factors relating to a deprived area (e.g. high population density, old street layout, poor education, low income, high unemployment) that could come together to create a road environment and road user behaviour and lead to an RTA. However, neither Christie (1995) nor Graham et al. (2005) establish a relationship between deprivation and RTAs which indicates to what extent the difference between deprivation at the accident site and at the casualty residence could have accounted for some of the excess risk of accidents in deprived areas. In other words, if deprivation is a major risk factor, then one would find that residents from relatively affluent areas would experience a higher proportion of their RTAs in less affluent areas.

In this paper, a district level study is presented for the 10 districts in Greater Manchester namely; Bolton, Bury, Manchester City, Oldham, Rochdale, Salford, Tameside, Trafford, Stockport, and Wigan (Figure 1). These areas have varying levels of deprivation, ethnicity, and varied road infrastructure. Despite the difference in almost all kinds of factors previously identified by Christie

(1995) and Graham et al. (2005), the main reasons behind the selection of these districts for this study were firstly, the availability of the information on residence of the casualties and secondly, to investigate the similarities in the experience of RTAs in these dissimilar districts.

Section 2 lists the type of data and software used in the study and the various quality issues. Section 3 describes the outcome of the comparative studies and the conclusions are given in final section.

2. Data and Software

2.1 Index of Multiple Deprivation

The Index of Multiple Deprivation (IMD) was produced by the Office of the Deputy Prime Minister (ODPM) in 2004 as a composite indicator of the socio-economic environment in England. IMD 2004 is derived by a combination of seven deprivation indices namely;

- income deprivation
- employment deprivation
- health deprivation and disability
- education, skills and training deprivation
- barriers to housing and services
- living environment deprivation, and
- crime.

Table 1 shows the weights assigned to the seven individual indices in deriving IMD 2004 (ODPM, 2004). The individual deprivation indices were based on a variety of data such as census statistics, air quality, immigration statistics, number of road traffic accidents etc., and the age of the datasets ranged from

1997 – 2003. IMD 2004 is summarised at areas the size of Super Output Area (SOA) Lower Layer, which has a minimum population of 1000 and a mean population of 1500. The boundaries of SOAs are based on the Standard Table Wards used in the 2001 Census[1]. It must be noted that IMD 2004 could also be derived at Ward level however in this work, the use of SOA level representation was found to be useful for visualising the local variation in the accident patterns. IMD 2004 data are available on the ODPM Neighbourhood Renewal website[1]. For presentation purposes, each IMD is assigned a rank, which is a value between 1 and 32482 for SOAs in England.

Figure 2 shows the frequency of the SOAs of the Greater Manchester Districts that fall within an IMD rank decile in comparison to the SOAs in the England. Figure 2 shows that the level of deprivation varies significantly across Greater Manchester. Manchester City has very high levels of deprivation, but Bury has a level of deprivation similar to the rest of England and further still Stockport stands out as the least deprived area.

2.2 Road Traffic Accident Data - STATS 19

Road traffic accidents for the years 1999-2003 were accessed from the STATS19 database on road accident records produced by the Local Police Authorities, which is then later processed and maintained by the UK Department for Transport (DfT). A variety of information related to the accident site, casualty and vehicle is stored in the STATS19 database. Not all information is available to the general public. For instance, although the STATS19 stores the postcode of the casualty's residence, it is withheld due to privacy reasons. However, for this

[1] http://www.communities.gov.uk/index.asp?id=1128440

study we were able to access the casualty residence postcode information as part of a project to evaluate the effectiveness of road safety interventions in deprived areas funded by the DfT. The relevant STATS19 parameters were the geographic location of the accident, postcode of casualty residence, age of casualty, type of casualty class (i.e. whether driver, passenger, pedestrian), and the type of vehicle(s) involved in the accident. The postcode was converted to a geographic coordinate using a lookup table maintained by the Post Office. The entire list of the parameters stored in the STATS19 database about an accident can be seen at http://www.stats19.org.uk.

In this study, two types of road casualties, namely car occupants and pedestrians were studied in three age groups viz. Child (a person under 16 years old), Adult (a person between 16-60 years old) and Older adult (a person older than 60 years). The choice of pedestrians in the current study may appear counter intuitive to the reader because it is well known that most pedestrian injuries occur nearby the casualty residences. The motivation behind the use of pedestrian injuries was to bring out a significant contrast to the car occupants injuries.

Accurate and complete records of the geographic coordinates of the casualty's residence are crucial for this study. Unlike other spatial modelling studies (e.g. see Chapters 12-15 in Longley et al., 1999), uncertainty arising due to incomplete geographical and non-geographical information in RTA records are not generally taken into account during the modelling of the spatial patterns of accidents. In this study, a visual comparison of the spatial patterns of the records with and without postcode information was carried out to identify any systematic bias that may arise due to incorrect, incomplete or missing records of the casualty residence postcodes. In a systematically biased set of records one would find cluster(s) of records with unusable postcode information. The comparison of the patterns was undertaken using spatial overlays, which will be discussed in more detail in the following section on the Geographic Information System (GIS).

2.3 Geographic Information System (GIS)

A Geographical Information System (GIS) is a software system for the visualisation, analysis and storage of spatial datasets. The ArcView® GIS developed by ESRI Inc., was used to perform two types of geospatial data processing in the current study. Firstly, the GIS was used to overlay geospatial datasets or themes over another to visualise any interesting spatial patterns. Secondly, a point-in-polygon type spatial query was used to retrieve various attributes of the overlapping datasets. A point-in-polygon query involves the test of the topological relationship "*d*oes *a point lie inside a given polygon*"? In traditional accident analysis the district codes attached to the accident records are used to identify which accidents took place within a district. However, in this study the point-in-polygon spatial query was used to identify accidents that took place within a district. Further, once a polygon (e.g. a SOA polygon) has been found to contain a point (e.g. an accident), usual relational queries can be carried out to find out the other attributes associated with the point(s) and polygons(s) (e.g. IMD rank of a SOA that contains a given point). The advantage of using the point-in-polygon spatial query in comparison to conventional relational queries rooted at accident records (stored at the Local Authority level scale) is that the spatial query allows the use of geographic coordinates for querying spatial information (e.g. IMD Rank) and hence can be used to retrieve information (e.g. IMD ranks) that may be stored at a sub-local authority level scale (e.g. SOAs).

3. Relationship between Deprivation at Accident Site and Casualty Residence

A case study of the methodology is presented using four of the ten Greater Manchester Districts (Bury, Manchester City, Salford, and Stockport). The reason behind the choice of these districts is to compare the geographical provenance in both deprived and affluent areas. In comparison to the national average level of deprivation, the levels of deprivation in the selected areas range from very high (Manchester City), higher than the national average (Salford), moderate (Bury), and low (Stockport).

3.1 Uncertainty in Casualty Residence

Figure 3 shows a comparison of the completeness of the casualty postcode information in the RTA records amongst the 10 Districts. Note that the four study areas have approximately similar levels of uncertainty. Figure 4 shows an overlay of the accident records with and those without the casualty postcode information. A visual inspection of the spatial distribution of the accidents in Figure 4 indicates that there are no obvious clusters of records with missing casualty residence postcode information in the study areas. Therefore, the subset of RTA records used in this study can be considered to be representative of the distribution of correct postcodes for the entire dataset.

3.2 A comparison between Deprivation at Accident Site and Casualty Residence

Owing to spatial autocorrelation, it is quite likely that locally adjacent SOAs will have similar IMD Ranks and thus may exaggerate minor differences in the level of deprivation. Therefore, the IMD ranks were aggregated into 10 classes (1-10) with a class interval of deciles of 32482 (which is the total range of the England IMD Ranks). Class 1 represents IMD Ranks ≤ 3248, Class 2 represent IMD Ranks > 3248 AND ≤ 6497 and so on.

The difference between the decile of the IMD Rank at the accident site (i_1) and the decile of the IMD rank at the casualty residence (i_2) was calculated to enable an investigation to be made of the distribution of these differences for pedestrians and car occupants of different ages. Negative ($i_1 - i_2$) values indicate that casualties resided in an area relatively more affluent than the accident site and vice versa for the positive ($i_1 - i_2$) values. Zero ($i_1 - i_2$) values indicate that casualties were injured in a SOA which had the same IMD rank decile as the SOA of their residence, which mostly happens if the RTA occurred in the SOA where the casualties live. Thus, ($i_1 - i_2$) = -9 would represent a case, where the IMD rank at the casualty residence > 29233 AND ≤ 32482, while the IMD rank at the accident site ≤ 3248. The hypothesis is that if the deprivation of the area in which the accident occurs is the predominant risk factor in accident occurrence then one would observe that, for most accidents, casualties will be from a relatively less deprived area than that of the accident site. In other words in most RTAs, ($i_1 - i_2$) would be negative. For the sake of brevity, casualties with negative ($i_1 - i_2$) values will be referred as Class 1 casualties, casualties with nil

($i_1 - i_2$) values will be referred as Class 2 casualties and casualties with positive ($i_1 - i_2$) values will be referred as Class 3 casualties.

Table 2 shows the number of pedestrian and car occupant casualties where the residence postcode information was available. It is clearly evident that Manchester City with the highest level of deprivation amongst the four districts also has the most number of casualties in all the casualty types, of which a large part are due to the fact that Manchester is a socio-economic hub in the region. The distinction between remaining districts is however not so clear despite quite different levels of deprivation, which is the first suggestion that the deprivation of the area itself is probably not the significant contributing factor in RTAs.

Figure 5 shows a plot of the ($i_1 - i_2$) for car occupants and pedestrians in different age groups. The bell-shape leptokurtic distribution centered at Class 2 with a small negative skew in all the four districts for all casualty classes suggests that the area deprivation at the accident site and at the residence did not have a major influence on the occurrence of these RTAs. All the districts display approximately similar trends, which suggest that the phenomena underlying the cause of RTAs were similar. Considering that the districts differ in almost all the influential factors in RTAs (e.g. Graham et al., 2005), the similarity in the distributions suggest that the relationships between RTAs and influential factors may not be straightforward to tease out. It is also important to note the number of RTAs decreases sharply with both an increase and decrease in the ($i_1 - i_2$) values, which shows that neither the residents of most deprived nor the residents of least deprived had many injuries in each other's neighbourhoods. Or, putting it another way, most injures to pedestrians of all ages occurred close to home or at least in areas similar to home. In fact, as a rather broad estimate based on Figure

5, it can be stated that the bulk of the casualties took place in an area which had a difference of ± ~3 IMD Rank deciles i.e. ± ~9744 IMD rank of the casualty's residence area IMD rank.

Further, in Figure 5 the highest proportion of casualties in all districts belonging to Class 2 comprised of child pedestrians. This observation is somewhat expected since it is well established that child pedestrians generally take place within a short distance of the residence (Ward et al., 1994). Manchester City has the largest number of casualties in both pedestrian and categories belonging to Class 2. The relatively higher number of casualties belonging to Class 2 in all casualty classes suggests that the road environments around the residences pose a significant threat. Therefore, the design of the road environment around residential neighbourhoods should involve a careful study of the road safety implications especially for vulnerable road users such as child pedestrians.

4 Conclusions and Future Work

The paper presented a methodology based on spatial analysis to visualise and quantify the relationship between the level of deprivation and the occurrence of pedestrian and car occupant injuries in four districts of Greater Manchester UK. Despite the differences in the socio-economic, demographic and traffic flow characteristics in the study areas, it has been shown that the majority of pedestrian and car occupant casualties took place in an area closer to the casualty's residence, which had the same level of the deprivation/affluence as their accident site. This suggests that the underlying phenomenon that influences the occurrence of casualties is same for all the study areas. The similarities between these quite distinct areas present a number of opportunities for future work. For instance, it will be interesting to model the trend of the positive and negative ($i_1 - i_2$) values to identify which type of distribution (e.g. exponential) matches closely with the values. In the present work, it was assumed primarily based on visual inspection that there was no systematic bias in the casualty residence postcode records. However, it will be essential to perform detailed modelling of the uncertainty arising due to lack of casualty residence postcode records in order to establish if there is likely to be under or over estimation.

References

Christie, N. (1995) 'Social, Economic and Environmental Factors in Child Pedestrian Accidents: A Research Overview', Transport Research Laboratory, Report No. 116.

Graham, D., Glaister, S., and Anderson, R. (2005) 'The effects of area deprivation on the incidence of child and adult pedestrian casualties in England', *Accident Analysis and Prevention*, **37**, 125-135.

Longley, P.A, Goodchild M.F., Maguire, D.J, and Rhind, D.W. (eds.) (1999) *Geographical Information Systems: Principles, Techniques, Applications and Management*, 2nd edition, John Wiley and Sons, Chichester.

Preston, B. (1972) 'Statistical Analysis of Child Pedestrian Accidents in Manchester and Salford', *Accidents Analysis and Prevention*, **4**, 323-332.

Ward, H., Cave, J., Morrison, A., Allsop, R., Kuiper, C., Willumsen, L. (1994) 'Pedestrian Activity and Accident Risk', AA Foundation for Road Safety Research, London.

White, D., Raeside, R., Barker, D. (2000) ‘Road Accidents and Children Living in Disadvantaged Areas: A Literature Review’, Scottish Executive Central Research Unit, Edinburgh.

Figure Captions

Figure 1 Greater Manchester Districts within England.

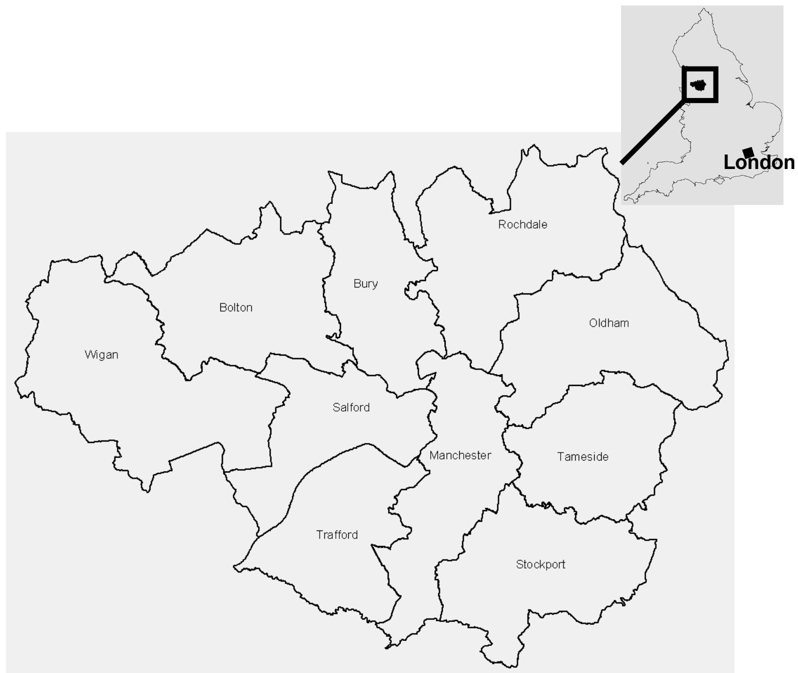
Figure 2 Variation of IMD Ranks in the super output areas of Greater Manchester Districts in comparison to the super out put areas of England. The IMD ranks have been classed into 10 intervals of 3248 each.

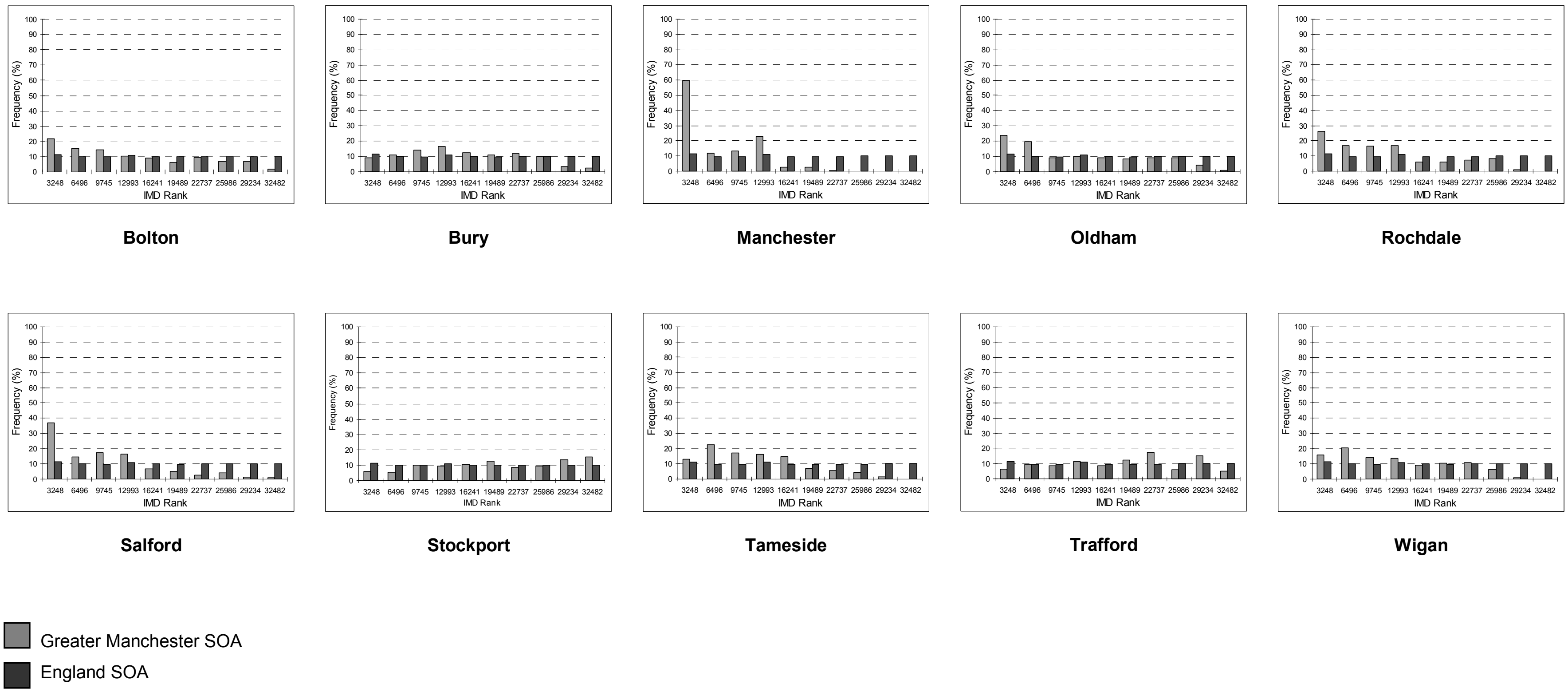
Figure 3 Variation in the storage of the casualty residence postcode information amongst Greater Manchester Districts.

Figure 4 Spatial distribution of accidents with and without postcodes. Grey dots are all records and black dots are records with casualty residence postcode information.

Figure 5 Difference between the IMD Rank at the accident site and IMD rank at the casualty residence. Note the similarities in all the trends despite significantly different levels of deprivation shown in Figure 2.

Figure 1


London
Rochdale
Bury
Bolton
Oldham
Wigan
Salford
Manchester
Tameside
Trafford
Stockport



Frequency (%)
IMD Rank
3248 6496 9745 12993 16241 19489 22737 25986 29234 32482
Bolton
Bury
Manchester
Oldham
Rochdale
Salford
Stockport
Tameside
Trafford
Wigan
Greater Manchester SOA
England SOA

Figure 2

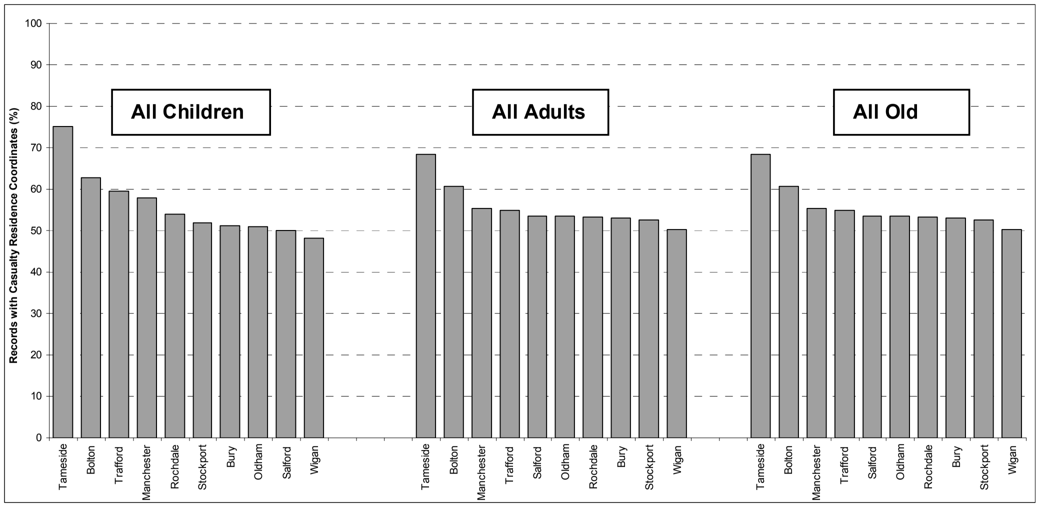
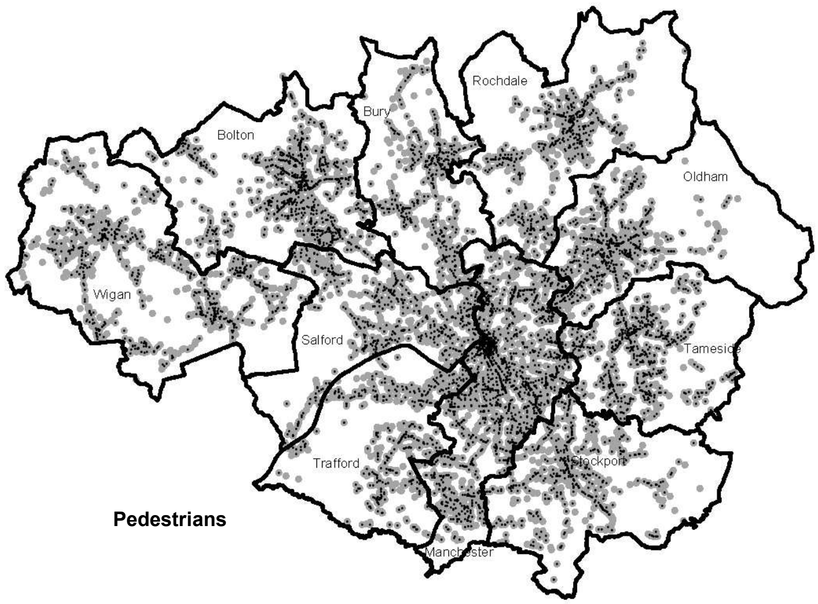
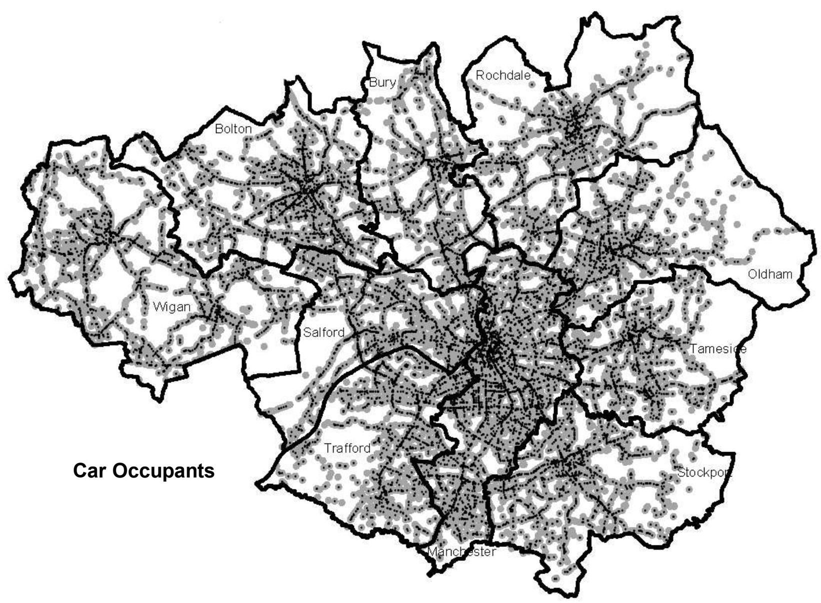

Records with Casualty Residence Coordinates (%)
0
10
20
30
40
50
60
70
80
90
100
All Children
All Adults
All Old
Tameside
Bolton
Trafford
Manchester
Rochdale
Stockport
Bury
Oldham
Salford
Wigan
Tameside
Bolton
Manchester
Trafford
Salford
Oldham
Rochdale
Bury
Stockport
Wigan
Tameside
Bolton
Manchester
Trafford
Salford
Oldham
Rochdale
Bury
Stockport
Wigan

Figure 3

Figure 4


Rochdale
Bury
Bolton
Oldham
Wigan
Salford
Tameside
Trafford
Stockport
Pedestrians
Manchester


Rochdale
Bury
Bolton
Oldham
Wigan
Salford
Tameside
Trafford
Car Occupants
Stockport
Manchester

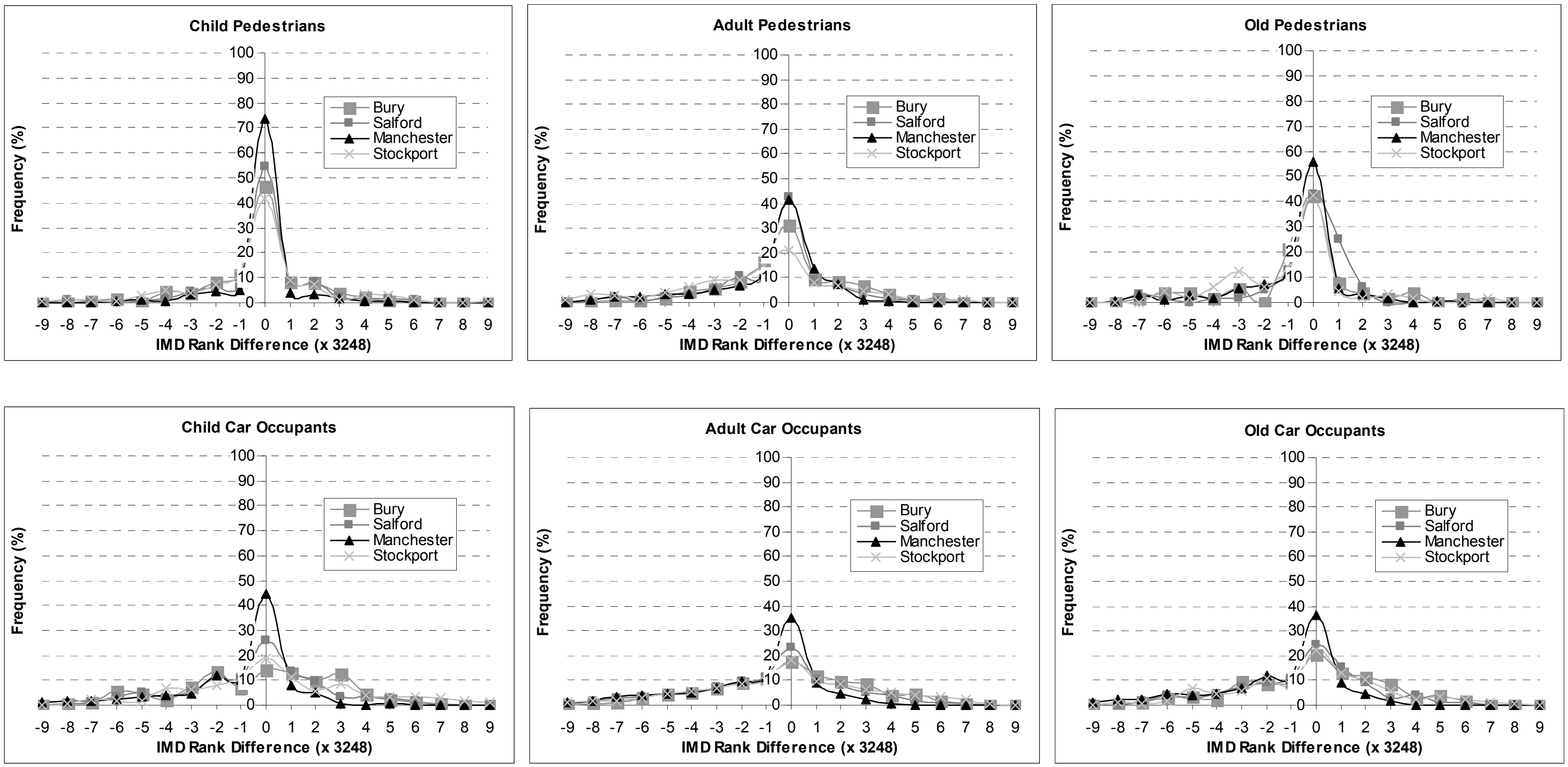

Child Pedestrians
Adult Pedestrians
Old Pedestrians
Child Car Occupants
Adult Car Occupants
Old Car Occupants
Frequency (%)
IMD Rank Difference (x 3248)
Bury
Salford
Manchester
Stockport

Figure 5

Table 1 Domain weights for the IMD 2004

Index	**Domain Weight**
Income Deprivation	22.5%
Employment Deprivation	22.5%
Health deprivation and disability	13.5%
Education, skills and training deprivation	13.5%
Barriers to housing and services	9.3%
Crime	9.3%
Living Environment and deprivation	9.3%

Table 2 Number of Pedestrian and Car Occupant casualties (1999-2003) with the residence postcode information.

Casualty Type	Salford	Manchester	Bury	Stockport
Child Pedestrian	245	671	184	190
Adult Pedestrian	203	1031	164	226
Old Pedestrian	63	182	52	64
Child Car Occupant	180	501	162	172
Adult Car Occupant	2296	5625	1699	1840
Old Car Occupant	166	456	167	164