Accident data study in support of development of autonomous emergency braking test procedures

PROJECT REFERENCE: LUEL 5989/6175

Undertaken for

Thatcham Research Centre

Insurance Institute of Highway Safety

Prepared by

J. Lenard and R. Danton

23 December 2010

Loughborough University Enterprises Limited

Loughborough University

THE QUEEN'S ANNIVERSARY PRIZES

FOR HIGHER AND FURTHER EDUCATION

Awarded to Loughborough University 2007
in recognition of research into vehicle, road and driver safety

Table of contents

1 Summary

There is currently considerable interest in the development of physical test procedures for autonomous braking systems and one consideration is that the test conditions should be realistic, i.e. they should resemble or have a known relationship to the circumstances under which real road accidents occur. In this context, the aim of the work presented in this report is to describe typical circumstances for pedestrian, rear-end and head-on impacts based on British accident data. The scope of this work is limited to the description of accident data and although this report is intended as a contribution to the development of test conditions, it contains no discussion of how the results from the accident data should be used for this purpose.

Two sources of information were available for analysis with the permission of the Department for Transport: the national accident database STATS 19 (2008) and the in-depth On-the-Spot database (2000–2009). The former provides a comprehensive coverage of the British road casualty population while the latter, compiled by research teams at the scene of accidents, provides a very detailed description of a representative sample of accidents (500 p.a.). Following the selection of cases from the two databases and the derivation of smaller datasets containing a reduced number of relevant fields, a data-mining technique known as cluster analysis was applied to partition the databases into groups of similar accidents, where the concept of 'similar' is provided with a mathematical definition. This computational procedure has the advantage of being fully transparent, objective and reproducible. It can also identify correlations or patterns among multiple parameters that are difficult or impossible to recognise in two- or three-dimensional tables and charts.

The source materials for cases from the On-the-Spot project (OTS)—photographs, videos, scene plans, sketches and text—were reviewed individually in detail with a view to specifying the position and movement of the road users in the seconds before impact as precisely as possible using techniques of accident reconstruction. Applied to 175 pedestrian accidents, 50 rear-end impacts and 50 head-on impacts, these case-by-case reviews constituted a considerable portion of the total resources applied to the work.

The extent to which STATS 19 (2008) and OTS (2000–2009) contain the same accidents cannot be determined exactly because the databases are anonymised; however the overlap is relatively small, involving approximately one-third of accidents that occurred in 2008 in the two OTS sample regions, South Nottinghamshire and Thames Valley. For this reason certain parallels that are discernible in the results of the cluster analyses are all the more

remarkable. For pedestrian accidents, both cluster analyses highlighted (a) a 'baseline' scenario where a pedestrian, visible to the driver, steps out from kerb in favourable light and weather conditions, (b) an otherwise similar situation where a smaller pedestrian emerges from behind a vehicle or object that at least partially obscures the driver's line of sight and (c) a scenario with a larger, adult pedestrian crossing the road in adverse meteorological conditions. For rear-end accidents, the most prominent scenarios were (a) striking a stationary vehicle at a roundabout or junction and (b) impacts on higher speed roads that could be related to congested, stop-go traffic conditions or to tailgating. For head-on accidents, STATS 19 highlighted collisions at junctions where one vehicle was often turning across the path of the other—this type of accident was not included in the OTS sample for analysis. Apart from this, both databases featured accidents on bends, accidents involving overtaking and accidents not involving junctions, bends or overtaking. The full details and essential results of this report are contained in Table 10 and Table 22 (pedestrian), Table 35 and Table 45 (rear-end), and Table 56 and Table 63 (head-on).

A limitation to the scope of these results is that they are based on British data. The frequency with which a certain event or combination of factors occurs is naturally dependent on the local road environment, vehicle fleet, driver characteristics and various social and legal factors. At a different level, the formation of clusters is determined in a substantial part by the fields on which accidents are compared. In this work, fields relevant to physical testing were chosen such as lighting, precipitation, vehicle speed, pedestrian crossing direction and vehicle separation. Not included were such things as the age and sex of the driver or the time of day of the accident, even though there could well be patterns in how these factors were correlated with other accident characteristics, for example a higher exposure of female drivers to pedestrian accidents involving children in the morning and afternoon 'school runs'. A further consideration relating to fields is that the number of fields that can be used in a cluster analysis is, or should be, limited by the number of cases. This was most applicable to the OTS rear-end and head-on analyses, each of which was based on 50 cases. Finally it should be noted not all of the relevant cases for STATS 19 rear-end and head-on accidents could be accommodated by the computer software and hardware used, 'memory overload' restrictions placing an upper limit of around 11,000 cases that could be processed. This was dealt with by an arbitrary (but reproducible) selection of cases for analysis. At the other extreme, resource limitations meant that the number of OTS rear-end and head-on cases was capped at 50, as just mentioned, and the possibility of other patterns emerging from a consideration of more cases cannot be excluded. For pedestrian accidents, on the other hand, it was technically possible to include all relevant cases from STATS 19 in the cluster analysis and resources were available to review and include the majority of relevant OTS cases for which sufficient information was available in the case files.

2 Introduction

Autonomous emergency braking (AEB) is one of a number of modern safety technologies designed to prevent or mitigate the severity of vehicle impacts. There is scope for considerable variation among the AEB systems installed on different makes and models of vehicles as each manufacturer can make independent decisions on which sensors to fit, what decision logic to program into the control units, and how and when to warn the driver and activate braking. For this reason there is an interest in conducting physical tests to assess performance, compare systems and inform consumers.

The setting of test conditions involves many considerations, one of which is to subject the vehicles to realistic accident conditions, i.e. circumstances that are encountered in real accidents. This provides the focus and scope of the work described in this report: to describe common accident scenarios based on factual data.

Three collision types are of specific interest: pedestrian, rear-end and head-on accidents. Typical conditions for each of these accident categories were derived separately from two data sources, STATS19 and OTS (described below), using a data mining technique known as cluster analysis, making a total of six analyses in all. A considerable amount of preparatory work was conducted on one of the databases, OTS, to quantify the position and velocity of the road users in the seconds before impact.

This report devotes a chapter to each of the three accident categories and each chapter contains three main sections: a cluster analysis of STATS19, a detailed review of OTS, and a cluster analysis of OTS. Before this, there is a single chapter describing the source databases and the common principles of the mathematical algorithm used to compute the accident clusters.

3 Materials and methods

3.1 Source databases

Two major sources of factual information about accidents in Britain were used in this work: the national accident database STATS 19 and the in-depth On-the-Spot study (OTS). Some key facts about these databases are provided in Table 1. Further information about OTS can be obtained from Cuerden et al (2008) and about STATS 19 from DfT (2004), both of which are available on the internet.

	STATS 19	OTS
Period	2008	2000–2009
Sample region	Great Britain	South Nottinghamshire and Thames Valley
Collected by	Police	Research teams at Loughborough Uni. and TRL
Purpose	National statistics	Detailed information to support casualty reduction programme
Method	Police reports	At-scene investigations
Inclusion criterion	Casualty on public road	Police attendance; rotating 8-hour shift (includes non-injury).
Number of accidents	171,000	500 p.a.

Table 1 Source databases STATS 19 and OTS

The amount of information in both STATS 19 and OTS exceeds the requirements and resources of the current work. For this reason a summary dataset containing relevant cases, a reduced set of variables and a simplified set of categories was prepared for each accident type from the source databases.

In addition there was an intermediary step for OTS, the in-depth database. The accidents selected for inclusion in the study were individually reviewed in detail and the pre-impact location and movement of the colliding parties were reconstructed from the evidence on record in the case files—including drawings, photographs and film—to achieve the highest level of confidence possible. In effect, whereas the summary datasets used for the cluster analyses of STATS 19 were derived directly from the source files by programmed computer logic, the summary datasets used for OTS were compiled by the reviewers based their assessment of the full range of materials contained in the OTS case files.

A note on the relationship between the STATS 19 and OTS databases. The national accident database STATS 19 effectively defines the whole reported road accident population for the year 2008. The in-depth accident database OTS contains a sample of cases from just two regions, South Nottinghamshire and Thames Valley, but over a greater period of time, 2000–2009. In addition, unlike STATS 19, OTS includes some non-injury accidents. There should consequently be some coverage of the same accidents, namely a third (roughly) of the casualty accidents that occurred in the two OTS sample regions in 2008; however this overlap constitutes a distinct minority of both databases. Furthermore, as an in-depth database, OTS contains more information about accidents than STATS 19, especially the quantitative information about velocity, location, injuries and causal factors that is based on accident investigation, reconstruction and follow-up data collection. The summary datasets prepared for the cluster analyses contain the most suitable fields of those available in each dataset. This means, for example, taking vehicle speed from OTS but speed limit, the best available proxy, from STATS 19. It is undesirable for technical reasons to include variables in a cluster analysis that are (necessarily) strongly correlated and it is also advised to restrict the number of variables according to the number of records. For these reasons, the datasets derived from STATS 19 and OTS contain only a partial overlap in (a) the variables used to describe the same type of accident and (b) the accidents covered. Each pair of cluster analyses undertaken for each of the three accident types can therefore be regarded as having largely independent sources of information, the main link being that OTS was designed to be representative of the accident population as far as possible within the scope of the study (with the deliberate exception of non-injury accidents).

3.2 Cluster analysis

The basic method used in this report to move from the accident data to formulation of accident scenarios is cluster analysis, in particular the hierarchical, ascending (agglomerative) variety. This works by grouping together the most similar records of a dataset, where the notion of similarity is defined mathematically. As applied in this report, each record describes an accident and so cluster analysis identifies groups of similar accidents. These groups or clusters have (by definition) common characteristics and can be interpreted as constituting accident scenarios.

The basic operation of the method can be illustrated by reference to a simple example. Figure 1 represents five points in a Euclidean plane. The relevant dataset in this case consists of five records, each of which is defined by an X and Y value. The intuitive notion of similarity is provided by the distance between the points, where closer together means more similar. Given the X and Y values of each point, the distance or (dis)similarity between any two of them is defined mathematically using elementary algebra.

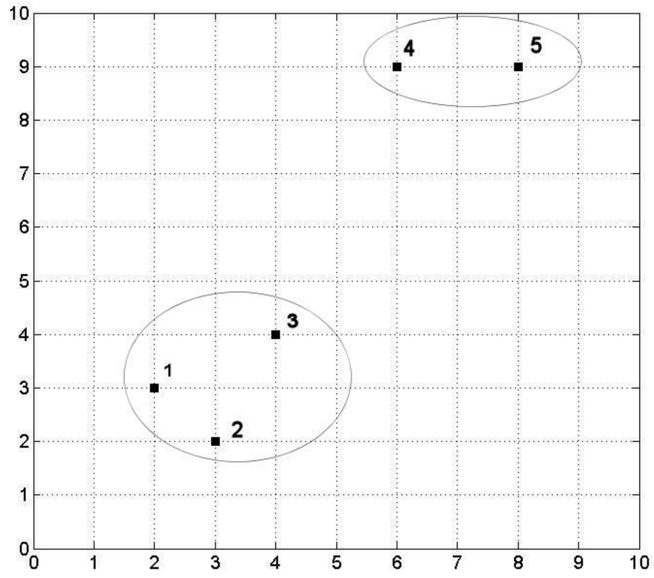


Figure 1 Illustrative dataset

The hierarchical, ascending method of cluster analysis begins by defining each point as a mini-cluster, thereby beginning with five clusters in this example. It then joins the two nearest points into a single cluster, reducing the number to four. The two nearest of these four clusters are then merged, reducing the number to three. And so the process loops until all of the points are united in a single cluster.

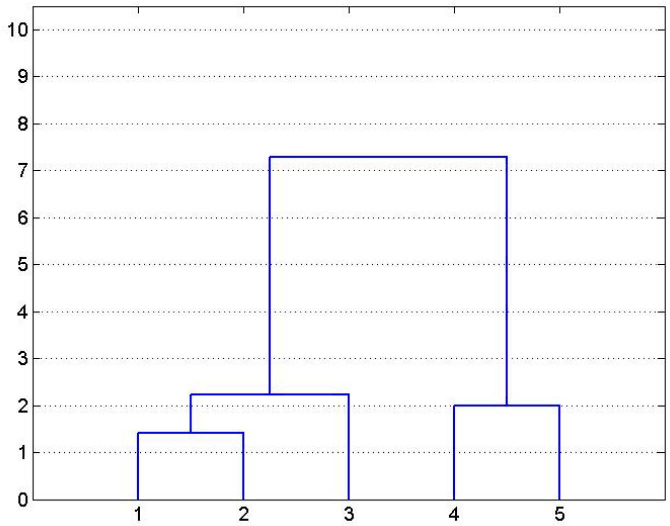


Figure 2 Illustrative dendrogram

This iterative process is conventionally represented in a dendrogram like that shown in Figure 2. Here the original five points (records) are labelled on the horizontal axis and each merger is represented by a horizontal line. The vertical axis represents the distance apart of points or clusters when they are merged. It can be seen that points 1 and 2 are the first to combine (distance 1.4 units), followed by points 4 and 5 (distance 2 units). Next the cluster formed by points 1 and 2 is merged with point 3, and finally the cluster 1-2-3 is merged with 4-5. The relatively long distance on the dendrogram preceding the merger of the clusters 1-2-3 and 4-5 indicates a ‘natural’ break between these two groups. Such breaks can be identified mathematically rather than visually and, due to the large datasets used in this work and correspondingly cluttered dendrograms, the mathematical approach proved more suitable for the STATS 19 and OTS accident analyses. An important point to note, however, is that the hierarchical cluster analysis provides every number of clusters from one for each case to one for the whole population. Each set of clusters is a valid representation of the dataset and it is up to the user to decide where to ‘cut of the tree’, i.e. to decide which set of clusters is most informative for the purposes of a particular research question or investigation.

This illustrative example uses points that can be easily visualised on a two-dimensional spatial plane. This works because each record (point) of the dataset has just two characteristics, an X and Y value. In contrast, the simplified accident datasets used in this work have 5–10 characteristics, including for example light conditions, weather, speed limit, vehicle speed and pedestrian location. This is too much to represent in a two- or three-dimensional space; however the idea of a distance or (dis)similarity between two accident records can be extended in a way that retains the analogy to the spatial distance between points in a multi-dimensional space. This is detailed below, following a consideration of the derived datasets used for the cluster analyses.

3.3 Derived datasets for cluster analysis

The limited datasets provided for cluster analysis needed to be carefully prepared and structured. As the sole carriers of information into the number-crunching process, they clearly play a major role in determining the clusters that are eventually obtained. The common principles of data preparation carried out for each of the six analyses described in this report can be illustrated by reference to the OTS pedestrian dataset.

Ordinal	Pedestrian severity	0.0 0.5 1.0	Slight Serious Fatal
Nominal	Light conditions	1 2	Daylight Darkness
Nominal	Weather	1 2	Fine Not fine
Nominal	Vehicle manoeuvre	1 2	Going ahead Turning
Ordinal	Pedestrian age-sex	0.00 0.33 0.67 1.00	Child 0–7 years Child 8–15 years Adult female Adult male
Nominal	Pedestrian movement	1 2 3	Crossing from left Crossing from right Stationary or along carriageway
Nominal	Pedestrian speed	1 2	Walking Running
Nominal	Line of sight (1 sec)	1 2	Not obstructed Obstructed
Scale	Vehicle travel speed	0–1	*scaled from km/h*
Scale	Change of speed to impact	0–1	*scaled from km/h*

Table 2 Dataset for OTS pedestrian analysis

This dataset contains ten fields or variables. Each of these is categorised as nominal, ordinal or scale. Books have been written on these distinctions; however the basic idea is that scale variables are continuous parameters measured in units such as kilometres per hour or metres, ordinal variables are categories that have a natural order such as size or severity, and nominal variables are simply categories without a natural order such as vehicle type or vehicle manoeuvre. Although seemingly abstruse, the concept of nominal, ordinal and scale variables is relevant to the mathematical definition of how similar two accidents are to each other, a topic that is taken up in the next section.

In the case of OTS, as mentioned above, the input datasets were prepared manually following individual case-by-case reviews. This was feasible because of the manageable number of cases (175 pedestrian, 50 rear-end and 50 head-on). For STATS 19, the national database containing thousands or tens of thousands of relevant cases, the datasets for input to cluster analysis were generated by computer, usually by combining the available categories or values for a selected variable into a smaller number of options. This is illustrated in Table 3 below. The pedestrian age-sex variable, which is intended to correlate with pedestrian size (height and weight), is an exception in being derived from two separate fields in STATS 19.

Daylight—lights present Daylight—no lighting Daylight—lighting unknown	Daylight
Darkness—lights lit Darkness—lights unlit Darkness—no lighting Darkness—lighting unknown	Darkness

Table 3 Simplification of field values

In order to compute the similarity between accident records, the cluster analysis algorithm requires the values of each variable to be expressed in numbers rather than words, e.g. 1 for ‘daylight’ and 2 for ‘darkness’. The categories of each nominal variable were assigned the numbers 1, 2, 3,… while the categories for ordinal variables and the values of scale variables were scaled to the range 0–1. This is illustrated in Table 4.

Field name	Field type	Numeric value	Field value
Light conditions	Nominal	1 2	Daylight Darkness
Vehicle manoeuvre	Nominal	1 2 3	Going ahead Turning Other
Speed limit (mph)	Ordinal	0.0 0.5 1.0	10–30 40–50 60–70
Pedestrian age-sex	Ordinal	0.00 0.33 0.67 1.00	Child 0–7 years Child 8–15 years Adult female Adult male
Vehicle speed	Scale	0.0 0.2 0.8 1.0	40 mph 50 mph 80 mph 90 mph
Following distance	Scale	0.000 0.125 0.500 0.750 1.000	3 m 4 m 7 m 9 m 11 m

Table 4 Standardised assignment of numbers to field values

The pattern for nominal and ordinal variables should be fairly apparent from these examples. Mathematically, if *n* is the number of categories, then the values of nominal fields are 1, 2, 3,… *n* and the values of ordinal fields are 0, 1/(*n*-1), 2/(*n*-1),… 1.

The standardisation of scale fields adopted for this work is more complicated as it depends on the maximum and minimum values in the database according to the formula $x'=(x-x_{min})/(x_{max}-x_{min})$ where x′ is the scaled value. Thus if 40 mph and 90 mph are the minimum and maximum speeds in a dataset, the standardised value of 50 mph is calculated as x’=(50-40)/(90-40)=0.2.

There are other ways to standardise nominal, ordinal and scale variables. The methods used here are geared to provided equal weight to each variable in the heterogeneous accident records in the sense of attributing a distance (dissimilarity) of 1 to the most different values of any field. The mechanism by which this works is described in more detail in the following section.

A final aspect of the formation of the dataset is the selection of cases. This is naturally different for pedestrian, rear-end and head-on accidents but also varies between STATS 19 and OTS for each given accident type due to the (necessarily) limited information available on the national database. For instance STATS 19 records the surface (front, left, right or rear) on which first impact was incurred, but leaves open whether other impacts occurred on these or other surfaces. It is known from in-depth studies that a significant proportion of vehicles incur multiple impacts and that the most severe impact, which would most naturally define the collision type, may occur on the second or subsequent event. Nonetheless STATS 19 is exceptional among national databases in its level of detail and the disruptive effect of multiple impacts probably blurs without materially changing the fundamental results. The selection criteria for each accident type and each database are described in detail in the relevant sections.

The algorithm for the agglomerative cluster method begins by calculating the distance (dissimilarity) between each pair of records in the dataset. For even just five records, there are ten pairs of records: 1-2, 1-3, 1-4, 1-5, 2-3, 2-4, 2-5, 3-4, 3-5 and 4-5. The distances are symmetrical, i.e. the distance from 1 to 2 is the same as from 2 to 1 and each record is zero distance from itself—this covers all possibilities. In general terms, for N records there are 1/2*N*(N-1) combinations. This means that for 1000 records there are 499,500 combinations. Off-the-shelf software packages such as Matlab or SPSS store this in memory which can lead to memory overload problems. It was found that around 10,000 to 11,000 cases was close to the maximum that could be processed using standard office computers and software, requiring up to six hours of processing time. For this reason it was necessary to reduce the number of cases in the STATS 19 datasets for rear-end and head-on accidents. This was done by sorting the files in a reproducible manner and deleting every second or third eligible case, depending on the reduction required. Details are provided in the relevant sections.

3.4 Algorithm for cluster analysis

The derivation of clusters from the datasets provided for analysis is a purely mathematical procedure once the algorithm has been specified. There are guidelines and reasons for specifying the computational algorithm in a certain way, but once this is decided, the groups follow with mathematical certainty from the input data. The effects of varying the algorithm are impossible to foresee and the strongest trends or patterns in the data appear to be relatively insensitive to the details of the computation, i.e. they continue to present themselves irrespective of variations in the technical specification of the algorithm. The intention in this section is to provide enough description of the common elements of the algorithm used for all six cluster analyses that an independent analyst could reproduce the same results from the same datasets. The datasets themselves are described in detail in the chapters devoted to each accident type.

The heart of algorithm lies in the mathematical definition of the dissimilarity or distance between accident records. In the earlier example of points in a two-dimensional space, spatial separation served as the measure of the similarity of points: the points were clustered on this basis. In fact there are three elements to a full computation:

● the distance between two values of a given variable
● the distance between two given records
● the distance between two given clusters.

Each of these requires an independent definition; however since a record is nothing more than a set of values and a cluster is nothing more than a set of records, each step of computation adds something new while at the same time building upon the previous stage.

These ideas and the exact technical specification of the algorithm can be stated in purely mathematical terms, but the most effective form of communication in this context may be by an illustrative example.

		Record 1		Record 2		Distance between values
Nominal	Light conditions	1	Daylight	1	Daylight	0.00
Nominal	Vehicle manoeuvre	3	Other	1	Going ahead	1.00
Ordinal	Speed limit	0.0	10-30 mph	1.0	60-70 mph	1.00
Ordinal	Pedestrian age-sex	1.00	Adult male	0.67	Adult female	0.33
Scale	Vehicle speed	0.2	50 mph	0.8	80 mph	0.60
Scale	Following distance	1.0	11 m	0.0	3 m	1.00
				Distance between records		3.93

Table 5 Distance between values and records

This example illustrates that for nominal variables, the distance or dissimilarity between two values is 0 if the values are identical and 1 if they are not identical; for ordinal and scale variables, assuming their values have been transformed into the range 0–1 as discussed in the previous section, the distance between two values is the (positive) numerical difference obtained by subtraction. On this treatment, the maximum distance or dissimilarity between two values of a given variable is 1 and the minimum distance, when two values are identical, is 0.

Also illustrated in Table 5 is that the distance between two records is computed as the sum of the distances between their constituent values. This is known as the city block, Manhattan or rectilinear distance. Although it is highly correlated with Euclidean distance ('as the crow flies') as used in the two-dimensional space above, the city block measure ('taxi driver in Manhattan') is not identical. This can be seen by consideration of two points with the Cartesian co-ordinates (0,0) and (3,4): the Euclidean distance is 5 whereas the city block distance is 7. The main rationale for choosing the city block measure is simplicity, given that the numerical values of nominal and ordinal fields in the accident records do not have the same standing as the specification of an exact location in a multi-dimensional Euclidean space.

The third independent element of the algorithm, which is not covered in Table 5, is how the distance between two clusters, i.e. two sets of records, is computed. Looking back again to Figure 1, there are several options. One is to say that two clusters are as close together as their nearest points; another is to say that they are as far apart as their furthest points. An intermediate option known as the 'unweighted pair-group method using arithmetic averages' was adopted for this work. This takes the average distance between each pair of records, one from each cluster. Where, as in the example, there are three points (1, 2, 3) in one cluster and two points (4, 5) in the other, there are six combinations of points: 1-4, 2-4, 3-4, 1-5, 2-5 and 3-5. The average of the distances between these six pairs of records is defined as the distance between the two clusters.

This concludes the description of computational algorithm, which should in principle be sufficient for the clusters computed for this report to be reproduced independently if the same datasets were provided as input. (The order of cases in the input dataset should make no difference.) It remains to state briefly how the number of clusters for each analysis was determined. As already explained, the hierarchical cluster analysis begins with one cluster for each record and ends with one cluster for the whole dataset. No particular set of clusters from the many to the one is right or wrong: each is a valid representation of the data. The question is rather the usefulness of a set of clusters for a particular purpose. Clearly neither extreme—one for each record or one for the whole population—is very meaningful. For the purpose of contributing to the design of AEB testing procedures, it is relevant to have a relatively small number of clusters that covers much of the population. To this end some supplementary code was written to assist in the identification of around 6 clusters to comprise about 75–80% of the population, including the fatal and seriously injured subpopulations. In conjunction with some further code to identify ‘natural’ gaps between the clusters, the number of clusters for each accident type and source database was chosen manually after inspection of several options identified programmatically.

4 Pedestrian accidents

4.1 Cluster analysis of STATS19

4.1.1 Selection of cases

The casualty file for STATS 19 (2008) contains information on 230,905 road users, among whom are 28,482 pedestrians. There is provision to nominate a vehicle with which each pedestrian interacted. These constitute the pool from which the cases of interest for pedestrian accidents were drawn.

Driver or rider	144941
Passenger	57482
Pedestrian	28482
Total	230905

Table 6 Road user casualties

The selection criteria for the inclusion of cases from a source database serve in practice as the definition of the accident type. The primary criteria for the selection of pedestrian accidents from STATS 19 were:

- passenger cars and taxis associated with a pedestrian casualty
- first point of impact on the front surface.

This yielded a total of 13,257 vehicles: 12,662 passenger cars and 595 taxis (Table 7).

	First point of contact						
	None	Front	Back	Right	Left	Unk.	Total
Pedal cycle	11	223	5	11	11	0	261
Motor cycle to 50 cc	13	157	2	10	23	0	205
Motor cycle 51-125 cc	15	305	4	24	37	1	386
Motor cycle 126-500 cc	5	118	3	10	15	0	151
Motor cycle over 500cc	14	251	3	14	37	0	319
Taxi/private hire car	84	595	80	117	232	0	1108
Car	1252	12662	1717	2124	4241	10	22006
Minibus (8-16 seats)	5	46	5	16	24	0	96
Bus or coach (17+ seats)	162	768	27	88	498	1	1544
Other motor vehicle	66	196	58	36	113	1	470
Other non-motor vehicle	0	15	3	0	0	0	18
Ridden horse	1	0	0	0	0	0	1
Agricultural vehicle	3	6	2	1	3	0	15
Tram	0	2	1	0	4	0	7
Goods vehicle to 3.5 t	91	542	253	123	339	0	1348
Goods vehicle 3.5-7.5 t	19	72	20	12	68	0	191
Goods vehicle over 7.5 t	40	125	46	21	103	1	336
Unknown	1	7	2	1	4	5	20
Total	1782	16090	2231	2608	5752	19	28482

Table 7 Vehicle type and impact surface for pedestrian accidents

A secondary filter was made of (a) vehicles that were parked or reversing and (b) records with any missing or unknown information in the key fields (Table 9). This resulted in a drop in the number of cases from 13,257 to 10,574, the main contributor being unknown pedestrian movement (2,263). Table 8 shows the pedestrian injury severity for the two groups, providing a check on the number of killed or seriously injured casualties lost in this secondary filter. The proportions are reasonably well balanced among the fatal, serious and slight categories and, as a practical matter, 13,257 was beyond the capacity of the computer hardware and software to process while 10,574 was possible, albeit at the upper limit. Rather than preserve these cases through the introduction of an ‘Unknown’ category for pedestrian movement and then have to filter the dataset anyway to avoid memory overload—thereby eliminating records with full information—the reduction in numbers was accepted.

	Filtered	Available
Fatal	240	319
Serious	2463	3022
Slight	7871	9916
Total	10574	13257

Table 8 Secondary filter of vehicles that struck pedestrians with front surface

4.1.2 Input dataset

The structure of the dataset derived from STATS 19 for pedestrian accidents to serve as the input to the cluster analysis is shown in Table 9. The parameters were chosen, as in all six analyses, (a) for their relevance to the formulation of physical testing procedures and (b) in response to the actual data. So, for instance, there are probably peak periods of the day when children suffer pedestrian accidents, e.g. on the trips to and from school; however the link to time of day was not considered relevant to the development of testing procedures. Secondly, if a category like ‘Adult male’ had very few numbers, the whole field ‘Pedestrian age-sex’ would in effect contract to the numerical range 0.00–0.67, lessening the contribution of this field to the dissimilarity between records compared to other fields that exploit the full range 0–1. There are many other ways in which the structure of a dataset is tweaked to suit its contents—this is part of the art of data mining to unearth patterns of interest—but apart from mentioning that this was a real factor in the design of the datasets used in this work, the details are not of foremost importance.

Ordinal	Pedestrian severity	0.0 0.5 1.0	Slight Serious Fatal
Ordinal	Speed limit (mph)	0.0 0.5 1.0	10–30 40–50 60–70
Nominal	Light conditions	1 2	Daylight Darkness
Nominal	Weather	1 2	Fine Not fine
Nominal	Vehicle manoeuvre	1 2 3	Going ahead Turning Other
Ordinal	Pedestrian age-sex	0.00 0.33 0.67 1.00	Child 0–7 years Child 8–15 years Adult female Adult male
Nominal	Pedestrian movement	1 2 3	Crossing from left Crossing from right Stationary or along carriageway
Nominal	Masked by vehicle	1 2	Not masked Masked

Table 9 Structure of dataset for pedestrian accidents (STATS 19)

4.1.3 Results

The outcome of the cluster analysis is shown in Table 10 at the level where the accident population was gathered into 23 groups. The characteristics of six clusters that comprise 75% of the population are shown in detail. Cells shaded in green indicate (a) that the distribution of numbers for the given field is significantly different from the distribution in the whole population (chi-square test to 99.5% significance) and (b) that the particular number highlighted is over-represented. To take an example, Cluster 1 contains 4,134 cases and all of these occurred in daylight compared to a distribution of 7,055 daylight and 3,519 darkness in the overall 10,574 population. The probability that this would happen by chance is less than 0.05% and the number 4,134 is over-represented. The figures on cluster representativeness are derived directly from pedestrian injury severity, representing the latter as row percentages. This is useful in showing for example that Cluster 3, which comprises 12% of the overall population, contains 23% of the pedestrian fatalities. It can therefore be recognised as a particularly dangerous scenario.

	Cluster							
	1	**2**	**3**	**4**	**5**	**6**	**7–23**	**Total**
Cluster representativeness (%)								
Slight	41	15	11	8	7	3	16	100
Serious	35	14	15	13	5	4	15	100
Fatal	24	4	23	19	3	14	14	100
Total	39	14	12	9	6	3	15	100
Pedestrian severity								
Slight	3209	1156	880	619	559	214	1234	7871
Serious	868	333	361	311	120	107	363	2463
Fatal	57	10	54	45	7	33	34	240
Total	4134	1499	1295	975	686	354	1631	10574
Speed limit (mph)								
10-30	3798	1453	1163	877	665	251	1500	9707
40-50	227	44	102	72	17	29	72	563
60-70	109	2	30	26	4	74	59	304
Total	4134	1499	1295	975	686	354	1631	10574
Light conditions								
Daylight	4134	1499	0	0	672	0	750	7055
Darkness	0	0	1295	975	14	354	881	3519
Total	4134	1499	1295	975	686	354	1631	10574
Weather								
Fine	3980	1498	920	708	684	281	683	8754
Not fine	154	1	375	267	2	73	948	1820
Total	4134	1499	1295	975	686	354	1631	10574
Vehicle manoeuvre								
Going ahead	4134	1499	1295	975	0	346	1009	9258
Turning	0	0	0	0	686	8	622	1316
Total	4134	1499	1295	975	686	354	1631	10574
Pedestrian age-sex								
Child 0-7 yrs	451	348	30	28	37	3	108	1005
Child 8-15 yrs	1425	630	239	153	77	31	420	2975
Adult female	1091	245	340	262	313	70	542	2863
Adult male	1167	276	686	532	259	250	561	3731
Total	4134	1499	1295	975	686	354	1631	10574
Pedestrian movement								
Crossing from left	2459	860	1295	0	431	0	963	6008
Crossing from right	1369	606	0	975	216	0	609	3775
Stationary or along	306	33	0	0	39	354	59	791
Total	4134	1499	1295	975	686	354	1631	10574
Masked by vehicle								
Not masked	4134	0	1295	975	686	353	880	8323
Masked	0	1499	0	0	0	1	751	2251
Total	4134	1499	1295	975	686	354	1631	10574

Table 10 Clusters for pedestrian accidents (STATS 19)

The highlighting of cells in green assists in four ways to interpret the clusters. Firstly, where all of the cases fall into a single category, the cluster can be thought of as “purely” something. For example in Cluster 1 all of the accidents occurred in daylight, all vehicles were going ahead and all pedestrians were not masked by a vehicle. As the starting point in building up the concept of a scenario based on this cluster, these characteristics are unambiguous. Secondly, where a category or natural group of categories is over-represented and constitutes a majority of the cases, it also naturally lends its character to the cluster. In Cluster 2 the vast majority of accidents occurred in a 10–30 mph speed zone where the pedestrian was either crossing from the left or from the right. Thirdly, where a category or natural group of categories is over-represented but constitutes a minority of the cases, this can be thought of as a tendency. In Cluster 6, serious and fatal casualties are strongly over-represented along with the higher speed limits 40–50 and 60–70 mph. It can therefore be understood as like the general population in these characteristics except with a significantly greater tendency towards higher injury outcomes in higher speed zones. Finally, where no cell is shaded, the column of numbers for a given characteristic is not significantly different from the overall population. This can be seen in the speed limit zones of Cluster 4.

The results in Table 10 are presented precisely and succinctly and it would not necessarily be informative to re-express these in words. A few ‘higher level’ observations may however be of interest. The two largest clusters, 1 and 2, mostly magnify the dominant characteristics of the overall population (slight, 10-30 mph, daylight, fine, going ahead and pedestrian crossing) with two exceptions, (a) an over-representation of children and (b) in cluster 2, the pedestrian being masked. Clusters 3 and 4, on the other hand, are weighted towards killed or seriously injured cases, occur in darkness with a tendency towards wet weather and adult males who are not masked, the really substantial difference between these two clusters being that the pedestrian was crossing from the left in one case and from the right in the other. Cluster 5 introduces a turning scenario at low speeds and low injury outcomes, mostly matching the dominant features of the overall population except for the over-representation of adults. Apart from the higher severity levels and speed zones in Cluster 6 just mentioned, it is worth noting that this group of accidents occurred in darkness with mostly adult men who were stationary in or moving along the carriageway. This is the only major cluster not dominated by pedestrian movement across the carriageway.

4.2 Detailed OTS case reviews

This section presents the results obtained from individual case reviews of 175 car-to-pedestrian accidents. The case reviews found that the predominant factor in pedestrian accidents was the pre-collision movement of the pedestrian and more specifically the pedestrian crossing action. The review process found that 59% (103) of the accidents occurred when the pedestrian was crossing the road from the nearside (left) of the road, 37% (64) with the pedestrian crossing from the offside (right) and 4% (8) when the pedestrian was either walking with the traffic or stepped into the carriageway without a known reason.

The pedestrian pre-impact movement in regards to walking or running was another important factor, as a running pedestrian would emerge in front of a car in a shorter time than a walking pedestrian, this giving less reaction and application time for the driver. The following chart (Figure 3) gives the correlation of pedestrian action in regards to crossing the road and general pedestrian movement of walking or running prior to impact.

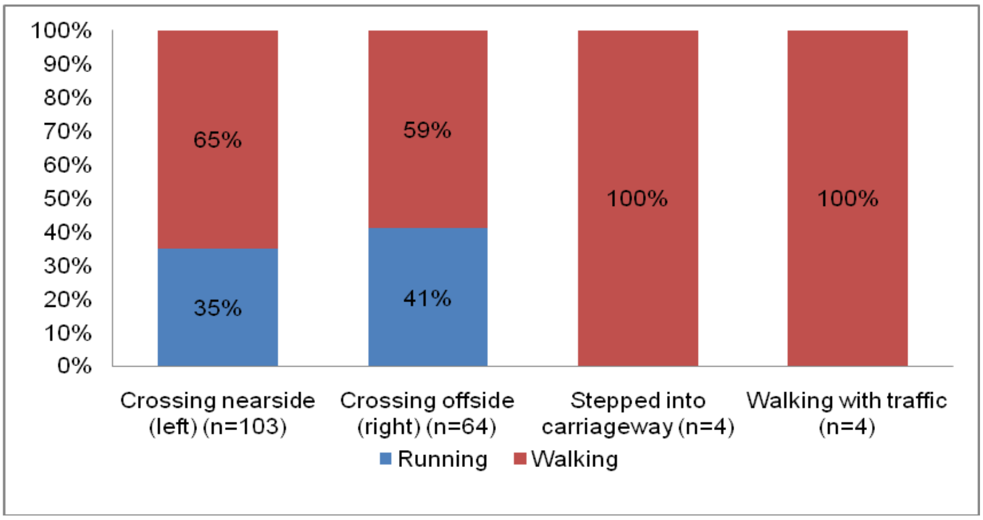


Figure 3: Pedestrian crossing action against pedestrian pre-impact movement type

In the pedestrian accidents reviewed 113 (64%) pedestrians were walking before the impact. If the pedestrian was crossing from the nearside (left), it was found that in 65% of those accidents the pedestrian was walking prior to the impact, 59% when crossing from the offside.

The reviewed pedestrian accidents predominantly occurred in an urban road environment 94% (165) with only 6% (10) occurring on rural roads. This is expected as the exposure of pedestrians in an urban environment is greater than in a rural environment.

The demographics of the pedestrians involved in the reviewed accidents are presented in the following tables.

	Female	Male	Unknown	Total
Crossing nearside (left)	30	73	0	103
Crossing offside (right)	29	34	1	64
Stepped into carriageway	3	1	0	4
Walking with traffic	1	3	0	4
Total	63	111	1	175

Table 11: Pedestrian crossing action by gender

The largest gender group involved in the pedestrian accidents were males crossing from the nearside (left) which formed 42% (73) of the accident sample, with males forming 63% (111) of the total sample.

	Fatal	Serious	Slight	Non-injury	Total
Crossing nearside (left)	5	33	64	1	103
Crossing offside (right)	4	15	43	2	64
Stepped into carriageway	0	3	1	0	4
Walking with traffic	1	3	0	0	4
Total	10	54	108	3	175

Table 12: Pedestrian crossing action by overall accident severity

Table 12 shows that in the sample of pedestrian accidents reviewed 64 (37%) of the accidents had a severity of killed or seriously injured (KSI). The following table gives injury data by the Maximum Abbreviated Injury Score (MAIS).

	MAIS							
	No injury	1	2	3	4	5	Unknown	Total
Crossing nearside (left)	3	44	25	16	6	3	6	103
Crossing offside (right)	3	28	14	6	3	2	8	64
Stepped into carriageway	0	1	1	2	0	0	0	4
Walking with traffic	0	1	2	0	0	0	1	4
Total	6	74	42	24	9	5	15	175

Table 13: Crossing action by pedestrian MAIS

Table 13 shows that in 80 (46%) of the accidents the pedestrian suffered a MAIS 2+ injury, with 50 (28%) of those injuries occurring when the pedestrian was crossing from the nearside (left).

Circumstances which can affect the nature of pedestrian accidents can also include the age of a pedestrian or more specifically the anthropometrics of the pedestrian. This not only affects the injury sustained by the pedestrian but can also affect the pre-impact situation, for example smaller pedestrians are susceptible to increased sight obstructions thus giving the driver less warning of the pedestrian's movement. In the reviewed accidents, the height of the pedestrian was not available for all accidents and therefore not presented in this report. The age range for all pedestrians is included and average heights for the age ranges are given in Table 14, displaying the heights of the 5th percentile, 50th percentile and 95th percentile of pedestrians in the UK (Pheasant 1998).

Age Group	Male (mm)			Female (mm)		
	5th%ile	50th%ile	95th%ile	5th%ile	50th%ile	95th%ile
0–5 yrs	940	1005	1100	925	1002	1080
6–10 yrs	1181	1280	1379	1174	1276	1378
11–15 yrs	1424	1558	1692	1420	1538	1656
16–18 yrs	1640	1746	1853	1523	1620	1716
19–25 yrs	1640	1760	1880	1520	1620	1720
26–45 yrs	1635	1745	1860	1515	1615	1715
46–65 yrs	1610	1720	1830	1495	1595	1695
66–80 yrs	1575	1685	1790	1475	1570	1670
Over 80	1515	1640	1765	1400	1515	1630

Table 14: Heights of UK population by age for male and female

Figure 4 below gives the distribution of pedestrian accidents by the age range of the pedestrian and the pre-impact movement of the pedestrian. This chart shows that the three common age ranges are 11–15, 26–45 and 19–25.

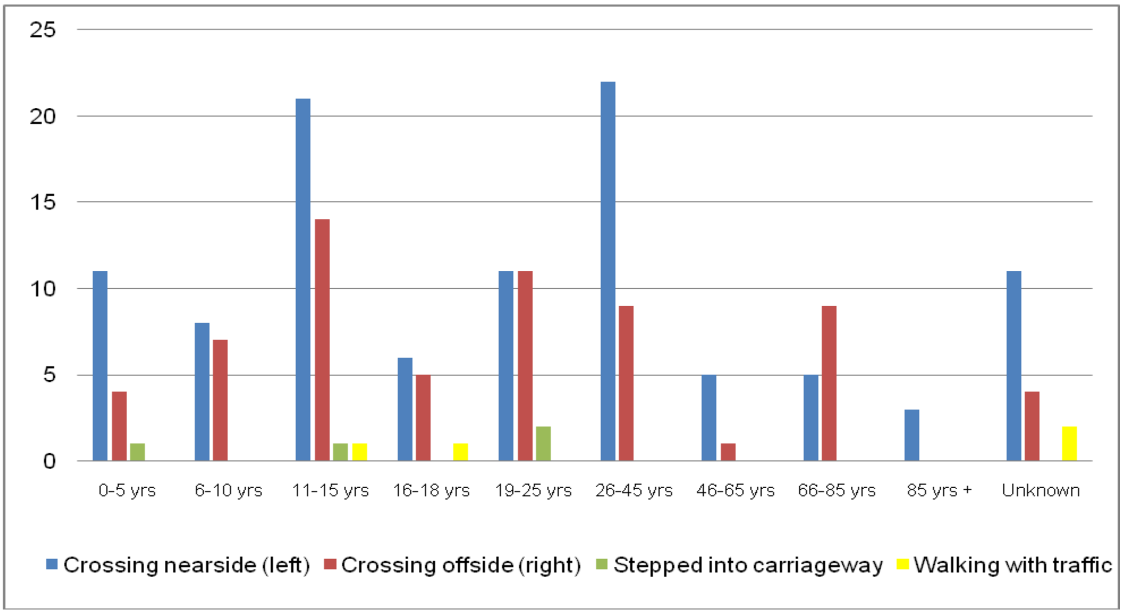


Figure 4: Pedestrian crossing action by pedestrian age

The same three age groups also have the highest recorded injuries for the reviewed sample. 34 (53%) of the pedestrians were recorded as killed or seriously injured (KSI). The older two of these three groups also accounted for 5 (50%) of the fatal accidents reviewed in this work, shown in Table 15.

	Fatal	Serious	Slight	Non-injury	Total
0–5 yrs	1	3	11	1	16
6–10 yrs	0	2	13	0	15
11–15 yr	0	10	27	0	37
16–18 yr	0	4	8	0	12
19–25 yr	1	8	13	2	24
26–45 yr	4	11	16	0	31
46–65 yr	0	2	4	0	6
66–85 yr	2	5	7	0	14
85 yrs +	1	1	1	0	3
Unknown	1	8	8	0	17
	10	54	108	3	175

Table 15: Pedestrian age rage by maximum reported injury severity

Road geometry is also a factor when reviewing accidents especially in regards to the obscuration of accident participants. If the car has been turning into or out of a junction then this has been recorded (Table 16), as the driver's viewpoint would be different compared to normal driving conditions on a straight road.

	Straight ahead		Turning at a Junction			
			Nearside (left)		Offside (right)	
Car direction of travel	152	87%	13	7%	10	6%

Table 16: Car direction of travel

The investigated pedestrian accidents largely occurred where the car was travelling straight ahead (87%) with a minority where the vehicle was turning into or out a junction just prior to impact (13%). The distribution of turning direction, to the nearside (left) or offside (right) was broadly similar with 7% (n=13) and 6% (n=10) of the accidents respectively.

The accidents where the vehicle was turning prior to impact also included the vehicle slowing prior to impact in order to perform the turning manoeuvre. The turning manoeuvre occurred 0–2 seconds before impact. Table 17 gives information relating to the time at which the vehicle turned into or out of a junction prior to impact.

	0 s	0.5 s	1 s	1.5 s	2 s	Total
Time of turning before impact	1	3	11	5	3	23

Table 17: Time for vehicle turning before impact (sec)

The results given in Table 17 show that in the majority, 11 (48%) of the turning cases the vehicle turned 1 second prior to impact. In one accident the vehicle was turning at the point of impact. Prior to the application of the turning manoeuvre this driving style is identical to the accidents where the vehicles were intending to go straight ahead.

The first point of impact between the car and pedestrian was recorded during the review process using a division of the front of the car into four equal zones across the bonnet's leading edge. Zone 1 was the first quarter of the front elevation from the left (nearside) corner, Zone 2 the following quarter to the centre of the vehicle, Zone 3 and 4 replicated this with Zone 4 starting at the right (offside) corner.

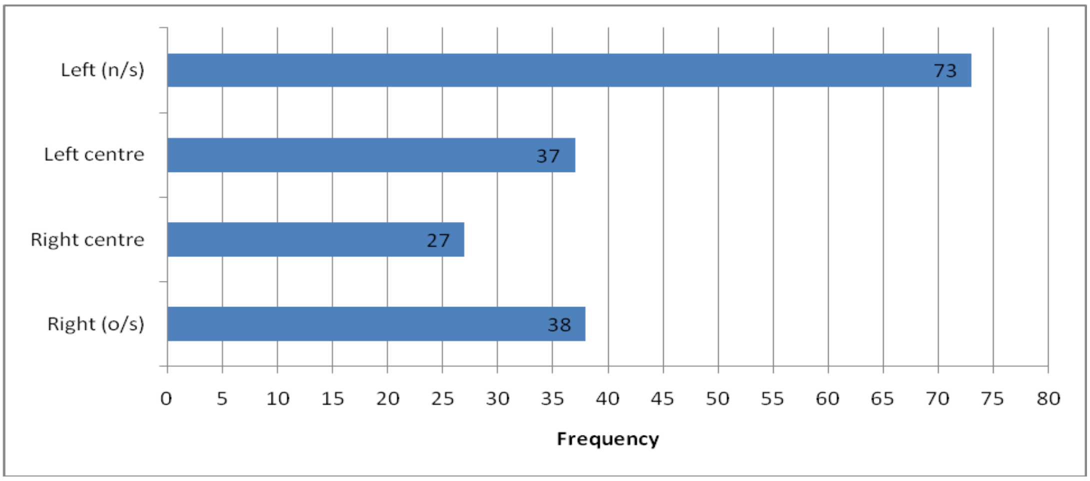


Figure 5: Lateral distribution for first point of contact between pedestrian and car

The results displayed in Figure 5 show that a large proportion of the first point of contact occurred on the front left quarter of the car (n=73, 42%). This is nearly twice as many as the first quarter on the right side of the vehicle (n=38, 22%). This is possible a feature of the crossing actions observed in the sample with 103 pedestrians crossing from the left compared to 64 crossing from the right. Collectively 111 (64%) of the pedestrians contacted the outer quarters of the front elevation compared to 64 (36%) contacting a central location of the vehicle.

Local and meteorological factors such as light conditions and precipitation can have an effect on the pre-impact movements as the pedestrian may be obscured due to poor lighting conditions or the driver may have reduced visibility due to glare from lights or precipitation on the windscreen. The distribution of pedestrian accidents by meteorological factors and pedestrian crossing action is given in Figure 6.

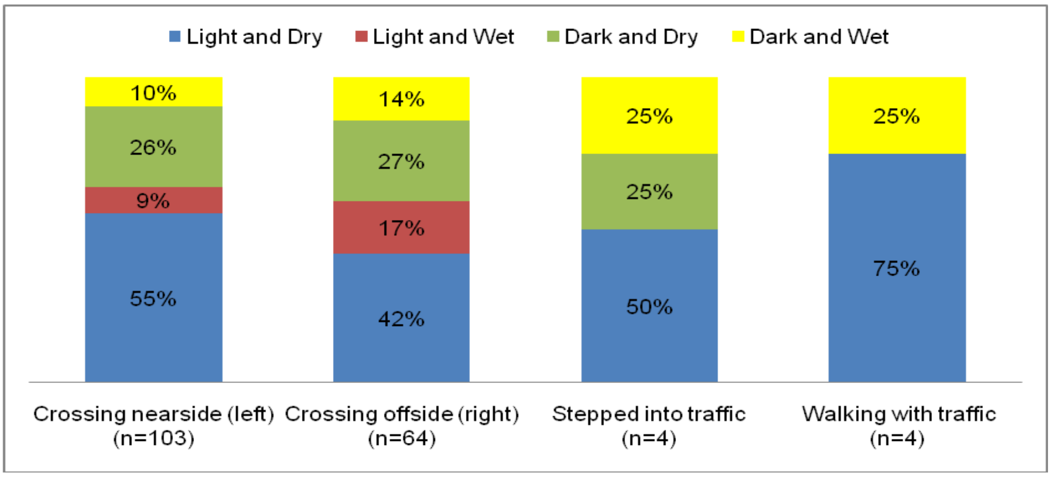


Figure 6: Pedestrian crossing action by meteorological conditions

The results are crosstabulated by the pedestrian crossing action and the meteorological conditions to identify any trends in the type of accident in adverse weather conditions. The majority of pedestrian accidents occurred in daylight and dry road conditions. In accidents where the pedestrian is crossing from the left, 55% occurred in daylight and dry conditions compared to nearly 10% occurring in adverse weather conditions of darkness and on wet roads.

	Daylight and Dry road	Daylight and Wet road	Darkness and Dry road	Darkness and Wet road	Total
Fatal	3	1	6	0	10
Serious	26	5	14	8	53
Slight	63	12	21	13	109
Uninjured	0	1	2	0	3
	92	19	43	21	175

Table 18: Road and light conditions by accident severity

Table 18 shows that road and light conditions by the accident severity of the reviewed pedestrian accidents, this shows that although accidents which occurred in darkness and dry roads accounted for 43 (26%) accidents, they provided 6 (60%) of the fatal accidents and 20 (31%) of the reported KSI accidents.

Whether the vehicle was braking at the point of impact or not was a factor recorded by the original investigation team. This affects the pre-impact movement and velocities of the accident participants, so was considered during the case review process and is displayed in Figure 7.

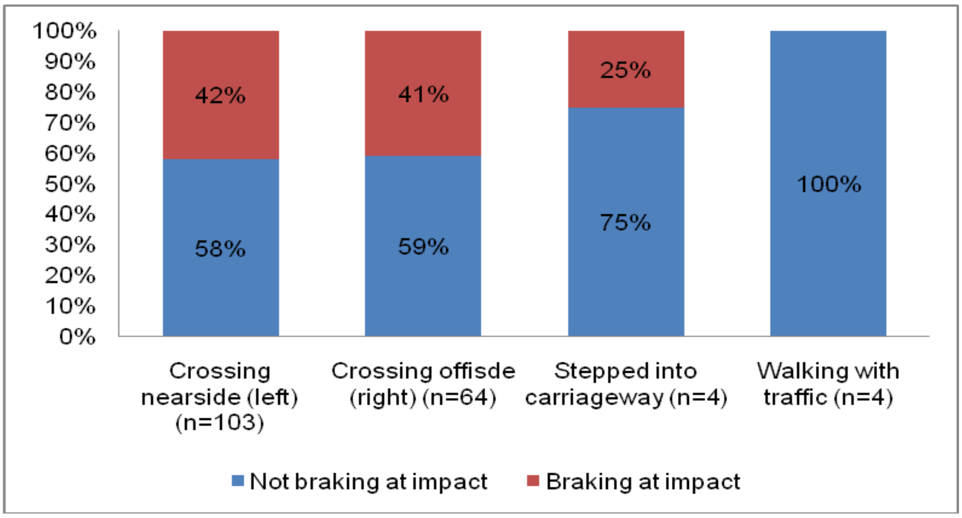


Figure 7: Pedestrian crossing action by braking input of driver at point of impact

The detailed reviews found that the vehicles were braking in 40% of the accidents in cases where the pedestrian was crossing from the nearside or offside. Figure 4 shows the proportion to be the same between nearside (left) and offside (right) crossing actions.

The following charts Figure 8 and Figure 9 show the pedestrian crossing action according to the impact and travel speeds which were established during this review process by additional case reconstructions or taken from the existing OTS data. The speed data was established where possible by reconstructing the available scene evidence or evaluating the vehicle damage. If no scene evidence was available then the travel and impact speeds were based on expert opinion and general travel speed or speed limit of the road at the accident site. The data is given as a percentage of the cases for the pedestrian crossing action in each speed range.

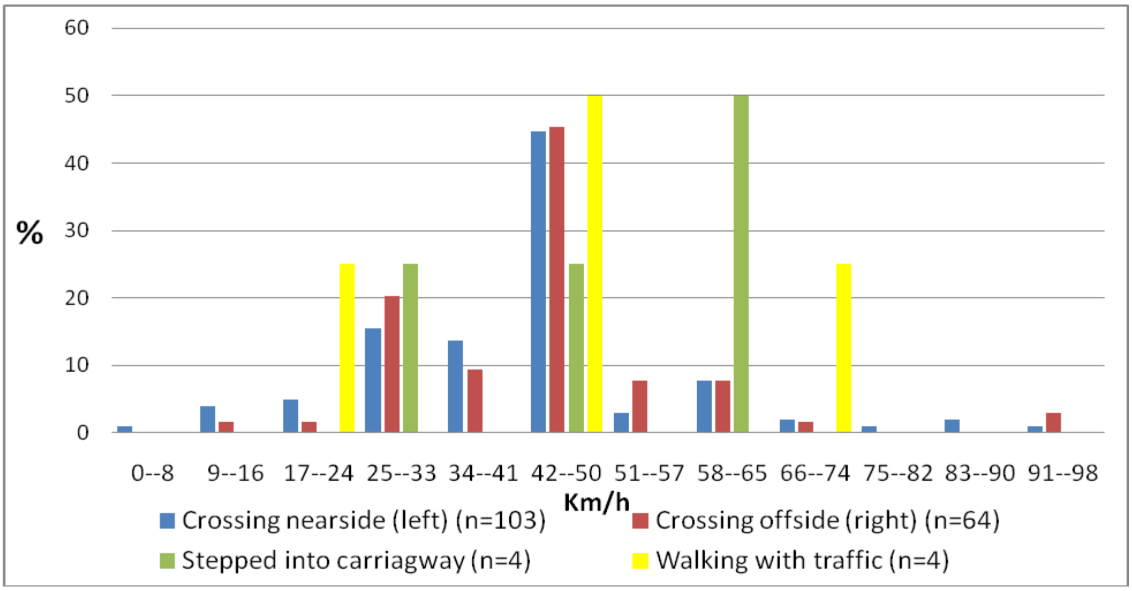


Figure 8: Travel speed of car by pedestrian crossing action

The travel speeds displayed in Figure 8 highlight that the greater proportion of cars are in the region of 42–50 km/h. For the pedestrian action ‘stepped into carriageway’, two of the four accidents had the travel speed in the region of 58–65 km/h.

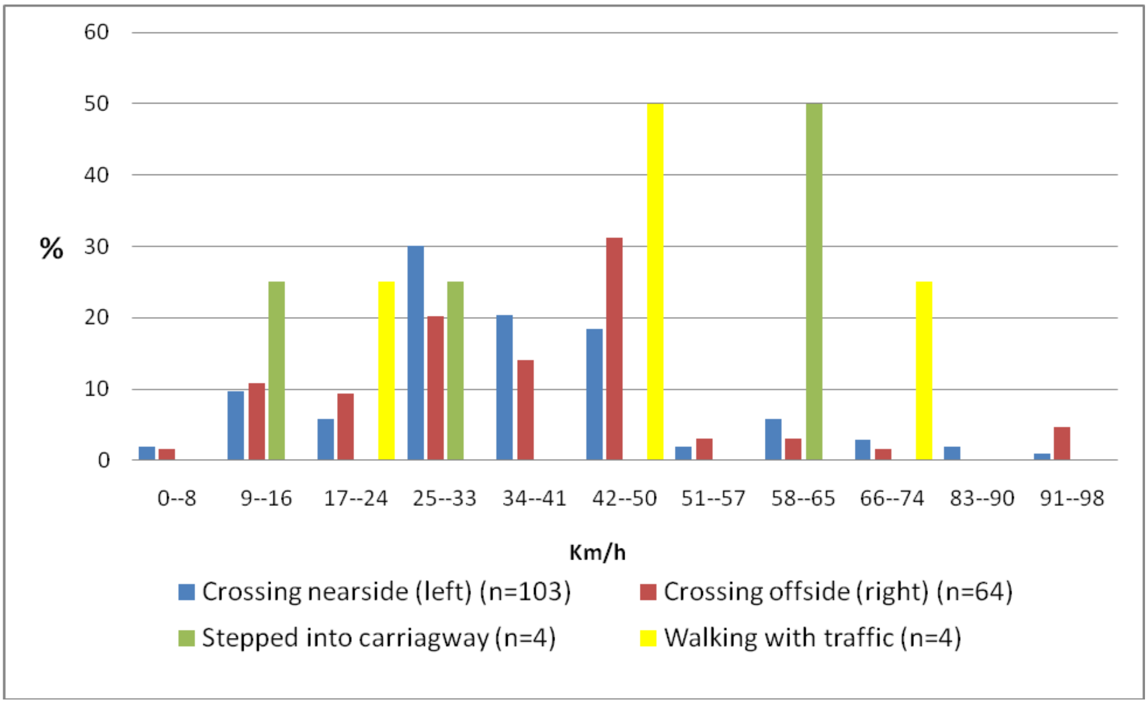


Figure 9: Impact speed of car by pedestrian crossing action

Figure 9 shows that a large proportion of the recorded travel speeds decreased to lower impact speeds suggesting that the driver has reacted and applied braking prior to impact.

The pre-impact movements of the car and pedestrian were reviewed and evaluated to establish any sight line issues, including roadside furniture, to understand if the pedestrian was obscured to the driver prior to impact, and to ascertain how long the driver would have been able to see the pedestrian prior to the impact.

The sight lines and potential objects or vehicles that could have obscured the vision of the pedestrian to the car driver were reviewed. This was conducted by evaluating the pre-impact movements of the pedestrian and car and assessing the sight lines at 0.5 second intervals. Potential sight obstructions included (not exclusively), other vehicles (parked or moving), vegetation, road infrastructure (light poles, telephone boxes), vegetation and road geometry. The pedestrian was classed as obscured if an object or another vehicle completely restricted the vision between the pedestrian and car. If it was believed that the car driver could not see the pedestrian fully due to obstructed vision then it was recorded at the specific time interval(s). In addition if the pedestrian could not see the front or side elevation of the car prior to impact a sight obscuration was also recorded, as provided the basis for assuming that the car driver could not see the pedestrian at the relevant point in time. The following section outlines the findings and common accident situations when the pedestrian was obscured at some point prior to impact, for each accident the five seconds preceding the impact was considered at half second intervals. The review process found that the pedestrians were obscured by objects and not visible to the driver at some point leading up to the accident in 132 cases, which is 75% of the accidents. The pattern of obscuration for the pedestrians varied across the cases: in some cases the pedestrian was obscured at four or five seconds before the impact and then continuously visible for the 3 seconds before the accident or their obscuration changed as the pedestrian or vehicle passed an object resulting in a change of the obscuration pattern.

The chart below shows the accident data split by the pedestrian crossing action and whether any obscuration had occurred before impact at any point in the five seconds leading to the accident. Figure 10 shows that when the pedestrian was crossing the carriageway from the nearside (left) the pedestrian was obscured at some point in the previous five seconds in 76% of the cases. This observation was higher with the pedestrians crossing from the offside (right) with 80% of pedestrian accidents.

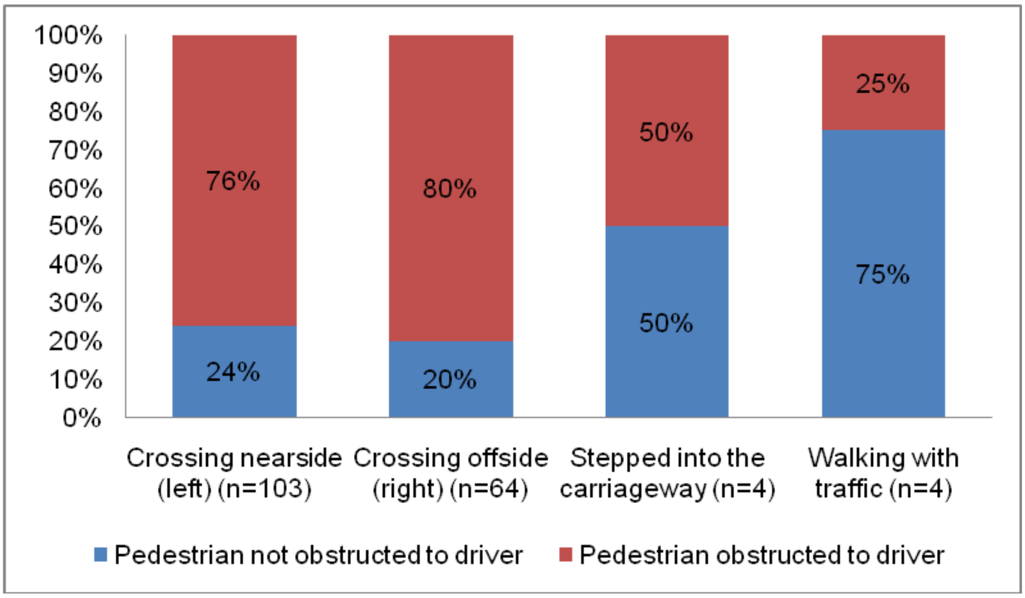


Figure 10: Pedestrian crossing action against pedestrian obscuration

Figure 11 shows the proportion of pedestrians who were obscured at 0.5 second intervals, up to five seconds before the impact, for all the 175 cases. As previously stated 75% of all the pedestrians were obscured at some point in the five seconds prior to the impact.

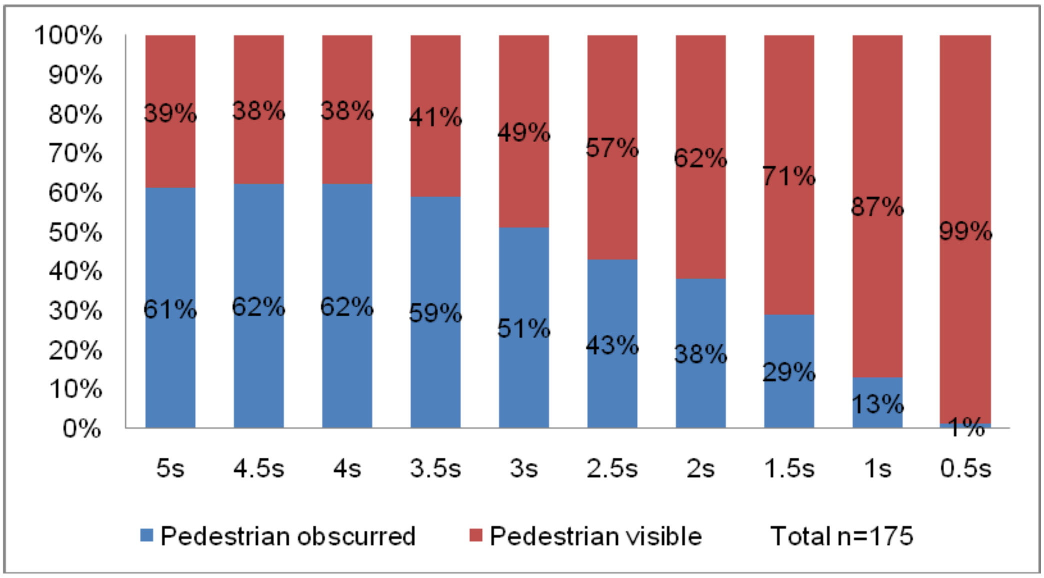


Figure 11: Percentage of pedestrians visible to the driver at set time intervals before impact

Figure 11 shows that at 2 seconds before the accident, 38% (66) of the pedestrians were not visible to the car driver. This proportion increases as the time before the impact increases. At 4 and 4.5 seconds before the accident, 62% of the pedestrians were not visible to the driver. In contrast, at 1 second before the accident only 13% of the pedestrians were not visible to the driver of the car.

As previously stated the anthropometric construction and particularly the height can influence the pre-impact movement or more specifically how it affects the driver's vision of the pedestrian prior to impact. Table 19 gives the distribution of pedestrian age as a function of being visually obstructed prior to impact to the driver. This shows that pedestrians under the age of 15 accounted for 38% (68) of the accident population and 42% (55) of the accidents where pre-impact obscuration was a factor.

	Pedestrian visible to the driver		Total
	Obscured	Not obscured	
0–5 yrs	5	11	16
6–10 yrs	5	10	15
11–15 yr	3	34	37
16–18 yr	3	9	12
19–25 yr	5	19	24
26–45 yr	9	22	31
46–65 yr	1	5	6
66–85 yr	5	9	14
85 yrs +	1	2	3
Unknown	6	11	17
Total	43	132	175

Table 19: Pedestrian age range by obscuration to the driver before impact

The following charts are scatter plots showing the relationship between the pedestrian and the cars as they approach the point of impact, giving their lateral and forward distances. The lateral distance is the movement to the left or right from the front centre point of the car, the forward distance is the longitudinal displacement from the front of the vehicle to the point of impact. The charts are split at 0.5 second intervals, 0.5 to 2 seconds before the accident, and then the time interval is increased to 1 second intervals up to 5 seconds before the accident. Each accident is represented on the charts and is recorded to show if the pedestrian is visible or not to the driver. The points encompassed by the blue boxed area broadly represent 75% of the points for the pedestrians crossing from the nearside and offside.

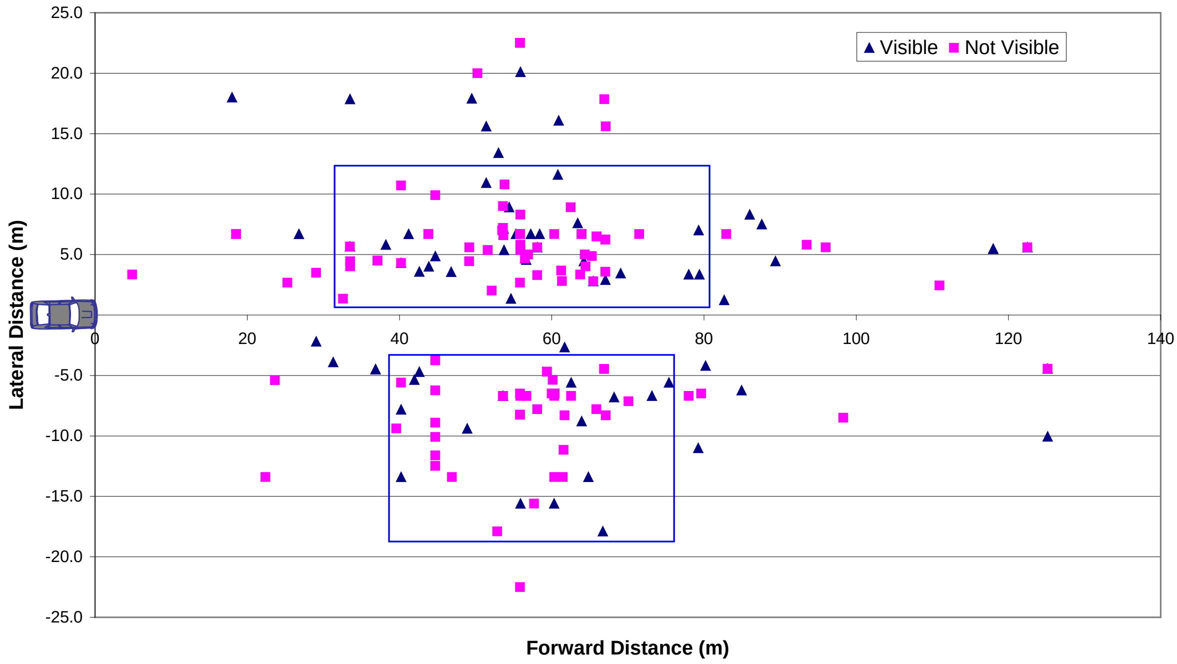


Figure 12: Distance to point of impact 5 seconds before impact (n=175)

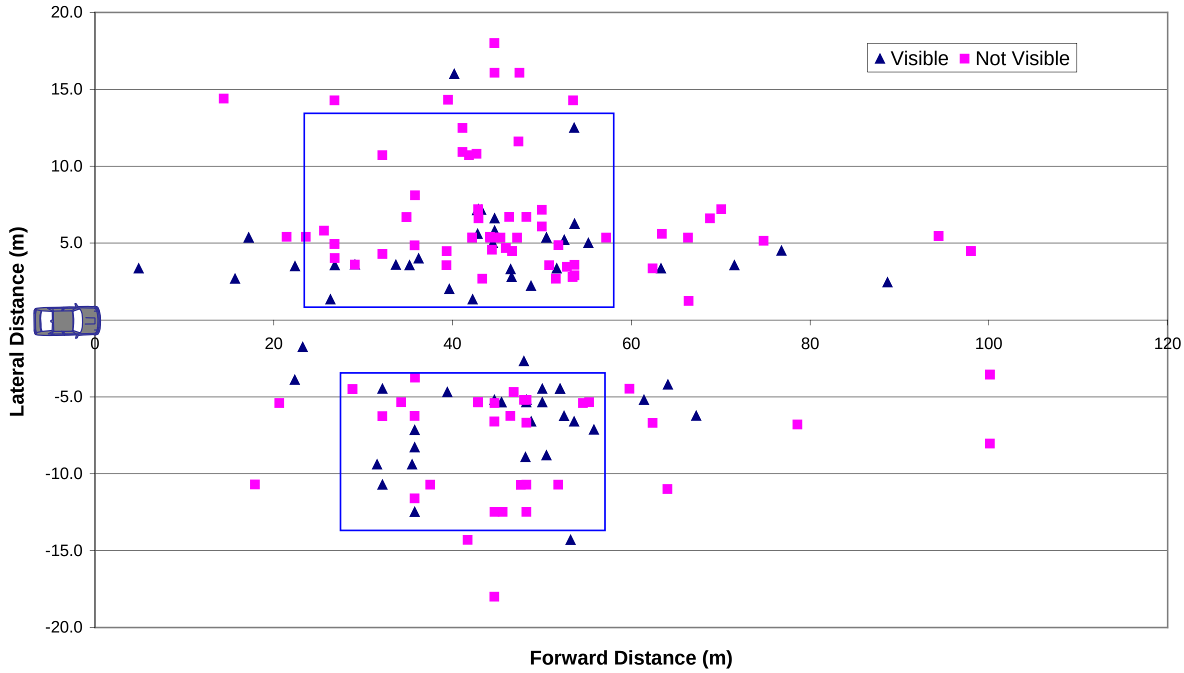


Figure 13: Distance to point of impact 4 seconds before the impact (n=175)

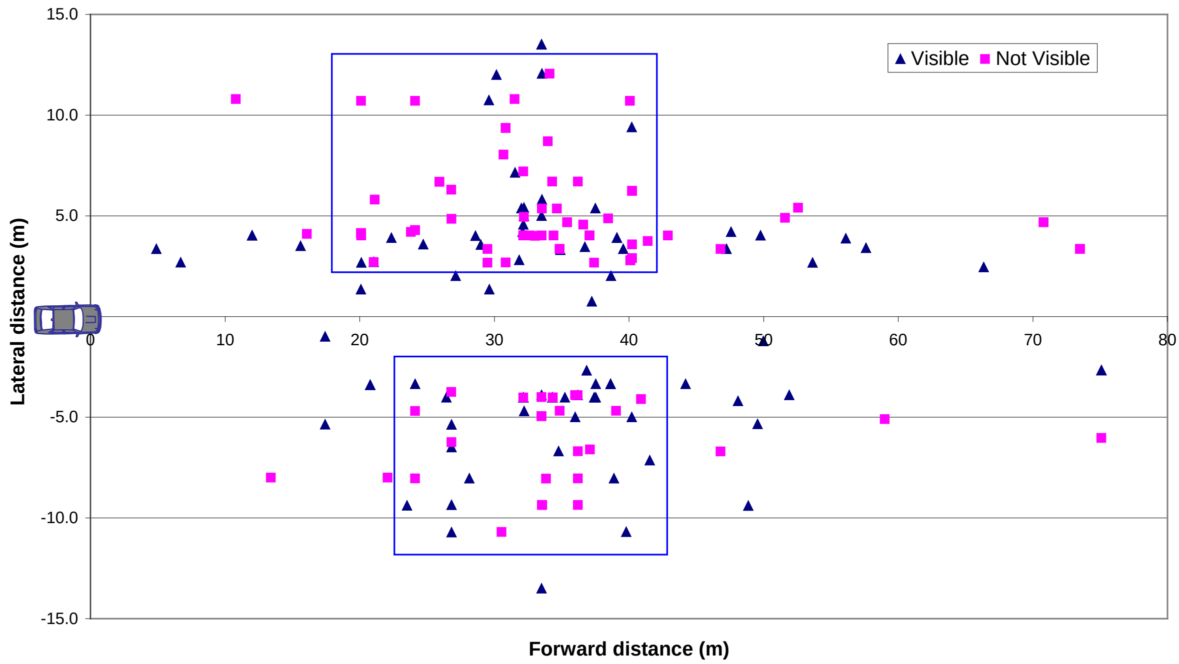


Figure 14: Distance to point of impact 3 seconds before impact (n=175)

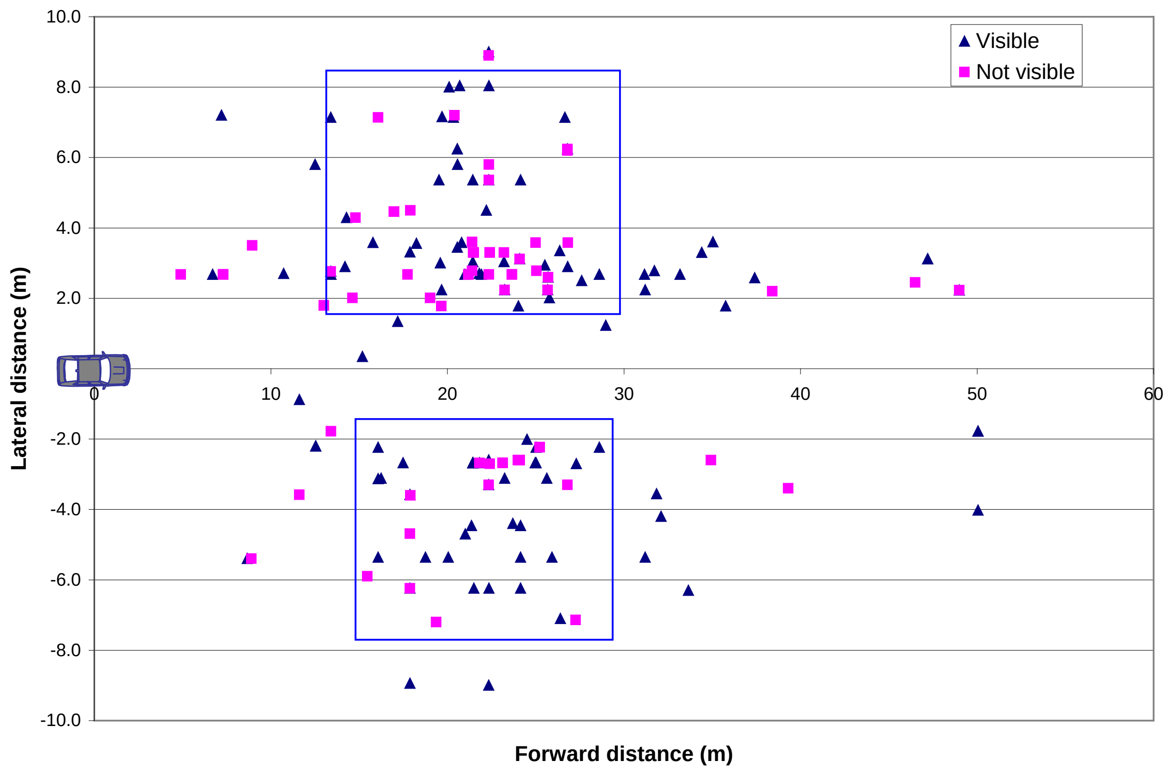


Figure 15: Distance to point of impact 2 seconds before impact (n=175)

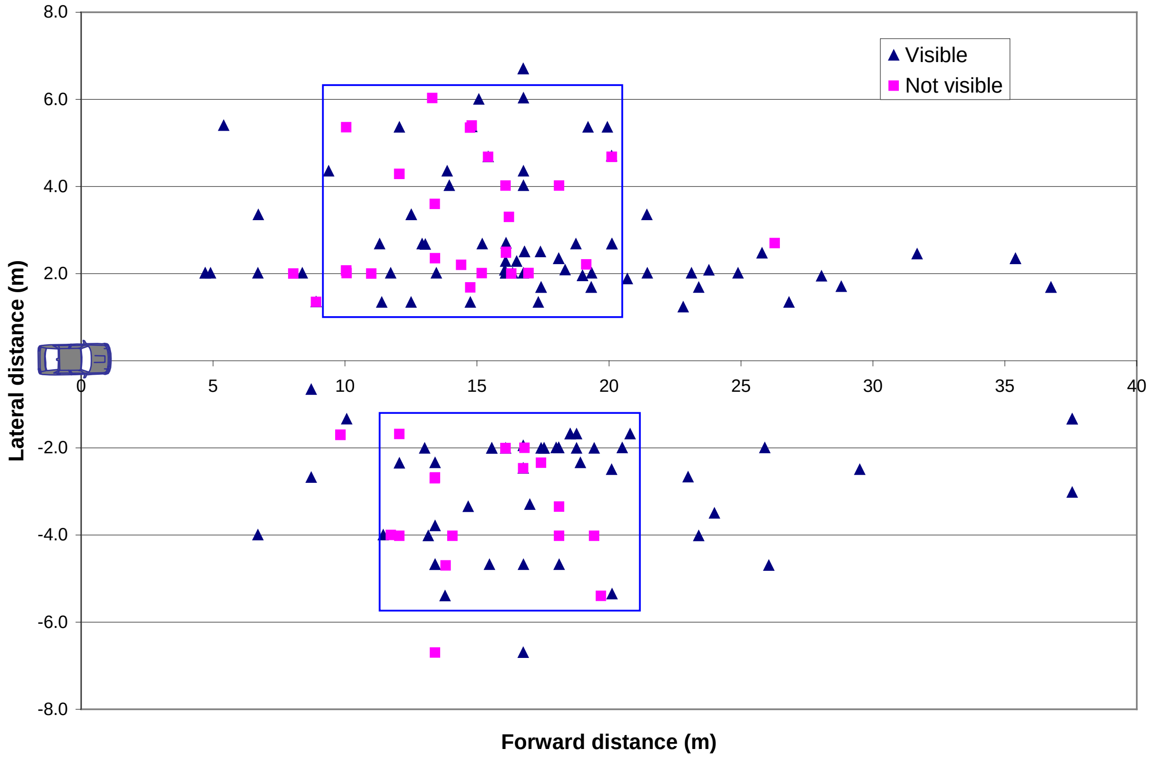


Figure 16: Distance to point of impact 1.5 seconds before impact (n=175)

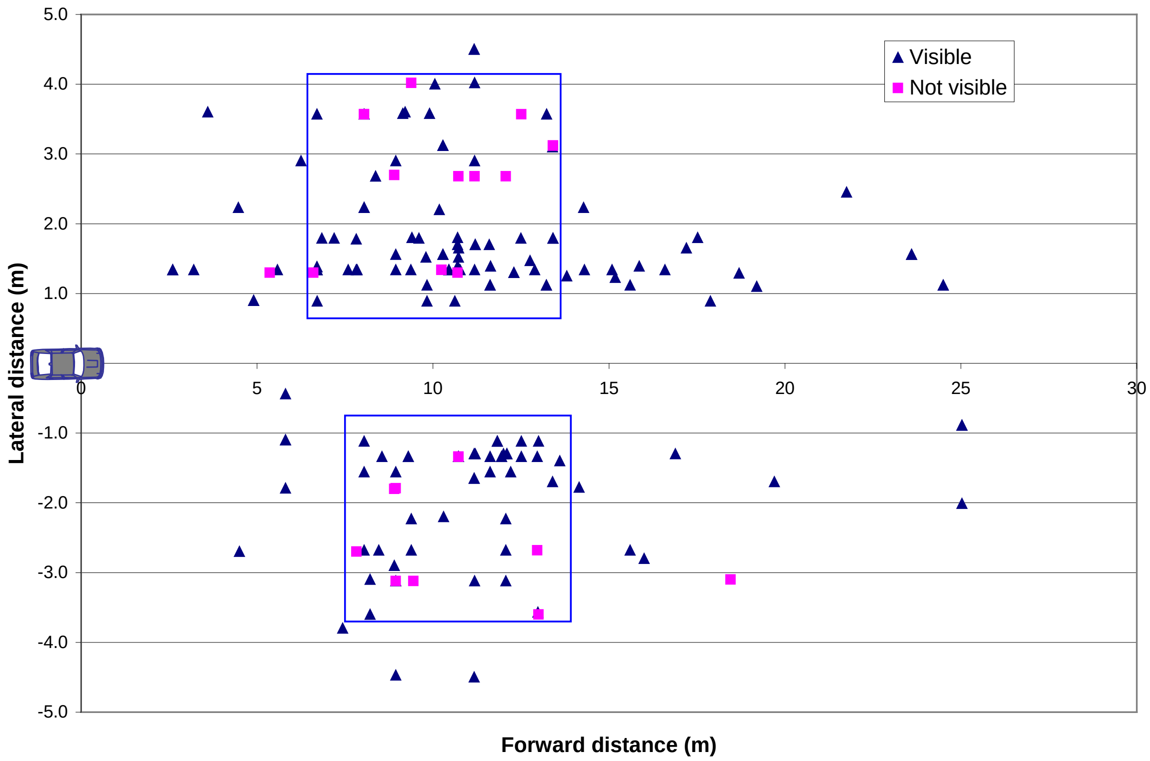


Figure 17: Distance to point of impact 1 second before impact (n=175)

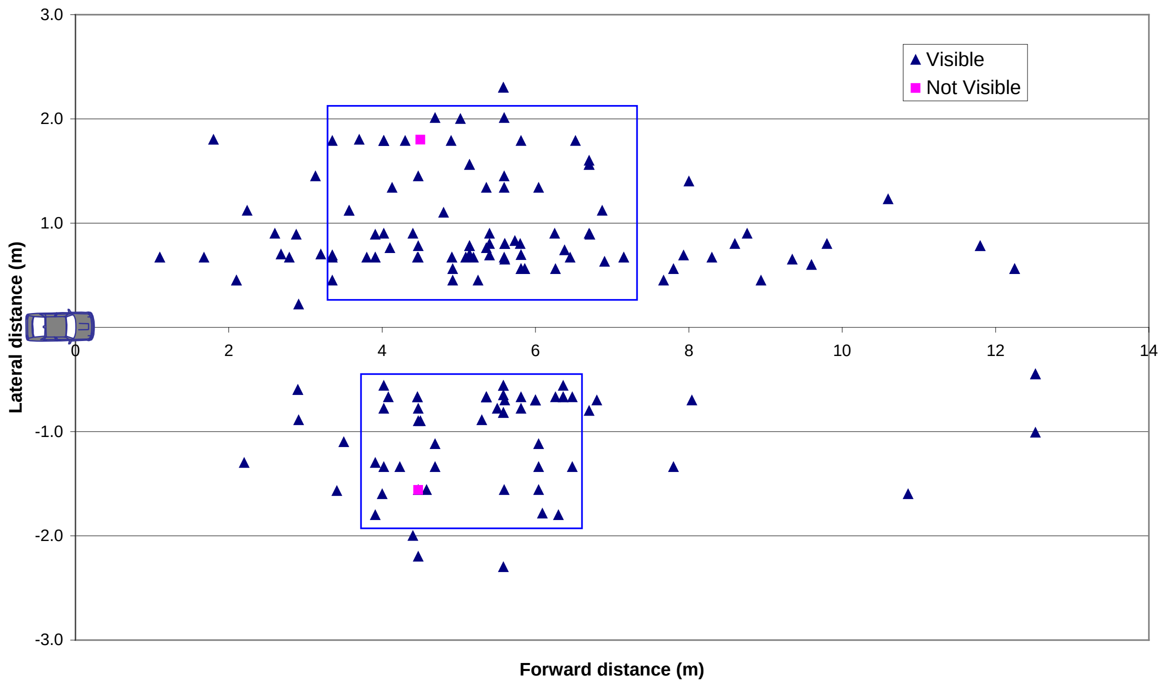


Figure 18: Distance to point of impact 0.5 second before impact (n=175)

The chart below shows the time when the pedestrian became constantly visible to the driver prior to the accident for the 132 accidents where pedestrian obscuration was an issue.

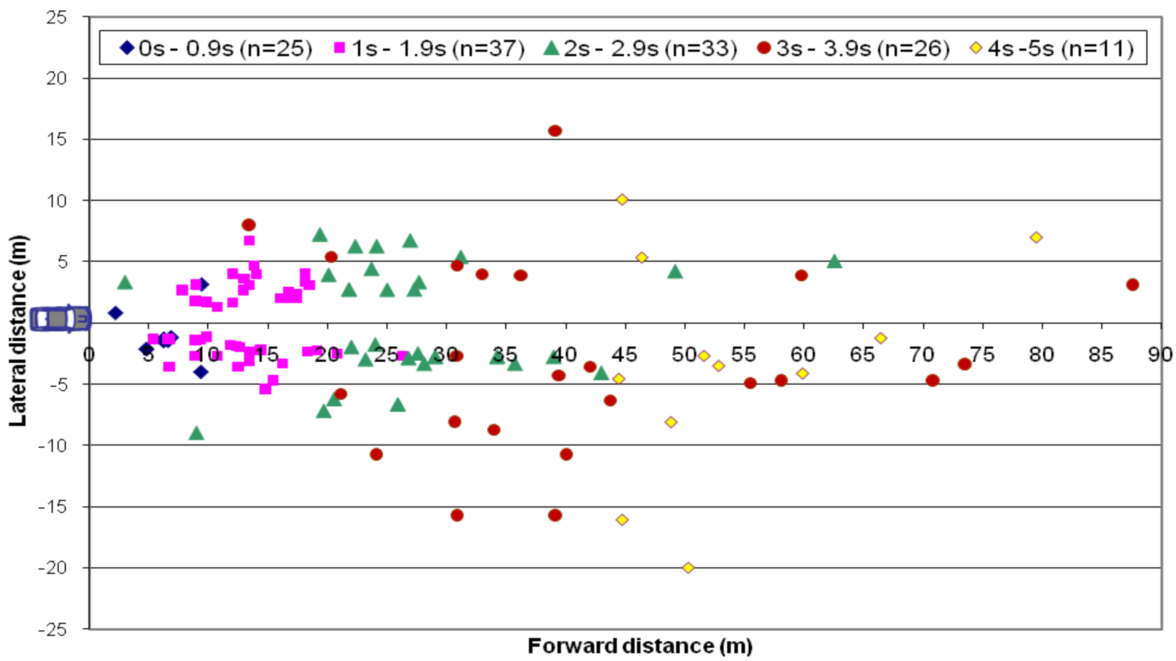


Figure 19: Point of constant visibility of the pedestrian to the car driver (n=132)

Figure 19 shows 1 second time intervals and the forward and lateral distances of when the pedestrian becomes constantly visible to the car driver. At 2 seconds before the accident, 70 (53%) pedestrians had become constantly visible to the car driver out of the 132 accidents with some obscuration. This means that at the 2 second time interval 113 (65%) of the total 175 pedestrians were visible prior to the impact. At 1 second before the accident, 81% of the obscured pedestrians were visible to the driver, which was 85% of the total pedestrian sample.

The following Figure 20 to Figure 25 shows the pre-impact pedestrian movement towards the car at 1 second intervals, the single points represent the forward and lateral distance the pedestrian is from the front centre point of the car and the set time intervals.

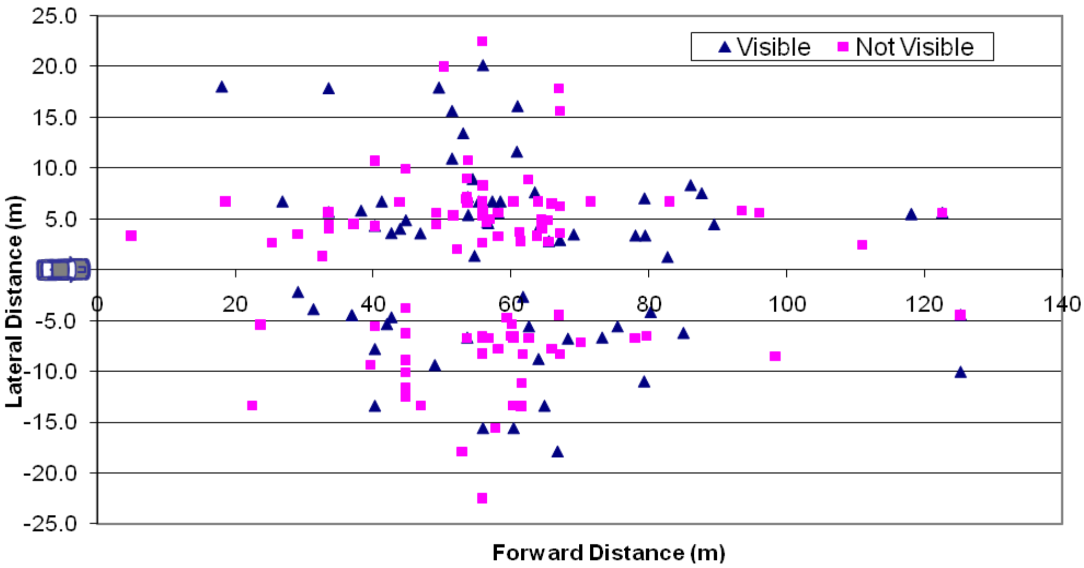


Figure 20: Distance to point of impact 5 seconds before impact (n=175)

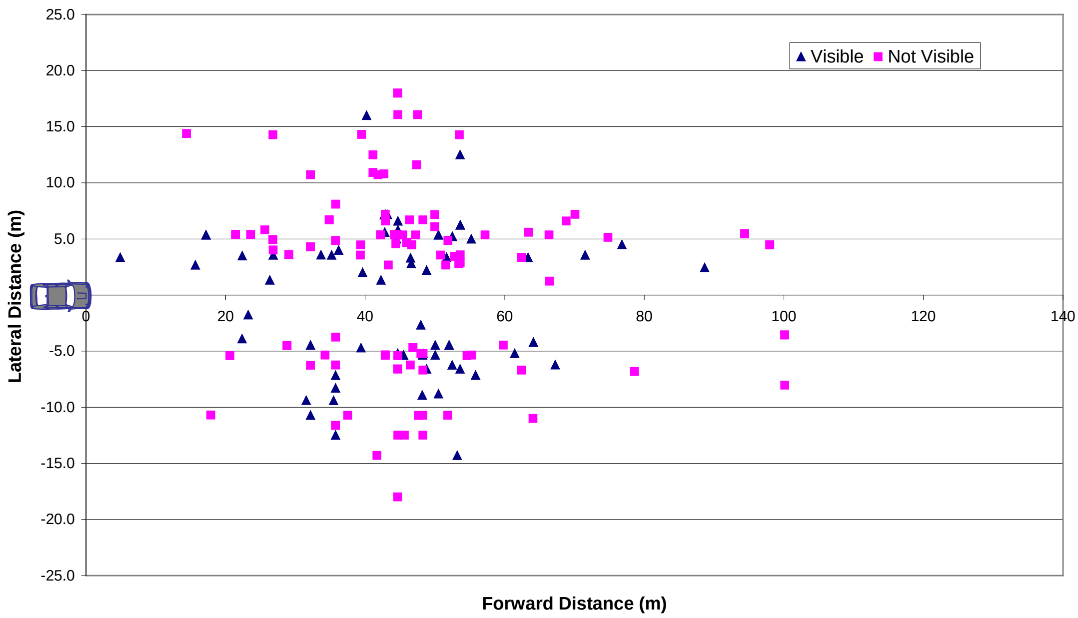


Figure 21: Distance to point of impact 4 seconds before impact (n=175)

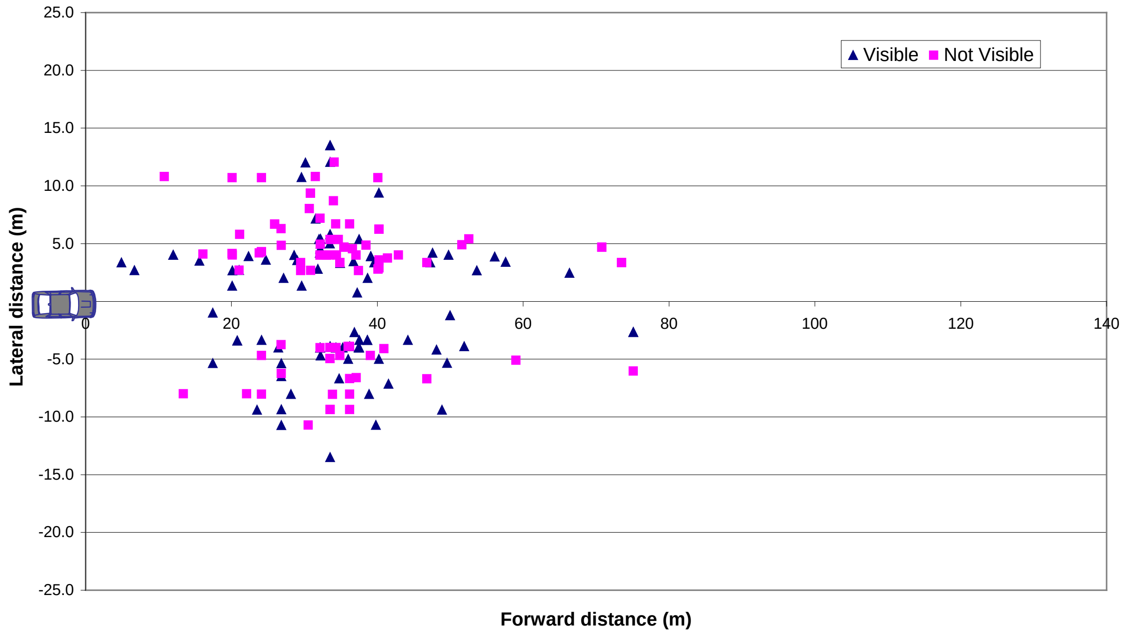


Figure 22: Distance to point of impact 3 seconds before impact (n=175)

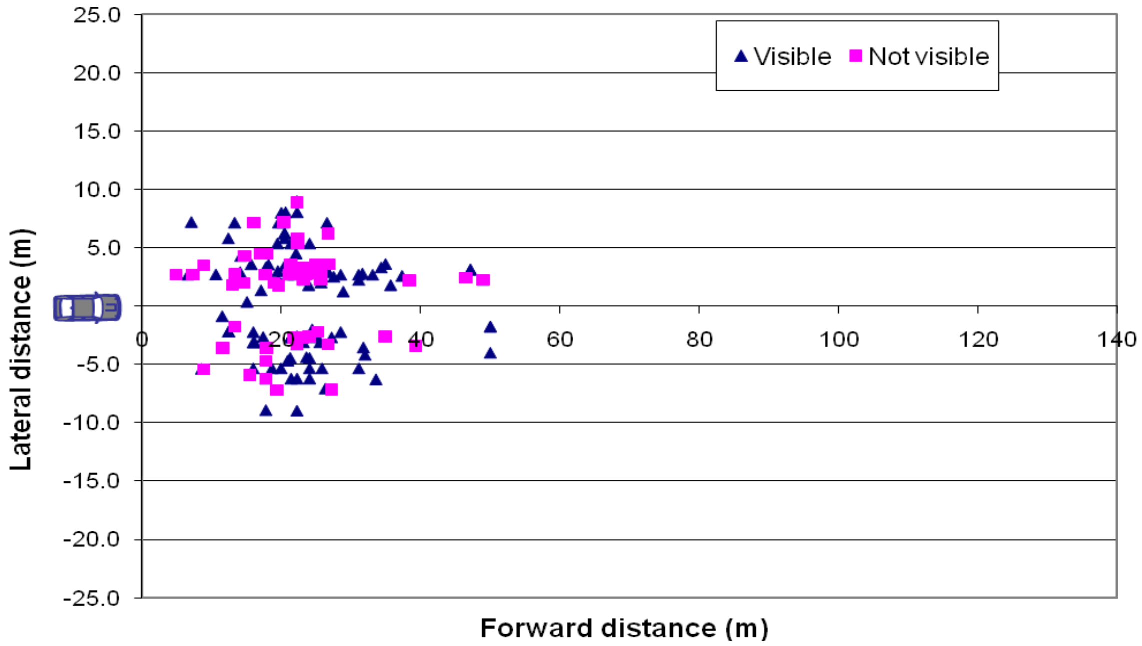


Figure 23: Distance to point of impact 2 seconds before impact (n=175)

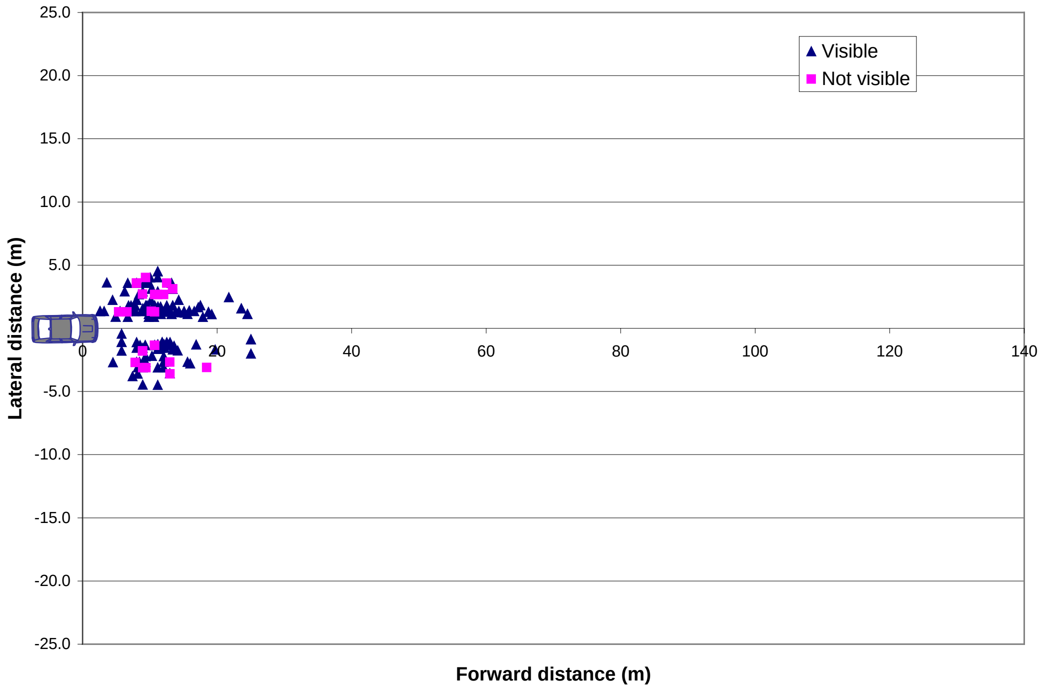


Figure 24: Distance to point of impact 1 seconds before impact (n=175)

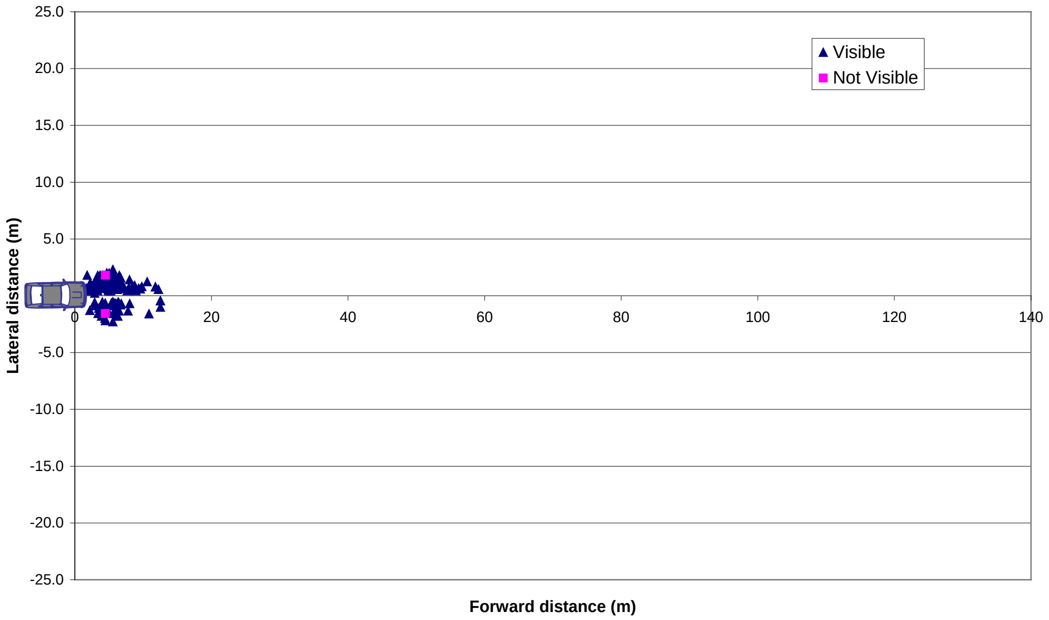


Figure 25: Distance to point of impact 0.5 seconds before impact (n=175)

Table 20 relates to the common causes of a sight obstruction between the driver of the car and the pedestrian. Although street furniture such as light poles and pedestrian barriers do cause issues for road users with sight lines, the predominant factors were parked vehicles, stationary vehicles in traffic and the road geometry.

	0–1 s	1–2 s	2–3 s	3–4 s	4–5 s	Total	%
Parked cars	9	14	5	5	0	33	25%
Road geometry	0	1	7	9	0	17	13%
Stationary cars	6	6	2	0	0	14	11%
Stationary bus	3	8	0	0	0	11	8%
Pedestrian barriers	1	3	5	0	0	9	7%
Vegetation	1	2	3	2	0	8	6%
Light poles	0	4	2	2	0	8	6%
High walls	0	3	3	0	0	6	5%
Moving cars	1	1	3	0	1	6	5%
Bus shelter	0	0	4	0	1	5	4%
Additional people (small group)	0	0	1	1	2	4	3%
Buildings	1	1	1	0	0	3	2%
Moving HGV	2	0	1	0	0	3	2%
Stationary HGV	1	2	0	0	0	3	2%
Telephone box	0	1	1	0	0	2	2%
Total	25	46	38	19	4	132	100%

Table 20: Causes of sight obstruction at one-second time intervals

Although many common accident features were identified within the sample of pedestrian accidents, there were differences in the duration of pedestrian obscuration and whether this was intermittent or constant. The following charts provide three examples of the typical obscuration patterns for the pedestrian accidents showing the movement trajectories. The chart shows the forward and lateral distance displacements from the point of impact at 0.5 second intervals. Each chart shows the pre-impact movement of the car and the pedestrian with a close-up showing the pre-impact movement towards the point of impact for the pedestrian and crossing direction.

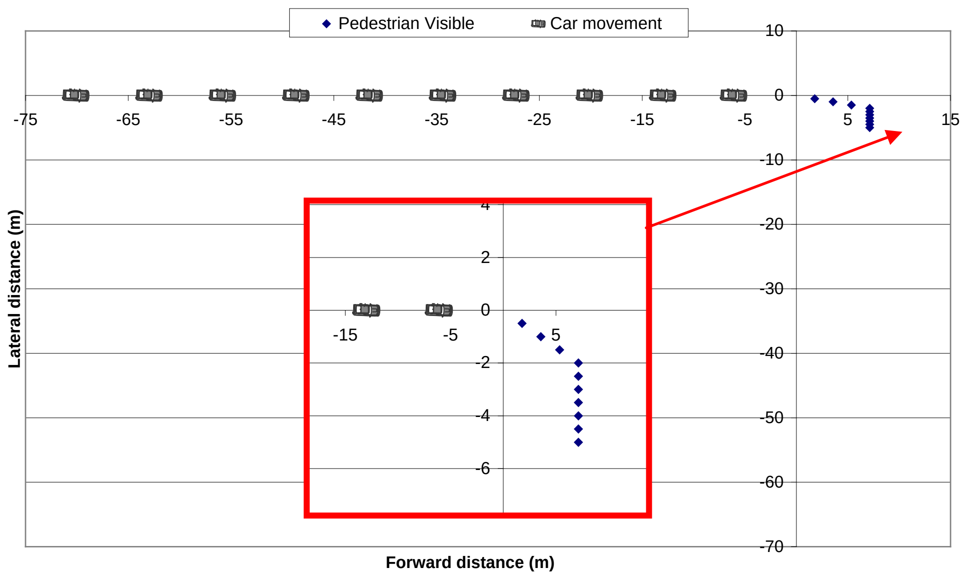


Figure 26: Pedestrian crossing from the offside—no obscuration

Example 1. Figure 26 shows a common accident situation where the pedestrian is crossing from the offside (right) with no obscuration to the vehicle. The pedestrian crossed the road at an angle prior to impact. This typical situation was observed in 22% (39) of the OTS detailed case reviews and contributed 22% of recorded KSI accidents.

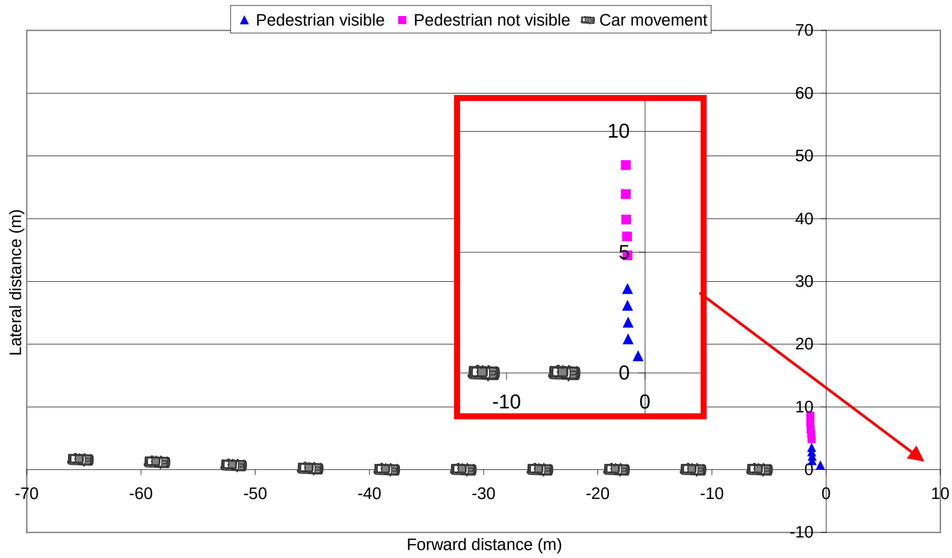


Figure 27: Pedestrian crossing from nearside with constant obscuration to 2.5s before impact

Example 2. Figure 27 shows another common accident situation where the pedestrian is crossing from the nearside (left) and is constantly obscured up to 2.5 seconds before the accident. This also shows how the car had travelled around a slight right bend 3 seconds before the accident. This accident situation was observed in 26% (45) of cases reviewed and contributed 30% of recorded KSI accidents.

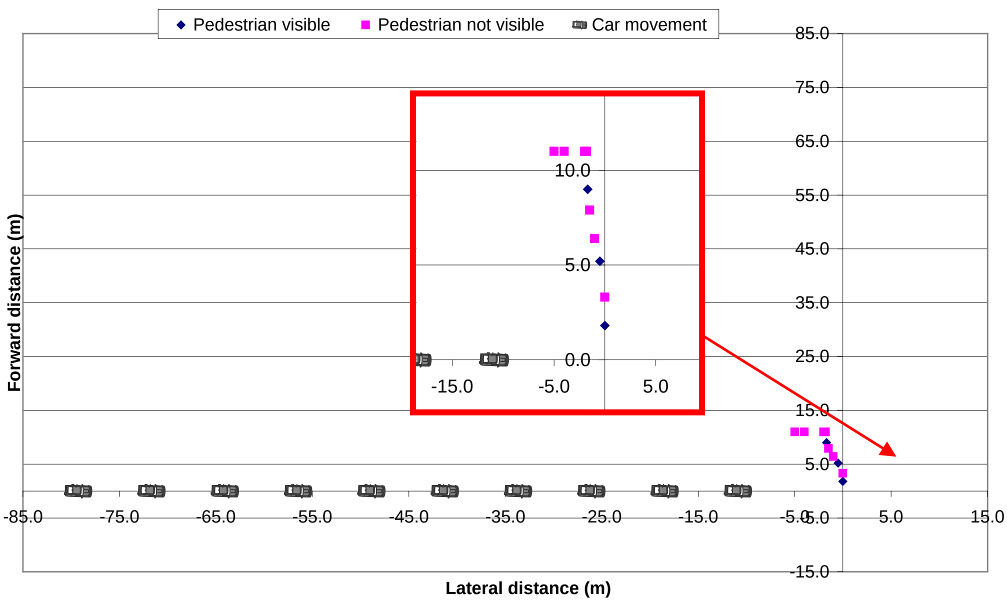


Figure 28: Pedestrian crossing from the offside with intermittent obscuration

Example 3. Figure 28 depicts a common situation where the pedestrian is crossing from the nearside (left) and is obscured to the vehicle intermittently. The pedestrian had travelled along the pavement not visible to the car driver; as the pedestrian crossed the road, visibility altered according to other obstacles on the road. This situation was observed in 16% (27) of cases reviewed and contributed 14% of recorded KSI accidents.

A random selection of twenty of the detailed case reviews have been included to show the pre-impact movement of the pedestrians in relation to the road environment including the footpath. In order to best represent the cases, an average footpath width and location has been used. The average footpath width was determined by reviewing the case details of the 20 cases in the example and deemed to be 2.6 metres. Three example charts have been chosen showing 1 second, 2 seconds and 5 seconds before impact.

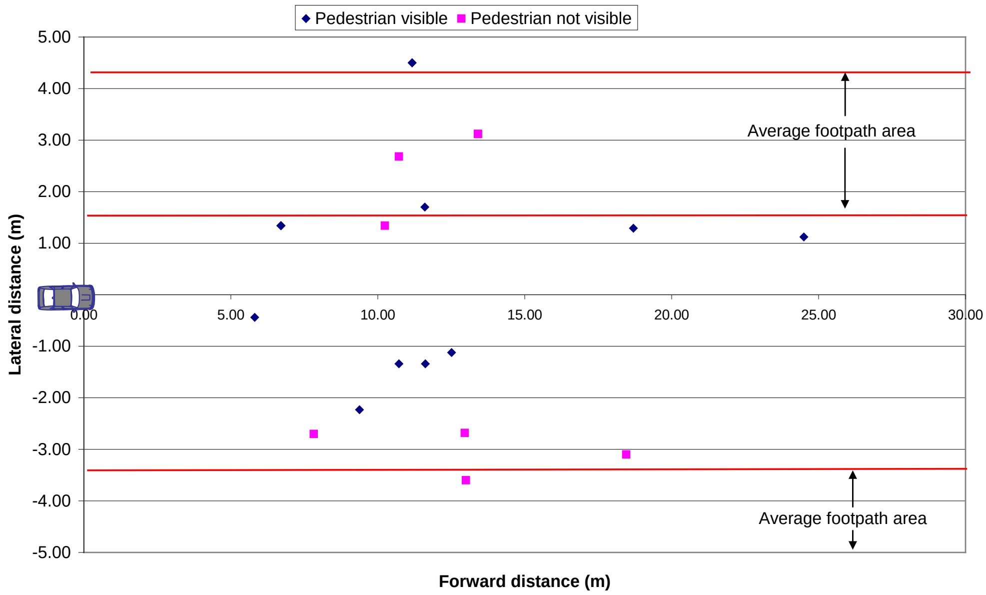


Figure 29: Distance to point of impact for collision participants 1 second before impact

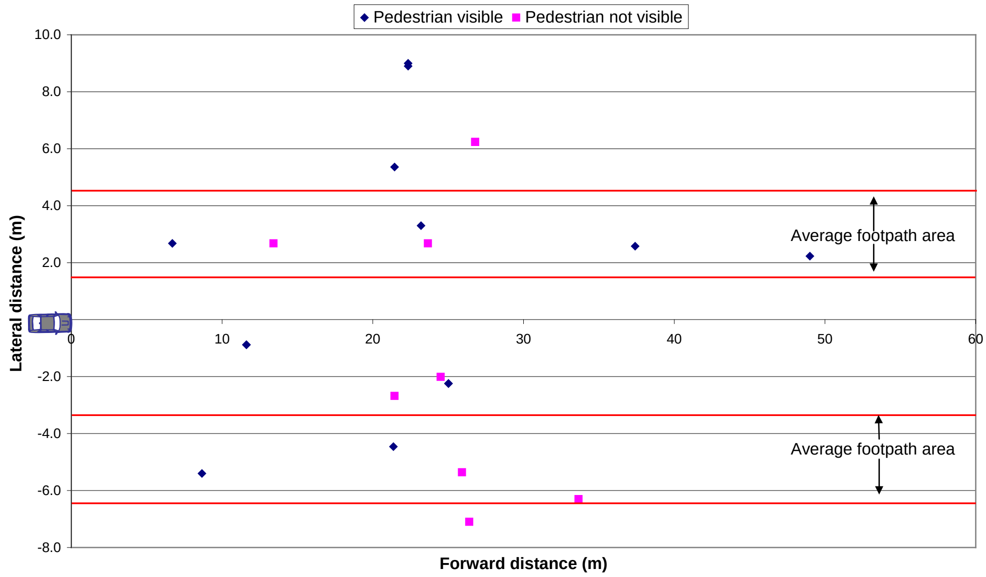


Figure 30: Distance to point of impact for collision participants 2 seconds before impact

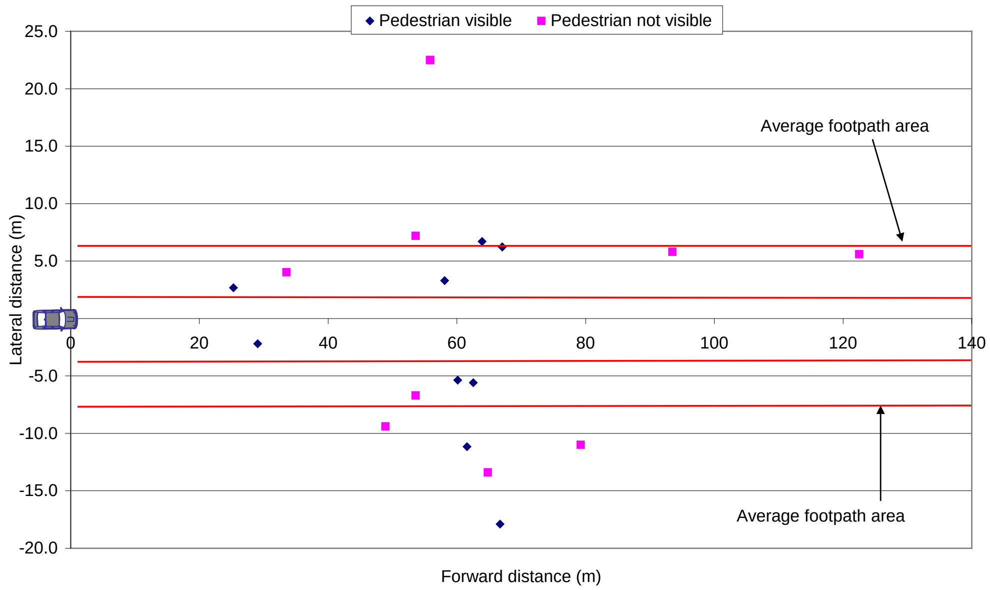


Figure 31: Distance to point of impact for collision participants 5 seconds before impact

It was observed during the review of the car-to-pedestrian accidents that in the majority of the cases the pedestrian was crossing from the nearside (left) of the road, which is most likely a characteristic of the UK road environment with the domestic fleet driving on the left-hand side. This manoeuvre places the pedestrian in front of the vehicle with relatively little time for driver reaction and application of braking. The review found that in the majority of the cases the pedestrian was obscured at some point in the five seconds leading to the accident, the most frequent cause was the pedestrian emerging from behind a stationary or parked vehicle. This scenario generally gave an obscuration pattern where the pedestrian was initially visible to the driver, albeit whilst the car was some distance from the point of impact. Then the pedestrian was obscured by the stationary object and finally the pedestrian became visible to the driver again between 0.5 and 2 seconds before the accident.

The travel speed was assumed to be at the speed limit of the road at the point of accident if no other evidence was available. In these situations photographic evidence and the general traffic conditions for the accident site provided information and the basis for the assumption of travel speed. This may result in an under- or over-estimation of the travel speed for these cases.

4.3 Cluster analysis of OTS

4.3.1 Selection of cases

All of the pedestrian cases from the detailed case reviews were included in the cluster analysis. The basic requirements were (a) that a passenger car should have struck a pedestrian with its front surface and (b) that the case files available at Loughborough University should contain enough evidence to enable a reasonable reconstruction of the location and position of the road users in the seconds before impact. The second requirement was generally not met for cases investigated outside of Loughborough's sample region in the first three years of the OTS because adequate supplementary case materials were not exchanged between the two data collection contractors in this period. There was also a proportion of pedestrian cases where very little or no physical evidence existed to collect, for example where a car struck a pedestrian lightly without leaving tyre marks or signs of impact and one or both road users left the scene of the accident before the research team arrived. It was estimated that the upper limit of suitable cases was less than 200 and, of these, 175 were selected at random and included in the study. The cluster analysis was performed on the same 175 cases as the detailed OTS case reviews in order to exploit the enhanced data including speed and obscuration information.

4.3.2 Input dataset

The dataset derived from OTS shares some variables with the STATS 19 dataset: pedestrian injury severity, light conditions, weather, vehicle manoeuvre, pedestrian age-sex and pedestrian movement. The STATS 19 field 'Masked by vehicle' was replaced by the more specific parameter 'Line of sight obstructed at 1 second before impact' which includes any source of obstruction, not just vehicles. Speed limit, which can serve as a proxy for vehicle speed in the absence of other information, was dropped in favour of vehicle travel speed immediately before the onset of emergency conditions. Change of speed to impact, reflecting the effect of braking, was also added.

Ordinal	Pedestrian severity	0.0 0.5 1.0	Slight Serious Fatal
Nominal	Light conditions	1 2	Daylight Darkness
Nominal	Weather	1 2	Fine Not fine
Nominal	Vehicle manoeuvre	1 2	Going ahead Turning
Ordinal	Pedestrian age-sex	0.00 0.33 0.67 1.00	Child 0–7 years Child 8–15 years Adult female Adult male
Nominal	Pedestrian movement	1 2 3	Crossing from left Crossing from right Stationary or along carriageway
Nominal	Pedestrian speed	1 2	Walking Running
Nominal	Line of sight (1 sec)	1 2	Not obstructed Obstructed
Scale	Vehicle travel speed	0–1	*scaled from km/h*
Scale	Change of speed to impact	0–1	*scaled from km/h*

Table 21 Structure of dataset for pedestrian accidents (OTS)

4.3.3 Results

The results of the cluster analysis of the OTS dataset are detailed at the level of 14 clusters, six of which cover 79% of the population of the dataset. The cells shaded in green indicate (a) that the distribution of numbers for the given field is significantly different from the distribution in the whole population (chi-square test to 95% significance) and (b) that the particular number highlighted is over-represented. This is the same as for the STATS 19 analysis in section 4.1.3 except that the statistical test is evaluated at 95% confidence instead of 99.5%. This level is better suited to the lower number of cases in OTS for providing an objective test of differences between the clusters and the overall population.

Cluster 1, the largest in the set comprising 29% of the population, has accidents in daylight involving vehicles going ahead and pedestrians walking. Other majority characteristics are fine weather and unobstructed line of sight at one second before impact. The mean travel speed was 43 km/h with a reduction of 7 km/h before impact. The range of these two parameters can be read off in Figure 32 and Figure 33. The distance of the pedestrian from the vehicle at one second before impact is shown in Figure 34. This parameter was however not included as part of the cluster analysis because it is highly correlated to vehicle travel speed—including it would have provided double weight to essentially the same information.

Cluster 2, the second largest, has an over-representation of children running from the left with a tendency to be obscured. This compares interestingly with the corresponding STATS 19 cluster. There are also parallels with the STATS 19 results in clusters 3 and 4, with the tendencies towards serious injury outcomes, darkness, wet weather and adults. Cluster 5 is the closest that a major cluster approaches to a turning scenario, involving children running across from the right side; the mean travel speed is 37 km/h with 11 km/h reduction in speed before impact. This is consistent with the STATS 19 turning scenario which has speed limits and injury outcomes at the lower end of the range. Two of the ten fatalities constitute cluster 6 which is too small to support any generalisations, but noteworthy for the very high vehicle speeds.

	Cluster 1	2	3	4	5	6	7–14	Total
Cluster representativeness (%)								
Slight or not injured	29	20	12	9	8	0	23	100
Serious	30	9	15	24	4	0	19	100
Fatal	20	0	40	10	0	20	10	100
Total	29	15	14	14	6	1	21	100
Pedestrian severity								
Slight or not injured	32	22	13	10	9	0	25	111
Serious	16	5	8	13	2	0	10	54
Fatal	2	0	4	1	0	2	1	10
Total	50	27	25	24	11	2	36	175
Light conditions								
Daylight	50	27	0	0	11	2	20	110
Darkness	0	0	25	24	0	0	16	65
Total	50	27	25	24	11	2	36	175
Weather								
Fine	45	23	25	9	9	2	22	135
Not fine	5	4	0	15	2	0	14	40
Total	50	27	25	24	11	2	36	175
Vehicle manoeuvre								
Going ahead	50	27	18	24	6	2	25	152
Turning	0	0	7	0	5	0	11	23
Total	50	27	25	24	11	2	36	175
Pedestrian (age-sex)								
Child 0-7 yrs	4	6	1	0	6	1	5	23
Child 8-15 yrs	12	12	2	1	5	0	15	47
Adult female	18	3	6	9	0	0	5	41
Adult male	16	6	16	14	0	1	11	64
Total	50	27	25	24	11	2	36	175
Pedestrian movement								
Crossing from left	29	27	25	7	0	1	14	103
Crossing from right	17	0	0	14	11	1	21	64
Other	4	0	0	3	0	0	1	8
Total	50	27	25	24	11	2	36	175
Pedestrian speed								
Walking	50	0	24	24	0	0	15	113
Running	0	27	1	0	11	2	21	62
Total	50	27	25	24	11	2	36	175
Line of sight (1 sec)								
Not obstructed	45	20	25	24	11	2	25	152
Obstructed	5	7	0	0	0	0	11	23
Total	50	27	25	24	11	2	36	175

	Cluster							
	1	**2**	**3**	**4**	**5**	**6**	**7–14**	**Total**
Vehicle travel speed (km/h)								
Mean	43	35	48	51	37	87	43	44
N	50	27	25	24	11	2	36	175
Change of speed to impact (km/h)								
Mean	-7	-6	-6	-7	-11	-7	-5	-7
N	50	27	25	24	11	2	36	175

Table 22 Clusters for pedestrian accidents (OTS)

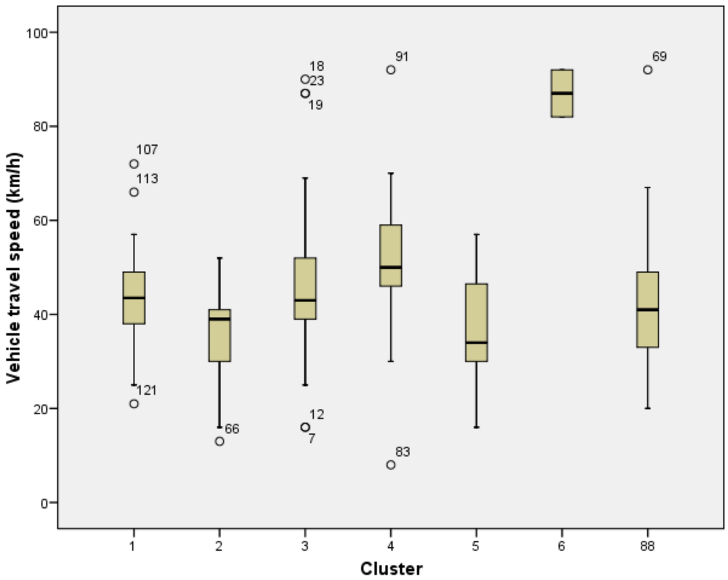


Figure 32 Vehicle travel speed (km/h)

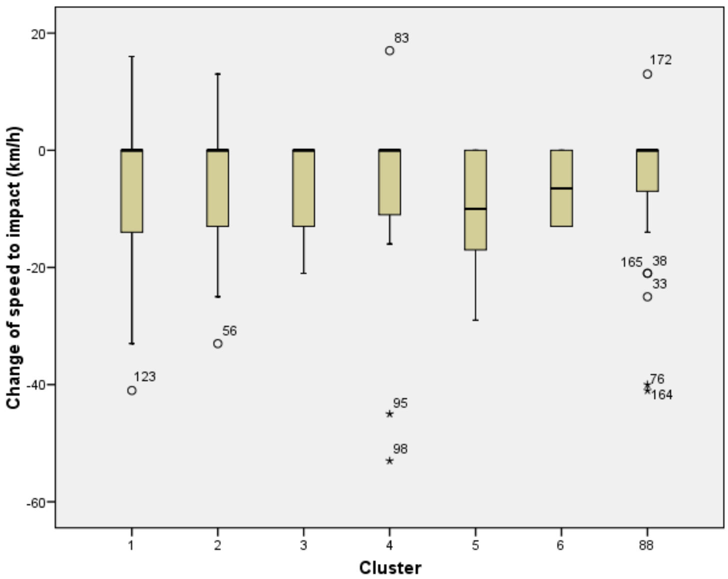


Figure 33 Change of speed to impact (km/h)

The fields contained in Table 22, Figure 32and Figure 33—pedestrian injury severity, light conditions, weather and so on—were used to define the similarity of accidents and in this sense constitute a set of core variables for these clusters. Each cluster is however nothing more than a group of accidents and can be described by any available fields, even where these were not used to form the groups. Table 23 and Figure 34 show two such parameters: the maximum injury level according to the Abbreviated Injury Score (AIS) and the distance between pedestrian and vehicle one second before impact.

	Cluster							
	1	**2**	**3**	**4**	**5**	**6**	**7–14**	**Total**
Pedestrian MAIS								
No injury	1	1	1	1	1	0	1	6
MAIS 1	21	18	10	6	4	0	15	74
MAIS 2	12	4	7	6	3	0	10	42
MAIS 3-6	8	4	7	8	2	2	7	38
Not specified	8	0	0	3	1	0	3	15
Total	50	27	25	24	11	2	36	175
Pedestrian distance 1 second before impact (m)								
Mean	11	9	12	13	9	23	12	11
N	50	27	25	24	11	2	36	175

Table 23 Further characteristics of OTS pedestrian clusters

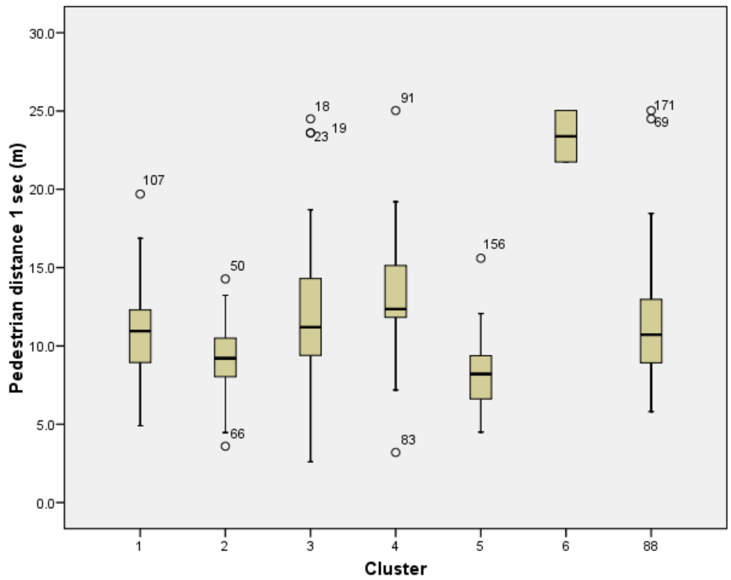


Figure 34 Pedestrian distance at 1 s before impact

4.4 Discussion

The decisive reason for using cluster analysis to identify groups or associated characteristics in the accident data was that the procedure is objective, reproducible and multivariate. It would not make sense to conclude this chapter on pedestrian accidents with a subjective interpretation of the data that attempts to override the factual findings presented in Table 10 and Table 22. The following observations should therefore be regarded as a postscript—a reflection on the results.

It would be ideal if the results of the independent cluster analyses of STATS 19 and OTS could be consolidated. In fact it would be remarkable if this were possible, considering the low overlap of cases, the different nature of the databases, and the limited number of fields that appear in identical form in the two datasets. Despite this, it is possible to perceive some striking parallels between the two sets of accident clusters. In the following three tables, elements of the four largest clusters from the two databases are summarised and placed side by side for comparison. These characteristics are then ‘amalgamated’ to produce what could be construed as consolidated accident scenarios.

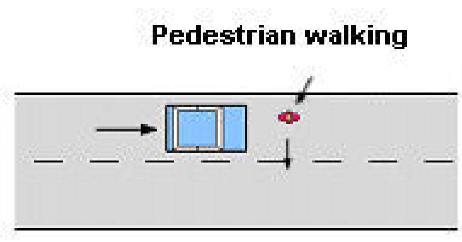
STATS 19 Cluster 1	OTS Cluster 1
39% of population	29% of population
Daylight	Daylight
Fine	Fine
Vehicle going ahead	Vehicle going ahead
10–30 mph speed zone	Speed 43 km/h
Children over-rep’d minority	Braking 7 km/h
Crossing, especially left	Crossing, especially left
Not masked	Walking
	Not obstructed
Consolidated features	
	Daylight Fine Vehicle going ahead Speed 43 km/h Braking 7 km/h Mid-size pedestrian Crossing from left Walking Not obstructed

Table 24 Baseline situation for pedestrian accidents

The set of characteristics of the largest clusters derived from STATS 19 and OTS mirror the most common features of the accident population, establishing a type of baseline scenario.

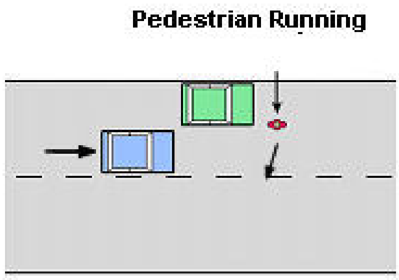
STATS 19 Cluster 2	OTS Cluster 2
14% of population Daylight Fine Vehicle going ahead 10–30 mph speed zone Children over-rep'd majority Crossing, especially left Masked by vehicle	15% of population Daylight Fine Vehicle going ahead Speed 35 km/h Braking 6 km/h Children over-rep'd majority Crossing from left Running Obstructed over-rep'd minority
Consolidated features	
	Daylight Fine Vehicle going ahead Speed 35 km/h Braking 6 km/h Small pedestrian Crossing from left Running Obstructed

Table 25 Small, obscured pedestrian

The set of characteristics from the second largest clusters differs from the first in having a smaller pedestrian who may be partially or fully obstructed from the line of sight of the driver and moving faster than walking pace.

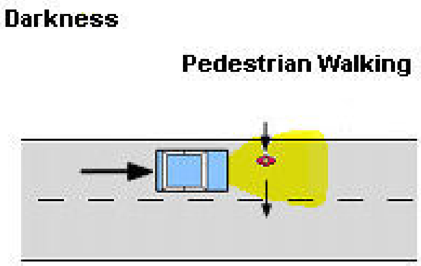
STATS 19 Clusters 3–4	OTS Cluster 3–4
21% of population (combined) Darkness Not fine over-rep'd minority Vehicle going ahead 10–30 mph speed zone Adult male over-rep'd majority Crossing, left and right Not masked	28% of population (combined) Darkness Fine/not fine Vehicle going ahead Speed 48–51 km/h Braking 6–7 km/h Adults Crossing, left and right Walking Not obstructed
Consolidated features	
	Darkness Not fine Vehicle going ahead Speed 50 km/h Braking 7 km/h Large pedestrian Crossing either direction Walking Not obstructed

Table 26 Large pedestrian in darkness and bad weather

The set of characteristics from the third and fourth largest clusters involves darkness and potentially wet conditions, with a large pedestrian crossing at walking pace from either side of the carriageway without sight obstruction.

4.5 Conclusion

The most common scenarios for pedestrian accidents identified in the STATS 19 and OTS databases are described in Table 10 and Table 22 respectively. These include a baseline scenario where a pedestrian steps out from the kerb without obstruction of the driver's line of sight, a similar situation except that the pedestrian is smaller and at least partially obscured, and a situation in adverse meteorological conditions with adult pedestrians. The derivation of these situations from the accident data using cluster analysis is objective and mathematically reproducible, also providing a clear definition of the proportion of the accident population represented by the scenarios.

5 Rear-end collisions

5.1 Cluster analysis of STATS19

5.1.1 Selection of cases

The vehicle file for STATS 19 (2008) contains information on 311,604 road users. The selection criteria for inclusion as case vehicles (the striking vehicle) in rear-end impacts were:

- cars and taxis with first point of impact on front
- collision partner a motorised vehicles with first impact to rear
- case vehicle and collision partner each made first impact with each other
- case vehicle and collision partner travelling to or from the same compass point direction
- no parked vehicles and case vehicle not reversing
- no unknown or missing information in key fields.

	None	Front	Back	Right	Left	Unk.	Total
Pedal cycle	869	8453	1924	3272	2273	6	16797
Motor cycle to 50 cc	372	2369	353	515	638	0	4247
Motor cycle 51-125 cc	509	3652	363	794	994	2	6314
Motor cycle 126-500 cc	267	1648	184	361	504	3	2967
Motor cycle over 500 cc	732	4880	530	1236	1519	2	8899
Taxi/private hire car	405	2338	863	804	733	1	5144
Car	11006	116029	46710	30919	26104	84	230852
Minibus (8-16 seats)	79	429	160	141	118	0	927
Bus or coach (17+ seats)	3724	2502	416	724	1007	2	8375
Other motor vehicle	514	1491	519	443	420	2	3389
Other non-motor vehicle	17	59	24	40	26	0	166
Ridden horse	47	21	18	22	6	0	114
Agricultural vehicle	104	196	112	168	64	0	644
Tram	2	10	1	2	8	0	23
Goods vehicle to 3.5 t	904	6635	2579	1827	1674	2	13621
Goods vehicle 3.5-7.5 t	236	1058	375	377	384	3	2433
Goods vehicle over 7.5 t	588	2671	950	1350	1044	4	6607
Unknown	13	35	4	7	9	17	85
Total	20388	154476	56085	43002	37525	128	311604

Table 27 Case vehicles: passenger cars with first point of impact to front surface

Cars (116,029) and taxis (2,338) with the first point of impact on the front surface were first selected from the entire database (Table 27). The vehicle type and first point of impact for the collision partners of these 118,367 are shown in Table 28. Motorised vehicles (excluding motorcycles) with the first point of impact to the rear were identified as suitable collision partners. This reduced the number to 30,130 with other cars (27,250) constituting the main type of vehicle struck in the rear.

	None	Front	Back	Right	Left	Unk.	Total
Pedal cycle	53	2417	1295	1210	1285	1	6261
Motor cycle to 50 cc	8	513	213	125	181	0	1040
Motor cycle 51-125 cc	9	738	241	208	252	0	1448
Motor cycle 126-500 cc	4	338	121	61	134	0	658
Motor cycle over 500 cc	9	916	345	216	295	0	1781
Taxi/private hire car	6	437	511	219	158	0	1331
Car	544	26845	27520	8923	6880	4	70716
Minibus (8-16 seats)	2	91	83	34	34	0	244
Bus or coach (17+ seats)	9	322	165	125	72	0	693
Other motor vehicle	9	240	193	113	72	0	627
Other non-motor vehicle	1	12	11	15	11	0	50
Ridden horse	0	10	7	3	3	0	23
Agricultural vehicle	3	61	62	52	25	0	203
Tram	0	2	0	2	1	0	5
Goods vehicle to 3.5 t	34	1191	1072	403	272	0	2972
Goods vehicle 3.5-7.5 t	8	189	151	78	54	0	480
Goods vehicle over 7.5 t	28	424	373	232	102	0	1159
Unknown	0	6	0	0	1	28669	28676
Total	727	34752	32363	12019	9832	28674	118367

Table 28 Collision partners: motorised vehicles with first point of impact to rear surface

A check that no third vehicles were involved in the first impacts of the case vehicle and its rear-ended collision partner reduced the number of cases to 27,142 (Table 28).

Case vehicle	27142
Other vehicle	2988
Total	30130

Table 29 First impact of collision partner

STATS 19 nominates compass points for the directions from which and to which a vehicle was heading. To capture the notion that the vehicles in a rear-end collision should have been generally travelling in the same line of traffic, it was specified that the colliding vehicles should share at least one direction from which or to which they were heading. This is shown in Table 30 where converging means that only the ‘to’ direction was shared, diverging that only the ‘from’ direction was shared, and same direction that they were heading both to and from the same compass points. This filter excluded 3,781 cases.

Same direction	20484	
Same direction—diverging	2369	
Same direction—converging	508	
Sub-total		23361
Other	3781	3781
Total	27142	27142

Table 30 Direction of movement of case vehicle and collision partner

A final filter in the selection of eligible cases was made of parked vehicles, any case vehicle that was reversing, and of any database record that contain missing or unknown information in the key fields for input to the cluster analysis. This resulted in a total of 22,384 case vehicles meeting the selection criteria for rear-end impacts (Table 31).

Included	22384
Excluded	977
Total	23361

Table 31 Exclusion of parked or reversing vehicles and cases with missing information

The capacity of the hardware and software used to carry out the cluster analysis was found to be limited to around 10,000 to 11,000 records, at which level memory storage errors were encountered or processing appeared to be continue indefinitely, e.g. over six hours. For this reason the 22,384 eligible vehicle records were sorted by their identification numbers in the database and every second record was removed. The remaining 11,192 were passed on to the cluster analysis.

Included	11192
Excluded	11192
Total	22384

Table 32 Filtering to half of eligible cases

Table 33 shows accident severity for the cases included or excluded at this final filter. There was no statistically significant difference between the groups for this or the other variables listed in Table 34 that were used in the cluster analysis (chi-square test to 90% significance).

	Selected		
	No	Yes	Total
Slight	10835	10844	21679
Serious	340	329	669
Fatal	17	19	36
Total	11192	11192	22384

Table 33 Comparison of included and excluded cases: accident severity

5.1.2 Input dataset

Eight descriptive parameters were included in the dataset for the analysis of rear-end accidents in STATS 19. Quantitative information on speed, following distance and location can only be obtained from an in-depth database; however speed limit is available as a proxy for speed and the information on junctions and vehicle manoeuvres enable a picture of accident circumstances to take form.

Ordinal	Accident severity	0.0 0.5 1.0	Slight Serious Fatal
Ordinal	Speed limit (mph)	0.0 0.5 1.0	10-30 40-50 60-70
Nominal	Junction detail	1 2 3	Not at junction Roundabout Junction
Nominal	Light conditions	1 2	Daylight Darkness
Nominal	Weather conditions	1 2	Fine Not fine
Nominal	Vehicle A manoeuvre	1 2 3	Going ahead Stopping, starting, held-up Turning
Nominal	Vehicle B manoeuvre	1 2 3	Going ahead Stopping, starting, held-up Turning
Nominal	Vehicle directions	1 2 3	Following Diverging Converging

Table 34 Structure of dataset for rear-end accidents (STATS 19)

The case vehicle, i.e. the car with frontal impact, is nominated as vehicle A and its collision partner, the motorised vehicle with rear impact, as vehicle B.

5.1.3 Results

The dendrogram produced by the cluster analysis was cut at a level where the population of the dataset (11,192) was formed into 18 groups. The largest six of these groups comprise 86% of the population and are described in detail in Table 34. The characteristics of the remaining clusters are shown in aggregate for completeness. The figures for the relative direction of movement of the vehicles as defined by ‘to’ and ‘from’ compass point directions are shown in Table 36. This parameter was not used in the cluster analysis to influence the formation of groups because it is highly correlated with the vehicle manoeuvre fields.

	Clusters							
	1	2	3	4	5	6	7–18	Total
Cluster representativeness (%)								
Slight	30	19	13	12	8	5	13	100
Serious	21	10	12	30	9	4	14	100
Fatal	21	0	0	53	0	11	16	100
Total	30	18	13	13	8	5	14	100
Accident severity								
Slight	3249	2025	1401	1293	890	523	1463	10844
Serious	69	32	41	100	28	12	47	329
Fatal	4	0	0	10	0	2	3	19
Total	3322	2057	1442	1403	918	537	1513	11192
Speed limit (mph)								
10-30	1966	1199	932	364	541	235	827	6064
40-50	582	355	236	190	171	117	274	1925
60-70	774	503	274	849	206	185	412	3203
Total	3322	2057	1442	1403	918	537	1513	11192
Junction detail								
Not at junction	1162	679	85	1403	0	363	633	4325
Roundabout	513	528	140	0	257	174	283	1895
Junction	1647	850	1217	0	661	0	597	4972
Total	3322	2057	1442	1403	918	537	1513	11192
Light conditions								
Daylight	2973	1932	1264	1007	747	359	518	8800
Darkness	349	125	178	396	171	178	995	2392
Total	3322	2057	1442	1403	918	537	1513	11192
Weather conditions								
Fine	3040	1983	1227	1096	840	0	833	9019
Not fine	282	74	215	307	78	537	680	2173
Total	3322	2057	1442	1403	918	537	1513	11192
Vehicle A manoeuvre (striking vehicle)								
Going ahead	3312	0	1225	1403	888	537	515	7880
Stopping, starting, held-up	0	2057	162	0	24	0	735	2978
Turning	10	0	55	0	6	0	263	334
Total	3322	2057	1442	1403	918	537	1513	11192
Vehicle B manoeuvre (struck vehicle)								
Going ahead	0	72	0	1403	918	46	145	2584
Stopping, starting, held-up	3322	1985	0	0	0	491	1088	6886
Turning	0	0	1442	0	0	0	280	1722
Total	3322	2057	1442	1403	918	537	1513	11192

Table 35 Clusters for rear-end accidents (STATS 19)

The cells highlighted in green indicate (a) fields for which the distribution of cases is significantly different from the population (chi-square test to 99.5% significance) and (b) cells in these fields that are over-represented. The results for cluster representativeness are derived directly from the numbers on accident severity, expressing the latter as row percentages.

Cluster 1, the largest group comprising 30% of the population, reflects the most common characteristics of the overall population: low accident severity, lower speed limits, daylight and fine weather with the case vehicle ‘going ahead’ and colliding with a vehicle in front that was ‘stopping, starting or held up’ at a junction. Cluster 2, comprising 18% of the population, is similar except for a tendency to occur at roundabouts with both vehicles stopping, starting or held up. Cluster 3 highlights accidents at junctions where the car in front is turning. Cluster 4, which comprises the same 13% of the population as cluster 3 but considerably more fatal and serious accidents, occurs away from junctions on mostly high speed roads with both vehicles coded as going ahead and some tendency to wet weather. The ‘vehicle directions’ field gives little indication that these accidents were directly related to merging or diverging manoeuvres.

	Clusters							
	1	**2**	**3**	**4**	**5**	**6**	**7–18**	**Total**
Vehicle directions								
Following	3223	2007	463	1386	869	525	1326	9799
Diverging	63	28	909	8	21	5	124	1158
Converging	36	22	70	9	28	7	63	235
Total	3322	2057	1442	1403	918	537	1513	11192

Table 36 Further characteristics of clusters for rear-end accidents (STATS 19)

5.2 Detailed OTS case reviews

This section outlines the results obtained from the detailed accident reviews of 50 OTS accidents involving car-to-car rear impacts. Each accident was examined from 5 seconds before the initial impact at 0.5 second intervals giving not only the closing distance between the vehicles but also the lateral distance relative to the centre of the vehicles involved. Where there is no evidence to suggest otherwise, it has been taken that the vehicles were travelling at the speed limit for the carriageway and at the point of braking for the case vehicle maximum braking capacity has been assumed. The lateral movement before impact is recorded to the nearest 0.5 metres.

As one may expect, the majority of the rear-impact accidents occurred when the vehicle in front has come to a complete stop and the following vehicle failed to stop in the same distance. A number of the accidents occurred in moving traffic although in reality this is more often than not ‘stop-start’ traffic when the speed of the following traffic is somewhat erratic.

The proportion of accidents by the general driving action by both vehicles involved is shown in Figure 35.

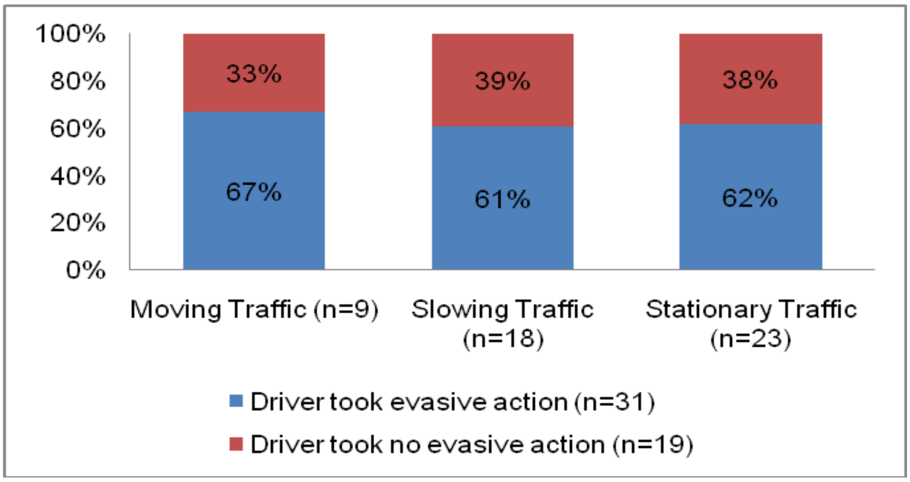


Figure 35: Struck vehicles driving action before impact

Figure 35 highlights the common accident scenarios where the vehicle in front is in stationary traffic (46%, n=23). Accidents occurring in moving traffic represent 18% (9) of the accident population.

For this work the term "stationary traffic" means where the struck vehicle is stationary or moving at less than 5 km/h. The term "slowing traffic" is where the struck vehicle is slowing down due to an obstruction or to perform a manoeuvre ahead of the case vehicle which does not react in time. The term "moving traffic" applies where the struck vehicle is in moving traffic, for example the struck vehicle is in moving traffic and the vehicle did not react and collided with the rear of the struck vehicle.

The accident configuration in regards to lateral overlap of the collision partners for the 50 OTS accidents is presented in Table 37 below. The overlap of collision damage was evaluated by recording direct contact damage across the front of the case vehicle, with the front elevation divided into four equal segments. The lateral area of damage is displayed by recording the percentage of damage from the front corners of the vehicle.

	Overlap from left (nearside) of vehicle		Overlap from right (offside) of vehicle		Total overlap	
25% overlap	4	8%	3	6%	**7**	**14%**
50% overlap	4	8%	4	8%	**8**	**16%**
75% overlap	10	20%	6	12%	**16**	**32%**
100% overlap	0	0	0	0	**19**	**38%**
Total	18	36%	13	26%	**50**	**100%**

Table 37: Damage overlap of collision partners

The largest proportion of the collisions had an accident configuration of 100% overlap (n=19, 38%); the second largest accident configuration had a 75% overlap with 32% (n=16) of the sample. The lowest category with only 14% (n=7) of the accident sample was a 25% overlap between accident participants. Table 37 shows the distribution of overlap for vehicles when 100% engagement between collision participants was not achieved. This shows the percentage of overlap and lateral location of damage. In 8% of the accidents the case vehicle only engaged the stuck vehicle by 25% located at the front left corner, a similar trend was observed for the right corner with 6% of accidents. Of the complete sample 36% (n=18) of the case vehicles had damage covering the left (nearside) ¾ of the front elevation.

The driver’s pre-impact action was evaluated in the review process and it was established whether the driver had performed an evasive action or not. The driver took evasive action in 31 accidents (62%) and the driver took no action in 19 accidents (38%). An evasive action included braking or steering by the driver prior to impact in an attempt to avoid the accident rather than as a part of normal driving.

Six accidents where the case vehicle was turning prior to impact were categorised according to driving conditions at the time; three were coded as moving traffic because the struck vehicle was slowing to turn into a side road; two accidents were in stationary traffic where the struck vehicle had stopped waiting for another vehicle to turn; and one accident was coded as slowing because the struck vehicle was slowing down to allow a vehicle to turn.

The results in the table below give the accidents according to road environment. The majority of the accidents (64%) occurred in an urban road environment. The 36% of accidents that occurred on a rural road generally encompassed the slowing traffic and moving traffic driving conditions.

	Driver took evasive action		Driver took no evasive action	
Rural road environment	12	24%	6	12%
Urban road environment	19	38%	13	26%
Total	31	62%	19	38%

Table 38: Evasive action by urban/rural environment

The injury severity of the reviewed accidents is shown in the table below and consisted of slight and non-injury accidents.

	Driver took evasive action		Driver took no evasive action	
Non-injury	14	28%	8	16%
Slight	17	34%	11	22%
Total	31	62%	19	38%

Table 39: Accident severity

Although rear accidents occur in longitudinal traffic with the vehicles in the same lane and travelling in the same direction, the road geometry still affects the perception and reaction time of the drivers involved. The table below presents the accidents according to the road geometry, showing that the majority (80%) of the accident population occurred on a straight road.

	Driver took evasive action		Driver took no evasive action	
Bend Left	5	10%	2	4%
Bend Right	1	2%	2	4%
Straight Road	25	50%	15	30%
Total	31	62%	19	38%

Table 40: Evasive action by road geometry

Figure 36 gives information regarding the meteorological conditions for the road environment prior to impact and if the driver took evasive action or not. The data is displayed to show the percentage of cases that occurred in four meteorological conditions and what proportion of accidents in those conditions where an evasive action was conducted or not.

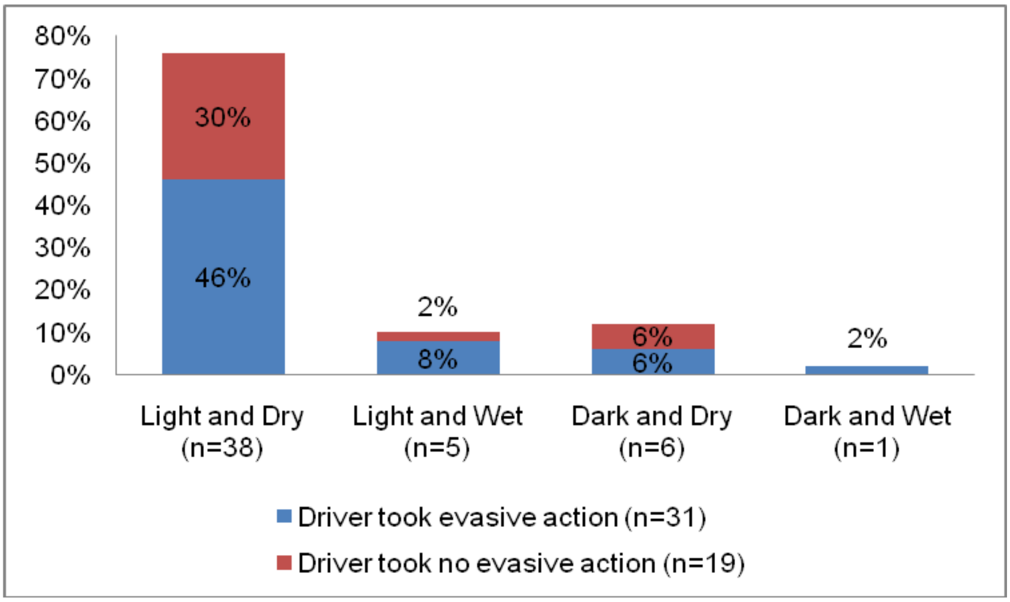


Figure 36: Carriageway conditions

The majority of rear impacts occurred during daylight hours when the carriageway surface was dry with 76% of the accidents fitting this condition, in 46% of those accidents the driver performed an evasive action. In comparison only 14% of accidents occurred in the dark with only 2% occurring in the dark and a wet road.

Sight obscuration between the drivers did not present itself as a significant causal factor in this accident group. The main causal factors identified included driver distraction, inadequate stopping distance between the vehicles and insufficient braking systems for the case vehicles.

Driver distraction	Driver took evasive action		Driver took no evasive action	
Yes	17	34%	11	22%
No	5	10%	4	8%
Unknown	9	18%	4	8%
Total	31	62%	19	38%

Table 41: Evasive action by driver distraction

Table 40 shows that although the driver performing no evasive action prior to the impact accounts for 38% of the accident population, the driver was deemed to be distracted prior to impact in 22% of the total accident population.

Figure 37 shows the travel and impact speeds for the striking (case) vehicle and struck vehicles for the 50 accidents reviewed.

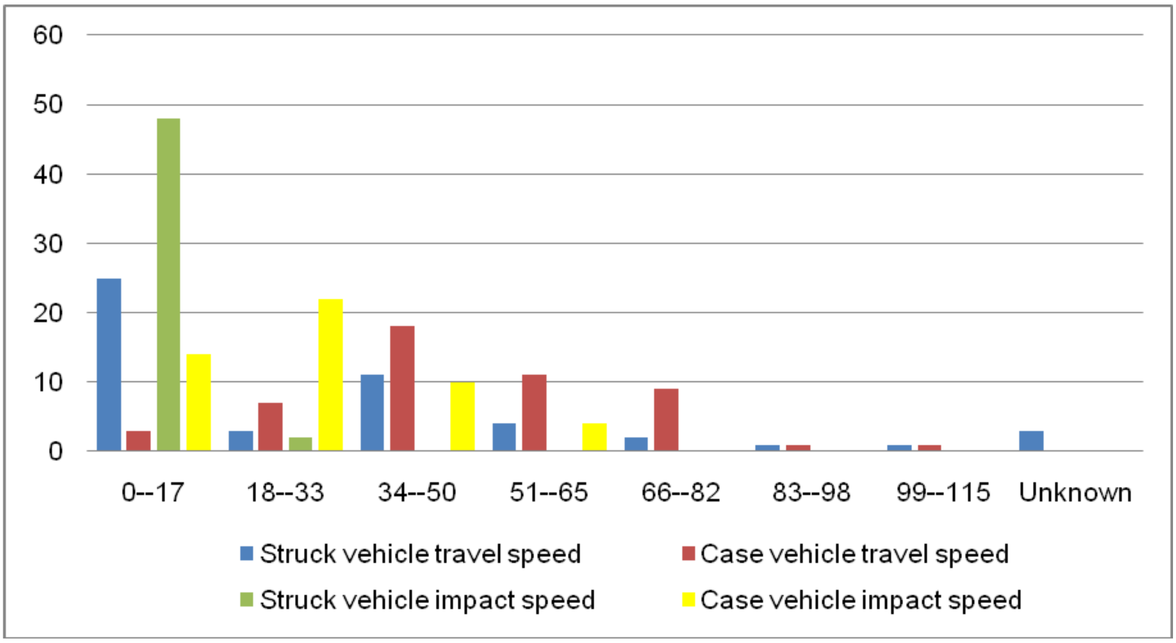


Figure 37: Vehicle impact and travel speeds for both vehicles involved in the collision

At the point of impact 48 (96%) of the struck vehicles were travelling below 17 km/h, with 25 (50%) of these vehicles also having a pre-impact travel speed of below 17 km/h, indicating no change in velocity of braking. In contrast, 38 (76%) of the striking (case) vehicles had a pre-impact travel speed of 34–82 km/h with 22 (44%) of these striking vehicles having an impact speed of 18–33 km/h demonstrating a change in velocity and braking prior to impact.

Figure 38 and Figure 39 below show the change in velocity for the case vehicle and struck vehicle respectively. The change in velocity from the recorded travel speed to the impact speed for each vehicle involved in the 50 car–to-car rear impacts.

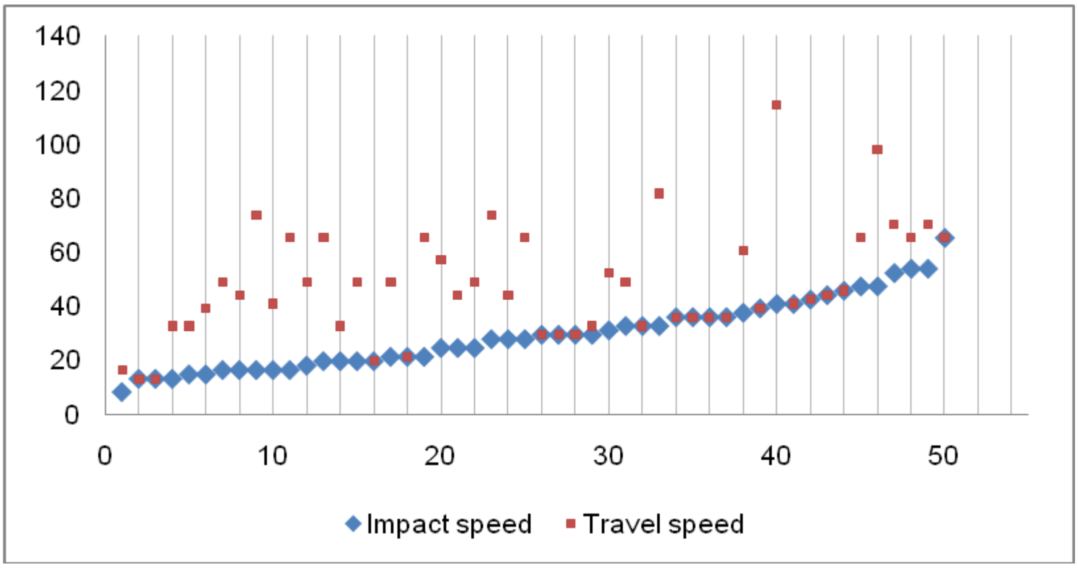


Figure 38: Showing change in travel speed to impact speed for the case vehicle (km/h)

Whether a vehicle braked before impact or not was assessed using the presence of physical scene evidence or vehicle damage. In the absence of evidence to the contrary, the travel speed was deemed to be equivalent to the speed limit of the road. Figure 38 shows that in the majority of the cases, 31 (62%), the case vehicle braked prior to impact reducing its speed at impact. The level of braking involved in those cases is hard to quantify as a result of a lack of scene evidence. Locked wheel marks were present in 10 of the 32 cases where braking had been applied indicating that full braking had been applied.

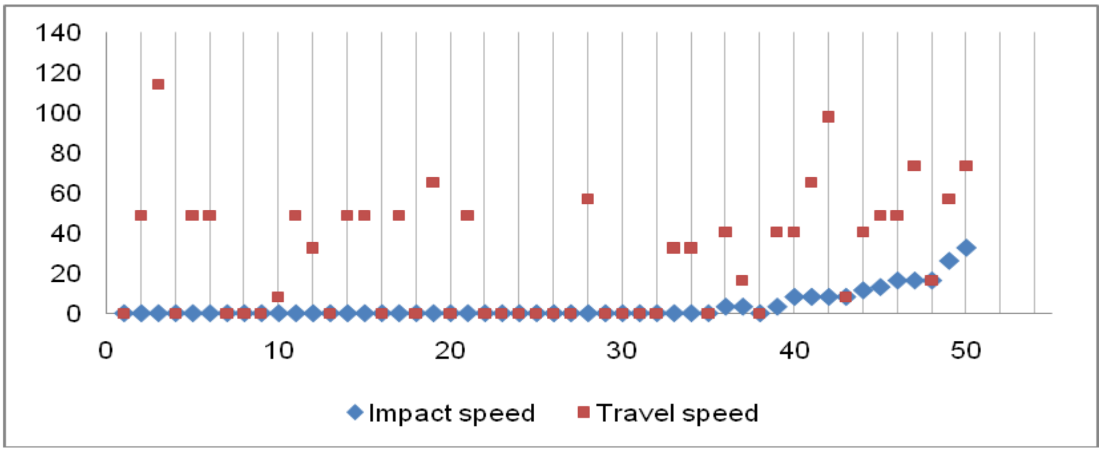


Figure 39: Change in travel speed to impact speed for the struck vehicle (km/h)

Figure 39 shows the change in velocity from travel speed to impact for the struck vehicle in the accident. This shows that there was a change in velocity in 27 (54%) of the accidents. This also concludes that in 23 of the accidents the struck vehicle was stationary prior to impact.

Figure 40 shows the average distance the case vehicle is from the point of impact at 0.5 second intervals from 5 seconds before the impact.

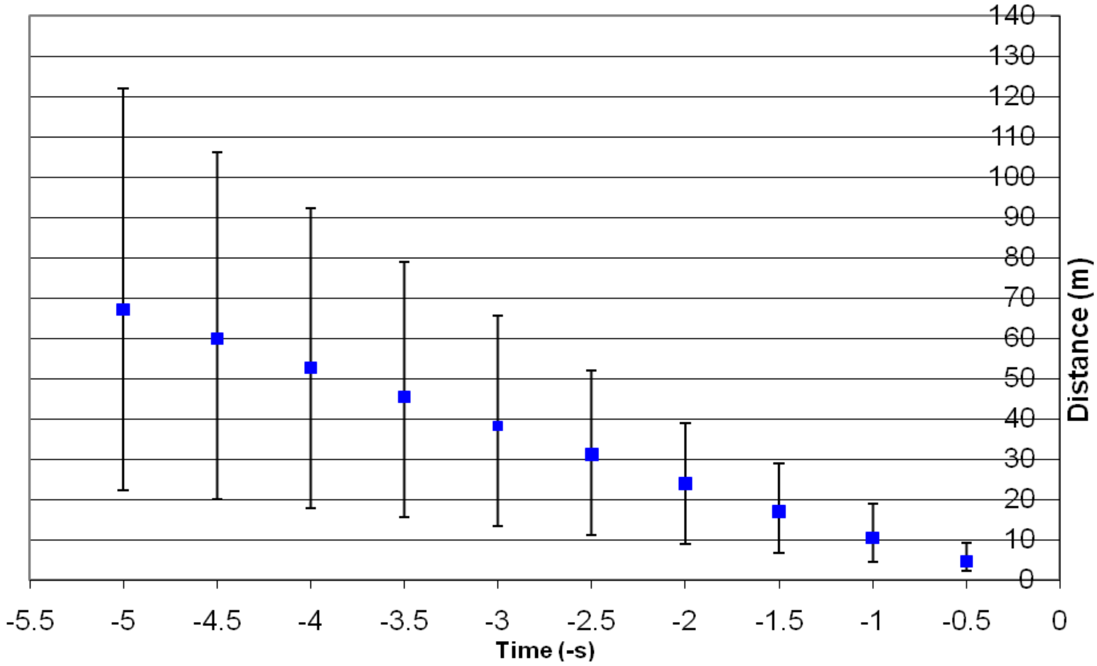


Figure 40: Distance of case vehicle from point of impact

This shows that 1 second before the impact the case vehicle is an average of 10 metres from the point of impact and ultimately the rear of the vehicle in front. The chart also shows the distribution of the distance for all vehicles with a maximum and minimum distance between the collision participants prior to the impact.

Figure 41 gives the average speed as travelled by the case vehicle prior to impact.

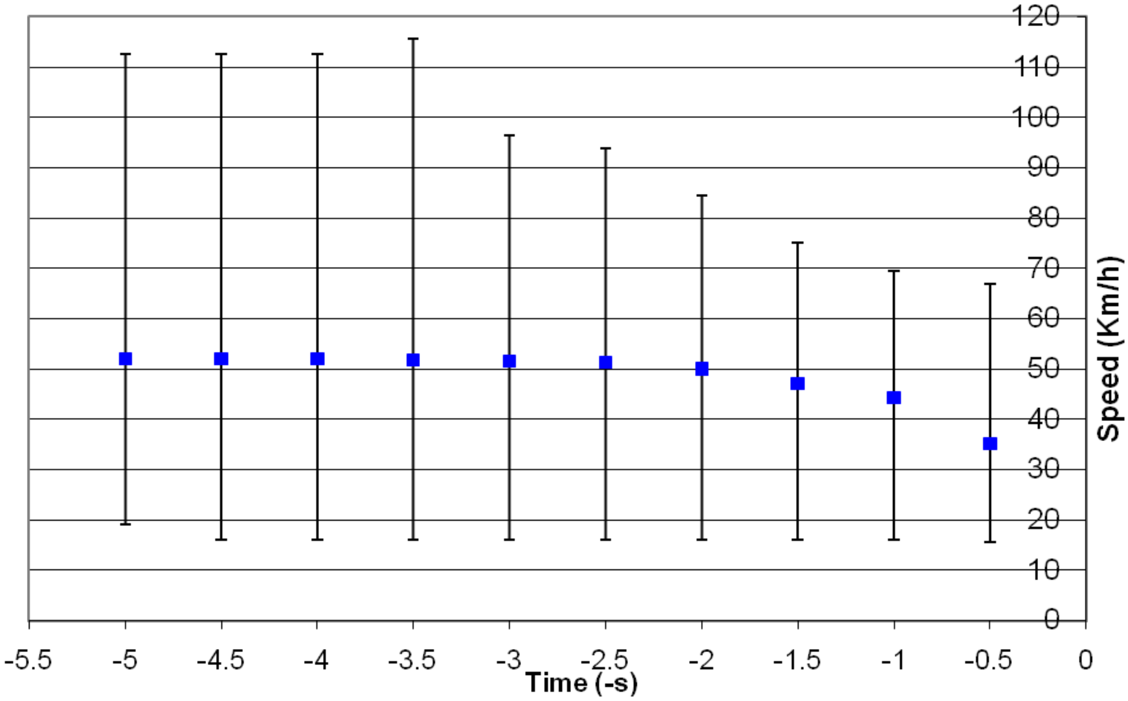


Figure 41: Speed of case vehicle before impact

The speed remains relatively constant at the average travel speed of 52 km/h before the driver has reacted and applied braking at approximately 2 seconds before the collision when the speed reduces to the impact speed.

Figure 42 shows the average headway distance between the accident participants. The headway distance is the distance between the two collisions partners (striking and struck vehicles) as they are moving and approaching the impact point. This is not to be confused with the distance to the point of impact, which is the distance from the vehicle concerned to point of impact which is independent of the location of the other vehicle.

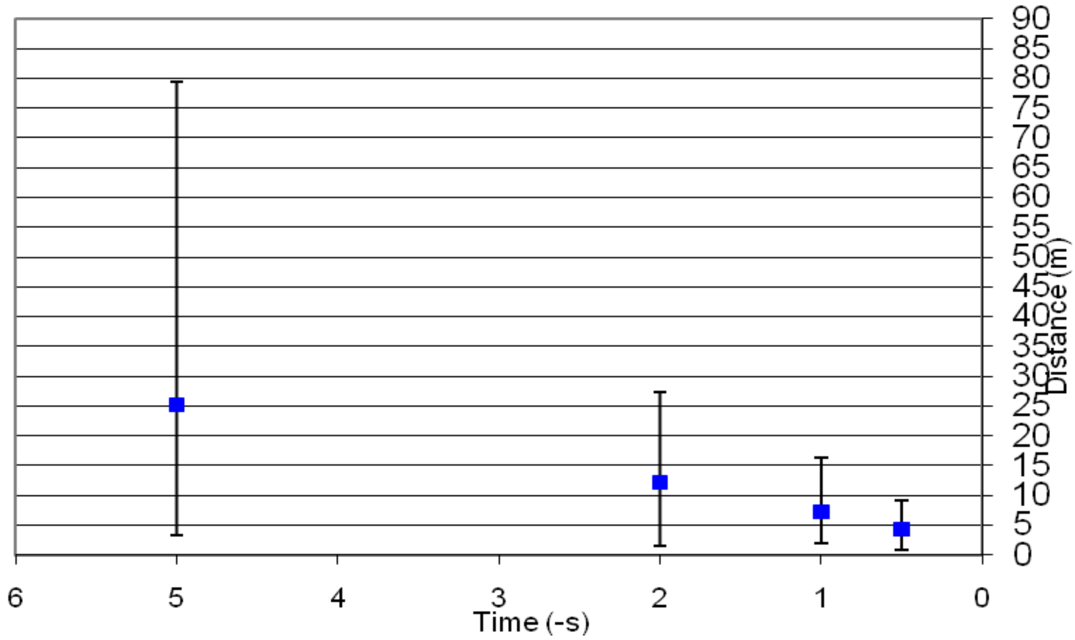


Figure 42: Headway distance between colliding vehicles before impact

The maximum and minimum distances are displayed to give the distribution of vehicles analysed. At 5 seconds before impact an average distance of 25 m was observed between accident participants with a maximum headway distance of 79 m and a minimum of 3.5 m. At 2 seconds before impact there was an average distance 12.5 m between the accident participants, with a maximum of 27.3 m and minimum of 1.4 m.

Figure 43 and Figure 44 show the headway distance between the accident participants where the driver of the case vehicle performed or did not perform an evasive action respectively.

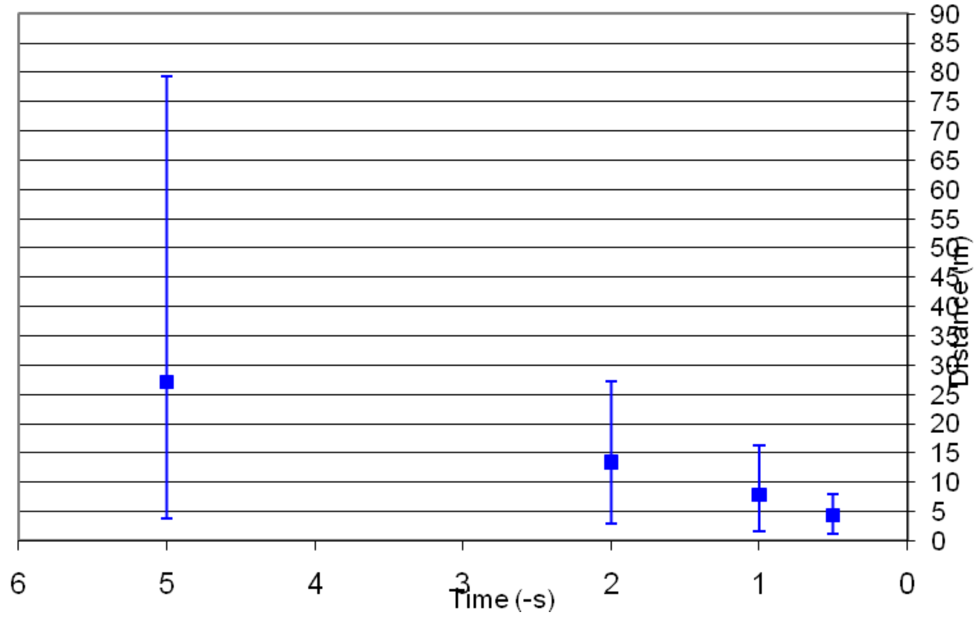


Figure 43: Headway distance before impact where driver took evasive action

These show the decrease in headway distance as a result of this action. If the driver had performed an evasive action then an average headway distance of 27 m was recorded compared to 24 m if no evasive action had been taken at 5 seconds before the impact. This decrease in headway was observed for 2 and 1 seconds before the impact but at 0.5 seconds before the impact the distance was similar at 4.4 m and 4.2 m for evasive action and no evasive action respectively.

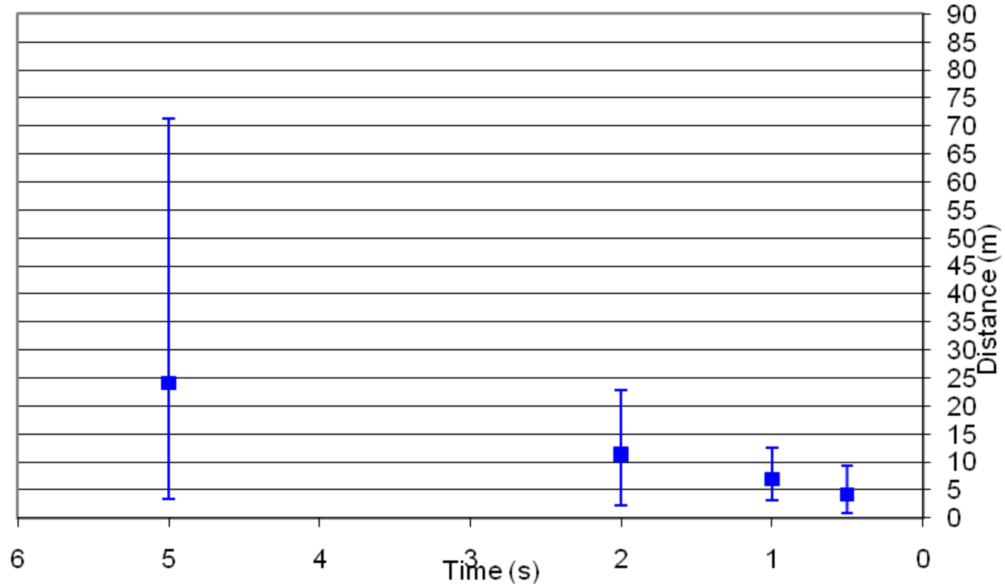


Figure 44: Headway distance before impact where driver did not take evasive action

The review of the 50 rear-end accidents found that the prevalent accident scenario occurred when there was stationary traffic ahead, although this group would also include junction restarts. With this type of accident scenario, typically the driver of the case vehicle reacted late, if at all, to the stationary vehicle ahead. Alternatively, in the junction restart accidents the driver believes the vehicle has moved off and is no longer an obstruction and therefore makes the decision not to react. The other common scenario in this accident group was slowing traffic, this is where the struck vehicle is slowing down to perform a manoeuvre, or due to an obstacle ahead, and the driver in the case vehicle does not react to the unexpected action and therefore only applies braking at the last moment. Although the human factors involved in the accidents were evaluated with regards to driver distraction, their evaluation is beyond the scope of this work.

5.3 Cluster analysis of OTS

5.3.1 Selection of cases

All of the vehicles from the detailed case reviews were included in the cluster analysis. To recap the process by which these were chosen, a random selection was made from all of the eligible cases in OTS provided adequate evidence from the scene of the accident had been collected and recorded. The criteria for case vehicles were:

- passenger cars
- first point of impact on front surface
- no interaction with pedestrians
- accident coded as “Rear end”.

	Collision sub-type								
Collision type	1	2	3	4	5	6	7	8	Total
A Overtaking and lane change	137	37	104	49	6	17	7	11	368
B Head on	52	14	31	14	36	63		8	218
C Lost control or off road (straight roads)	221	271	134					8	634
D Cornering	358	348	45					15	766
E Collision with obstruction	92	9	60	1	6			9	177
F Rear end	190	24	11	294	68	75		14	676
G Turning versus same direction	3	40	8	16	39	19		2	127
H Crossing (no turns)	243								243
J Crossing (vehicle turning)	285	3	9					7	304
K Merging	64	71	6					3	144
L Right turn against	12	212	1					1	226
M Manoeuvring	34	9	43	12	1		4	11	114
N Pedestrians crossing road	117	84	1	2	1	1	5	6	217
P Pedestrians other	5	1	5	2	1	5		7	26
Q Miscellaneous	1	4	1	7			16	15	44
Total	1814	1127	459	397	158	180	32	117	4284

Table 42 Collision category for all accidents (OTS)

All accidents on the OTS database are categorised according to the system in Table 42. The sub-type numbers have a different meaning for each main category A to Q. There were 676 accidents classified as rear-end from the population of 4,284 accidents on the database.

	Back	Front	Left	Right	Top	Bottom	Other	Total
Car	912	2598	796	1121	16	79	478	6000
LGV	75	210	43	62	2	13	50	455
HGV	29	183	48	77	4	15	66	422
Bus	12	55	13	15		9	15	119
Motorcycle	14	194	54	55	2	9	102	430
Pedal cycle	11	49	34	40		1	29	164
Other	2	3	4	5			6	20
Unknown	7			2		38	8	55
Total	1062	3292	992	1377	24	164	754	7665

Table 43 Vehicle type by first impact side

The population of 4,284 accidents involves 7,665 vehicles. Of these, there were 2,598 cars that made first impact on the front surface, as shown in Table 43.

A total of 480 vehicles were cars that made first impact on the front surface, were in accidents classified as rear end, and did not strike a pedestrian. Following a check that the collision partner was struck in the rear, the quota of 50 vehicles was selected randomly from accidents investigated in phases 2 and 3 of the OTS project (2003–2009) and used for both the detailed case reviews and the cluster analysis.

5.3.2 Input dataset

The dataset derived from OTS for rear-end impacts contains seven parameters to describe the circumstances of the accidents. The four scale variables are drawn exclusively from the detailed case reviews and reconstructions. Seven dimensions for 50 cases is a high number and, depending on the data, could have resulted in 'over-fitting' of the information in the dataset; however the coherency of the results validates the set up. The rationale for using 'change of speed to impact' rather than the perhaps more natural 'speed at impact' was to avoid the strong correlation between the pre-emergency travel speed and the speed at impact. The information conveyed is the same, but sidesteps some technical obstacles. Vehicle A refers to the case (striking) vehicle and vehicle B to the collision partner struck in the rear.

Ordinal	Accident severity	0.0 0.5 1.0	Slight or non-injury Serious Fatal
Nominal	Light conditions	1 2	Daylight Darkness
Nominal	Weather	1 2	Fine Not fine
Scale	Veh A travel speed	0–1	*Scaled from km/h*
Scale	Veh A change speed to impact	0–1	*Scaled from km/h*
Scale	Veh B travel speed	0–1	*Scaled from km/h*
Scale	Following distance (1 sec)	0–1	*Scaled from m*

Table 44 Structure of dataset for rear-end accidents (OTS)

5.3.3 Results

The results of the cluster analysis of the OTS dataset are detailed at the level of 8 clusters, three of which cover 82% of the population of the dataset. The cells shaded in green indicate (a) that the distribution of numbers for the given field is significantly different from the distribution in the whole population (chi-square test to 95% significance) and (b) that the particular number highlighted is over-represented. This is the same as for the STATS 19 analysis except that the statistical test is evaluated at 95% confidence instead of 99.5%. This level is better suited to the lower number of cases in OTS for providing an objective test of differences between the clusters and the overall population.

	Cluster				
	1	2	3	4–8	Total
Cluster representativeness (%)					
Slight or non-injury	58	12	12	18	100
Serious	0	0	0	0	0
Fatal	0	0	0	0	0
Total	58	12	12	18	100
Accident severity					
Slight or non-injury	29	6	6	9	50
Serious	0	0	0	0	0
Fatal	0	0	0	0	0
Total	29	6	6	9	50
Light conditions					
Daylight	29	6	0	7	42
Darkness	0	0	6	2	8
Total	29	6	6	9	50
Weather					
Fine	29	6	6	3	44
Not fine	0	0	0	6	6
Total	29	6	6	9	50
Vehicle A travel speed (km/h)					
Mean	41	54	39	73	48
N	29	6	6	9	50
Vehicle A change of speed to impact (km/h)					
Mean	-15	-6	-9	-44	-19
N	29	6	6	9	50
Vehicle B speed at impact (km/h)					
Mean	0	19	5	3	7
N	29	6	6	9	50
Following distance : 1 sec (m)					
Mean	8	4	6	8	3
N	29	6	6	9	50

Table 45 Clusters for rear-end accidents (OTS)

Cluster 1, which comprises a full 58% of the population, contains accidents that occurred in daylight and fine weather with the struck vehicle stationary or near stationary in all cases (mean speed 0.3 km/h). The striking vehicles were travelling with a speed of 41 km/h at a following distance of 8 m and slowed by 15 km/h before impact. These are all mean values—the spread of values is shown in Figure 45 to Figure 48. Clusters 3 and 4 each contain 6 members, from which it is unsafe to draw strong generalisations; however common characteristics are that both vehicles are moving at impact and the initial following distance is centred around 4–6 m.

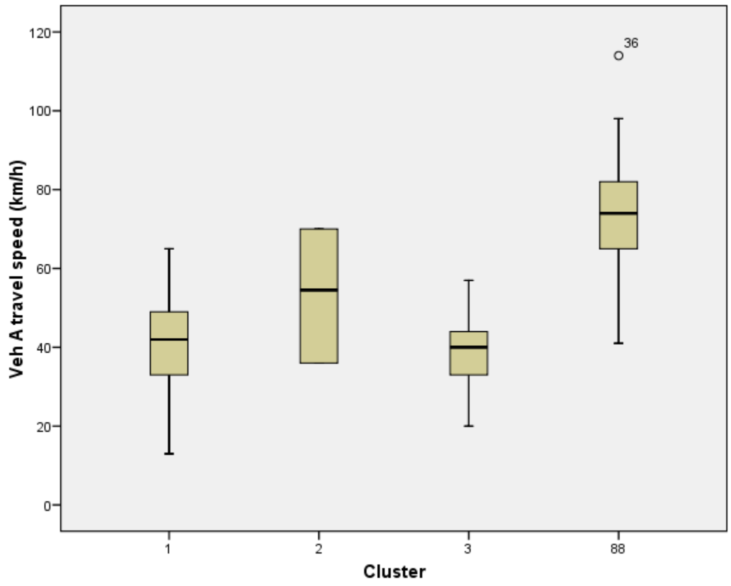


Figure 45 Pre-emergency travel speed of striking vehicle (frontal impact)

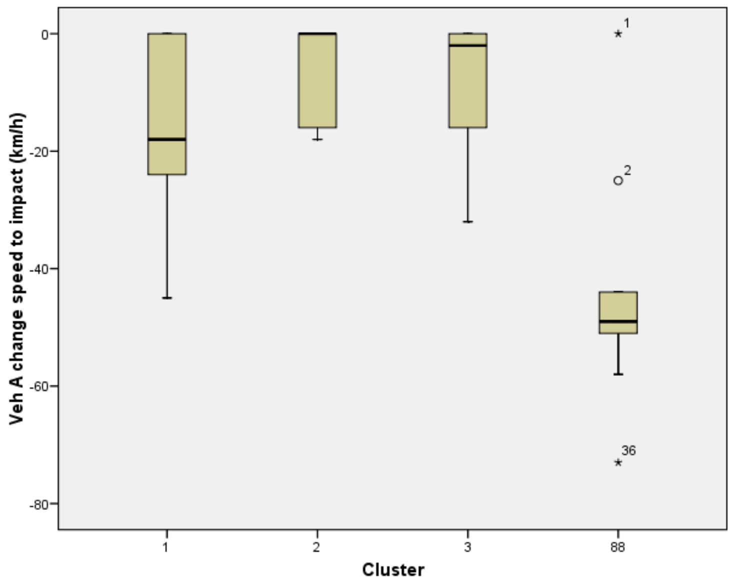


Figure 46 Change of speed to impact of striking vehicle

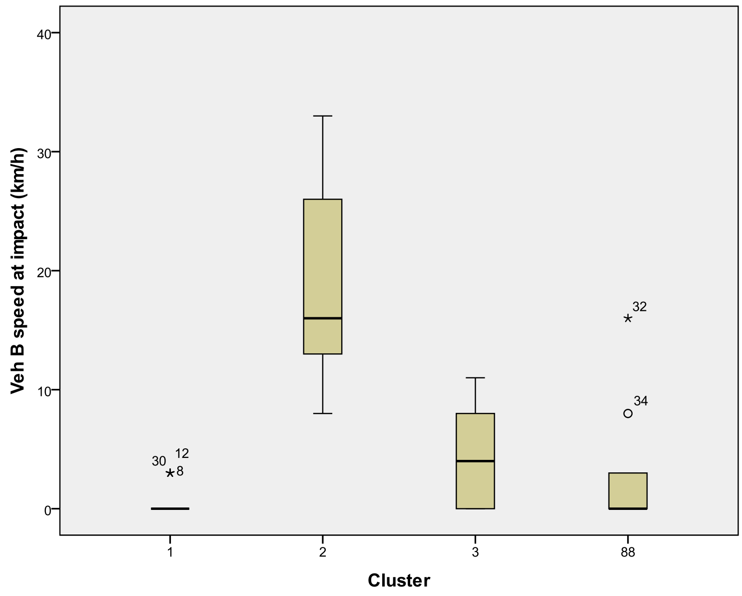


Figure 47 Speed at impact of struck vehicle (rear impact)

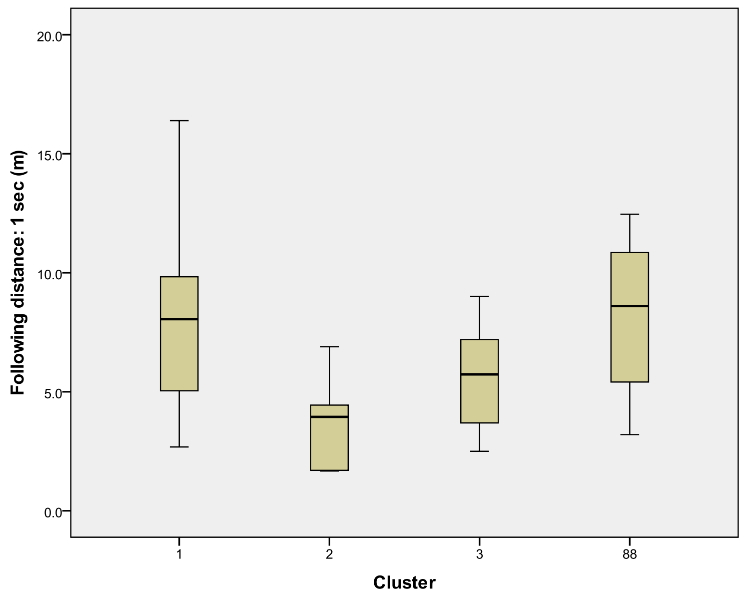


Figure 48 Following distance at 1 second before impact

5.4 Discussion

The results of the two independent analyses of rear-end accidents derived from STATS 19 and OTS can be compared and contrasted—not forgetting the point stressed in section 4.4 that this is a subjective reflection on objective results. As with pedestrian accidents, it is possible to perceive some striking parallels.

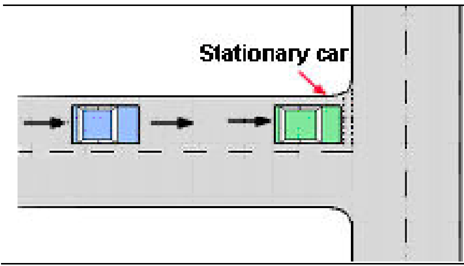
STATS 19 Clusters 1–3	OTS Cluster 1
61% of population (combined) Daylight Fine weather Roundabouts and junctions Lower speed zones Struck vehicle stopping, starting, held-up or turning	58% of population Daylight Fine weather Travel speed 41 km/h Braking 15 km/h Struck vehicle stationary at impact Following distance 8 m
Combined features	
	Daylight Fine weather Roundabouts and junctions Travel speed 41 km/h Braking 15 km/h Struck vehicle stationary at impact Following distance 8 m

Table 46 Rear-end accidents: stationary vehicle

It is not possible to specifically identify stationary vehicles in STATS 19, but the emphasis in STATS 19 clusters 1–3 on ‘stopping, starting and held-up’ at roundabouts and junctions suggests low speed on the part of the struck vehicle. The shared characteristics of these three largest clusters appear consistent with OTS cluster 1, and the closeness of the representativeness figures is remarkable, both here and in the next scenario.

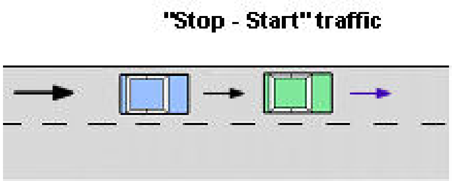
STATS 19 Cluster 4	OTS Cluster 2
13% of population Daylight Fine weather Not at junction Majority 60–70 mph speed zone Both vehicles going ahead	12% of population Daylight Fine weather Travel speed 54 km/h Braking 6 km/h Struck vehicle 19 km/h at impact Following distance 4 m
Combined features	
	Daylight Fine weather Not at junction Travel speed 54 km/h Braking 6 km/h Struck vehicle 19 km/h at impact Following distance 4 m

Table 47 Rear-end accidents: moving traffic

STATS 19 cluster 4 portrays vehicles going ahead on a high-speed road in daylight and fine weather. It is not possible to directly extract any information about the density of traffic—although this could perhaps be partially inferred from contributory factors or from the day of week and time of day—but it would not be unreasonable to hypothesize that stop-go traffic or tailgating could be associated with rear-end impacts. OTS clusters 2 and 3 both offer situations consistent with moving traffic, the salient difference being that cluster 3 is a darkness scenario. Putting STATS 19 cluster 4 alongside OTS cluster 2 does not quite determine a stop-go or tailgating scenario, but is at least highly suggestive of the broader concept of a rear-end impact in moving traffic.

5.5 Conclusion

The most common scenarios for rear-end accidents identified in the STATS 19 and OTS databases are described in Table 35 and Table 45 respectively. These include a situation where the vehicle in front is stationary at impact and a situation where both vehicles are moving on a relatively high speed road. For the latter, 'moving traffic' scenario, the information extracted from STATS 19 is consistent with both tailgating at full driving speed and with an accident in congested, stop-go traffic.

6 Head-on collisions

6.1 Cluster analysis of STATS19

6.1.1 Selection of cases

The vehicle file for STATS 19 (2008) contains information on 311,604 road users. The selection criteria for inclusion as case vehicles in head-on impacts were:

- cars and taxis with first point of impact on front
- collision partner a motorised vehicles with first impact to front
- case vehicle and collision partner each made first impact with each other
- case vehicle and collision partner travelling to or from opposite compass point directions
- no parked or reversing vehicles
- no unknown or missing information in key fields.

It can be recognised from these criteria that the collision partner of a case vehicle can also itself be a case vehicle. This was not the case for rear-end accidents where the first impact to the case vehicle and its collision partner had to be on the front and rear surfaces respectively.

	None	Front	Back	Right	Left	Unk.	Total
Pedal cycle	869	8453	1924	3272	2273	6	16797
Motor cycle to 50 cc	372	2369	353	515	638	0	4247
Motor cycle 51-125 cc	509	3652	363	794	994	2	6314
Motor cycle 126-500 cc	267	1648	184	361	504	3	2967
Motor cycle over 500 cc	732	4880	530	1236	1519	2	8899
Taxi/private hire car	405	2338	863	804	733	1	5144
Car	11006	116029	46710	30919	26104	84	230852
Minibus (8-16 seats)	79	429	160	141	118	0	927
Bus or coach (17+ seats)	3724	2502	416	724	1007	2	8375
Other motor vehicle	514	1491	519	443	420	2	3389
Other non-motor vehicle	17	59	24	40	26	0	166
Ridden horse	47	21	18	22	6	0	114
Agricultural vehicle	104	196	112	168	64	0	644
Tram	2	10	1	2	8	0	23
Goods vehicle to 3.5 t	904	6635	2579	1827	1674	2	13621
Goods vehicle 3.5-7.5 t	236	1058	375	377	384	3	2433
Goods vehicle over 7.5 t	588	2671	950	1350	1044	4	6607
Unknown	13	35	4	7	9	17	85
Total	20388	154476	56085	43002	37525	128	311604

Table 48 Case vehicles: passenger cars with first point of impact to front surface

Cars (116,029) and taxis (2,338) with the first point of impact on the front surface were first selected from the entire database (Table 49). The vehicle type and first point of impact for the collision partners of these 118,367 are shown in Table 49. Motorised vehicles (excluding motorcycles) with the first point of impact to the front were identified as suitable collision partners. This reduced the number to 29,802 with other cars (26,845) constituting the main vehicle type struck on the front.

	None	Front	Back	Right	Left	Unk.	Total
Pedal cycle	53	2417	1295	1210	1285	1	6261
Motor cycle to 50 cc	8	513	213	125	181	0	1040
Motor cycle 51-125 cc	9	738	241	208	252	0	1448
Motor cycle 126-500 cc	4	338	121	61	134	0	658
Motor cycle over 500 cc	9	916	345	216	295	0	1781
Taxi/private hire car	6	437	511	219	158	0	1331
Car	544	26845	27520	8923	6880	4	70716
Minibus (8-16 seats)	2	91	83	34	34	0	244
Bus or coach (17+ seats)	9	322	165	125	72	0	693
Other motor vehicle	9	240	193	113	72	0	627
Other non-motor vehicle	1	12	11	15	11	0	50
Ridden horse	0	10	7	3	3	0	23
Agricultural vehicle	3	61	62	52	25	0	203
Tram	0	2	0	2	1	0	5
Goods vehicle to 3.5 t	34	1191	1072	403	272	0	2972
Goods vehicle 3.5-7.5 t	8	189	151	78	54	0	480
Goods vehicle over 7.5 t	28	424	373	232	102	0	1159
Unknown	0	6	0	0	1	28669	28676
Total	727	34752	32363	12019	9832	28674	118367

Table 49 Collision partners: motorised vehicles with first point of impact to front surface

A check that no third vehicles were involved in the first impacts of the case vehicle and its collision partner reduced the number of cases to 26,445 (Table 50).

Case vehicle	26445
Other vehicle	3357
Total	29802

Table 50 First impact of collision partner

STATS 19 nominates compass points for the directions from which and to which a vehicle was heading. To capture the notion that the vehicles in a head-on collision should have been generally travelling in opposite directions, it was specified that the ‘from’ direction of at least one vehicle should match the ‘to’ direction of the other. This is shown in Table 51 where ‘opposite direction’ means that the ‘from’ direction of both vehicles matched the ‘to’ direction of the other, ‘opposite direction—pass’ means that only one ‘from’ direction matched the ‘to’ direction of the other vehicle and that the paths of the vehicles did not cross (e.g. vehicle A from south to north and vehicle B from north-east to south), and ‘opposite direction—cross’ means that only one ‘from’ direction matched the ‘to’ direction of the other vehicle and that the paths of the vehicles did cross (e.g. vehicle A from south to north and vehicle B from north--west to south). This filter excluded 8,398 cases.

Opposite direction	10597	
Opposite direction—pass	916	
Opposite direction—cross	6534	
Sub-total		*18047*
Other	8398	
Total	26445	

Table 51 Direction of movement of case vehicle and collision partner

A final filter in the selection of eligible cases was made of any parked or reversing vehicles and of any database record that contain missing or unknown information in the key fields for input to the cluster analysis. This resulted in a total of 14,081 case vehicles meeting the selection criteria for head-on impacts (Table 52).

Included	14081
Excluded	3966
Total	18047

Table 52 Exclusion of parked or reversing vehicles and cases with missing information

The capacity of the hardware and software used to carry out the cluster analysis was found to be limited to around 10,000 to 11,000 records, at which level memory storage errors were encountered or processing appeared to be continue indefinitely, e.g. over six hours. For this reason the 14,081 eligible vehicle records were sorted by their identification numbers in the database and every third record was removed. The remaining 9,387 were passed on to the cluster analysis.

Included	9387
Excluded	4694
Total	14081

Table 53 Filtering to two-thirds of eligible cases

Table 54 shows accident severity for the cases included or excluded at this final filter. There was no statistically significant difference between the groups for this or the other variables listed in Table 55 that were used in the cluster analysis (chi-square test to 90% significance).

		Included		
	Yes	No	Yes	Total
Slight	3715	3739	3706	11160
Serious	838	831	849	2518
Fatal	140	124	139	403
Total	4693	4694	4694	14081

Table 54 Comparison of included and excluded cases: accident severity

6.1.2 Input dataset

Eight descriptive parameters were included in the dataset for the analysis of rear-end accidents in STATS 19. Quantitative information on speed, separation distance and location can only be obtained from an in-depth database; however speed limit is available as a proxy for speed and the information vehicle manoeuvre enable a picture of accident circumstances to emerge.

Ordinal	Accident severity	0.0	Slight
		0.5	Serious
		1.0	Fatal
Ordinal	Speed limit (mph)	0.0	10-30
		0.5	40-50
		1.0	60-70
Nominal	Junction detail	1	Not at junction
		2	Roundabout
		3	Junction
Nominal	Light conditions	1	Daylight
		2	Darkness
Nominal	Weather conditions	1	Fine
		2	Not fine
Nominal	Vehicle A manoeuvre	1	Going ahead
		2	Overtaking
		3	Going ahead: bend
		4	Turning
Nominal	Vehicle B manoeuvre	1	Going ahead
		2	Overtaking
		3	Going ahead: bend
		4	Turning
Nominal	Vehicle directions	1	0110 Opposite direction
		2	0112 Opposite—pass
		3	0210 Opposite—cross

Table 55 Structure of dataset for head-on accidents (STATS 19)

The case vehicle is nominated as vehicle A and its collision partner, also with a frontal impact, as vehicle B. Bearing in mind that the collision partner of a case vehicle can itself be a case vehicle, the number of head-on accidents is less than the number of case vehicles (9,387).

6.1.3 Results

The dendrogram produced by the cluster analysis was cut at a level where the vehicle population of the dataset (9,387) was formed into 15 groups. The largest five of these groups comprise 86% of the population and are described in detail in Table 56. The characteristics of the remaining clusters are shown in aggregate for completeness. The figures for the relative direction of movement of the vehicles as defined by ‘to’ and ‘from’ compass point directions are shown in Table 57. This parameter was not used in the cluster analysis to influence the formation of groups because it is highly correlated with the vehicle manoeuvre fields.

	Cluster						
	1	2	3	4	5	6–15	Total
Cluster representativeness (%)							
Slight	34	21	16	10	5	14	100
Serious	17	31	17	11	10	14	100
Fatal	2	38	17	9	17	16	100
Total	30	23	16	10	6	14	100
Accident severity							
Slight	2532	1575	1178	706	376	1054	7421
Serious	292	521	289	178	173	234	1687
Fatal	6	107	48	25	47	46	279
Total	2830	2203	1515	909	596	1334	9387
Speed limit (mph)							
10-30	2020	830	365	405	180	723	4523
40-50	345	301	156	118	97	159	1176
60-70	465	1072	994	386	319	452	3688
Total	2830	2203	1515	909	596	1334	9387
Junction detail							
Not at junction	0	2203	1515	678	550	364	5310
At junction	2830	0	0	231	46	970	4077
Total	2830	2203	1515	909	596	1334	9387
Light conditions							
Daylight	2198	1582	1515	674	0	819	6788
Darkness	632	621	0	235	596	515	2599
Total	2830	2203	1515	909	596	1334	9387
Weather conditions							
Fine	2569	2203	1076	3	382	789	7022
Not fine	261	0	439	906	214	545	2365
Total	2830	2203	1515	909	596	1334	9387
Vehicle A manoeuvre							
Going ahead	2825	2199	0	901	0	267	6192
Overtaking	5	4	0	7	10	446	472
Going ahead: bend	0	0	1515	1	586	621	2723
Total	2830	2203	1515	909	596	1334	9387
Vehicle B manoeuvre							
Going ahead	779	1838	107	746	65	591	4126
Overtaking	69	190	20	75	16	55	425
Going ahead: bend	67	135	1388	81	515	371	2557
Turning	1915	40	0	7	0	317	2279
Total	2830	2203	1515	909	596	1334	9387

Table 56 Clusters for head-on accidents (STATS 19)

The cells highlighted in green indicate (a) fields for which the distribution of cases is significantly different from the population (chi-square test to 99.5% significance) and (b) cells in these fields that are over-represented. The results for cluster representativeness are derived directly from the numbers on accident severity, expressing the latter as row percentages.

Cluster 1, which constitutes 30% of the 9,387 vehicles, portrays accidents at junctions where the case vehicle is going ahead and the collision partner is mostly turning, often across the path of the case vehicle (Table 57). This group of accidents tends to occur in daylight and fine weather conditions, in lower speed zones, and to involve slight injury. Cluster 2, which constitutes 23% of the population, has an over-representation of serious and fatal casualties and accidents in higher speed zones. These accidents occur away from junctions where the case vehicle is going ahead and the collision partner is mostly going ahead or overtaking. Accidents on bends are represented in clusters 3 and 5 which represent 16% and 6% of the population respectively. These two clusters differ most in that the cluster 3 occurs in daylight and cluster 5 in darkness.

	Cluster						
	1	**2**	**3**	**4**	**5**	**6–15**	**Total**
Vehicle directions							
Opposite	839	2096	1456	846	557	916	6710
Opposite—3-pt pass	171	51	34	16	27	94	393
Opposite—3-pt cross	1820	56	25	47	12	324	2284
Total	2830	2203	1515	909	596	1334	9387

Table 57 Further characteristics of clusters for head-on accidents (STATS 19)

6.2 Detailed OTS case reviews

This section outlines the results for 50 accidents from the OTS accident database relating to head-on accidents. Each accident was examined from 5 seconds before the initial impact at 0.5 second intervals giving not only the closing distance between the vehicles but also the lateral distance relative to the centre of their original carriageway and the point at which the vehicles came into sight of each other.

Where there is no evidence to suggest otherwise, it has been taken that the vehicles were travelling at the speed limit for the carriageway and at the point of braking it has been assumed the vehicles would have been at maximum braking capacity. The lateral distance given will be slightly less accurate than the closing distance and this is therefore taken to the nearest 0.5 m.

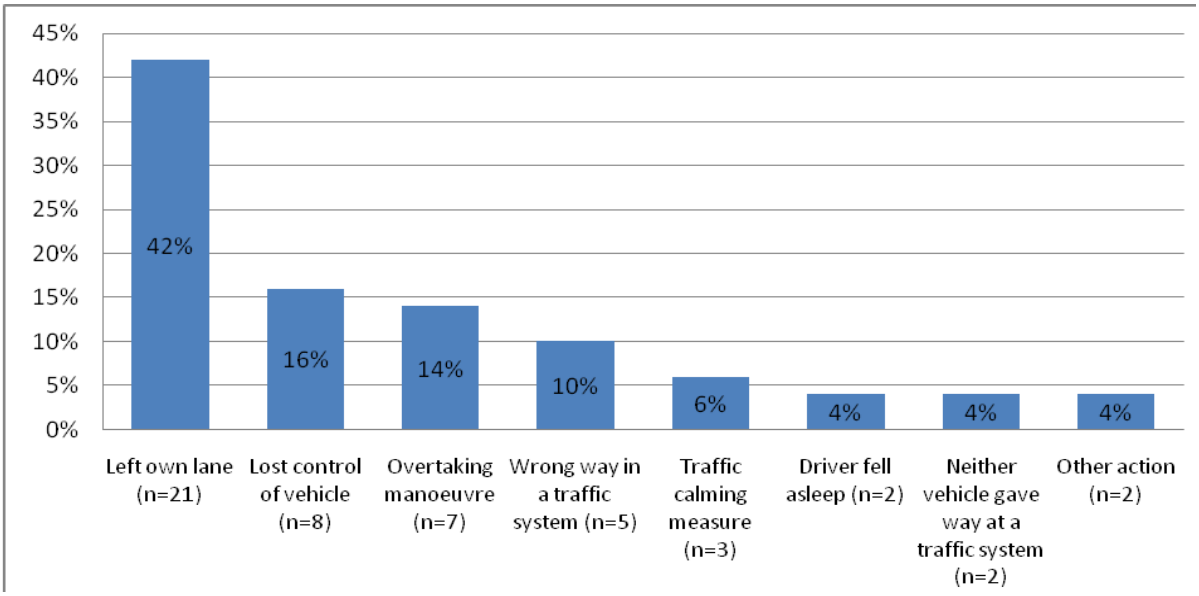


Figure 49: Accident configuration and type

From the sample of 50 head-on accidents a greater number (42%) of these accidents were as a result of one of the accident participants leaving their own lane and entering the path of another vehicle. This action includes drifting into the opposing lane as a result of going into a bend too quickly.

The driver losing control of the vehicle was the second highest group with 16% of the accidents. This group encompassed accidents where the driver lost complete control of the vehicle due to excessive speed or inappropriate speed for the conditions.

The third common cause for head-on impacts was a result of vehicles attempting to overtake on single carriageway roads (14% of all accidents). One may expect a higher proportion of overall accidents to be attributed to attempted overtaking manoeuvres, however this report only focuses on those resulting in a collision with an oncoming vehicle.

For cases selected from the OTS data, 60% were within an urban environment where on average the posted speed limit is lower than that found in rural areas.

Driving Action	Rural road		Urban road	
Left own lane	13	26%	8	16%
Overtake	1	2%	6	12%
Loss of control	4	8%	4	8%
Wrong way in a traffic system	1	2%	4	8%
Traffic calming measure	0	0%	3	6%
Neither vehicle gave way	0	0%	2	4%
Fell asleep	1	2%	1	2%
Other	0	0%	2	4%
Total	20	40%	30	60%

Table 58: Driving action by urban/rural environment

The overall case injury for the 50 accidents is reported in the table below; due to the nature of these accidents a greater closing speed for the accident participants tends to be involved resulting in higher severities than in the rear impact accidents. For the 50 accidents reviewed, 14 accidents (28%) had at least one occupant reported as killed or seriously injured.

Driving Action	Fatal	Serious	Slight	Non-injury	Total
Left own lane	1	6	10	4	21
Loss of control	1	2	4	1	8
Overtake	0	1	5	1	7
Wrong way in a traffic system	1	1	3	0	5
Traffic calming measure	0	0	1	2	3
Fell asleep	0	1	1	0	2
Neither vehicle gave way	0	0	2	0	2
Other	0	0	1	1	2
Total	3	11	27	9	50

Table 59: Accident severity

As previously stated the main cause of head-on accidents was loss of control resulting in the case vehicle leaving its own lane and colliding with another vehicle, Figure 50 shows that 52% of the accidents occurred in close proximity to a bend with the bend having an influence on the loss of the control of the vehicle.

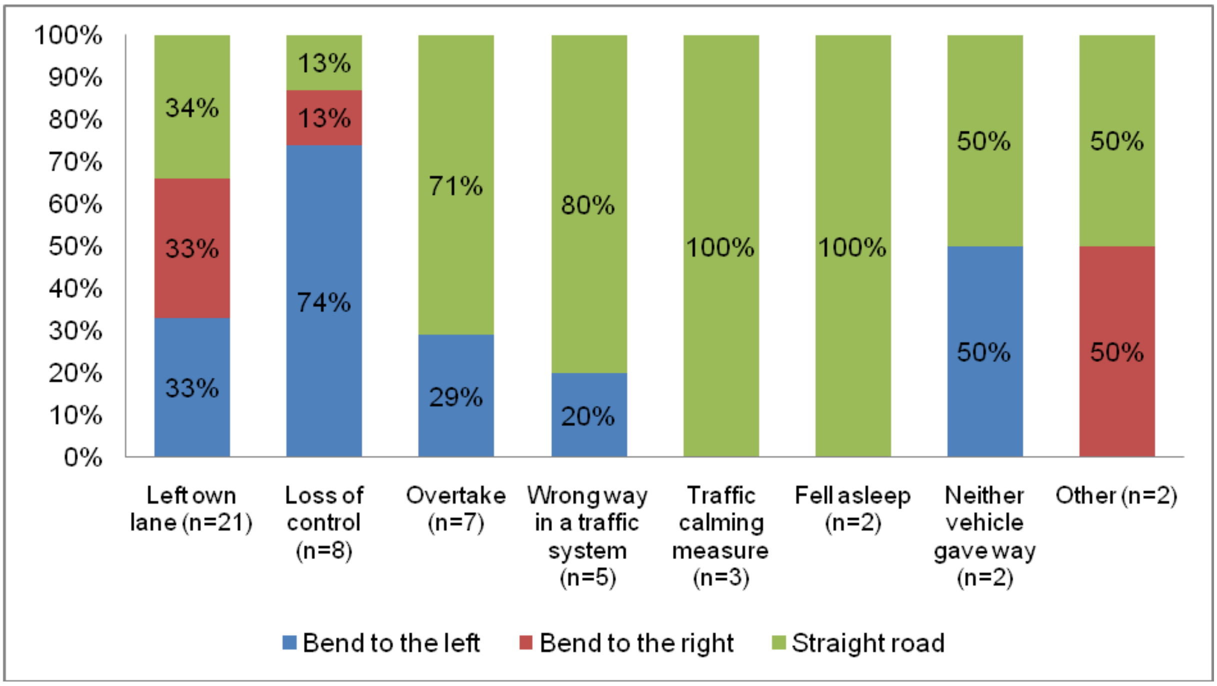


Figure 50: Road geometry (% of cases) by accident configuration group

In the majority of analysed cases the accident participants had reacted to the accident risk and were braking at the point of impact. Figure 51 and Figure 52 show the velocity reduction from the travel speed to impact speed for case vehicles and struck vehicles respectively.

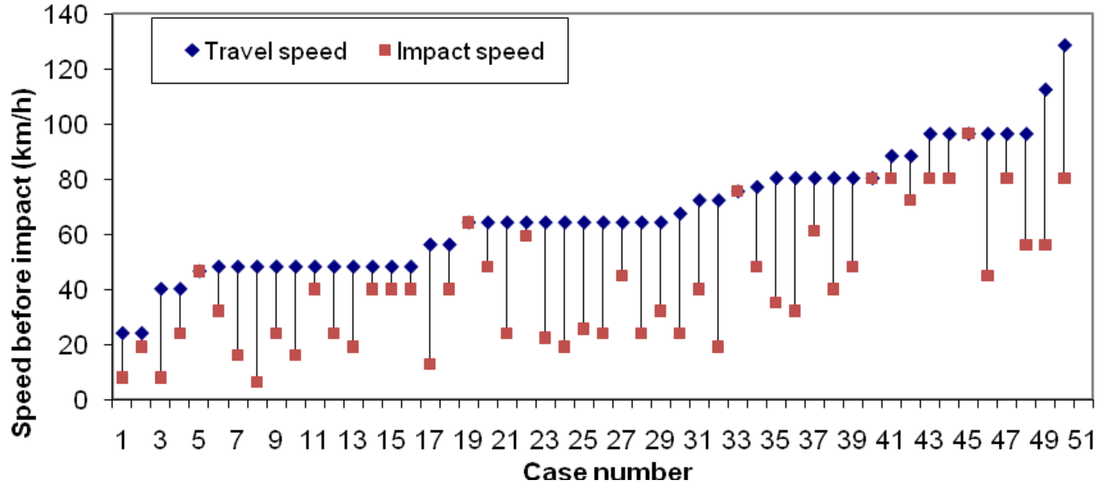


Figure 51: Speed reduction from travel speed to impact speed for case vehicles

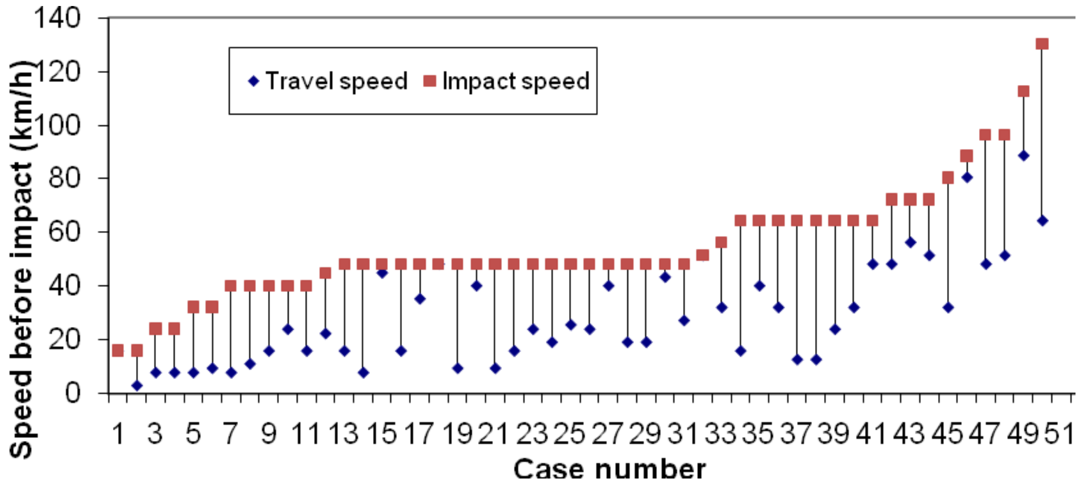


Figure 52: Speed reduction from travel speed to impact speed for struck vehicles

The linear areas of the travel speed points represent the different speed limits imposed on the British roads. These linear areas, especially from the case vehicle graph, may be somewhat over-emphasised due to the investigator relying on the vehicle travelling at the speed limit for the road where any other method for speed calculation was unavailable.

One of the main differences between the speed reductions of the two types of accident participant is at the lower end of the speed range where the case driver has failed to reduce their speed before impact as much as the struck vehicle drivers who have reduced their travel speed considerably before the collision.

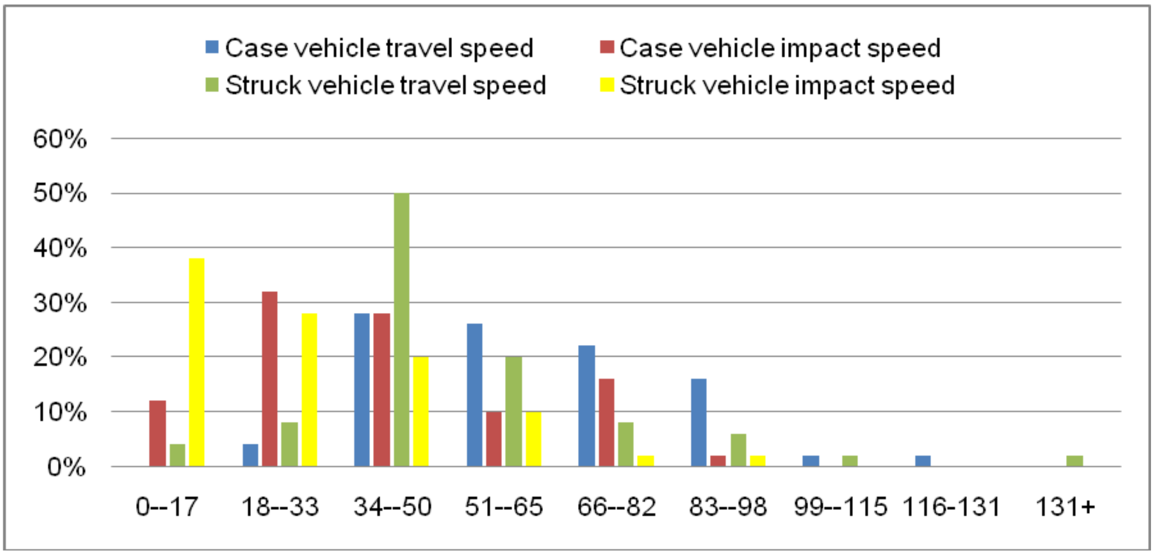


Figure 53: Percentage of case vehicles and struck vehicles by speed range (km/h).

Figure 53 gives a breakdown of the impact and travel speeds for the accident participants by speed. The travel speed for the case vehicle tends to be higher than that of the case vehicle, where 24 (48%) of the case vehicles had a travel speed of 51 to 82 km/h compared to the struck vehicle which has 25 (50%) of the vehicles travelling from 34 to 50 km/h.

Although it is difficult to quantify the amount of braking the road user applied prior to impact, Figure 51 and Figure 52 show that braking was applied due to the observed speed reduction. It is known that 15 (30%) of the case vehicles achieved complete wheel lock up, leaving locked wheel marks at the scene of the accident. Maximum braking may have been achieved by other vehicles in the sample however no physical evidence was identified for those cases.

The environmental conditions at the time of the accident can have implication for braking efficiency or general visibility afforded to the drivers involved. Figure 54 displays the results pertaining to road conditions, light conditions and accident configuration by the percentage of cases falling into each category.

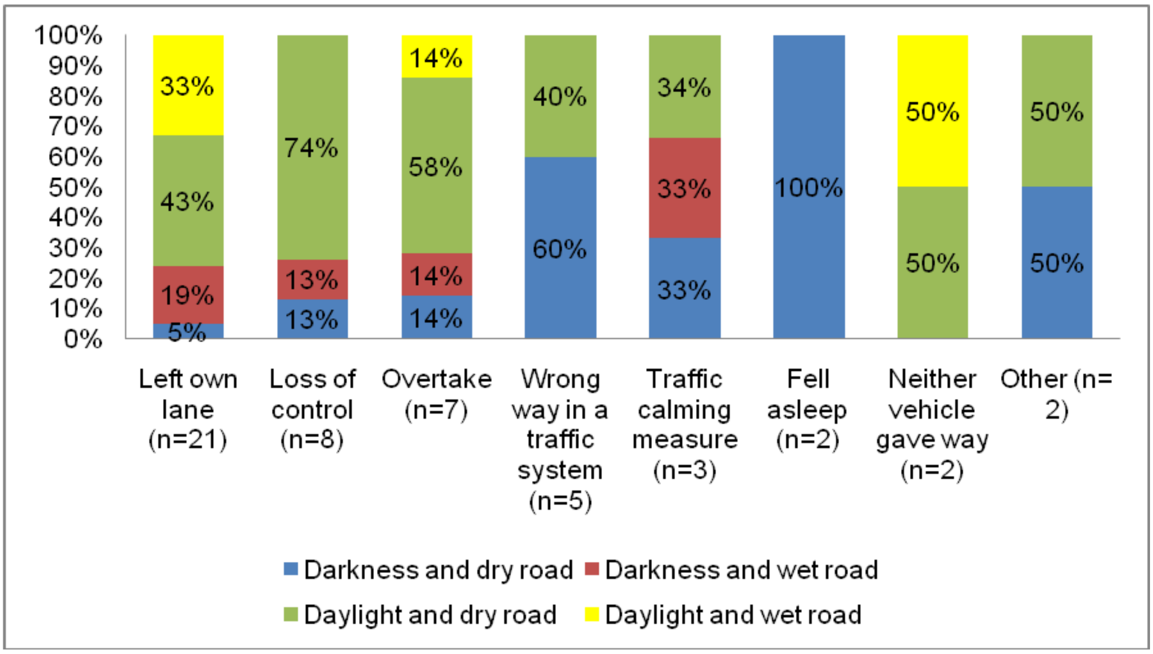


Figure 54: Environmental conditions at the time of the accident by accident configuration

The higher percentage (48%) of loss of control accidents within the sample of head-on accidents would appear to occur in the more favourable driving conditions of daylight on dry roads. This would indicate the loss of control is more attributed to excessive velocity as opposed to adverse road conditions as indicated in Figure 54.

Of the two accident groups “Left own lane” and “Loss of control”, which total 28 accidents, 22 occurred on a bend with the majority being left-hand bends (Figure 50).The distribution of vehicles fitted with electronic stability control (ESC) and anti-lock braking (ABS) for the 28 accidents is given in Figure 55.

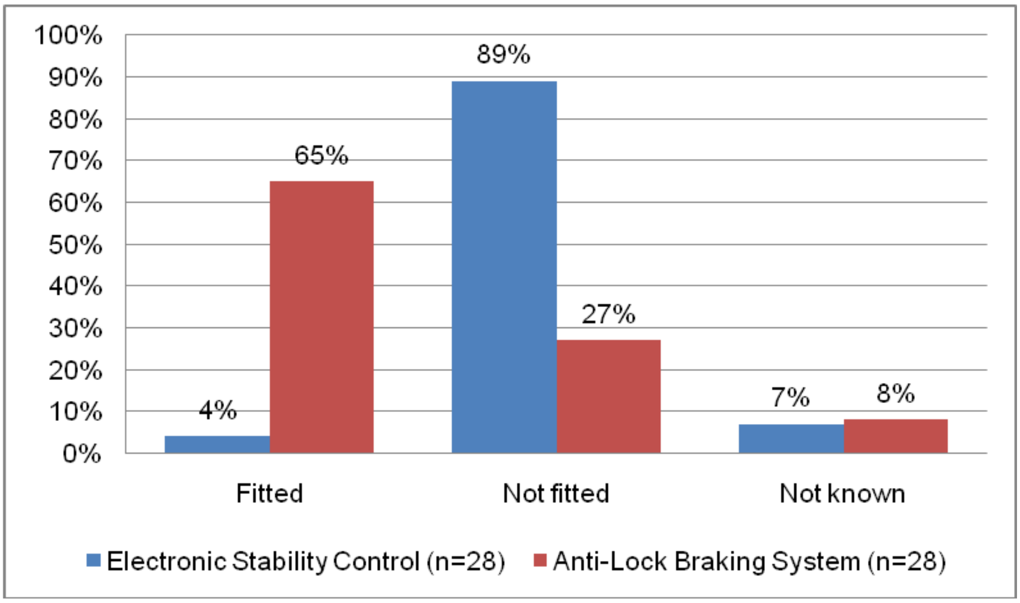


Figure 55: Active safety system fitment

Only1 (4%) vehicle in the 28 accidents was fitted with ESC compared to 19 (65%) of vehicles being fitted with ABS.

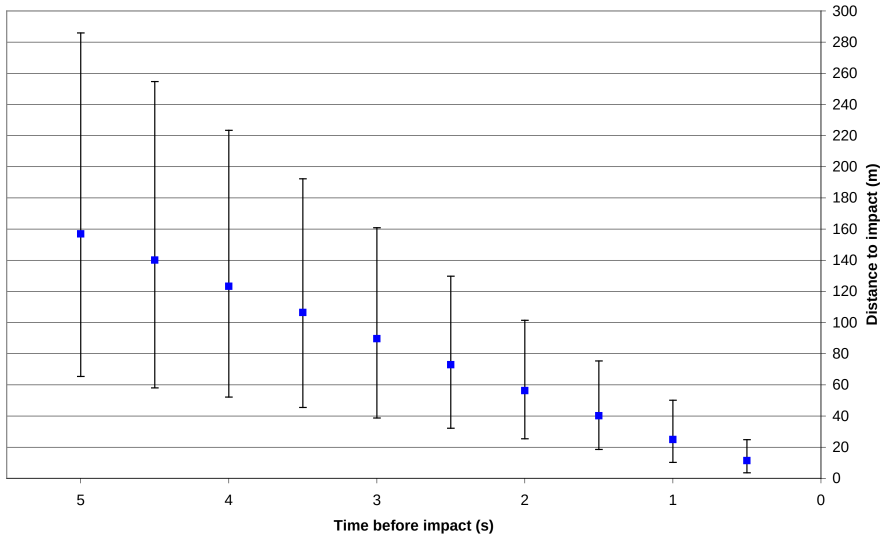


Figure 56: Closing distance between collision participants before impact

Figure 56 gives the average closing distance (headway distance) between accident participants at 1 second before the collision is 25 meters. At 5 seconds before the accident the average distance is 157 meters further highlighting the greater closing speeds involved with this type of accident scenario. The maximum and minimum headway distances have been included.

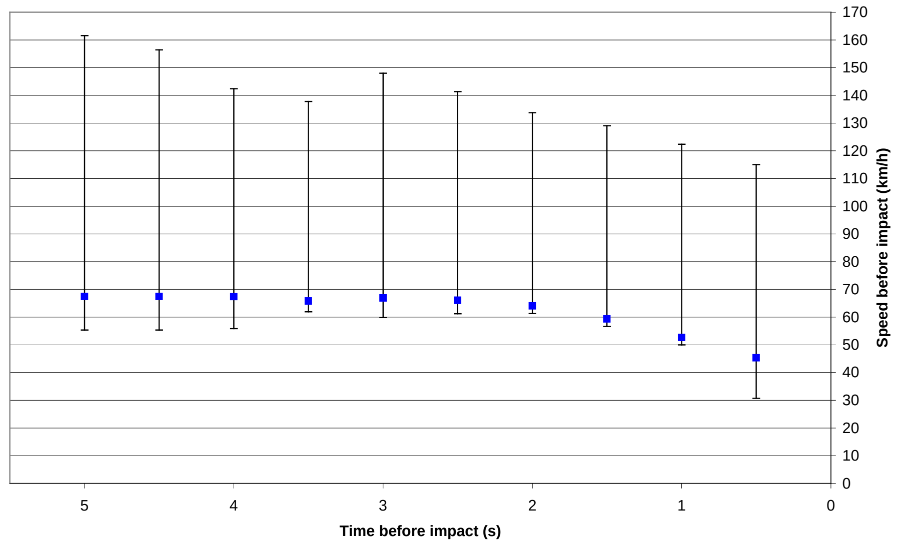


Figure 57: Velocity of the case vehicle before impact

The average speed for the case vehicle is given over the 5-second period leading to the accident. Figure 57 shows the average, maximum and minimum velocities for the case vehicle at each time interval.

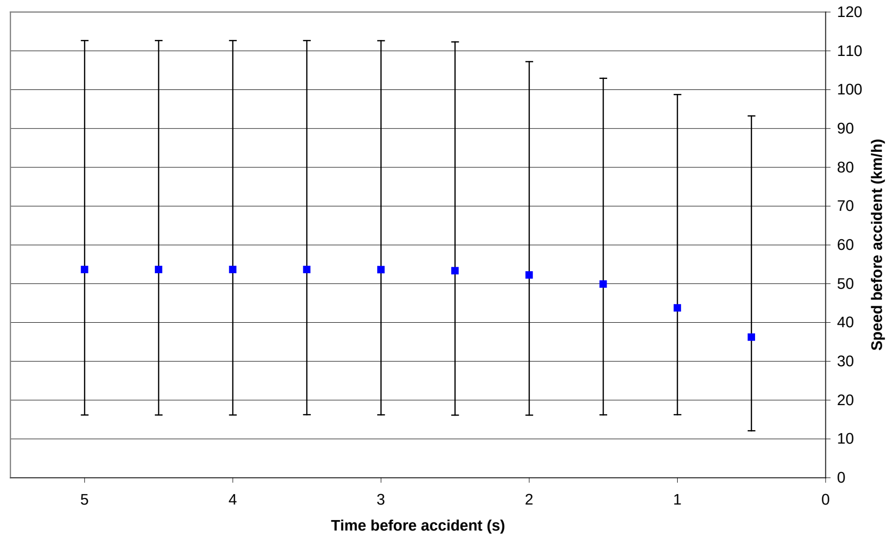


Figure 58: Velocity of struck vehicle before the impact (-s)

Figure 58 shows the average, maximum and minimum velocities for the struck vehicle prior to impact for five seconds leading up to the accident.

This group of accidents present a number of challenges in regards to the reconstruction process and establishing a time line prior to the impact. As the vehicles were travelling towards each other and generally at higher speeds than the other two accident groups (pedestrian and rear-end), the distances involved were larger and the accuracy for the distance plots was consequently reduced. As this analysis was based on the OTS data which has been collected since the year 2000, the vehicles involved were largely not fitted with active safety systems such as electronic stability control and lane departure warning, both of which have subsequently become common in the UK vehicle fleet. As the prevalent accident scenario for the collision group was loss of control on a bend at moderate speeds, (48% of accidents up to a speed of 82 km/h), it is feasible that active safety systems could have reduced the risk of this accident group. This would leave the focus on the overtaking accident scenario, which featured in 16% of the accidents reviewed. As that is a risk-taking manoeuvre by the driver, any system implemented to reduce this accident group would have to warn the driver or otherwise prevent them starting the initial overtaking manoeuvre.

6.3 Cluster analysis of OTS

6.3.1 Selection of cases

All of the vehicles from the detailed case reviews were included in the cluster analysis. These 50 vehicles were chosen at random from all of the eligible cases in OTS provided adequate evidence from the scene of the accident had been collected and recorded. The criteria for case vehicles were:

- passenger cars
- first point of impact on front surface
- no interaction with pedestrian
- accident coded ‘Head on’ or ‘Overtaking and lane change: head-on’.

Collision type	Collision sub-type 1	2	3	4	5	6	7	8	Total
A Overtaking and lane change	137	37	104	49	6	17	7	11	368
B Head on	52	14	31	14	36	63		8	218
C Lost control or off road (straight roads)	221	271	134					8	634
D Cornering	358	348	45					15	766
E Collision with obstruction	92	9	60	1	6			9	177
F Rear end	190	24	11	294	68	75		14	676
G Turning versus same direction	3	40	8	16	39	19		2	127
H Crossing (no turns)	243								243
J Crossing (vehicle turning)	285	3	9					7	304
K Merging	64	71	6					3	144
L Right turn against	12	212	1					1	226
M Manoeuvring	34	9	43	12	1		4	11	114
N Pedestrians crossing road	117	84	1	2	1	1	5	6	217
P Pedestrians other	5	1	5	2	1	5		7	26
Q Miscellaneous	1	4	1	7			16	15	44
Total	1814	1127	459	397	158	180	32	117	4284

Table 60 Collision category for all accidents

All accidents on the OTS database are categorised according to the system in Table 60. The sub-type numbers have a different meaning for each main category A to Q. There were 255 accidents classified as head-on (including overtaking and lane change cases) in the population of 4,284 accidents on the database.

	Back	Front	Left	Right	Top	Bottom	Other	Total
Car	912	2598	796	1121	16	79	478	6000
LGV	75	210	43	62	2	13	50	455
HGV	29	183	48	77	4	15	66	422
Bus	12	55	13	15		9	15	119
Motorcycle	14	194	54	55	2	9	102	430
Pedal cycle	11	49	34	40		1	29	164
Other	2	3	4	5			6	20
Unknown	7			2		38	8	55
Total	1062	3292	992	1377	24	164	754	7665

Table 61 Vehicle type by first impact side

The population of 4,284 accidents involves 7,665 vehicles. Of these, there were 2,598 cars that made first impact on the front surface, as shown in Table 61.

A total of 187 vehicles were cars that made first impact on the front surface, were in accidents classified as head on (category B, 161; category A2, 26), and did not strike a pedestrian. Following a check that the collision partner was struck on the front surface and that the direction of the impact force was 12 o’clock, i.e. directly frontal, the quota of 50 vehicles was selected randomly from accidents investigated in phases 2 and 3 of the OTS project (2003–2009) and used for both the detailed case reviews and the cluster analysis.

6.3.2 Input dataset

The dataset derived from OTS for head-on impacts contains five parameters to describe the circumstances of the accidents. The two scale variables are drawn exclusively from the detailed case reviews and reconstructions.

Ordinal	Accident severity	0.0 0.5 1.0	Slight or non-injury Serious Fatal
Nominal	Vehicle A manoeuvre (case vehicle)	1 2 3 4	Going ahead Going ahead: overtaking Going ahead: bend Other
Nominal	Vehicle B manoeuvre (collision partner)	1 2 3 4	Going ahead Going ahead: overtaking Going ahead: bend Other
Scale	Closing speed	0–1	Scaled from km/h
Scale	Closing speed at impact	0–1	Scaled from km/h

Table 62 Structure of dataset for head-on accidents (OTS)

6.3.3 Results

The results of the cluster analysis of the OTS dataset are detailed at the level of 9 clusters, three of which cover 78% of the population of the dataset. The cells shaded in green indicate (a) that the distribution of numbers for the given field is significantly different from the distribution in the whole population (chi-square test to 95% significance) and (b) that the particular number highlighted is over-represented. This is the same as for the STATS 19 analysis except that the statistical test is evaluated at 95% confidence instead of 99.5%. This level is better suited to the lower number of cases in OTS for providing an objective test of differences between the clusters and the overall population.

	Cluster				
	1	2	3	4–9	Total
Representativeness (%)					
Slight or non-injury	46	20	14	20	100
Serious	17	75	0	8	100
Fatal	0	0	0	100	100
Total	36	32	10	22	100
Accident severity					
Slight or non-injury	16	7	5	7	35
Serious	2	9	0	1	12
Fatal	0	0	0	3	3
Total	18	16	5	11	50
Vehicle A manoeuvre					
Going ahead	0	16	0	2	18
Overtaking	0	0	5	2	7
Going ahead on bend	18	0	0	3	21
Other	0	0	0	4	4
Total	18	16	5	11	50
Vehicle B manoeuvre					
Going ahead	0	16	5	5	26
Overtaking	0	0	0	3	3
Going ahead: bend	18	0	0	3	21
Other	0	0	0	0	0
Total	18	16	5	11	50
Closing speed (km/h)					
Mean	127	130	127	112	124
N	22	7	7	10	50
Closing speed at impact (km/h)					
Mean	77	74	47	73	72
N	22	7	7	10	50

Table 63 Clusters for head-on accidents (OTS)

The largest group, Cluster 1, which contains 36% of the 50 vehicles in the dataset, describes head-on accidents on a bend where the mean closing speed before braking was 127 km/h, reducing to 77 km/h at impact. The scatter of these values is shown in Figure 59 and Figure 60. Cluster 2, with 32% of the population, is similar except that the colliding vehicles are not on a bend. Cluster 3, which was the only remaining group to contain at least five members, shows the case vehicle overtaking. While the mean initial closing speed between the colliding vehicles was quite similiar in Cluster 3 to the other two clusters, there was considerably more braking or slowing down in the overtaking scenario, from 127 to 47 km/h.

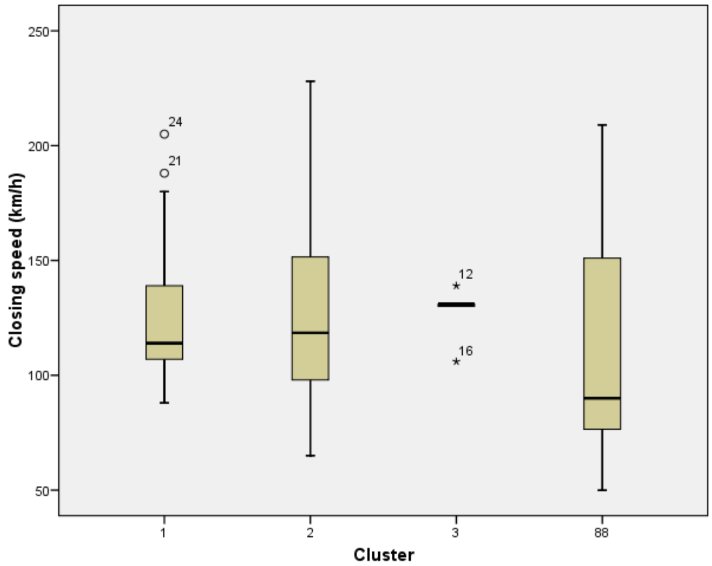


Figure 59 Closing speed (km/h)

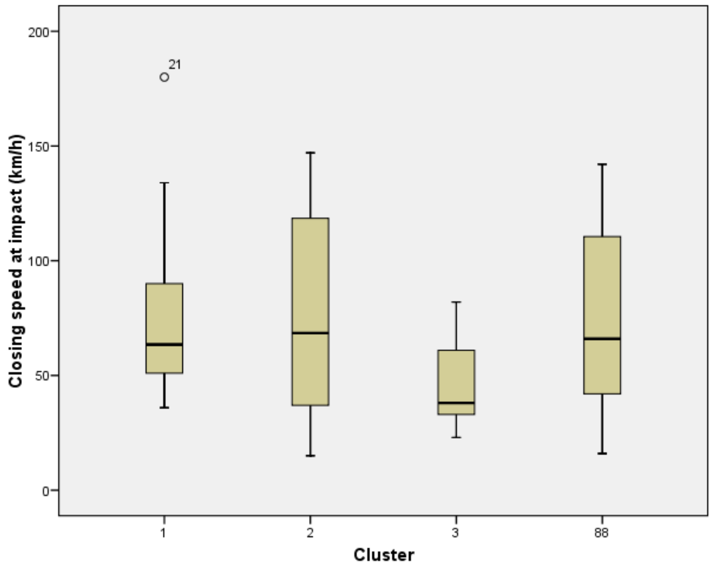


Figure 60 Closing speed at impact (km/h)

6.4 Discussion

It will be recalled that the largest clusters in the analysis of STATS 19 were accidents at junctions, accidents away from junctions and accidents on bends, to summarise these in very general terms. Overtaking was mixed in among these groups without forming a distinct cluster of its own. The analysis of OTS highlights head-on collisions firstly on bends and secondly with both vehicles going straight ahead, with an overtaking scenario in third place. The major difference here—the prominence of accidents at junctions in STATS 19—arises from a divergence in the selection criteria for the OTS and STATS 19 datasets. The 50 OTS cases were selected entirely from the ‘head-on’ categories (B and A2) of Table 60 while accidents at junctions would often be most naturally placed into other categories (e.g. J or L). An interest in frontal impacts at junctions arose in the course of the work undertaken for this report and was more easily accommodated in the STATS 19 analysis, which was largely based on computer programming, than in the OTS analysis, which was more heavily dependent on labour-intensive case reviews and accident reconstruction. Putting aside accidents at junctions, it is clear from both sets of results that accidents on bends and accidents involving overtaking represent common scenarios. An examination of the STATS 19 contributory factors file (and other fields) could throw light on the manner by which vehicles collide when away from junctions and bends and while not overtaking, which is also a common scenario. The same applies to OTS Cluster 2: having identified that both vehicles going ahead not at a bend and not overtaking is a common scenario, it would be informative in future work to describe the critical elements of real accidents in further detail so that these could be reproduced in a physical test if desired.

6.5 Conclusion

The most common scenarios for head-on accidents identified in the STATS 19 and OTS databases are described in Table 56 and Table 63 respectively. These include accidents at junctions, especially where the path of the collision partner crosses the path of the case vehicle, accidents on bends, accidents involving overtaking, and accidents not involving any of these factors. The last of these would benefit from further description in future work if key elements are these accidents were to be incorporated in a test procedure. Accidents at junctions were mostly excluded from the OTS dataset by the requirement that the impact was categorised as ‘head on’. The inclusion of additional OTS cases in an extended analysis would enable a more direct comparison with the STATS 19 clusters and provide more scope for reflecting the diversity of frontal impacts.

References

R. Cuerden, M. Pittman, E. Dodson and J. Hill, The UK On-the-Spot accident data collection study—Phase II report, Road Safety Research Report No. 73, Department for Transport, London, February 2008.
http://www.dft.gov.uk/pgr/roadsafety/research/rsrr/theme5/onthespotaccident.pdf

Department for Transport (DfT), Instructions for the completion of road accident reports with effect from 1 January 2005, October 2004.
http://www.dft.gov.uk/collisionreporting/Stats/stats20.pdf

S. Pheasant, Bodyspace – Anthropometry, Ergonomics and Design. 1998.

Acknowledgements

The OTS project was funded by the Department for Transport and the Highways Agency. The On-the-Spot investigations were carried by teams at the Transport Research Laboratory (TRL) in Berkshire and the Vehicle Safety Research Centre (VSRC) at Loughborough University. The project was not possible without the help and support of many individuals, especially the Chief Constables of Nottinghamshire and Thames Valley Police Forces and their officers.

The views expressed in this report are those of the authors and do not necessarily reflect those of any other party.

The Department for Transport does not guarantee the accuracy or completeness of information inferred from its data. It cannot accept liability for any loss or damages of any kind resulting from reliance on the information or guidance this document contains.