

# Typical pedestrian accident scenarios for the development of autonomous emergency braking test protocols

James Lenard<sup>a</sup>, Alexandro Badea-Romero<sup>b</sup>, Russell Danton<sup>c</sup>

<sup>a</sup>*Loughborough University, Loughborough, Leics LE11 3TU, United Kingdom*

<sup>b</sup>*INSIA, Technical University of Madrid (UPM), 28031 Madrid, Spain*

<sup>c</sup>*GBB (UK) Ltd., Crown House, Bridgewater Close, Burnley BB11 5TE, United Kingdom*

---

## Abstract

An increasing proportion of new vehicles are being fitted with autonomous emergency braking systems. It is difficult for consumers to judge the effectiveness of these safety systems for individual models unless their performance is evaluated through track testing under controlled conditions. This paper aimed to contribute to the development of relevant test conditions by describing typical circumstances of pedestrian accidents. Cluster analysis was applied to two large British databases and both highlighted an urban scenario in daylight and fine weather where a small pedestrian walks across the road, especially from the near kerb, in clear view of a driver who is travelling straight ahead. For each dataset a main test configuration was defined to represent the conditions of the most common accident scenario along with test variations to reflect the characteristics of less common accident scenarios. Some of the variations pertaining to less common accident circumstances or to a minority of casualties in these scenarios were proposed as optional or supplementary test elements for an outstanding performance rating. Many considerations are incorporated into the final design and implementation of an actual testing regime, such as cost and the state of development of technology; only the representation of accident data lay within the scope of this paper. It would be desirable to ascertain the wider representativeness of the results by analysing accident data from other countries in a similar manner.

*Keywords:* pedestrian, autonomous emergency braking, accident scenario, test configuration, cluster analysis

---

## 1. Introduction

Autonomous emergency braking (AEB), also referred to as advanced or automatic emergency braking and collision-imminent braking, is one of several mod-

---

*Email addresses:* j.a.lenard@lboro.ac.uk; j.a.lenard@datarye.com (James Lenard), alexandro.badea@upm.es (Alexandro Badea-Romero), russell@gbbuk.com (Russell Danton)  
*URL:* [orcid.org/0000-0002-5084-0484](https://orcid.org/0000-0002-5084-0484) (James Lenard)

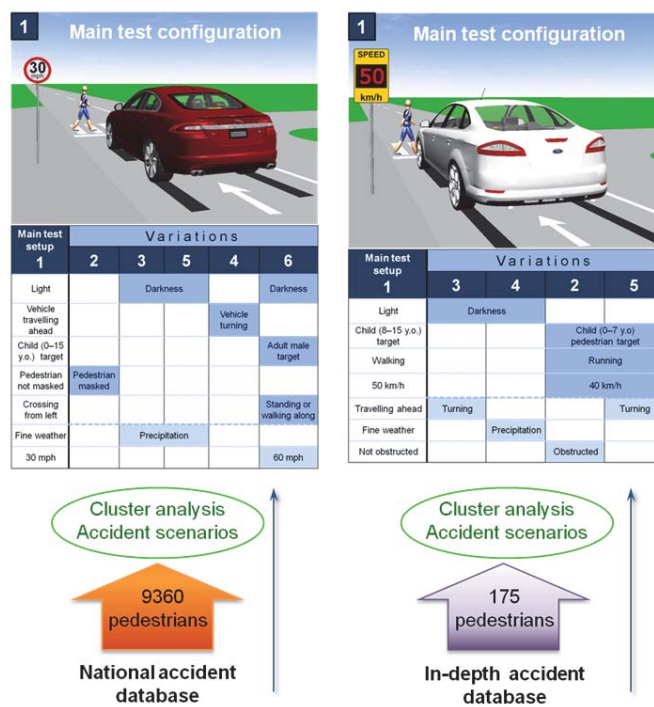


Figure 0: Graphical abstract.

ern safety systems that are offered as standard or optional equipment on an increasing proportion of new vehicles in the mass market. It is designed to avoid accidents or mitigate the severity of impacts by applying the brakes automatically when a collision is imminent, in some cases following a warning to the driver. There is scope for considerable variation among AEB systems depending on the number and type of sensors, the decision logic programmed into the control unit, if and when the driver is alerted, how and when braking is initiated and other factors. How these elements combine to determine the effectiveness of a particular system is not easy to anticipate for consumers interested in making a purchase or for regulators, insurers, researchers and indeed the manufacturers themselves. To address this difficulty, preparations have been underway in Europe (Aparicio et al., 2013; Hulshof et al., 2013; Zecha et al., 2013) and elsewhere (Searson et al., 2014) to introduce track testing to evaluate performance, compare systems and promote the penetration of this active safety technology into national vehicle fleets.

As part of these preparations, accident data has been used to evaluate the potential benefit of active safety technologies for the protection of pedestrians and other vulnerable road users. One of the first available technologies, the brake assist system, proved to offer a potential reduction of impact speed and head injury risk but limited avoidance capability (Badea-Romero et al., 2013). In the case of pedestrian sensor systems, studies suggest that early detection is possible in most common scenarios (Huang et al., 2008). In general it is agreed that the implementation of pedestrian detection and autonomous braking as part of an integrated active and passive system will have an important impact for the protection of vulnerable road users (Badea-Romero and Lenard, 2013; Searson et al., 2014) and this will help to tackle the problem of pedestrian accidents commonly caused by human error (Habibovic and Davidsson, 2012; Habibovic et al., 2013). Based on this evidence, governments and regulators, manufacturers, insurance companies and consumer rating organisations are already encouraging the implementation of pedestrian detection and AEB systems. The assessment of AEB for pedestrian crash avoidance is scheduled to be phased in by EuroNCAP from 2016 (EuroNCAP, 2013).

The identification of common accident scenarios provides a baseline for the design of relevant test protocols (Fleury and Brenac, 2001; Lenard et al., 2011). In addition test procedures have to be defined adequately for implementation and practicality while reconciling many potentially conflicting requirements. To mention just a few, a test procedure should be (a) reliable in the sense that

Table 1: Overview of the national STATS19 accident database and the in-depth On-the-Spot study.

	STATS19	On-the-Spot
Period	2010	2000–2010
Sample region	Great Britain	South Notts and Thames Valley
Purpose	National statistics	Detailed information for casualty reduction programmes
Source	Police reports	Research teams (at-scene)
Sample criterion	Road user casualty	Police attendance (rotating eight hours)
Accidents	154414	4744

the same vehicle receives the same result when retested at the same or different site, (b) sensitive in the sense that better or worse performances are reflected in better or worse results, (c) appropriately difficult so that the results for a batch of test vehicles are reasonably spread over the range of possible scores given the cost and state of art of the technology at a certain time, (d) objective, (e) non-destructive and (f) comprehensible and useful for private and corporate consumers. To these may be added (g) valid, that the test measures what it is supposed to measure, effectiveness in avoiding or mitigating the severity of accidents that occur on the public road network.

A natural way to promote validity is by simulating the circumstances of real accidents in track testing. The specific aims of this study were to contribute to the development of AEB test procedures by (a) describing typical scenarios for pedestrian accidents based on empirical data and (b) illustrating how test configurations can be related to accident scenarios.

## 2. Material and methods

Two sources of information about road accidents in Great Britain were used in this work, the national accident database STATS19 and the in-depth On-the-Spot (OTS) research study. The main characteristics of these two sources are summarised in Table 1. STATS19 is compiled by the Department for Transport each year from police reports (Dep, 2011). OTS was run for the Department for Transport and the Highways Agency by two research institutes from 2000 to 2010 to collect in-depth information about a representative sample of road accidents based on approximately 500 at-scene investigations per year in two

regions (Cuerden et al., 2008). The latest versions of the databases available for the analysis were STATS19 2010 and OTS from 2000 to mid-2009.

The method used to move from accident data to formulation of accident scenarios was cluster analysis, specifically the agglomerative or hierarchical ascending method (Kaufman and Rousseeuw, 1990). This works by progressively grouping together the most similar records of a dataset. The resulting clusters have common characteristics and can be construed as accident scenarios in this context. The foremost merits of this data-mining technique are that multiple variables are taken into account simultaneously, the procedure and results are objective and reproducible, the representativeness of the clusters is clearly defined and the basic principles of the method can be readily understood.

The database fields used in the cluster analyses of STATS19 and OTS were selected for their relevance to a test procedure for autonomous emergency braking in consultation with an industry group (Lenard et al., 2011). An AEB system is engineered to optimise the sensitivity of sensors, field of view, image processing speed and decision logic against the pre-impact location, speed of movement, trajectory and physical size of a design set of pedestrians. Detrimental ambient light conditions affect perception of the pedestrian (Sullivan and Flannagan, 2007) and could diminish the functionality of certain types of sensors. It is relevant whether vehicles are turning or proceeding straight ahead prior to impact and whether the line of sight from the driver to the pedestrian is obscured by intervening vehicles or roadside objects. Information on the initial speed of the vehicles and the extent to which drivers are able to brake before impact is relevant to the choice and effectiveness of systems that warn the driver, reinforce avoidance actions initiated by the driver or operate fully autonomously. Fields related to these aspects of pedestrian accidents were sought out in preference to others that might be relevant to the construction of accident scenarios in other contexts, such as the sex of the driver or the time of the accident, but would not be reflected in the setup of a physical test.

The primary selection criteria for the inclusion of cases from STATS19 were

- pedestrian injury,
- passenger car associated with pedestrian,
- first point of impact on front surface of vehicle,
- vehicle not parked or reversing.

Cases with unknown information in relevant fields were excluded from the final analysis.

The primary selection criteria for the inclusion of cases from OTS were

essentially the same as STATS19 provided that sufficient evidence had been collected from the at-scene investigation to enable a quantitative reconstruction of the accident including the speed of the vehicle. Resources allowed for a detailed review of 175 accidents, assessed beforehand as close to the maximum number of cases available. Eligible cases were included in the study starting with the most recent and working back towards the earliest. The fields used in the OTS cluster analysis were extracted directly from the case files.

Side impacts were filtered from the study to accommodate the current capability of AEB systems which are designed to detect pedestrians in front of the vehicle within a range defined as the driving corridor (Broggi et al., 2009). In addition it has been shown that head injuries incurred by pedestrians in side impacts are mainly caused by impact with the road or ground rather than impact with the vehicle (Badea-Romero and Lenard, 2013), indicating the limitations of safety technology for this form of interaction between vehicles and vulnerable road users.

The measure of injury severity used in the cluster analyses categorised casualties as fatal, serious or slight according roughly to a police assessment of death within thirty days, admission to hospital as an in-patient, or medical treatment as an out-patient (Dep, 2011). These were elements of both the national and in-depth databases. The AAAM abbreviated injury scale (AIS) was coded for the in-depth OTS database using detailed medical records (Ass, 2008).

An ascending cluster analysis begins by nominating each case as a cluster, then progressively merges the most similar clusters in an iterative process until the whole set of cases is encompassed in a single cluster. Neither extreme is of interest and the decision where to terminate the process depends on the usefulness of a set of clusters for a particular purpose. In the context of the design of testing procedures, it was relevant to have a relatively small number of clusters that covered a relatively large proportion of the population. A guideline was set for around six clusters to comprise about 75–80% of the whole dataset and of the killed or seriously injured sub-population. The final number of groups was chosen with the aid of an inconsistency coefficient (The MathWorks Inc, 2012) that identified steps in the iterative process where there were larger gaps between the clusters.

The algorithm for computing similarity of clusters required specification at three levels:

- (a) At field level, the algorithm was set to compute a distance (dissimilarity) in the range 0–1 for any two values of a field with 0 signifying identity and 1

signifying maximum difference. For nominal fields, the distance between two values was always 0 or 1, depending on whether the characteristic was the same or different in two cases. Otherwise the range was set to span 0–1 in equal increments for ordinal parameters and continuously for scale parameters.

- (b) At record level, the distance between two cases was defined as the sum of the distances between their field values, i.e. the city block or Manhattan distance.
- (c) At cluster level, the distance between two groups of cases was defined as the average of the distances between each pair of records in the groups, an application of the average linkage method.

These technical specifications were implemented in MATLAB (R2012b) using the built-in functions *pdist*, *linkage*, *cluster* and *inconsistent*. Any software implementing the logic of the algorithm described should derive the same clusters. The order of cases in the input dataset makes no difference.

Following the partition of the two datasets into clusters, an assessment was made of how a test could be configured to account for the characteristics of the largest group in each. After this had been dealt with, the main features that would need to be varied or introduced to cover smaller accident scenarios were identified. Although no fixed rule was applied, the aim was not to neglect characteristics that constituted more than about 10% of a cluster. While the derivation of the accident clusters from the source datasets was purely a mathematical procedure, the construction of test configurations corresponding to the accident scenarios involved an element of judgement and in this sense could be done in slightly different ways based on the same accident results. Further details of the reasoning applied are provided in Section 4.

### 3. Results

#### *Accident scenarios*

The STATS19 database had records on 25,845 pedestrian casualties. There is provision to nominate the vehicle that struck or otherwise interacted with each pedestrian. Uninjured pedestrians are not reported and the same vehicle may be nominated for two or more pedestrians. The filtering of cases for the cluster analysis was carried out in steps as illustrated in Fig. 1. The majority of pedestrian casualties (20,892) were associated with passenger cars, including taxis and private hire cars. The front end of the vehicle was the point of first

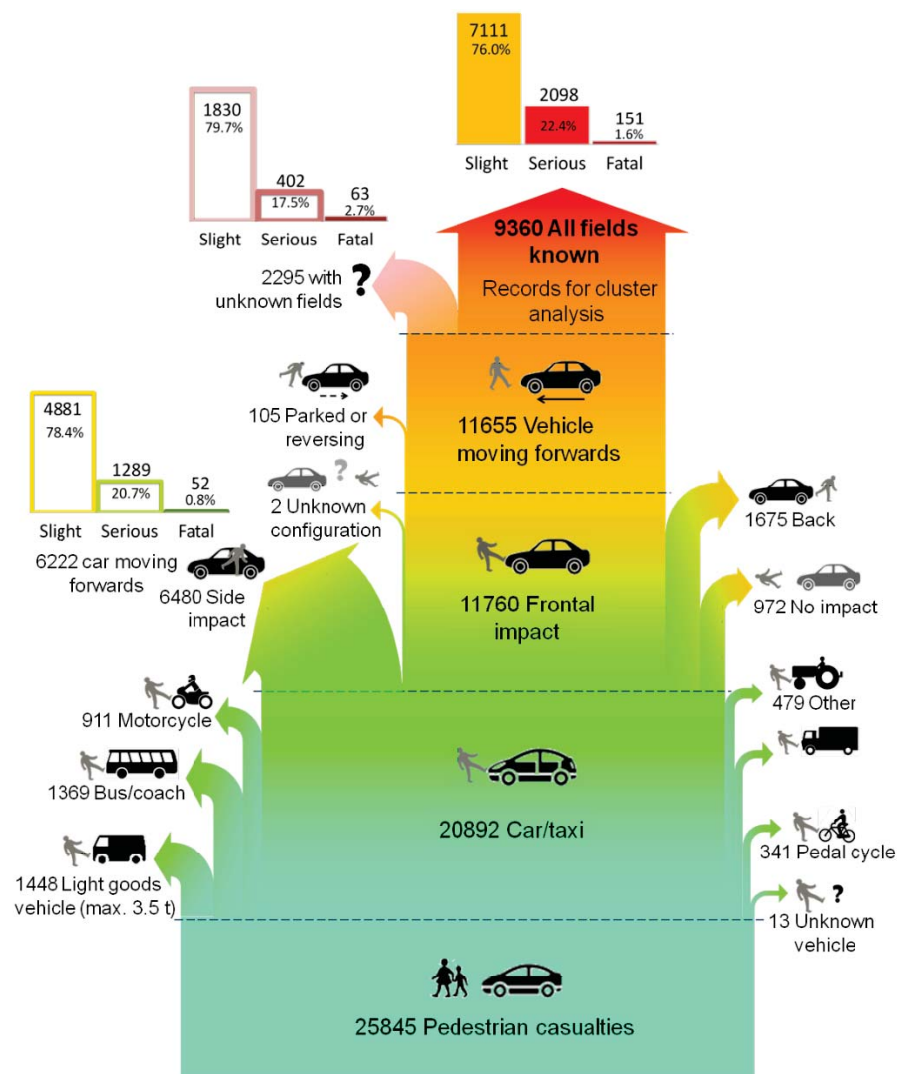


Figure 1: Selection of relevant cases from national accident database STATS19 (2010): pedestrians injured in frontal impacts with passenger cars travelling forwards. Distribution of injury severity shown for pedestrian accidents with full information (top), missing information (upper left) and side impacts with vehicles moving forwards (mid left).



impact for 56% (11,760) of pedestrians associated with passenger cars. A small proportion of these vehicles (105) were parked or reversing, leaving 11,655 cases that met the principal selection criteria for the target population: pedestrians injured in frontal impacts with passenger cars travelling forwards.

A brief review of the 6480 side-impact cars excluded from the main STATS19 analysis (4368 left, 2112 right) revealed a higher proportion of stopped vehicles and a shift downward in the distribution of injury severity compared with frontal impacts as shown in Fig. 1. The most marked differences in manoeuvre for vehicles moving forwards were that side-impact vehicles were less often turning right and more often overtaking a static vehicle or turning left.

Cases with unknown or missing information in one or more of the eight fields used for the cluster analysis of STATS19 were filtered from the dataset. This reduced the selection by 2295 records, most of which (1967) arose from uncertainty about the pedestrian’s movement on the road; the pedestrian’s age and sex (269) and precipitation (219) were also sometimes unknown. The remaining 9360 cases were carried into the cluster analysis. This group had a distribution of slight, serious and fatal casualties of 76.0%, 22.4% and 1.6% respectively compared to 79.7%, 17.5% and 2.7% in the excluded cases.

The cut-off point for the cluster analysis of the national STATS19 dataset was set at a level where the population was partitioned into 22 groups. The characteristics of six clusters that together comprise 86% (8030) of the total (9360) are detailed in Table 2 in individual columns; for completeness the remaining sixteen clusters 7–22 are shown in aggregate.

Cluster 1 was the largest STATS19 group across all injury categories, comprising 41% of all casualties, 38% of all serious-or-fatal cases and 30% of all fatalities as detailed in Table 3. These figures were derived as row totals from Table 2. The concentration of serious and fatal casualties was variable: cluster 6 for example represented only 3% of all cases but contained 15% of all fatalities.

The OTS dataset was derived directly from relevant case files that contained information in the required fields. No subsequent filtering of cases with missing or unknown field values was therefore required. The cut-off point for the cluster analysis of the in-depth OTS dataset was set at a level where the population was partitioned into fourteen groups. The characteristics of six clusters that together comprise 79% (139) of the population (175) are detailed in columns in Table 4.

The three clusters with a higher quotient of serious or fatal pedestrian casualties (3, 4, 6) tended towards a higher speed distribution as shown in Fig. 2.

Table 2: Pedestrian accident scenarios (N=9360) for national accident database STATS19 (2010). Corresponding elements of the main test configuration are highlighted in cluster 1 (green) with test variations (orange) and sub-variations (light shading) in the smaller clusters.

	Cluster							
	1	2	3	5	4	6	7-22	Total
<i>Accident severity<sup>[O]</sup></i>								
Slight	3021	993	717	522	639	175	1044	7111
Serious	801	297	303	218	123	91	265	2098
Fatal	45	3	31	23	5	23	21	151
<i>Speed limit (mph)<sup>[O]</sup></i>								
10-30	3608	1254	957	698	744	191	1219	8671
40-50	201	34	69	57	20	24	56	461
60-70	58	5	25	8	3	74	55	228
<i>Light conditions<sup>[N]</sup></i>								
Daylight	3867	1293	0	0	767	0	615	6542
Darkness	0	0	1051	763	0	289	715	2818
<i>Weather<sup>[N]</sup></i>								
Fine	3401	1293	790	564	677	222	919	7866
Not fine	466	0	261	199	90	67	411	1494
<i>Vehicle manoeuvre<sup>[N]</sup></i>								
Going ahead	3867	1293	1051	763	0	289	889	8152
Turning	0	0	0	0	767	0	441	1208
<i>Pedestrian age-sex<sup>[N]</sup></i>								
Child 0-15 yrs	1831	905	192	176	127	15	349	3595
Adult female	1023	192	293	193	343	38	360	2442
Adult male	1013	196	566	394	297	236	621	3323
<i>Pedestrian movement<sup>[N]</sup></i>								
Cross from L	2593	775	1051	0	483	0	513	5415
Cross from R	1274	518	0	763	249	0	387	3191
Other	0	0	0	0	35	289	430	754
<i>Masked by vehicle<sup>[N]</sup></i>								
Not masked	3864	0	1051	763	767	289	722	7456
Masked	3	1293	0	0	0	0	608	1904
Total	3867	1293	1051	763	767	289	1330	9360

Field type: nominal (N), ordinal (O), scale (S).

Table 3: Representativeness of national accident clusters (STATS19) according to injury severity.

	Cluster							Total %
	1	2	3	4	5	6	7-22	
Fatal	30	2	21	3	15	15	14	100
Serious or fatal	38	13	15	6	11	5	13	100
All severity	41	14	11	8	8	3	14	100

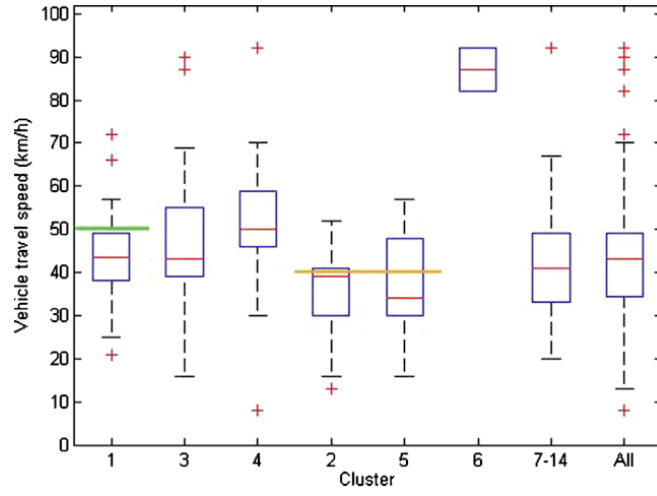


Figure 2: Vehicle travel speed prior to braking for pedestrian accident clusters derived from in-depth On-the-Spot database. Speed of main test configuration 50 km/h (green line) varied to 40 km/h for clusters 2 and 5 (orange line).

Table 4: Pedestrian accident scenarios (N=175) for in-depth On-the-Spot database. Corresponding elements of the main test configuration are highlighted in cluster 1 (green) with test variations (orange) and sub-variations (light shading) in the smaller clusters.

	Cluster							
	1	3	4	2	5	6	7-14	Total
<i>Pedestrian severity</i> <sup>[O]</sup>								
Slight or not injured	32	13	10	22	9	0	25	111
Serious	16	8	13	5	2	0	10	54
Fatal	2	4	1	0	0	2	1	10
<i>Light conditions</i> <sup>[N]</sup>								
Daylight	50	0	0	27	11	2	20	110
Darkness	0	25	24	0	0	0	16	65
<i>Weather</i> <sup>[N]</sup>								
Fine	45	25	9	23	9	2	22	135
Not fine	5	0	15	4	2	0	14	40
<i>Vehicle manoeuvre</i> <sup>[N]</sup>								
Going ahead	50	18	24	27	6	2	25	152
Turning	0	7	0	0	5	0	11	23
<i>Pedestrian (age-sex)</i> <sup>[O]</sup>								
Child 0-7 yrs	4	1	0	6	6	1	5	23
Child 8-15 yrs	12	2	1	12	5	0	15	47
Adult female	18	6	9	3	0	0	5	41
Adult male	16	16	14	6	0	1	11	64
<i>Pedestrian movement</i> <sup>[N]</sup>								
Crossing from left	29	25	7	27	0	1	14	103
Crossing from right	17	0	14	0	11	1	21	64
Other	4	0	3	0	0	0	1	8
<i>Pedestrian speed</i> <sup>[O]</sup>								
Walking	50	24	24	0	0	0	15	113
Running	0	1	0	27	11	2	21	62
<i>Line of sight (1 sec)</i> <sup>[N]</sup>								
Not obstructed	45	25	24	20	11	2	25	152
Obstructed	5	0	0	7	0	0	11	23
<i>Vehicle travel speed (km/h)</i> <sup>[S]</sup> , described in Fig. 2								
<i>Change of speed to impact (km/h)</i> <sup>[S]</sup> , described in Fig. 3								
Total	50	25	24	27	11	2	36	175

Field type: nominal (N), ordinal (O), scale (S).

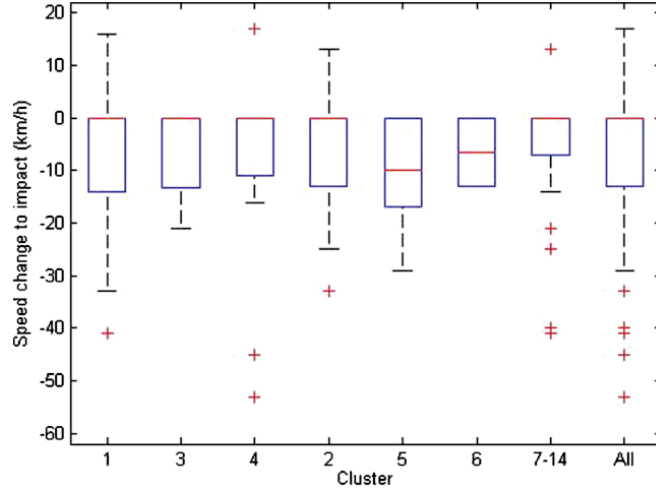


Figure 3: Change of speed to impact for pedestrian accident clusters derived from in-depth On-the-Spot database.

Table 5: Representativeness of in-depth accident clusters (OTS) according to injury severity expressed as percentage of whole dataset.

	Cluster							
	1	2	3	4	5	6	7-14	Total %
Fatal	20	0	40	10	0	20	10	100
Serious or fatal	28	8	19	22	3	3	17	100
All severity	29	15	14	14	6	1	21	100

The median value of the change of speed from normal driving to impact was zero for the larger clusters 1–4 as seen in Fig. 3; in these cases there was effectively no braking. In many cases there was braking (negative speed change) and in some cases an increase of speed prior to impact—this could occur for example on accelerating out of a turn.

Cluster 1 was the largest OTS group overall (29%) and contained more serious-or-fatal cases (28%) than the other clusters (Table 5). The proportion of fatalities in OTS is stated for completeness but the sample size (10) is too small to provide a smooth distribution. Further information on the injury severity of pedestrians in the OTS dataset was available through the maximum AIS severity (MAIS) presented in Table 6. At least 46% (80) of pedestrian casualties incurred MAIS 2+ injuries and of these almost half (38) incurred MAIS 3+ injuries.

Table 6: Maximum injury severity (MAIS) for pedestrians in clusters derived from in-depth On-the-Spot database.

	Cluster							Total
	1	3	4	2	5	6	7–14	
No injury	1	1	1	1	1	0	1	6
MAIS 1	21	10	6	18	4	0	15	74
MAIS 2	12	7	6	4	3	0	10	42
MAIS 3–6	8	7	8	4	2	2	7	38
Unknown	8	0	3	0	1	0	3	15
Total	50	25	24	27	11	2	36	175

#### *Test configurations*

Two sets of test configurations were derived independently from the two sets of accident scenarios, one based on the national STATS19 dataset and the other based on the in-depth OTS dataset. In both cases the largest cluster, cluster 1, which for STATS19 comprised over 40% of cases (3867 of 9360) and for OTS almost 30% of cases (50 of 175), was used to define a baseline test setup and the smaller clusters in each group were accounted for by varying certain elements of the main test. Cluster 1 in each dataset also contained the highest number of serious and fatal casualties (where this could be reliably ascertained), otherwise a different approach may have been necessary.

For the national database STATS19, the main test configuration, illustrated in Fig. 4, has a passenger car travelling straight ahead at 30 mph (50 km/h) that encounters a child-sized pedestrian target (0–15 years) crossing from the left (near-side) kerb in sight of the driver in daylight and fine weather, corresponding to the scenario described in Table 2. The first variation of the main test is based on cluster 2, the second largest accident scenario, where the pedestrian target is masked from view of the driver. The second variation of the main test is based on clusters 3 and 5 where darkness replaces daylight and, in an additional sub-variation, precipitation could be introduced to simulate weather that is not fine. The third variation of the main test is based on accident cluster 4 where the vehicle is turning rather than travelling straight ahead. The fourth and final variation is based on cluster 6, where a male-sized pedestrian target is standing or moving along the carriageway in darkness; an additional sub-variation of this test could be to have the vehicle is travelling at a higher speed.

For the in-depth database OTS, the main test configuration, shown in Fig. 5, has a passenger car travelling straight ahead at 50 km/h that encounters a

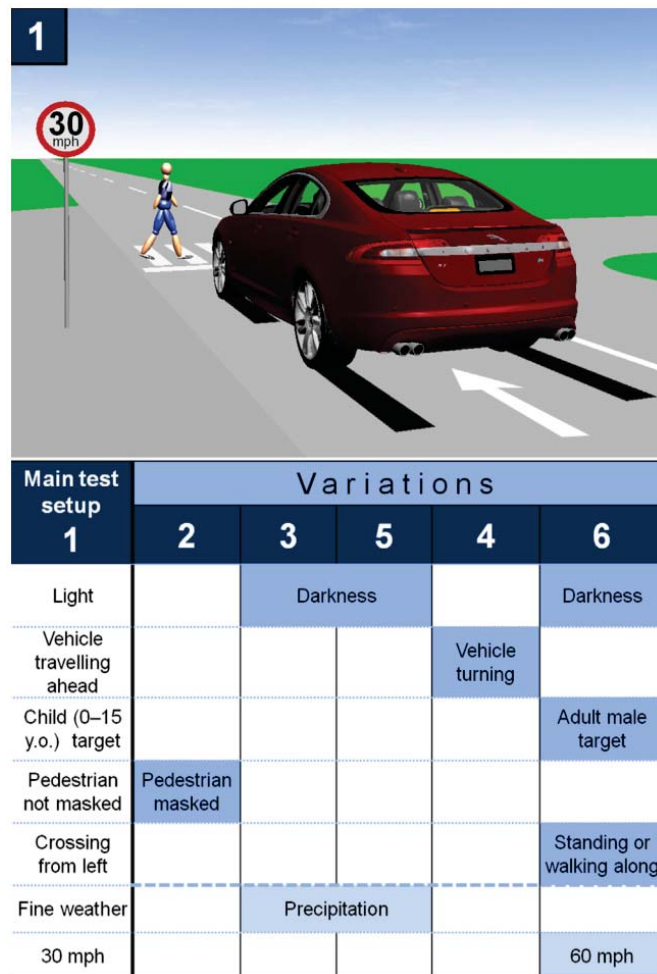


Figure 4: Main AEB test configuration (1) from national STATS19 database with four variations: pedestrian masked (2); darkness with precipitation as sub-variation (3 and 5); vehicle turning (4); adult male target on road in darkness with higher speed as sub-variation (6).

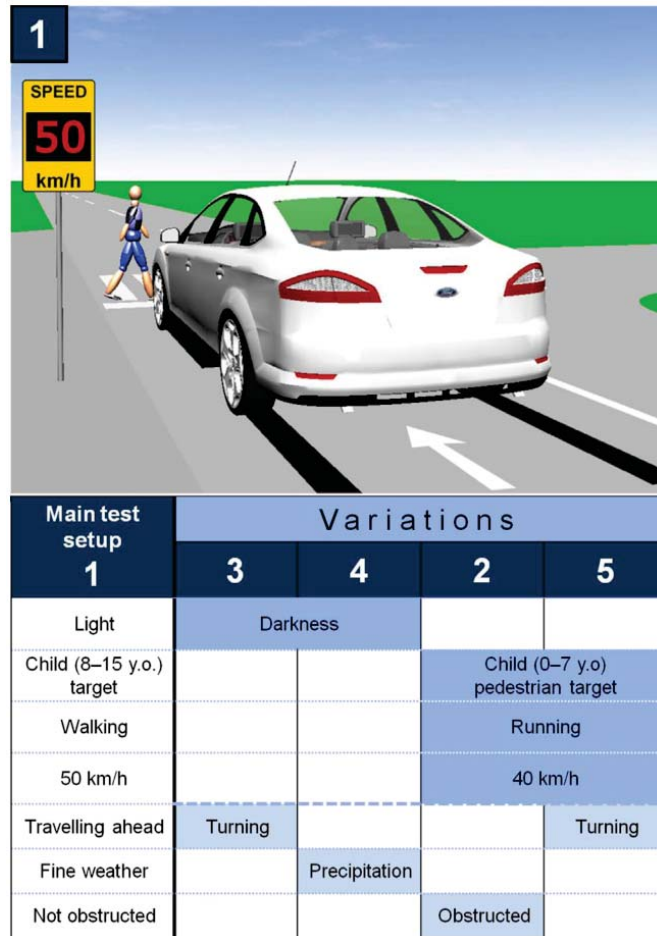


Figure 5: Main AEB test configuration (1) from in-depth On-the-Spot database with two variations: darkness with vehicle turning or precipitation as sub-variations (3 and 4); vehicle approaching at 40 km/h, small child running target with sight obstruction or vehicle turning as sub-variations (2 and 5).



child-sized pedestrian target (8–15 years) walking across from the left (near-side) kerb in sight of the driver in daylight and fine weather, derived from the scenario described in Table 4. The first variation of the main test is based on clusters 3 and 4 where darkness replaces daylight. Two additional sub-variations of this test are possible, one in which the vehicle is turning and one in which precipitation is present. The second and final variation of the main test is based on clusters 2 and 5 where the pedestrian target is reduced in size and increased in speed to represent a young child (0–7 years) running across the road and the vehicle speed is reduced to 40 km/h. Two additional sub-variations of this test are possible, one in which the pedestrian target is obscured from view and one in which the vehicle is turning.

#### 4. Discussion

There was a strong similarity between the largest accident scenarios in the STATS19 and OTS databases and the main test configurations based on these. The similarity of the accident scenarios was not guaranteed to occur because the formation of groups using cluster analysis is sensitive to the fields used and these were not identical in the two databases. There was also no overlap of cases in the accident samples, STATS19 coming from the whole of Great Britain in 2010 and OTS coming from just two sample regions in the period 2000–2009. The results therefore indirectly support the representativeness of the OTS sample and the robustness of the method of analysis.

The derivation of accident clusters from the source databases was a reproducible mathematical computation. The choice of fields to include in the source datasets and the technical specification of the algorithm determine the final results but are not predictable in advance. No attempt was made to manipulate these parameters in order to harmonise the results for STATS19 and OTS or otherwise influence the results.

It was not an analogous mechanical computational procedure to obtain the test configurations. It was assumed that the model of a baseline test with a minimal number of variations to account for less common accident conditions would be more appropriate than specifying a complete new set of test conditions for each individual accident scenario. The size of the clusters, with one scenario that was considerably larger than the others in both STATS19 and OTS, happened to be compatible with this approach. Knowing that a testing body needs to take into account more factors than just the accident data, as

mentioned above, the main purpose for deriving test configurations was to illustrate a sound method of working with the accident clusters rather than to produce a set of final recommendations. The priority of the test variations from purely the accident point of view is expressed by the relative size of the smaller clusters and distribution of serious and fatal casualties across these scenarios. The reasoning behind the choice of the test configurations is detailed in the following paragraphs.

For STATS19, cluster 1 was relatively large, containing 41% of the cases (3867 of 9360). Most characteristics of a track test were clearly suggested: speed, light, weather, vehicle manoeuvre and visibility of pedestrian target. The size of the pedestrian target was less apparent, with a distribution of 0.47, 0.26 and 0.26 for children (1831), women (1023) and men (1013) respectively. On the basis that a sensing system that detects a target of a given size will detect larger targets but not necessarily smaller targets, specification of a mid-sized target (adult female) could fail to provide protection for the children that constituted 47% of this cluster, therefore a child-sized target was nominated. The second issue was pedestrian movement, crossing from the left or crossing from the right. A single test cannot do both but replicating tests from the right and from the left seemed like a dubious use of resources. The accident data demonstrated that it is essential for an AEB sensing system to operate symmetrically. It is suggested that this should be established independently without doubling up on the number of tests, e.g. by the submission of technical data from the manufacturer's in-house development and testing programme.

STATS19 clusters 3 and 5 share the characteristic of darkness as the major divergence from cluster 1. Two other differences, namely a higher proportion of larger pedestrians and crossing from the right (cluster 5), were already covered by the small pedestrian target in the baseline test and the general requirement for symmetry. Precipitation, i.e. weather 'not fine', was present among 25% of the cases in these clusters (450 of 1814). While this proportion was too large to be neglected, it was also debatable whether it compelled precipitation as a test condition. As a compromise, it is suggested that clusters 3 and 5 could be represented by varying the baseline test to occur in darkness and to award bonus points or a top rating if the emergency braking system can also cope with precipitation, a test 'sub-variation'.

In STATS19, cases of the vehicle turning were concentrated in cluster 4 without other notable deviations from the main test configuration and so formed a natural third test variation. Finally, cases of the pedestrian standing or mov-

ing along the carriageway were concentrated in cluster 6 which also featured darkness, a predominance of adult males and a slight shift towards higher speed zones. This formed a fourth test variation with the higher speed treated as an optional sub-variation.

Turning to the accident scenarios based on OTS, cluster 1 was relatively large, containing 29% of the cases (50 of 175). The selection of most characteristics for the main test followed the logic applied to STATS19 with the outcome that pedestrian target size was based on an older child (8–15 years) to cover 92% of cases (46 of 50). The suggested vehicle speed of 50 km/h was close to the 75th percentile of cluster 1 as marked in Fig. 2 and fit with a preference to match the test speed to a nominal speed limit or to a current pedestrian impact test.

OTS clusters 3 and 4 occurred in darkness and this constituted the first variation of the baseline test conditions. Precipitation and turning vehicles occur in a minority of cases in these clusters and could be considered as supplementary sub-variations. The second variation of the main OTS test followed from clusters 2 and 5 which required coverage of a faster moving and smaller pedestrian target (running child, 0–7 years old). The vehicle travel speed tended to be lower in these clusters and was nominally set at 40 km/h, between the 50th and 75th percentiles. Obstructed line of sight and turning vehicles occurred in a minority of cases and could be considered as supplementary sub-variations.

The change of speed from emergency braking had a median value of zero for OTS clusters 1–4 as shown in Fig. 3, indicating that many pedestrian impacts occurred before the driver had time to react. Even in cluster 5, a turning scenario where braking would occur in normal driving, the upper quartile value was zero. It is very difficult after the event to infer whether a driver reacted as fast as humanly possible. The automatic intervention of typical AEB systems aims to improve on the reaction of the driver, even when the collision is inevitable and the benefit of the system is injury mitigation rather than prevention. The development of advanced detection systems will further improve the robustness of the target-recognition algorithms and boost reaction time in particularly challenging scenarios such as small, fast moving targets emerging from visual obstructions (Broggi et al., 2009). For such extreme scenarios, the test configurations corresponding to STATS19 cluster 2 (Fig. 4) and OTS cluster 2 (Fig. 5) could be considered to not only rate the avoidance capability but also to assess the reduction of the collision speed and the modification of the impact point, elements that indicate the reduction of the severity risk (Badea-Romero

et al., 2013).

## 5. Conclusion

Common accident scenarios were successfully generated from two databases using cluster analysis, an objective and computationally reproducible technique. The two sets of clusters showed strong commonalities even though computed from independent sets of accidents. It proved reasonably straightforward to define AEB test conditions corresponding to the accident scenarios by matching the largest accident cluster to a main test and introducing variations to the main test to account for characteristics of the smaller accident clusters. This approach, which is neither mechanical nor guaranteed to work for any set of clusters, struck a good balance between minimising the number of proposed tests and covering the greatest diversity of accident circumstances. It is intended to ascertain the wider representativeness of the results by analysing accident data from other countries in a similar way.

## Acknowledgements

The On-the-Spot project was funded by the Department for Transport and the Highways Agency. Funding for analyses of OTS and STATS19 was received from Thatcham Motor Insurance Repair Research Centre (UK) and the Insurance Institute for Highway Safety (USA). The views expressed in this paper are those of the authors and do not necessarily represent those of any other body.

## References

- A. Aparicio, J. Bargalló, and S. Baurès. Development of a methodology and the test tools for the evaluation of pedestrian detection systems. In *Proceedings of the 23rd International Technical Conference on the Enhanced Safety of Vehicles*, number 13-0422 in ESV technical paper, pages 1–6, Seoul, May 2013. NHTSA.
- The Abbreviated Injury Scale 2005: update 2008*. Association for the Advancement of Automotive Medicine, Des Plaines, 2008.
- A. Badea-Romero and J. Lenard. Source of head injury for pedestrians and pedal cyclists: striking vehicle or road? *Accident Analysis and Prevention*, 50:1140–50, 2013. doi: 10.1016/j.aap.2012.09.024.

- A. Badea-Romero, F. J. Páez, A. Furones, J. M. Barrios, and J. L. de Miguel. Assessing the benefit of the brake assist system for pedestrian injury mitigation through real-world accident investigations. *Safety Science*, 53:193–201, 2013.
- A. Broggi, P. Cerri, S. Ghindoni, P. Grisleri, and H. Jung. A new approach to urban pedestrian detection for automatic braking. *IEEE transactions on Intelligent Transportation Systems*, 10(4):594–605, 2009.
- 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 73, Department for Transport, London, 2008.
- STATS 20: Instructions for the completion of road accident reports from non-crash sources*. Department for Transport, London, September 2011. URL [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/230596/stats20-2011.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/230596/stats20-2011.pdf).
- EuroNCAP. EuroNCAP rating review. Technical report, European New Car Assessment Programme, July 2013. URL <http://www.euroncap.com/en/for-engineers/technical-papers/>. Accessed May 2015.
- D. Fleury and T. Brenac. Accident prototypical scenarios, a tool for road safety research and diagnostic studies. *Accident Analysis and Prevention*, 33(2): 267–76, 2001.
- A. Habibovic and J. Davidsson. Causation mechanisms in car-to-vulnerable road user crashes: implications for active safety systems. *Accident Analysis and Prevention*, 49:493–500, 2012.
- A. Habibovic, E. Tivesten, N. Uchida, J. Bårgman, and M. L. Aust. Driver behavior in car-to-pedestrian incidents: an application of the driving reliability and error analysis method (DREAM). *Accident Analysis and Prevention*, 50: 554–65, 2013.
- S. Huang, J. Yang, and F. Eklund. Evaluation of remote pedestrian sensor system based on the analysis of car–pedestrian accident scenarios. *Safety Science*, 46(9):1345–55, 2008.
- W. Hulshof, I. Knight, A. Edwards, M. Avery, and C. Grover. Autonomous emergency braking test results. In *Proceedings of the 23rd International Technical Conference on the Enhanced Safety of Vehicles*, number 13-0168 in ESV technical paper, pages 1–13, Seoul, Korea, May 2013. NHTSA.
- L. Kaufman and P. J. Rousseeuw. *Finding groups in data: an introduction to cluster analysis*. John Wiley and Sons, Belgium, 1990.
- J. Lenard, R. Danton, M. Avery, A. Weekes, D. Zuby, and M. Kühn. Typical pedestrian accident scenarios for the testing of autonomous emergency

- braking systems. In *Proceedings of the 22nd International Technical Conference on the Enhanced Safety of Vehicles*, number 11-0196 in ESV technical paper, pages 1–9, Washington DC, June 2011. NHTSA. URL <http://www-esv.nhtsa.dot.gov/Proceedings/22/files/22ESV-000196.pdf>.
- D. J. Searson, R. W. G. Anderson, and T. P. Hutchinson. Integrated assessment of pedestrian head impact protection in testing secondary safety and autonomous emergency braking. *Accident Analysis and Prevention*, 63:1–8, 2014.
- J. M. Sullivan and M. J. Flannagan. Determining the potential safety benefit of improved lighting in three pedestrian crash scenarios. *Accident Analysis and Prevention*, 39(3):638–47, 2007.
- The MathWorks Inc. Matlab (R2012b) statistics toolbox documentation. Matlab “help” facility, 2012.
- S. Zecha, G. Jürgens, and P. Quittenbaum. Innovative test methods and facilities for predictive pedestrian protection. In *Proceedings of the 23rd International Technical Conference on the Enhanced Safety of Vehicles*, number 13-0377 in ESV technical paper, pages 1–8, Seoul, Korea, May 2013. NHTSA.