



**Feasibility study for equipment database and
exposure data evaluation**

Final Report

**Undertaken on behalf of the
Department for Transport**

DfT Framework Ref: SO802/V6

Prepared by

Ruth Welsh

Transport Safety Research Centre

Loughborough University

May 2009

vsrc

Vehicle Safety Research Centre





**THE QUEEN'S
ANNIVERSARY PRIZES
2007**

Awarded to the Vehicle Safety Research Centre

Contents

EXECUTIVE SUMMARY 1

1. INTRODUCTION 3

2. FITMENT DATA 6

2.1 Make, Model, Variant based data sources..... 8

2.2 VIN based data sources 13

2.3 Other data sources 17

2.4 No data available..... 19

2.5 Outstanding replies..... 19

2.6 Summary..... 19

3. EXPOSURE DATA..... 22

3.1 Government Organisations 23

3.2 Commercial Organisations 29

3.3 Vehicle Manufacturers 31

3.4 Insurance Companies and Thatcham 32

3.5 Knowledge Gaps..... 32

3.6 Summary..... 33

4. LINKING FITMENT, EXPOSURE AND ACCIDENT DATA BASES 35

5. DISCUSSION AND NEXT STEPS 39

EXECUTIVE SUMMARY

This report outlines the approach undertaken and the results obtained for a study to assess the feasibility of establishing a safety equipment database and suitable measures of exposure of safety equipment within the passenger car fleet. The work has been led by the TTSC at Loughborough University with additional support from WSP.

At present, information on the fitment of technologies within passenger cars is not readily available as a central resource. It is therefore difficult to establish whether any relationship exists between accident data and the effectiveness of specific technologies. A knowledge of safety equipment fitted to vehicles in the fleet, the exposure of the technology on the road and appropriate accident data are prerequisites in order to determine the true effectiveness of safety features in the future once sufficient fleet penetration has occurred. Methods for linking fitment, exposure and accident data with any additional requirements are discussed in the report.

Government organisations, commercial organisations, research bodies, motor manufacturers and the insurance industry have been contacted, together with a review of the available literature sources, in order to ascertain the level of current data and its suitability for meeting the Department's future analytical requirements.

Feasible approaches for developing a fitment database have been established. This can be either done identifying vehicles according to their make/model/variant and production model, by using the Vehicle Identification Number (VIN) as an identifier or by a combination of both. Some caution would be required to minimise any inaccuracies relating to variations within a production year but fleet coverage should be possible. Coordination with the Department's Technology Watch Programme will be required upon building a fitment database in order to ensure that technologies are classified in a similar generic manner.

A suitable exposure database is less feasible with the currently available resources. Whilst as measure of the composition of the fleet is possible and some mileage data accessible for cars undertaking MOT examinations, data relating to the road type and

the driver characteristics (available as estimates in the National Travel Survey) are not currently linked to the vehicle specification. This data is required in order to assess technologies designed for a particular environment or for a target sector of the driving population. Modified or additional data collection activities would be required in order to establish a sufficient exposure database.

Further work is now required in order to determine the most feasible approach out of those identified and to determine ways in which to collect the data that is needed but not currently available. This would depend upon accessibility of the data in practice and the cost of the data. Pilot studies should be undertaken to establish further the usability of the promising resources beyond that intimated during the enquiries, to identify any barriers when the data is applied and to assess the analytical power of the data.

1. INTRODUCTION

This report outlines the approach undertaken and the results obtained for a study to assess the feasibility of establishing a safety equipment database and suitable measures of exposure of safety equipment within the passenger car fleet. The study was undertaken on behalf of the DfT, led by the Transport Safety Research Centre (TTSRC) at Loughborough University and supported by WSP. Both the TSRC and WSP are within the supply chain for the DfT Framework consortium that is led by Faber Maunsell.

At present, information on the fitment of technologies within passenger cars is not readily available as a central resource. It is therefore difficult to establish whether any relationship exists between accident data and the effectiveness of specific technologies. In order to assist the Department to make informed policy decisions about the support of new technologies or conversely action to mitigate any risks they might pose, this study explores the potential for collating databases that identify the safety equipment fitted to each vehicle in the fleet and the on road use of the vehicle. A knowledge of safety equipment fitment by make and model together with the exposure of each on the road is essential in order to determine the true effectiveness of safety features in the future once sufficient fleet penetration has occurred.

One of the eventual aims of the Department is to be able to evaluate the effectiveness of drivers assistance systems and new vehicle technologies and the contribution they make towards accident and injury mitigation for car occupants. The fast rate at which such systems are developing means that there is often no formal assessment of the safety claims or the possible risks, for example from increased driver distraction or workload. There is also a risk that the un-coordinated introduction of new technologies ignores issues of compatibility and could lead to systems with potentially significant safety benefits being overlooked in favour of those that are considerably more marketable.

The current Technology Watch project being conducted by TRL aims to inform the Department of existing and forthcoming technologies, their functionality and claimed effectiveness. However proving the effectiveness is not possible with the current data resources available. This study assesses the feasibility of collecting the data required

to make such effectiveness studies of the systems currently highlighted within the Technology Watch programme and to add new data as future systems are notified through Technology Watch. It is important that a standard classification of system functionality is applied to both the Technology Watch and the Fitment database.

A previous workshop co-ordinated by the DfT identified a number of key research actions relating to the assessment of active safety systems. The actions fell into two work phases. The first phase relates to information gathering. The second phase considers the development and application of methodologies to predict the effectiveness and introduced risks of new systems using, amongst others, the information gathered in the first phase. This feasibility study relates to the information gathering activity in phase 1.

The objectives of the study are therefore to assess the feasibility of developing two basic information sources to support accident analysis work. These resources are

- A standard recording system to identify the safety equipment fitted to each model of vehicle; and
- Information on the on-road use of different models of data (exposure data) to enable calculation of accident risks.

For the fitment data evaluation two identifiers have been considered. The first relies upon only knowledge of the make, model, variant and manufacture date of the vehicle under consideration. The second uses codes potentially stored within the Vehicle Identification Number (VIN) to establish the safety equipment a vehicle is fitted with. The latter approach is advantageous since it uniquely identifies each vehicle and hence has the potential to build a highly accurate database. The VIN approach also deals with the issue of optional equipment fitment more conclusively than a make/model approach.

For the exposure data consideration has been given to the level of any information available and its applicability to future research questions. For example, some data may be available at a vehicle level (mileage), some at a road level (proportion of travel on motorway, A road etc.) and some at a driver level (age and gender). Recording driver level information will require appropriate data protection considerations, however this type of data is currently made available in other

databases used extensively for Governmental research projects such as the STATS 19 data. Similar protocols to those applied to the STATS 19 data could be employed for the current purpose. The requirement of the data will depend upon the question under consideration.

The work undertaken for the feasibility study has been split into tasks. The first, allocated to the TSRC, considers available sources of information relating to fitment data whilst the second, allocated to WSP considers the exposure data. For each task agreed organisations were approached either via email, telephone or in person in order to establish what existing data sources are available. This report details the organisations approached together with the responses received. The study has aimed to identify the resources currently available and to identify gaps in the data that would require new data collection activities to be developed. At the kick off meeting and the interim project meeting it was reiterated by the client that it is not the intention that the feasibility study should determine the most appropriate method for collating a fitment and exposure data base, but should identify all of potential data sources so that further investigations can be undertaken.

Section 2 of the report covers the fitment data activity and section 3 covers the exposure data activity. In section 4 links between the fitment data, the exposure data and the accident data bases are discussed and the ability of the current data sources to answer the anticipated future research questions is discussed. Section 5 draws conclusions including recommendations for the next steps that need to be made.

2. FITMENT DATA

The fitment database would ideally list by either VIN or make, model, variant and production year the safety systems present. A fitment database derived from make/model listings is likely to have inaccuracies based around optional fitment and changes to the specification within a production year. The resulting data base would give a make model variant and model year listing. A fitment database derived from the VIN has the advantage that each vehicle is uniquely identified, supplies accurate data but can also be linked to other datasets through the make and model. It should be noted that, upon consultation with some manufacturers, their opinion was that a VIN based dataset would be preferential. Both methods for sourcing the information required were explored in the study.

The current Technology Watch Programme is tasked with identifying existing and forthcoming technologies and informing the Department regarding the system functionality and potential effectiveness based upon the available literature. It is important that the classification method used by Technology Watch and any fitment database is coordinated to allow direct comparison between claimed and proven system effectiveness. During the course of the feasibility study a representative from the Technology Watch programme was consulted and made aware of the current activity. Further consultations will be appropriate upon development of the fitment database.

The scope of the fitment database should include equipment that comes as either standard or optional fit to the vehicle at the time of purchase. It does not include nomadic after sales devices.

The organisations and resources considered for the fitment data assessment were as follows;

- Literature
 - Web sites
 - Brochures
 - Auto Magazines
 - Glass's Guide

- Dealers and Franchisers
- Manufacturers
 - Volvo Car Corporation
 - Ford of Europe
 - Toyota
 - Renault
 - Vauxhall
- Thatcham Motor Insurance Repair Research Centre
- JATO Net - Automotive Intelligence on the Internet
- European Council for Automotive R & D (EUCAR)
- EuroNCAP
- European Automobile Manufacturers' Association (ACEA)
- Vehicle and Operator Services Agency (VOSA)
- Society of Motor Manufacturers and Traders Ltd (SMMT)
- Driver and Vehicle Licensing Agency (DVLA)
- Vehicle Certification Agency (VCA)

Additionally consideration was given to ways in which to access information stored within the VIN and the link between the VIN and the Vehicle Registration Mark VRM.

The information available from each of the resources listed above is discussed in the following sections. Section 2.1 details information available based on a make/model/variant identification, Section 2.2 shows the potential for using the VIN as an identifier, Section 2.3 lists the organisations where there is the potential for limited information that could be used to enhance or validate a database, Section 2.4 lists those organisations where no information is available and Section 2.5 lists the outstanding replies. Section 2.6 provides a summary of the fitment data activity.

Where data is potentially available the following points are considered;

- The method for data acquisition
- The accuracy of the data
- The sufficiency of the data to meet the research needs including any gaps in the data.

2.1 Make, Model, Variant based data sources

2.1.1 Published literature review

Method for data acquisition

Data relating to safety equipment fitment are readily accessible through a number of literature sources. These are for example;

- Web sites
- Sales Brochures
- Auto Magazines
- Independent publications e.g. Glass's Guide

These can be reviewed and a data base developed manually by inputting relevant information into a suitable spread sheet. This method has been employed in previous studies in particular for a review of ESC effectiveness. Glass's Guide provides the most comprehensive and central source of information which can be complemented using the other literature sources.

There is an indication that the information contained within the hard copy of Glass's guide may be available in electronic format since their Automotive Industry Branch are able to offer tailored analyses including discussion for projects that currently fall outside of Glass's data range. Further enquiries have been made but no reply has been received as yet.

Accuracy of data

Although Glass's Guide is a comprehensive resource, it is accepted that inaccuracies in the recording exist. This has been confirmed by the motor industry and other research organisations. The extent of these inaccuracies are as yet unknown but enquiries have been made in order to quantify the likely errors. Cross referencing with manufacturers websites and car magazines can improve the accuracy and in cases of ambiguity direct consultation with the manufacturer is possible. Glass's guide indicates at what point in time optional fitment becomes standard but it does not give a rate of optional fitment uptake.

Potential of data source

This method of collecting data is guaranteed to provide a basic fitment database since the source material is readily available and does not require collaboration or co-operation from other organisations. The resource has the potential to cover all makes and models in the fleet thus building a comprehensive database. The method does not differentiate for optional fitment uptake. Estimates would need to be made based upon sales figures obtained from an alternative source (e.g. JATO Net see section 2.1.4)

Although inaccuracies exist, once these are quantified and suitable estimates for optional fitment made, the method would produce a database that would be sufficient for the research need.

The resources required for this method are not expensive. However the method is extremely laborious and hence costly in terms of man power. The data can be open to inaccuracy and the basic method does not currently provide an automated approach for updating the data as new technologies emerge. Should it emerge that an electronic database is available then the method would be substantially less laborious but more costly. It is likely that some reclassification of the data would be required in order to be compatible with the Technology Watch Listings.

Further notes

A further activity based around this method is currently being undertaken by Thatcham. This has been used in order to previously develop and update their ESC safety rating system which rates a vehicle according to whether ESC is standard across the model range, is optional or is not present. It is believed that this is now being extended to cover other safety features. Thatcham indicate that they believe the method to be reliable since they have additional data sources from within the manufacturers. They also indicated that there would be potential to share the information that they are currently collating.

2.1.2 Car franchisers / dealers

A number of car franchisers / dealers were visited to establish the detail of information available.

Method for data acquisition

Sales representatives were not able to provide any more information than that available in the show room brochures, a part of the method covered in 2.1.1. No central electronic database with safety equipment listed against make, model and variant was apparent.

Accuracy of data

Although the brochures would provide relatively accurate information no accommodation can be made for optional fitment and this would need to be estimated from alternative sources (see 2.1.1)

Potential of data source

This resource was not explored any further since no additional information or different presentation of data was available beyond that achievable through a published literature review.

2.1.3 Motor manufacturers

Five motor manufacturers were approached for the purposes of the study. These were;

- Volvo Car Corporation
- Ford of Europe
- Toyota
- Renault
- Vauxhall

Responses were received from Ford, Toyota and Renault.

Method for data acquisition

In general, for information on a make/model/variant basis manufacturers also referred to web sites and brochures in a similar approach to that outlined in 2.1.1 / 2.1.2.

Accuracy of data

Ford in particular indicated that inaccuracies would be incurred using this method.

Potential of data source

This resource, for the purpose of a database built purely around make and model rather than VIN, was not explored any further since no additional information or different presentation of data was available beyond that achievable through a review of available literature outlined in 2.1.1.

2.1.4 JATO Net

During the course of discussions with Ford and Renault, both suggested exploring a web based resource called JATO Net (<http://www.jatonet.com>). JATO Net aims to deliver automotive data via an on-line accessible and searchable data source concerned with vehicle pricing, specifications and volumes. The specification data cover 44 countries Queries can be made by searching on vehicle make and model, by searching on vehicle features or by viewing recent changes to vehicles. The volumes data covers 50 countries giving sales and registration figures. The vehicle specification is discussed in this section. The potential for JATO Net to provide exposure data is discussed in section 3.

Method for data acquisition

1. JATO Net methodology

The process of data collection is similar to that outlined above for the 'literature' resource. Additionally details from EuroNCAP, though not researched, are injected directly into the database from the EuroNCAP database. New models are added to the database as they enter the market place. Data is collected before the vehicles are released for sale and data is available as soon as the price list is released. For the UK data there is already a link between the car specification data and the sales volumes with option take up data.

The variables included in the data set are reviewed annually; generally 20 or more new variables are added with none being removed. New variables can be considered according to the clients needs. Currently 1626 vehicle specifications are listed. There is scope to ensure that the data available are classified to be compatible with the definitions used by Technology Watch and the data can be supplied in any required format. The data automatically cover all manufacturers with a sales volume of 25 units / annum or more, but can be driven by clients requiring the data. Retrospective data are available for vehicles no longer in production. Optional fitment is coded for each make / model and option take up is collected as part of the sales figures.

2. Acquisition of JATO Net database

The JATO Net data base can be accessed on-line once a subscription has been taken. Typically access to the entire data available on-line requires a subscription of ~ £30K per year. Alternatively only parts of the data can be accessed at a reduced cost, or JATO Net are able to perform ad-hoc analyses on behalf of the client for a fixed cost.

Accuracy of data

Quality checks are made of the data. A point system is used to rate the importance of each individual variable and a weighted error rate is calculated based upon which variables were in error. Thus, though inaccuracies can occur these are to some extent quantifiable. The rates are available to clients for expert comment. Optional fitment uptake is accommodated by various links within the fitment and exposure elements of the JATO Net data.

Potential of data source

The JATO Net data base offers a great deal of potential to supply fitment data sufficient for the Department's research needs. Since data are available relating to both standard fit and optional take up there is the potential to access a comprehensive data set. However, a further meeting with JATO Net and the DfT would be beneficial in order to fully establish the level of flexibility of the data provision; this should also include a trial analysis. Liaison is also recommended between JATO Net and a representative from the Technology Watch programme should this data source be pursued further.

2.2 VIN based data sources

The Vehicle Identification Number (VIN) is a 17 digit alpha/numeric code fitted to all UK registered vehicles. It is a mix of manufacturer, SAE and ISO elements. These elements are broken down into: the World Manufacturer Identifier (WMI), the Vehicle Descriptor Section (VDS) and the Vehicle Identifier section (VIS).

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17		
WMI			VDS						VIS									
Manufacturer identifier			Vehicle attributes						Check digit	Model year	Plant code	Sequential number						
Managed by SAE Denotes country of origin (box 1) and manufacturer (boxes 2 and 3)			Managed by manufacturer Information varies, may include: Model Body Style Engine Size Platform Gross Weight Restraint information No of doors						Calculated value This is calculated based on the sum of digits and weighted values for letters	Code given in ISO 3779	Managed by manufacturer Unique to each vehicle, may simply be a sequential number, or may include set numbers within a manufactures year indicating: Engine size Transmission Options fitted etc							

The **WMI** is managed by SAE and denotes the country of origin and the manufacturer.

VDS: For some manufacturers, e.g. Mercedes Benz, the VDS contains Model and Variant information, which could be cross referenced with Glass's Guide/manufacturer information to indicate the standard equipment fitted.

VIS: It may be possible to decode an option fitted to a single model, by a single manufacturer. E.g. for a 2006 Ford Focus the option of a handling pack may be indicated by the sequential number within the VIS. Unfortunately for the majority of manufacturers VIS is simply a sequential number in order to bring the VIN up to the required 17 digits.

The VIN can be used either directly by decoding information contained within the VDS to identify the make/model and then refer to a literature source, to or by using the VIN as an identifier for a more detailed vehicle specification held by the manufacturer. Both approaches are discussed in this section.

2.2.1 VIN Decoders / Manuals

Method for data acquisition

A VIN guidebook is researched and published yearly by the International Association of Auto Theft Investigators (IAATI). This book provides information on decoding Vehicle Identification Numbers. IAATI have confirmed that it is possible to decode safety technology fitted as an option to certain vehicles by limited manufacturers but not all manufacturers. The IAATI VIN guidebook is not available in any electronic format.

This method for obtaining fitment data would rely upon VIN listings being available for the fleet which is feasible since the DVLA data also contains the VIN. The VIN can be decoded to give the make and model however this is already available from the DVLA together with the VIN. Since only a very small number of elite car manufacturers record safety fitment within the VIN, a make/model approach as listed in the previous section would need to be applied to determine safety fitment. Thus, VIN decoders offer no additional potential beyond a make/model approach.

Accuracy of data

The data acquired through a VIN decoder would be accurate where the manufacturer records sufficient information within the VIN.

Potential of data source

This method offers no additional potential than the data that would be available from the DVLA. Using VIN decoders to establish the make and model and then apply a literature resource would in fact be far more labour intensive than a direct approach from the DVLA.

2.2.2 Enhanced data linked by VIN

Consultation with Ford of Europe has indicated that it is feasible for the manufacturer to provide listings of safety equipment by VIN. Ford keep an in-house enhanced VIN database where the VIN has 80 digits rather than the 17 recorded on the vehicle. This enables considerable in-depth information for each vehicle to be stored, primarily for the purposes of product recall. FORD believe that other manufacturers would also have such databases available. This has been confirmed by Renault.

There is a current activity being undertaken by TUV in Germany. Consultation with TUV revealed that fitment data is purchased from a number of manufacturers including

- Ford
- Volvo
- BMW
- Daimler
- Toyota
- Honda
- General Motors
- Volkswagen

The scope of the data base being co-ordinated by TUV covers the following technologies;

- cornering light
- adaptive cruise control
- adaptive brake lights
- airbag / (seat) belt (pre-) tensioner
- hill hold assistant (hill launch assist)
- autonomous trailer stabilisation control
- auto hold
- anti-lock braking system
- emergency brake (currently for heavy vehicles)
- automatic headlight levelling
- automatic light / headlamp assist
- downhill assist
- brake assist
- bending light
- electric power steering
- electro-hydraulic controlled pusher and trailing axle (heavy vehicles)
- electro mechanic parking brake (EPB)
- electronically-controlled braking system
- electronic differential lock
- electronic stability program
- main beam assistant

- speed limiter
- cruise control
- trailer stabilisation control
- (bus) stop brake (brakes automatically while a door is open)
- tilt stabilisation control
- mechatronic wedge brake / electro hydraulic brake
- pre safe brake
- lane keeping assist
- lane departure warning system with break intervention
- lane changing assist
- traction control
- active steering
- roll-over protection (active)

Method for data acquisition

There are two potential ways in which to access data from the enhanced VIN. There is potential to share the data that is currently being purchased by TUV. Further enquiries have been made regarding the feasibility and cost that would be associated with collaborating with TUV but TUV have stated that they would need to know more specifically the data required before a response can be given. In order that all the databases were compatible, some reclassification of the data would be required in order to be consistent with the Technology Watch classification system. This would be a one off activity that should not be overly costly.

Alternatively a new data acquisition process could be initiated and tailored specifically to the Department's needs. It is likely some manufacturers would be willing to contribute to a new database, but there would be an associated cost.

Accuracy of data

In cases where the manufacturer was able to provide VIN based fitment data the data would be extremely accurate and reliable but there would be gaps in the data for those manufacturers that were unable to contribute.

Potential of data source

Depending upon the level of co-operation within the industry / potential to collaborate with TUV, this data source has the potential to build an accurate data base but it is likely that it will need to be complimented by a make/model approach to build a fully

comprehensive data base since it is unlikely that all manufacturers would contribute. The resulting data base would be sufficient for the Department's research needs.

Since the information would be provided in an electronic format the data base would be easily established and maintained. Manual additions would be necessary for missing data.

Follow up work is now required to determine further the potential for collaboration with TUV and the associated cost.

2.3 Other data sources

2.3.1 Vehicle Certification Agency (VCA)_

Within Europe, two systems of type approval have been in existence for over 20 years. One is based around EC Directives and provides for the approval of whole vehicles, vehicle systems, and separate components. The other is based around ECE (United Nations) Regulations and provides for approval of vehicle systems and separate components, but not whole vehicles.

Type approval is the confirmation that production samples of a design will meet specified performance standards. The specification of the product is recorded and only that specification is approved.

Automotive EC Directives and ECE Regulations require third party approval - testing, certification and production conformity assessment by an independent body. Each Member State is required to appoint an Approval Authority to issue the approvals and a Technical Service to carry out the testing to the Directives and Regulations. An approval issued by one Authority will be accepted in all the Member States.

VCA is the designated UK Approval Authority and Technical Service for all type approvals to automotive EC Directives and ECE Regulations.

VCA tend to deal with crash, braking and emissions performance data. Some of this fitment data (for instance electronic stability control) is held within reports but these

reports are confidential and cannot be released otherwise industry confidentiality would be infringed. Other EU authorities have their own Type approval processes and it is not guaranteed that a European or British car will have British type approval, for instance Rover could go to Germany for type approval of their systems.

Method of data acquisition

Not applicable

Accuracy of data

Not applicable

Potential of data source

Due to the scope of information held at VCA and issues of confidentiality, this data source would not be able to contribute effectively to a fitment data base.

2.3.2 EuroNCAP

Initial enquiries have been made with EuroNCAP. EuroNCAP do not hold a suitable fitment data base but clearly regularly make detailed examinations of new production vehicles. EuroNCAP would be interested in contributing to a future data base and co-operating where possible. Further discussion would be required to establish a process for information sharing should this data source be considered of sufficient use to the fitment data base.

Method of data acquisition

Not applicable

Accuracy of data

Not applicable

Potential of data source

EuroNCAP has the potential to make a useful contribution to a fitment data base but would not be able to be the sole contributor or maintain such a database. Information gathered during a detailed vehicle examination could be used to validate information acquired through other sources.

2.4 No data available

The following organisations do not hold information relating to safety equipment fitment;

- EUCAR
- DVLA
- DfT STATS Branch

2.5 Outstanding replies

The following organisations have yet to respond to requests for information;

- ACEA
- SMMT
- VOSA

2.6 Summary

Although some organisations / groups did not reply to the request for information, those that did resulted in feasible approaches being identified. The potential to build a fitment data base, either on a make / model basis or a VIN basis or a combination of both has been established.

Literature sources such as Glass's guide could be used to manually build a database based upon the make/model/variant and production year, this can be complemented by manufacturers information available through brochures, websites and from motoring magazines. This approach has some level of inaccuracy and does not appear to adequately accommodate estimates for optional fitment; additional information from alternative sources (potentially JATO Net) would be required to estimate optional fitment uptake rates. The resources required are not expensive, but the method is labour intensive. Additional cost would be incurred to estimate optional fitment take up.

JATO Net is a readily available on-line resource that holds a comprehensive vehicle specification listing. Both standard and optional fitment can be determined from the

data and estimates of data errors are available. There is potential to tailor data collection to the requirements of the DfT. The data can either be bought via a subscription in any required data format, or JATO Net are able to provide ad-hoc analyses.

The VIN based approach would be best achieved directly from the manufacturer since VIN decoders would be extremely limited in their ability to give any information regarding the safety equipment fitment.

A make model approach has the advantage that the method would include all makes and models in the fleet, however a VIN based approach is likely to be more accurate. Gaps in the fleet (where decoders or manufacturer co-operation is not forthcoming) can be filled using the literature.

It should be noted that the consensus among those asked from research organisations and the motor industry was that a database built around the VIN would be the favourable approach.

Based upon the information currently available the table below summarises the advantages and disadvantages of the 4 potential methods for building a fitment data base; Literature, JATO Net, VIN decoder, VIN link to manufacturer databases.

Whilst the feasibility study has shown that data sources exist and how they could be utilised to build a fitment database, further exploration is now required to identify the exact levels of co-operation that could be expected from the various organisations and what the associated costs would be.

	Make/Model sources		VIN sources	
	Literature	JATO Net	VIN decoder	VIN link to manufacturer
Fleet	Comprehensive	Comprehensive	Extremely	Some

coverage			Limited	manufacturers may not contribute
Optional fit	No	Yes	Extremely Limited	Yes
Readily available	Yes	Yes	No	No
Accuracy	Inaccuracies	Inaccuracies	Accurate	Accurate
Electronic source	No	Yes	No	Yes
Sufficient for research need	When combined with other method	Yes	No	When combined with other method
Update process	Manual	Provided by JATO Net	Manual	Upon request

3. EXPOSURE DATA

In order to analyse the accident data and form conclusions relating to the 'absolute' as opposed to 'relative' risk of accident involvement and hence determine more precisely the effect a given safety feature is having, exposure data is required. It is important to have exposure data not only for vehicles fitted with a given system but also those without the system in order to make system effectiveness estimates.

Exposure data can be collected at a number of levels. In its most basic form this could simply be the number of a particular vehicle in the fleet at any one time. This can then be refined to measure the distance travelled (km) by make and model and then further segregated by characteristics such as age and gender of driver or road type (e.g. rural/urban, motorway/non-motorway etc.) Different analyses will require exposure data measured at different levels and with different base units. For example, an analysis of the effectiveness of lane departure warning systems would ideally require the exposure data to include the road type. An analysis of Alcolock would benefit from knowledge of the driver age and gender.

For the purposes of this study, the feasibility of collecting data relating to the following measures of exposure has been explored together with some additional points as listed below;

- Make and model of the vehicle
- Vehicle mileage data
- Distance travelled and road type data
- Age and gender of the driver
- Sales figures and rental figures by make/model on an annual basis
- Whether the data is available for other European countries
- Potential for contribution to an exposure data base.

The following organisations were approached;

- DVLA
- DfT Stats Branch (including National Travel Survey)

- Highways Agency NTCC
- Highways Agency ID
- Highways Agency Traffic Technology Division
- Highways Agency SSR
- VOSA
- JATO Net
- Traffic Link
- Traffic Master
- WSP roaming car data project
- WSP car fleet
- Centrico
- FISITA
- SMMT
- Insurance companies (Norwich Union, Zurich)
- Car manufacturers/traders (Ford, Honda, VW, Volvo, Mazda, Vauxhall)
- Thatcham

The findings from the exposure feasibility study are presented sections 3.1, 3.2, 3.3 and 3.4. The organisations approached have been categorised as either Government organisations, commercial organisations, motor manufacturers or insurance companies/institutes. For each organisation where a reply was received comments are made according to the type of data available from the resource, the availability of the data and the suitability of the data for the Departments research needs. Section 3.5 discusses gaps in the knowledge and section 3.6 provides a summary of the exposure data activity.

3.1 Government Organisations

3.1.1 DVLA

Type of Data

DVLA holds a record of the make and model of all vehicles registered in Great Britain including vehicle fleets. Records are also held for vehicles in Northern Ireland that have been registered through the Automatic First Registration and Licensing (AFRL) scheme. This accounts for about 90% of NI new registrations. The vehicles within the

data can be identified by the VIN. DVLA do not hold vehicle mileage information, these data are held by VOSA as part of the MOT database and are not currently linked to DVLA's vehicle records. The potential to link these data would need to be assessed.

Two distinct data bases exist, one with the vehicle as the base unit, the other with the driver as the base unit. It is not possible to identify the age of the vehicle keeper. The data covers the UK only, no information is kept relating to other European countries.

Availability of Data

Much of the relevant data held by the DVLA is already passed on to the DfT Stats Branch and so may be readily available. Any relevant data that is not currently supplied to DfT could be obtained by direct contribution from the DVLA provided the end use of the data was fully understood and confidentiality issues were suitably addressed. Data could be supplied in a format to suit the Department's requirements.

Suitability of Data

The data would be sufficient to provide exposure information relating to fleet constitution and for mileage covered (older cars). Importantly a direct link could be made via the VIN to a potential fitment data base. Mileage data would however only be available for cars that are required to take an MOT test. Estimates for the mileage would need to be made for new cars.

3.1.2 Highways Agency (HA)

Various individuals were contacted within the HA, concentrating on three services within the HA, which were;

- Safety Standards and Research (SSR)
- National Control Centre (NTCC)
- Information Directorate (ID)
- Traffic Technology Division / Traffic Operations (TO)

No response was received from the SSR and no relevant information is held by the ID. TO are not in a position (due to work commitments) to contemplate becoming involved in developing a database.

Type of Data

Some data could potentially be available from NTCC through a combination of road side sensors and Automatic Number Plate Recognition (ANPR). This is termed "Classified Vehicle Flow".

Currently Classified Vehicle Flow is the flow (expressed in vehicles per hour) of a particular classification of vehicles travelling along a Link during a 5 minute monitoring period. Flow may be derived directly from monitoring equipment on the link or calculated from measurements made on adjacent links. Vehicles are classified according to their length as follows:

- Vehicles less than 5.2 metres long
- Vehicles between 5.2 and 6.6 metres long
- Vehicles between 6.6 and 11.6 metres long
- Vehicles over 11.6 metres long

The vehicle classification is done by MIDAS loops so there is no way it could go any further other than calculating vehicle length. ANPR could potentially provide such information if it had some kind of interface with DVLA (to identify vehicle make from the registration number), however at present ANPR only focuses on a single lane (usually the middle one) and could only provide a snapshot.

Availability of Data

Though there is potential to link various sources to generate some of the required data (vehicle flow) the data is not currently available and a substantial amount of collaboration and set up would be required to determine whether the links mentioned above would be possible and if so for them to be out into practice.

Suitability of Data

The data potentially available from the HA would be limited in its analytical capability. Since the data are sampled they would provide only *estimates* of traffic flow and so the sampling and subsequent weighting issues would need to be considered. If the method could be applied to all road types, it would however provide some estimated information relating to the respective flow of specific makes and models on different road types. A potential DVLA link would also provide information for identified vehicles on mileage for older vehicles and also details of the driver age and gender.

Data protection issues would need to be considered if personal data were collected and, since the data represent only a sample, appropriate issues would be apparent.

Further exploration of the potential of a combination of vehicle flow estimates, ANPR and a link to the DVLA is required in order to fully assess the feasibility of this approach.

3.1.3 DfT Statistical Branch

Type of Data

As indicated by the DVLA, the Stats branch at the DfT hold data giving the make and the model of the vehicle. Data regarding new registrations of vehicles is also held. Very basic information relating to the registered keeper of the vehicle is available, this includes the gender but not the age. The driver is not however always the same person as the registered keeper. The data held by the Stats branch covers vehicles registered in Great Britain. No other European data is held.

In addition to the DVLA data, the DfT commissions a continuous household survey of personal travel, the National Travel Survey (NTS). As part of this survey, each household member is interviewed to collect information on their socioeconomic, demographic and travel characteristics. They are then asked to record details of their trips over a 7 day period. The NTS collects information from over 8,000 households each year, comprising over 19,000 individuals. The survey used to be carried out by the Office for National Statistics but, since 2002, it has been contracted to the National Centre for Social Research (NatCen). A weighting strategy was introduced to the NTS a couple of years ago, to improve the representativeness of the results. It is worth noting that the NTS is primarily designed to measure long term trends at a national level and is not suitable for measuring change from one year to the next or for producing local estimates.

As a part of the NTS respondents are asked to supply the VRM for each household vehicle; this information is provided by most but not all respondents. The VIN is not collected and it is unlikely that it would be practical to ask respondents to provide this. Where accurate VRMs are collected, these are linked to the DVLA data so that additional vehicle details can be attached, thus reducing the burden on NTS

respondents. (Where respondents do not want to supply their VRM, they are asked some additional questions about their vehicle during the interview). Information collected on household vehicles includes the fuel type, vehicle age, engine size and annual mileage (as estimated by the main driver).

The NTS questionnaire is reviewed each year and a more thorough review of the vehicle questionnaire is planned this year. There is therefore some scope to add additional questions. However, new questions can only be added if some existing questions are dropped, to ensure the burden on respondents is kept to a reasonable level. Moreover, only certain types of questions are suitable for inclusion on the NTS. As the survey is designed to measure long term trends, new questions are only added if they address issues which the Department is likely to want to measure in the long term.

In relation to exposure, the NTS only holds details of the vehicles used by respondents during their 7 day travel week if they are household vehicles. So, for example, if a respondent records a trip as a passenger in a non-household vehicle, the details of that vehicle would not be known as these would not have been collected during the interview.

Availability of Data

The availability of the DVLA data held by the Stats branch for subsequent research purposes would be subject to the nature of the specific data required and the end user. Historically the Stats branch have provided enhanced accident data, including the make and model of the vehicle, to research organisations acting on behalf of the Department. Suitable end user agreements are issued and no personal information has been released. Requests for access to the DVLA data held by DfT Stats are considered on a case-by-case basis. However, any requests for personal or sensitive information would need to be cleared with DVLA before their release was agreed. It is likely that this process can be extended beyond the accident data to other data sets such as the National Travel Survey data.

The NTS branch can produce ad hoc analyses on request, providing the sample sizes are large enough to produce reliable results. An edited extract of the raw NTS data is also deposited at the UK Data Archive to allow researchers to carry out their own analyses. A number of variables are removed from this extract to minimise the

risk of disclosure. However, some of these variables are sometimes made available to external researchers if they can demonstrate their need for this extra level of detail; if they satisfy the Department that they will treat these data in accordance with data protection legislation and if they sign an end user licence. The Department does not receive the names and addresses of respondents from the survey contractor and information at this level of detail would never be released to external users.

Suitability of Data

The DVLA data held at the DfT Stats branch is in essence a reduced version (no mileage information) of that held by the DVLA. The VIN is not carried over to the DfT, but the VRM is available. Linking via the VRM to the VIN is possible but inaccuracies occur and typically only around 80% of cars are matched. The data is therefore less adequate than that held at the DVLA.

The data collected within the National Travel Survey would not be adequate at present however, the variables collected are currently under review, particularly those related to the vehicle. There is therefore potential for some discussion within the department to establish whether there is scope for data collection that would give adequate exposure data for the Department's future research needs. It is evident from discussions with the National Travel Survey representative that if new data is to be collected then it is seen as a replacement for some other data that is no longer required. This, along with sampling issues, would need careful consideration.

3.1.4 VOSA

No reply was received from VOSA.

Type of Data

n/a

Availability of Data

n/a

Suitability of Data

n/a

3.2 Commercial Organisations

3.2.1 JATO Net

A description of JATO Net is given in section 2.

Type of Data

The JATO Net data base contains volumes data sourced from SMMT. This is available for over 40 countries. The data represent wholesale figures with no detail about the vehicle purchaser.

Availability of Data

The data is readily available via subscription to the JATO Net service; alternatively JATO Net are able to prepare reports and conduct analyses.

Suitability of Data

The data would provide details of the fleet constitution only and so has only limited use for the future research requirement. As it stands, the JATO Net database offers no more than can be sourced (for Great Britain) from the DVLA.

3.2.2 Trafficlink

Trafficlink is an independent provider of real time traffic and travel information to broadcasters, business and government. Trafficlink have been aggregating traffic and travel information for 10 years. Some data is supplied to the Highways Agency for operational purposes and to the DfT's Transport Direct travel portal.

Type of Data

Vehicle make and model is available for a section of vehicles. It is not clear how comprehensive this sample would be. Estimates of vehicle fleet sizes could be provided (approximately 60,000 vehicles monitored each month). Information relating to distance travelled is not currently available but there is potential to generate this data with an associated development cost. The data is primarily available for the UK but some information is gathered relating to UK based vehicles travelling abroad. No data is available regarding the age or gender of the driver.

Availability of Data

Trafficlink are in a position to supply any data that is relevant subject to any confidentiality agreements that would need to be checked.

Suitability of Data

The data would provide only estimates of the fleet and mileage data would need to be developed. As such the data is less comprehensive and suitable than that available from the DVLA or from the DfT's Stats branch.

3.2.3 Traffic Master

Traffic Master records and supplies data from a network of 7,500 traffic sensors covering 8,000 miles including all of the UK's motorways and 95% of its trunk roads. Whilst data relating to make, model and distance travelled is stored the data cannot be passed on.

3.2.4 VCA

The VCA does not hold any exposure data.

3.2.5 FISITA

No response was received from FISITA

3.2.6 SMMT

No response was received from SMMT

3.2.7 WSP car fleet / WPS roaming data

Type of Data

Make and model data is available for cars within the WSP fleet. No data relating to mileage, road type or agenda gender are stored. No information is available through the WSP roaming data.

Availability of Data

Limited data can be made available in excel format.

Suitability of Data

The data available would be insufficient for the Departments research requirement.

3.2.8 CENTRICO

[CENTRICO](#) is a Euro Regional Project implementing Intelligent Transport Systems (ITS) to improve the efficiency of the Trans European Road Network (TERN) in order to keep people and goods moving. The CENTRICO project is supported by the DG TREN of the European Commission. One of the sources of contact recommended contacting the project manager of this project in regards to the queries of exposure data. The project manager from Austria did initially try to contact other member states however was unable to get feedback from colleagues in regards to the exposure data queries.

3.3 Vehicle Manufacturers

The following motor manufacturers were contacted;

- VW
- Ford Motor Company
- Vauxhall
- Honda
- Volvo
- Mazda

In general there was a lack of replies from the manufacturers. Although a response was received to initial enquiries and assurances that the questions posed would be looked in to, none of the manufacturers came back with any further information despite follow up calls and emails.

It is thought that whilst some demographic information relating to the market population may be stored by marketing teams, other exposure data would not be available and so any data would be limited. Co-operation with the manufacturers is more crucial for the fitment database.

3.4 Insurance Companies and Thatcham

Norwich Union

Norwich Union responded that Thatcham were the best source of contact because the insurance companies pass on such information to them. Norwich Union are not in a position to supply any exposure data to a database.

Zurich

Zurich reported that they did not hold any such information and could not provide any contribution to the potential database.

Thatcham

Thatcham were not forthcoming with any information passed to them by the insurance industry. They mentioned a database that exists in the USA; the HLDI insurance data base which contains policy holder information (gender/age/location/VIN) and claims history. The UK used to have a similar database but it lost support from insurers in 1984. There is therefore a model that could be reapplied in the UK if the Association of British Insurers were able to contribute. It would be beneficial to arrange a meeting with appropriate stake holders and DfT to explain the purposes the data would be used for. Assurance regarding anonymity and potential benefits to the ABI would be required. This exposure data would not give mileage, but with the VIN it could be linked to DVLA and MOT data held by VOSA.. This type of data source would have the added benefit of being able to identify the non-injury accidents through the claims history, not available on the current accident data bases. Again, data protection issues would need to be considered.

3.5 Knowledge Gaps

Considering all of the resources listed in the previous sections, the following gaps in the available data that would be required in order to meet the future research requirements are apparent.

- Mileage data for pre MOT cars
- Road type
- Definitive gender of the driver (not necessarily the keeper)

- Driver age

Mileage data for newer cars could be roughly approximated from the data for older cars or could be collected and estimated by additions (collecting the make and model information) to the National Travel Survey data. It may be possible to make approximations from vehicle service records, but the manufacturers did not respond to queries for data.

In theory the NTS could obtain the make and model for NTS household vehicles but we do not currently hold this on our main database as we do not need this data at this level of detail. The NTS does not currently have information on the road types used and this would require additional data collection. Age and gender are recorded in the travel survey but the link to the vehicle would need to be made.

Data protection issues would need to be addressed for personal data.

3.6 Summary

The feasibility of establishing an exposure database is limited and many gaps in the desired data exist, specifically with regard to driver characteristics, the road type upon which travel occurs and the distance travelled (newer cars).

The table below summarises the data **currently available** from the different sources. The most promising and available source of data is that held by the DVLA, much of which is currently also held by the DfT Stats Branch.

Potential for generating the required information has been established but needs further investigation at this stage. Additional information could be collected as part of the National Travel Survey though the sample size may need to be increased in order to have a sufficient confidence in the data for analytical purposes. The potential for amendments to the National Travel Survey is still being assessed.

A data base built upon policy holder information with claims included would have the benefit of providing data relating to non-injury accidents. This is currently a gap in the accident data knowledge. The potential to develop such a database, previously held in the UK but now obsolete, would need further assessment.

	Make/Model	Mileage	Road Type	Driver Age	Driver Gender
DVLA	Fleet	Only with link to VOSA – older cars	No	No	Keeper gender
HA	No	No	No	No	No
DfT Stats	Fleet	No	No	No	Keeper gender
National Travel Survey	Potentially	Sample based	No	Sample based	Sample based
VOSA	No	No	No	No	No
JATO Net	Fleet	No	No	No	No
Trafficlink	Limited	No	No	No	No
Traffic Master	No	No	No	No	No
VCA	No	No	No	No	No
FISITA	No reply				
SMMT	No reply				
WSP fleet	Limited	No	No	No	No
CENTRICO	n/k	n/k	n/k	n/k	n/k
Motor Manufacturers	No reply				
Insurance Companies	No	No	No	No	No

4. LINKING FITMENT, EXPOSURE AND ACCIDENT DATA BASES

In order to fully assess the effectiveness of safety systems a common identifier is needed across the fitment, the exposure and the accident databases. This could be either the make/model/variant and production period of the vehicle or it could be the VIN. Both alternatives are discussed in this section together with gaps in the knowledge that would hinder the power of future analyses. The data sources indicated for each approach are those that show the most potential/ease of acquisition based upon the findings in the previous sections.

Linking by make, model, variant and production period

The following table indicates a possible approach for building databases based upon the make/model/variant and production period. Where shortfalls in any of the data sources exist, these are highlighted in the column 'Additional Requirements'.

		Data source	Additional Requirements
Fitment data		JATO Net	Attention to any errors in the data / categorisation of technologies in parallel with Tech Watch
Exposure data	Fleet composition	DVLA/Stats Branch	Caution required with correct definition of production period.
	Mileage	DVLA link to MOT data at VOSA	Only for post MOT vehicles. Estimates required for newer vehicles (e.g. service history data / extrapolation from older vehicles)
	Road type	New data collection process.	New variables required to include vehicle type and larger sample required in NTS, or complete new data collection activity required.
	Driver gender	National Travel Survey or other new data collection process.	New variables required to include vehicle type and larger sample

			required in NTS or new data collection activity required.
	Driver age	National Travel Survey or other new data collection process.	New variables required to include vehicle type and larger sample required in NTS or new data collection activity required.
Accident data		CCIS/STATS19/OTS etc	All list make/model and variant (~80% of STATS 19) but more accuracy required regarding the production period.
Non injury accidents		Not fully covered by existing accident data	New data collection process required.

The main issue with a make/model/variant and production period approach is the correct categorisation of the vehicles with the information available. Some data sets record first year of registration or approximate year of manufacture based upon the licence plate and as a result some inaccuracies may result when matching the fitment database with the exposure and accident data.

Linking by VIN

		Data source	Additional Requirements
Fitment data		Direct from manufacturer	May require 'top up' for missing manufacturers using make/model approach. Need to ensure compatible technology classification with Tech Watch.
Exposure data	Fleet composition	DVLA/Stats Branch	Make/model link required for some data, the VIN gives this with the model year.
	Mileage	DVLA link to MOT data at VOSA	Only for post MOT vehicles. Estimates required for newer vehicles (e.g. service history data / extrapolation from older vehicles) Make/model link required for some data,

			the VIN gives this with the model year.
	Road type	New data collection process.	New variables required to include VIN and larger sample required in NTS, or complete new data collection activity required.
	Driver gender	National Travel Survey or other new data collection process.	New variables required to include VIN and larger sample required in NTS or new data collection activity required.
	Driver age	National Travel Survey or other new data collection process.	New variables required to include VIN and larger sample required in NTS or new data collection activity required.
Accident data		CCIS/STATS19/OTS etc	VIN would need to be matched with fitment and exposure data bases before being removed from the data. A link to the VIN from the VRM would be needed in some cases. All list make/model and variant (~80% of STATS 19) but more accuracy required regarding the production period.
Non injury accidents		Not fully covered by existing accident data	New data collection process required.

The VIN has the advantage of being a completely unique vehicle identifier but an entirely VIN based system is unlikely to be achieved. It would however produce a highly accurate method for those parts of the fleet where the VIN can be used to produce a fitment database. The alternative approach of using the make/model would be required to fill the gaps in the fitment data. VIN is either available, can be sourced retrospectively or added as an additional collection requirement for the other data sources. From the VIN the make/model/variant and model year can be identified allowing for the link to those parts of the fitment data where the VIN is not available.

This section has given two possible approaches for building data that would be sufficient to assess the effectiveness of safety systems. Neither approach however should be considered independent of the other since there is, in most cases, a direct link from the VIN to the make/model/variant and model year of the vehicle. The actual method employed in the future, which could be a combination of make/model, VIN, existing and new data, will depend largely on the associated costs and ease with which potential sources can be turned into actual sources of information.

5. DISCUSSION AND NEXT STEPS

At present it appears feasible to set in place a process to develop a fitment database. This could essentially be in any required data format. The level of accuracy would vary depending on the data sources chosen. Some require more co-operation with external bodies than others.

Although data relating to the composition of the fleet is readily available and some information relating to mileage accessible, there are substantial gaps in the current data sources relating to the on road use of vehicles and the driver characteristics. Certain technologies will have a functionality that is designed particularly for mitigation of accidents on specific road types or layouts and hence information relating to the exposure of vehicles to different environments is required to enable effectiveness studies. Similarly some technologies may be aimed at a target driver population and hence knowledge of the driver demographics is required to assess the effectiveness of these technologies. Thus, new data collection activities are required in order to build an exposure data base with the levels of information required to address the future research objectives.

There would be benefit in carrying out further work to

- Hold a workshop, to include members from all relevant organisations, in order to review the findings and discuss the advantages and disadvantages associated with the different available data.
- Develop a test analysis for JATO Net in order to fully appreciate the potential of the data source; an evaluation of electronic brake assist could be a good candidate as the systems are becoming relatively common in the fleet.
- Establish further the potential for collaboration with the TUV fitment data and the associated cost.
- Conduct comprehensive consultations with the industry (possibly through SMMT though there has been no reply to any enquiries from SMMT) to establish level of co-operation and cost for manufacturers to supply VIN and

fitment listings into a central database aside from the activity being undertaken by TUV.

- Consider further methods for increasing the levels of exposure data. There is potential within the National Travel Survey but this would require negotiation and further discussion between the Department and the Stats Branch.
- Determine a methodology for generating mileage data for vehicles in the first three years of their life. This could in part be addressed by additions to the National Travel Survey but the sample size would be small. Extrapolation from older vehicles may be required or a new data source such as service history information.
- Consider a method for generating a new specific mileage data source for vehicles of particular interest in the fleet.
- Consider a method for accessing information relating to non injury accidents currently not covered by the accident data. This would involve consultation with the insurance industry.
- Consider steps to assess the potential difficulties with collecting/recording personal information.
- Ascertain the costs associated with accessing the different data sources / developing new data collection processes.
- Conduct a pilot study to determine the viability of the potential data sources and establish a firm method for generating the required data bases.