THE CORRELATION BETWEEN THE CONSUMERS' ASSOCIATION SECONDARY SAFETY RATING SYSTEM AND THE FOLKSAM CAR SAFETY RATINGS

P.F.Gloyns, S.J.Rattenbury, Vehicle Safety Consultants Ltd.,
A.Kullgren, A.Lie, C.Tingvall, Folksam Research,
J.Edwards, R.Hill, The Association for Consumer Research

ABSTRACT

The CA/VSC Secondary Safety Rating System is a predictive method for estimating a new vehicle's performance in the mix of accidents it will predictably encounter. It is published via consumer organisations to allow purchasers to take secondary safety into account when making their buying decisions. About 200 models of vehicle have been inspected and assessed using this system over the past 8 years.

The Folksam Car Safety Ratings deduce the relative safety of different models of car retrospectively, using insurance and police accident data. The results are given wide publicity, again with the intention of influencing consumer decisions.

A subset of 22 car models, common to both rating systems, was selected and the correlation between the predicted safety performance and the actual experience in accidents was examined. This correlation is shown to be high.

The various approaches to safety rating used around the world are discussed and reviewed for their strengths and weaknesses.

INTRODUCTION

The CA/VSC Secondary Safety Rating System for cars (1) is the only system in the world that attempts to predict the secondary safety performance of new cars in the mix of accidents they will predictably encounter. As such it is intended to allow the car buyer to take safety into account when making a buying decision. This in turn, apart from benefitting the individual consumer, introduces market forces into the area of secondary safety design.

The Folksam Car Rating System (2) attempts to deduce the relative safety of different designs of car retrospectively by analysis of insurance and police accident data in Sweden. The need to accumulate sufficient accident experience prior to analysis necessarily causes a delay of several years between the introduction of a new model and its evaluation on this rating system. However, there is the clear intention to influence buying decisions, and indeed change the way in which safety is viewed within car companies.

It should be stressed that the definition of safety performance is a very complex issue, and that there is room for much discussion on what constitutes the most meaningful measure of this quality. Whilst there are substantial differences in the detailed nature of what is being measured in the two rating systems considered here, there is a shared aim of trying to guide consumers and car manufacturers towards safer vehicle designs. There is interest both in the extent to which the advice offered by the two systems correlates, and also in exploring the scope for predictive and retrospective approaches complementing each other in the future.

It is believed that this is the first time that the correlation between a predictive system, aimed at rating vehicles in the whole population of accidents likely to be encountered, has been compared with a retrospective accident based system in this way.

DISCUSSION OF VARIOUS APPROACHES TO RATING OF SECONDARY SAFETY

There are two fundamental approaches to the provision of secondary safety information on cars. The first attempts to PREDICT how a new design will behave when released into the traffic environment. This prediction can be based on a variety of methods, ranging from a single frontal crash test with instrumented dummies, as used with the New Car Assessment Program (NCAP) in America for example, to the application of field accident experience as used in the CA/VSC Secondary Safety Rating System. A second type of approach seeks to deduce RETROSPECTIVELY how a vehicle is actually performing on the basis of accident experience with the model in question. Examples of this approach are the Highway Loss Data Institute (HLDI) ratings in America and the Folksam ratings in Sweden.

It is worth reviewing the strengths and weaknesses of these two basic approaches. It will be seen that both have their positive points, and it is not the case that one approach should be used to the exclusion of all others. Rather, it is apparent that the best results can be obtained in the future when predictive and retrospective systems can work in harmony.

Positive Aspects of Predictive Systems

1) Ideally they should be capable of use during the vehicle design process, as well as being used to assess the vehicle in production form.

2) Because they can be applied to new models as they are launched onto the market, predictive ratings have the potential for producing the greatest effect on the vehicle mix on the roads. They therefore have the greatest potential for encouraging vehicle manufacturers to give high priority to safety design.

3) People can actually reduce the risk of sustaining unnecessary injury by avoiding less safe vehicles.

4) By indicating that some named designs fall short of the level a consumer could reasonably expect, particularly in comparison to competitive models, they encourage manufacturers to avoid such design to minimise product liability risks.

5) Results are free from complications introduced by factors associated with the type of person who chooses to buy and drive a particular model of car, and other issues of exposure.

Negative Aspects of Predictive Systems

1) The technique used for vehicle assessment must relate very closely to real accident conditions. If it does not, there is a risk that design will be optimised to produce good results on a rating system without associated good performance in accidents.

2) If destructive testing is used as the whole or part of a rating system, the costs increase rapidly. Nevertheless, relevant destructive testing should be considered to improve predictive systems.

3) There are difficulties associated with making a suitable allowance for the influence of vehicle weight.

Positive Aspects of Retrospective Systems

1) By using the results of actual accidents with human occupants some of the difficulties associated with crash testing and interpretation of dummy experience are bypassed.

2) Depending on the nature of the data used, there is the potential for commenting on vehicle performance in the entire mix of accident types, including complex situations which are difficult to reproduce under test conditions.

3) If the data used is collected for other purposes anyway, the additional cost of analysis to produce rating information can be relatively small.

Negative Aspects of Retrospective Systems

1) The results necessarily only become available some years after a new car model is launched. Results on high selling cars become available sooner than those on lower volume models.

2) There are real problems in eliminating driver and vehicle use characteristics entirely from accident outcome data.

3) There are difficulties in making the analysis suitable for commenting of all occupants, not just the driver, and all accident types.

4) There are major difficulties in explaining the reasons behind the differences in apparent safety between two cars of similar weight for example, if only injury data and superficial accident data is available.

Problems of Rating Systems based on Crash Tests and Dummy Outputs

1) Because costs are high, there is a need to conduct the minimum number of tests, usually only one. This must introduce problems of repeatability.

2) Injury predictions based on dummy outputs are critically influenced by the detailed design of the specific area of the vehicle contacted in a single test. Since the characteristics of the steering wheel or knee impact area can change dramatically if the impact site is moved a few centimetres, or the contact speed is changed slightly, dummy measurements from a single test provide a poor measure of the predicted overall safety of a vehicle in the full mix of accident circumstances.

3) The practical constraints on the number of tests that can be conducted, imposed by cost, mean that only a very limited simulation of the range of conditions encountered in accidents can be undertaken.

4) Whole vehicle crash tests against a fixed barrier provide no measure of the benefits or disadvantages available to car occupants as a result of differing vehicle mass.

Problems Common to any Rating System

1) A suitable definition of what is implied by the term "safety" is needed. Decisions have to be made about the relative importance given to preventing death, serious injury, disabling injury and minor injury.

2) Decisions have to be made about what restraint usage rates and conditions are appropriate. Should the intention be to indicate the level of performance anticipated if all restraints are used, or if restraints are used at the rates actually observed in the field.

DIFFERENCES OF APPROACH BETWEEN FOLKSAM AND CA/VSC SYSTEMS

Before considering the analyses that follow, it is important to understand the substantial differences between the bases of the two rating systems.

CA/VSC System	**Folksam System**
Based on UK accident data	Based on Swedish accident data
Based on prevention of fatal and serious injury with heavy emphasis on fatal injury	Two measures available - one relates to risk of injury of any severity; the other relates to risk of 10% disabling injury or fatal outcome, with no distinction between these two outcomes
Comments on safety of vehicle for all occupant positions	Mainly confined to injuries to adults in front seat
Anticipates full spectrum of accident conditions	Mainly based on car to car accidents

CA/VSC System (cont.)	Folksam system (cont.)
Based on actual rates of restraint usage seen in UK accidents	Excludes cases in which front seat occupants are known to be unrestrained
Based on specific model in range	Based on complete range of particular model, all engine options and body styles.

It can readily be appreciated that with the fundamental differences outlined above, it would be unrealistic to expect identical results from the two systems, even if both were perfect within their own terms of reference.

COMPARISON OF OUTPUTS OF CA/VSC AND FOLKSAM SYSTEMS

Basic Description of Methods

The CA/VSC assessment is based on a detailed inspection of the vehicle, with some dismantling. Over 50 features of the car are rated for design quality against the current state of knowledge. The rating for these individual assessments is then combined in a way which reflects the relative importance of each area in the whole spectrum of accidents. This produces a Raw Score which is the assessment of the overall secondary safety design quality. Then a correction for vehicle weight is applied to produce the Corrected Score, which is the predicted performance of the vehicle overall in accidents.

The Folksam injury risk rating is based on paired comparisons of two car collisions from police data where the outcome in both vehicles is considered. The injury severity rating is assessed by a sub-sample of accidents reported within the third-party liability insurance and includes all kinds of accidents. Detailed medical information is available and is used to calculate the risk of death or disability.

In some cases, more than one model in a model range had been assessed on the CA/VSC system. In that situation, where possible the saloon or hatchback variants were selected for comparison with the Folksam data rather than estates, and petrol engine versions were chosen rather than diesels. Then the most recently assessed model was chosen which matched the Folksam model year range. All cars included in the analysis had been fitted with rear seat belts.

Material For Analysis

Twenty-two vehicle models were selected which were common to both rating systems. Selection was controlled solely by the need for sufficient numbers of a given model to exist within Folksam's accident files for statistically reliable analysis. The models chosen are shown in table 1.

Table 2 shows the data used for the analysis that follows. The data is in the same order as table 1.

Table 1: Models Used for the Comparison

Specific Model (CA/VSC DATA)	Year Assessed (CA/VSC DATA)	Model Range (FOLKSAM DATA)	Model Years (FOLKSAM DATA)	No. in Sample (FOLKSAM DATA)
Audi 100 2.0 E	1989	Audi 100	83-90	116
Citroen BX	1986	Citroen BX	All	42
Fiat Regatta 85 S	1987	Fiat Ritmo	All	48
Fiat Uno 45	1986	Fiat Uno	All	39
Ford Escort 1.4 L	1986	Ford Escort	81-90	317
Ford Fiesta 1.1 L	1984	Ford Fiesta	76-89	215
Ford Granada Ghia	1985	Ford Scorpio	All	67
Ford Sierra 2.0 GL	1988	Ford Sierra	All	233
Mazda 626 GLX	1986	Mazda 626	83-88	90
Mercedes 190 E	1984	Mercedes 190	All	51
Nissan Micra GL	1984	Nissan Micra	All	107
Vauxhall Cavalier L	1987	Opel Ascona	82-88	117
Vauxhall Nova 1.3 L	1988	Opel Corsa	All	53
Vauxhall Belmont L	1986	Opel Kadett	85-91	167
Renault 5 TR	1988	Renault 5	85-91	37
Saab 900	1988	Saab 900	All	670
Saab 9000i	1987	Saab 9000	All	87
Volvo 340 GL 1.7	1986	Volvo 300	All	425
Volvo 240 GL Est	1989	Volvo 240	All	2026
Volvo 740 GLE	1984	Volvo 700	All	620
VW Golf C 5-dr	1986	VW Golf/Jetta	84-91	245
VW Polo C	1986	VW Polo/Derby	All	101

Table 2: Different Measures Used in the Analysis of Correlation between CA/VSC and Folksam Data.

CA/VSC DATA: Raw Score	CA/VSC DATA: Corrected Score	CA/VSC DATA: Vehicle Weight	FOLKSAM DATA: Injury Risk	FOLKSAM DATA: Injury Severity	RANKED ORDER: Corrected Score	RANKED ORDER: Vehicle Weight	RANKED ORDER: Injury Risk	RANKED ORDER: Injury Severity
210.1	220.8	1250	1.19	0.0536	4	4	12	7
130.8	114.2	900	1.24	0.0508	22	13	14.5	4
182.7	161.3	920	1.10	0.0517	13	12	10	5
157.0	121.8	710	1.82	0.0983	21	21	20	20
172.9	148.2	870	1.38	0.0897	17	15	17	19
166.6	133.0	755	1.63	0.0717	18	19.5	19	17
151.5	157.7	1230	0.98	0.0706	15	5	6	15
174.7	169.8	1095	0.99	0.0535	11	8	7	6
165.8	157.8	1055	1.23	0.0603	14	9	13	12
263.6	261.0	1130	1.02	0.0745	2	7	8	18
172.2	130.6	676	2.57	0.2390	20	22	22	22
191.6	178.4	1015	0.87	0.0487	8	10	4	3
213.1	174.9	799	1.55	0.0573	9	17	18	9
190.8	165.5	890	1.32	0.0686	12	14	16	14
165.4	132.1	755	1.24	0.0558	19	19.5	14.5	8
212.7	211.6	1140	0.94	0.0714	6	6	5	16
251.3	271.2	1305	0.64	0.0186	1	2	1	1
201.2	184.1	983	1.12	0.0683	7	11	11	13
194.8	214.5	1349	0.85	0.0578	5	1	3	10
213.6	226.6	1269	0.80	0.0440	3	3	2	2
185.4	156.7	845	1.08	0.0583	16	16	9	11
212.5	170.8	765	2.14	0.1155	10	18	21	21

The variables used in the analysis can be explained briefly as follows:

RAW SCORE - A measure from the CA/VSC system of the sophistication of secondary safety design. Vehicle weight is not taken into account. A higher Raw Score indicates better safety design.

CORRECTED SCORE - The measure of secondary safety from the CA/VSC system which includes the correction for the vehicle weight. A higher Corrected Score indicates better predicted secondary safety performance.

VEHICLE WEIGHT - The kerb weight for the particular UK model inspected in kilograms.

RISK OF INJURY (R) - A measure of the risk of sustaining injury of any severity derived from the Folksam system based on the paired comparison. The ratio between the number of injured drivers and passengers (over 18 years of age) in the model under consideration is related to the number of injured occupants in the other vehicle thus standardizing for accident severity and exposure. A higher figure indicates poorer secondary safety performance.

RISK OF DEATH OR DISABILITY (INJURY SEVERITY - Z) - A combination of R and the predicted number of fatal and disabling injuries based on the AIS assessment of the injuries. A higher figure indicates poorer secondary safety performance.

RANKED ORDER - The position in which a given model of vehicle appears in an ordered list of all 22 vehicles considered. The lists are arranged so that rank order 1 corresponds to the vehicle with the highest secondary safety and rank order 22 corresponds to the lowest. In the case of vehicle weight, the heaviest vehicle is given rank order 1 down to 22 for the lowest weight vehicle.

It should be noted that within the CA/VSC system higher scores are associated with vehicles having better secondary safety. Within Folksam's data, higher numerical values of R and Z are associated with greater risk of injury, and thus reduced levels of secondary safety.

RESULTS

Published Data

The most basic comparison apparent to the consumer is whether the published advice given by the two systems is comparable. Table 3 shows the currently available published results. It will be seen that a subset of 17 of the 22 models considered in the main analysis have been the subject of published information (references 3 and 4). It should be noted that while the Folksam published data is based on driver injury experience only, the more recent unpublished data used for the bulk of the analysis which follows includes data for drivers and adult passengers.

Table 3: Comparison of Published Information from the CA/VSC and Folksam Systems

CA/VSC STAR RATING	FOLKSAM RESULTS BETTER THAN AVERAGE	AVERAGE	WORSE THAN AVERAGE
8	Saab 9000		
7	Audi 100 Saab 900 Volvo 240 Volvo 700		
6	Opel Ascona	Opel Corsa Volvo 300	
5	Fiat Regatta Ford Sierra	Ford Escort Mazda 626 Opel Kadett VW Golf VW Polo	
4		Ford Fiesta	Nissan Micra

It is clear that there is some consistency between the two systems. All the 8 and 7 Star cars in the CA/VSC system are classified in the Folksam Best group. No 6 or 5 Star vehicles appear in the Folksam Worst group and no 4 Star vehicles appear in the Folksam Best classification.

Ranked Order Comparisons

Throughout this section, Spearman's Rank correlation method has been used to examine the data ranked on the basis of different measures available within the two rating systems. For reference, with 22 observations and 20 degrees of freedom, using a one-tailed test, a significance level of 5% is indicated by a coefficient of 0.3608 and a level of 1% is indicated by a coefficient of 0.4975.

There is a highly significant correlation between the CA/VSC Corrected Score and the Risk of Injury measure (R) from the Folksam system ($r = 0.6643$, significant at the 1% level). It is also clear that vehicle weight and Folksam R score are very highly correlated ($r = 0.8727$), significant at the 0.1% level). However, vehicle weight is itself correlated with the measure of sophistication of secondary safety design available within the CA/VSC rating system, CA/VSC Raw Score ($r = 0.4605$, significant at the 5% level).

The relationship between vehicle weight and Raw Score is a reflection of the current situation in which the larger, more expensive, and heavier cars tend to be those with the best secondary safety design. This does not have to be the case, and indeed it could be argued that in the future, if cars are downsized to reduce their environmental impact, lighter cars will need more sophisticated secondary safety design to compensate for the increased risk of injury that would be associated with a simple reduction in mass. However, at present, the fact that heavier cars tend to have better secondary safety design is important in understanding how to interpret analyses of what appear superficially to be comparisons of vehicle weight and safety. Positive influences attributed to vehicle weight should be viewed more correctly as benefits from the effects of both increased vehicle mass and improved secondary safety design. This point has been overlooked by many studies in the past, and leads to an over-estimate of the benefits of increased vehicle mass.

Regression analyses

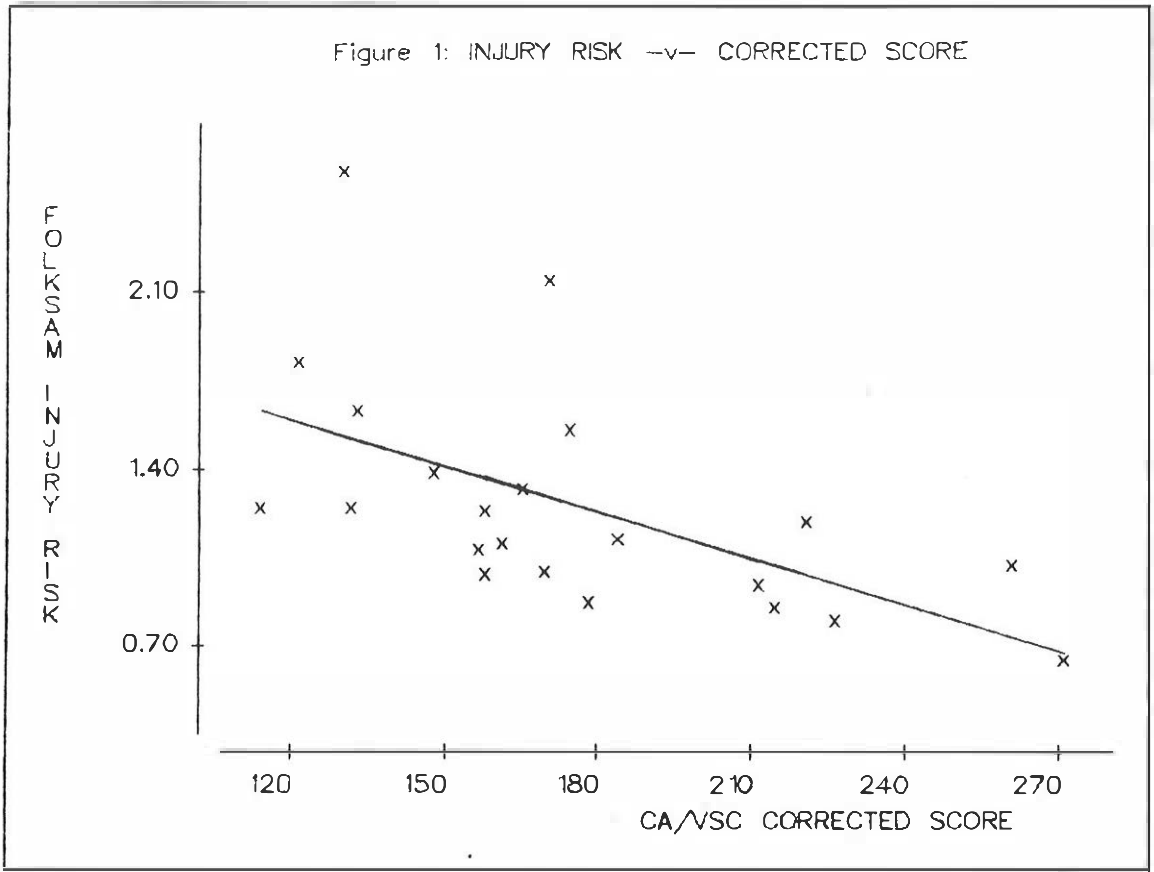
A further way to examine the data is by using regression analyses on the data. Again for reference, using a one-tailed test, a significance level of 5% is indicated by a |correlation coefficient| > 0.3598 and significance at 1% is shown by a |coefficient| > 0.4921 . Because the CA/VSC system predicts secondary safety some years before results are available for similar models from the Folksam system, the regression models use the CA/VSC data as independent variables to predict the Folksam results as dependent variable.

The CA/VSC Corrected Score is again highly correlated with Folksam's risk of injury measure (R) as illustrated in figure 1 (r = -0.5727, significant at the 1% level). The CA/VSC Corrected Score is also significantly correlated with Folksam's risk of death and disability measure (Z) (r = -0.3813, significant at the 5% level). There is a consistent tendency for CA/VSC Corrected Score to be better correlated with Folksam's R score than the Z score.

The regression analysis also confirms the significant linking between vehicle weight and quality of secondary safety design as measured by the CA/VSC Raw Score (r = 0.4299, significant at the 5% level). To explore this point a little further, all the data from the CA/VSC system has been used to look at the relation between CA/VSC Raw Score and weight for a group of 173 vehicles. Regression analysis confirms the trend apparent with the 22 vehicles used in the bulk of the current paper (r = 0.555, significant at the 0.1% level).

DISCUSSION OF RESULTS

It can be seen that the degree of correlation established was generally higher in those analyses that compare predictive results with the R factor, risk of injury, than the Z factor, severity of injury. The Z factor, which gives equal weight to injuries that result in 10% or greater disability and death, treats the injury outcome in a particular way. Injuries such as neck strain are given considerable emphasis, due to the high incidence of disability following such trauma. On the other hand, injuries such as rib fractures or splenic wounds, if they do not result in death, tend to be ignored as they rarely contribute to disability. This skewing of the treatment of injury data is quite acceptable, but reflects a certain way of judging the outcome. The CA/VSC predictive system skews the treatment of injury data



in a different way. In this instance, only serious and fatal injuries are considered, and considerable weight is given to measures capable of reducing the risk of death. Again this is acceptable, but of necessity gives a different view of a given set of injury outcomes. Ideally, a single measure of secondary safety would be desirable, but this is a complex area where no single approach necessarily has all the merit.

It is also important to recall that the Folksam system primarily seeks to comment on car performance in car to car accidents. One would expect vehicle mass to have a greater influence on this subset of accidents than in all accidents, including car to rigid object and car to heavy goods vehicle impacts. The term "vehicle mass" is used here to differentiate it from "vehicle weight" which has been discussed previously as including secondary safety design sophistication influences.

Because "vehicle weight" has been shown to include both true "vehicle mass" effects and some contribution of secondary safety design quality, regression models using both vehicle weight and Raw Score will give undue emphasis to vehicle weight. Further analysis will be undertaken in the future to try and establish the true "vehicle mass" effects, without any contribution from secondary safety design.

SCOPE FOR FUTURE DEVELOPMENTS OF THE CA/VSC AND FOLKSAM SYSTEMS

Both systems are capable of further improvement. For example, the predictive CA/VSC system in its present form is relatively poor at discriminating between bodyshells with differing abilities to control passenger compartment intrusion in frontal impacts. The addition of an accident-like whole vehicle frontal impact test could improve this aspect. In other areas, simple destructive component tests may be helpful, and some progress is already being made in this field. Within the Folksam system, important developments are in hand to improve the quality of vehicle data collected. This has the potential for greatly enhancing the understanding of sources of injury, and may help to provide explanations for differences which are observed. In the future, transfer of information between the two rating systems could allow further refinement of the predictive approach. The ability to monitor the relationship between predicted and actual outcome is seen as most important.

There is a clear responsibility on those who would provide consumer advice in this area to refine and monitor the reliability and relevance of their rating criteria. This is particularly important in view of the rapidly rising interest amongst both consumers and vehicle manufacturers in safety related design.

CONCLUSIONS

1) There are significant correlations between the outputs of the CA/VSC Secondary Safety Rating System and the Folksam Rating method. These correlations are apparent in both the simplified information published for consumer guidance, and within the more detailed measures available within both systems.

2) There are significant correlations between vehicle weight and the outputs of the Folksam Rating Method.

3) There are significant correlations between vehicle weight and the sophistication of secondary safety design as measured by the CA/VSC Raw Score. This is a feature of the current car market structure rather than an absolute and necessary relationship. Studies that overlook this relationship are at risk of over-stating the true benefits associated with increasing vehicle mass.

4) There are possibilities for the predictive and retrospective rating systems used by CA/VSC and Folksam respectively to be developed further. Future co-operation will benefit both approaches, and lead to refinements in the guidance provided to consumers.

REFERENCES

1. THE SECONDARY SAFETY RATING SYSTEM FOR CARS - First Revision. The Association for Consumer Research, London. November 1989. ISBN 0 85202 4290.

2. FOLKSAM CAR MODEL SAFETY RATING 1989-1990. Folksam Research, Stockholm. 1989

3. GUIDE TO NEW AND USED CARS - WHICH? The Consumers' Association, London, June 1991

4. SAFE AND DANGEROUS CARS 1989-90. Folksam Research, Stockholm, 1990.