TITLE: A framework for assessing the relative effectiveness of vehicle crashworthiness measures through data simulation.

AUTHORS: Amanda Delaney, Stuart Newstead and Max Cameron
Monash University Accident Research Centre

ABSTRACT

The measurement of relative vehicle occupant injury risk and injury severity is useful for assessing the safety performance of particular vehicle models and for monitoring design improvements in vehicles over time. Internationally, five key measures of vehicle crashworthiness have been developed that attempt to measure the risk of injury or serious injury to a vehicle driver involved in a two-car collision. This paper aims to extend existing theoretical comparisons of the safety measures by examining the relative effectiveness of each measure using simulated crash data with known properties.

A theoretical framework is developed that describes the process of injury data generation in terms of injury risk and injury severity. Application of the safety rating methods to simulated crash data derived from this framework enables quantification of the ability of each rating method to accurately estimate known vehicle safety performance. This is not possible through a theoretical statistical comparison alone. Multiple simulation runs have been conducted to investigate the effects of different relationships between injury risk and impact severity on the effectiveness of the safety rating measures. Results of the analysis highlight the efficiencies, deficiencies and biases of each of the rating methods under various input scenarios.

INTRODUCTION

The measurement of relative vehicle occupant injury risk and injury severity is useful for assessing the safety performance of particular vehicle models and for monitoring design improvements in vehicles over time. Internationally, five key measures of vehicle crashworthiness have been developed that attempt to measure the risk of injury or serious injury to a vehicle driver involved in a two-car collision. However, the formulation of the rating measures and the complexity of the injury data generation process limit the ability to quantitatively assess the effectiveness of each rating method based on theoretical models. Therefore, it has been necessary to extend existing theoretical comparisons of the safety measures by examining the relative effectiveness of each measure using simulated crash data with known properties. This paper describes the development of a theoretical framework to describe the process of injury data generation and the simulation of crash data derived from it. Each of the safety rating methods are applied to the simulated data and the results of the analysis are then presented and used to highlight the efficiencies, deficiencies and biases of each of the rating methods under various input scenarios.

THE PROCESS OF INJURY DATA GENERATION

Injury Risk and Crash Severity

There is a clear relationship between injury risk to human vehicle occupants and impact severity in a crash. This relationship is commonly formalised in the literature as a probability of injury to a vehicle occupant as a function of the crash impact severity. Regardless of the impact severity measures or injury scale used, differential performance between vehicles in protecting their occupants in the event of a crash is reflected in different relationships between the probability of injury and crash impact severity. This concept is illustrated for two different hypothetical vehicle models in Figures 1 and 2 below.

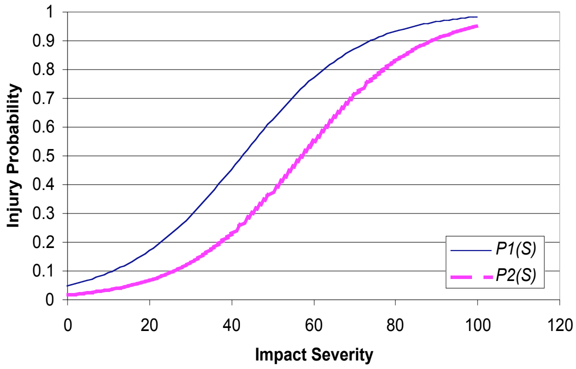


Figure 1. A hypothetical relationship between probability of injury and crash impact severity.

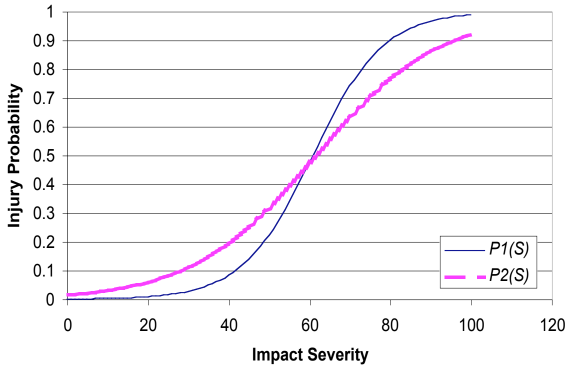


Figure 2. Alternative relationships between the probability of injury and crash impact severity.

Both figures show hypothetical injury probability curves, $p_1(S)$ and $p_2(S)$, as a function of crash impact severity for two vehicles. As would be expected, the probability of injury increases monotonically for each vehicle with increasing crash impact severity. However, it is clear from these figures that an injury probability curve will be unique to each vehicle model with no necessarily specific relationship between the curves for different vehicle models apart from each being monotonic. For example, in Figure 1 the occupant of vehicle 1 has a proportionately lower risk of injury than the occupant of vehicle 2 across all impact severities. In contrast, in Figure 2, the relative injury probability between the occupants of vehicle 1 and 2 is dependent upon crash impact severity. Here, the occupants of vehicle 2 have higher probability of injury than those of vehicle at crash impact severities up to 60, and relatively lower injury probabilities at impact severities above 60. Figure 2 may reflect reality if, for example, vehicle 2 were better optimized in terms of design to protect its occupants in high impact severity crashes than in low impact severity crashes relative to vehicle 1.

In addition to impact severity, injury risk in a crash is determined by factors that are beyond the design of the vehicle itself. These include characteristics of the occupant whose injury outcome is being assessed, such as occupant age, gender, height and weight. They also include characteristics of the crash such as impact point on the vehicle and collision partner (for example another vehicle or fixed object, type of other vehicle or fixed object). Seating position in the vehicle will also determine injury probability. However, in all that follows only one vehicle occupant is considered from the same seating position. In practice this is typically the driver.

To denote the dependency of the injury probability curves on impact severity as well as, say, c other non-vehicle related factors, let $x_1, x_2, \ldots, x_c$ be measures of the other factors. The injury probability function for vehicle model, m, is denoted as

$$p_m(s, x_1, x_2, \ldots, x_c) \ldots \text{(Eqn.1)}$$

This form will be used below and, for the moment, no specific assumptions are made about the specific functional form of the function given in equation 1.

The Distribution of Crashes by Severity and Other Factors

The distribution of crash frequency by injury severity is another important consideration in understanding how observed crash injury data arise. Figure 3 shows hypothetical probability distributions of crashes by crash impact severity for two different vehicle models. The two probability distribution functions, $f_1(S)$ and $f_2(S)$, show the proportion of crashes occurring at each crash severity for each of the two vehicles. Vehicle 1 has a higher proportion of crashes at lower impact severities whilst vehicle 2 has a higher proportion of crashes at higher impact severities. In practice, these types of distributions might arise if, for example, vehicle 1 has higher exposure and hence more crashes in low speed environments (e.g. densely populated urban areas) whilst vehicle 2 has higher exposure and hence more crashes on high-speed open roads.

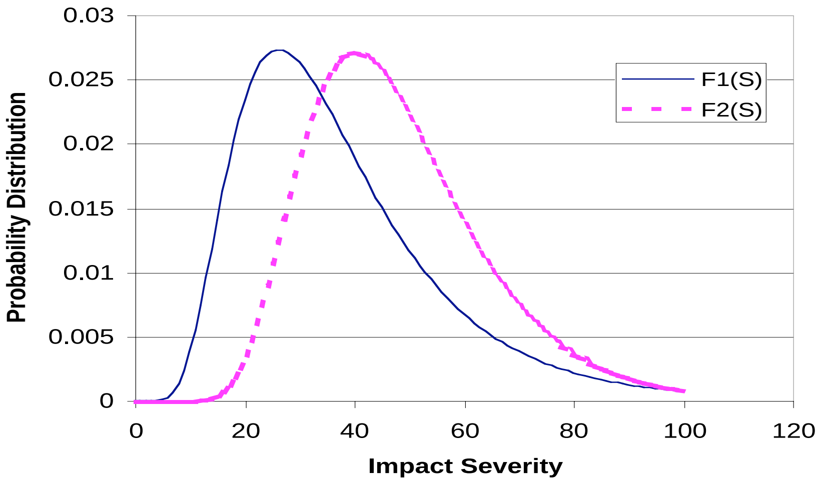


Figure 3. Probability distribution function of crashes by injury severity

It has been noted above that the injury risk in a crash is determined by factors in addition to crash impact severity that are beyond the design of the vehicle itself. It is therefore relevant to extend the definition of the probability distribution function for crashes to reflect the distribution of crashes not only by severity but also on the other factors determining injury probability. As above, for the c non-vehicle related factors, let $x_1, x_2, \ldots, x_c$ be measures of the other factors. The probability distribution function of crashes for vehicle model, m, jointly dependent on impact severity and the c factors is then denoted

$$f_m(s, x_1, x_2, \ldots, x_c) \ldots \text{(Eqn.2)}$$

As in the univariate case, the integral of the extended multivariate probability distribution function across its domain is equal to unity. That is,

$$\int_s \int_{x_1} \cdots \int_{x_c} f_m(s, x_1, \ldots, x_c)\, ds\, dx_1 \ldots dx_c = 1 \ldots \text{(Eqn.3)}$$

The integrals associated with each factor are taken across all possible values of the factor.

Ideal Vehicle Safety Rating Measures

In understanding the mechanism of crash injury data generation, it is possible to propose a vehicle safety rating system for consumer information that measures relative vehicle safety. The most comprehensive rating information that could be presented would be an estimate of the full curve *p(s)* for each vehicle model. However, in reality, it is unlikely that consumers would be able to readily interpret injury probability curves in making decisions about a safe vehicle to purchase. Further, if confronted with the curves given in Figure 2 it would be difficult to make an assessment of the relative vehicle safety. Clearly, the comparison would have to be made in light of the likelihood of crashing at various levels of impact severity. Therefore, a single summary statistic that encapsulates the "average" safety performance of each vehicle is proposed to enable consumers to rank or classify each vehicle relative to the others.

The summary statistic is a weighted average injury probability, p_{WA}, defined as:

$$p_{WA} = \int_{s=0}^{\infty} p(s) f(s) ds \ \ldots \text{(Eqn.4)}$$

This statistic has the advantage of not requiring an arbitrary choice of integral bound and being of finite value because of the properties of *f(s)*. Because *f(s)* is possibly different for each vehicle model, to make comparison between each vehicle model a function of only the injury probability curve of the vehicle *p(s)*, a standardised form of *f(s)* to be used in calculating p_{WA} for each vehicle model could be chosen. The most logical choice for a standardised *f(s)* would be the average across all vehicles appearing in the available crash data. If the relative injury probability curves between vehicles looked like those in Figure 1, the ultimate choice of f(s) would probably have little bearing on the relative average rating, p_{WA}, estimated between vehicles. In contrast, if the injury probability curves look like those in Figure 2, the relativity of the resulting ratings will be a strong reflection of the vehicle fleet from which the ratings are derived through determining the standardised form of *f(s)* used.

Extending the rating of Equation 4 to consider the *c* non-vehicle factors affecting injury outcome is relatively straightforward and would result in the weighted average injury probability in Equation 5.

$$p_{WA} = \int_S \int_{x_1} \cdots \int_{x_c} p(s, x_1, \ldots, x_c) f(s, x_1, \ldots, x_c) ds dx_1 \ldots dx_c \ \ldots \text{(Eqn.5)}$$

Again, to make the average ratings by vehicle model comparable on a common basis, a standard form of $f(s,x_1,\ldots,x_c)$ could be used for each vehicle. As before, this standard form could represent the joint average profile severity and non-vehicle factors represented in the available data. Similar comments about the relative vehicle ratings and the average profile chosen again apply.

The vehicle safety ratings to be assessed against the ideal safety rating presented above were developed by the following organizations: Monash University Accident Research Centre (Cameron et al., 1994 Newstead et al., 2004), Folksam Insurance, Sweden (Gustafsson et al., 1989), The Department for Transport, UK (Craggs and Wilding, 1995) and The University of Oulu, Finland (Huttula et al., 1997). The precise forms of the five measures are provided in Appendix A.

SIMULATION METHODOLOGY

Due to the complexity of both the data generation process discussed above as well as the formulation of many of the ratings methods, the theoretical comparison of each rating system based on the data generation framework is limited in the extent to which it can define the efficacy and accuracy of each of the rating systems. Therefore, data simulation, using the framework for injury data generation described above, is used to assess the ability of each rating system to capture the underlying mechanism of injury data generation. The simulation study has been carried out in a number of steps as follows.

- Set known forms of *p(s)* and *f(s)* for a range of hypothetical vehicle models. For reasons of simplicity five hypothetical vehicle models are used in each simulation scenario and the vehicle fleet is assumed to consist of these five hypothetical vehicle models only. The specific form of *p(s)* and *f(s)* chosen for each vehicle model and the relationship between the forms for each vehicle model are chosen explicitly to attempt to expose the potential biases noted from previous the theoretical analyses.

- Using the process of injury data generation detailed above, the primary determinants of injury outcome are used to simulate individual injury outcome data for each hypothetical vehicle model. The simulated data were generated for each of the scenarios using the simulation software Arena (version 3.01). In order to produce sufficiently reliable estimates of crash injury outcomes for each of the five vehicle models, 1001 crashes are generated within each simulation. Each simulation generates injury data for the occupant of the focus and opposing vehicles as shown in Table 1 below.

  Table 1. Outputs of simulation model.

	Opposing Vehicles	
Focus Vehicle	Injured	Not Injured
Injured	N_{iik}	N_{ink}
Not Injured	N_{nik}	N_{nnk}

  It is noted that within each simulation run the focus vehicle remains constant and the distribution of opposing vehicles is determined by the distribution of vehicle models within the vehicle fleet. Therefore, five simulations of each of the scenarios is run to obtain a table equivalent to Table 1 for each vehicle model as the focus vehicle.

- Each existing rating system is applied to the simulated output data to estimate ratings for each hypothetical vehicle model. Estimated ratings from each system are then compared to the ideal safety rating developed above. Definitions of the five rating systems applied are provided in Appendix A.

Multiple simulation runs have been carried out to investigate the effects of different relationships between the forms of the *p(s)* and *f(s)* curves generated for the hypothetical vehicle models.

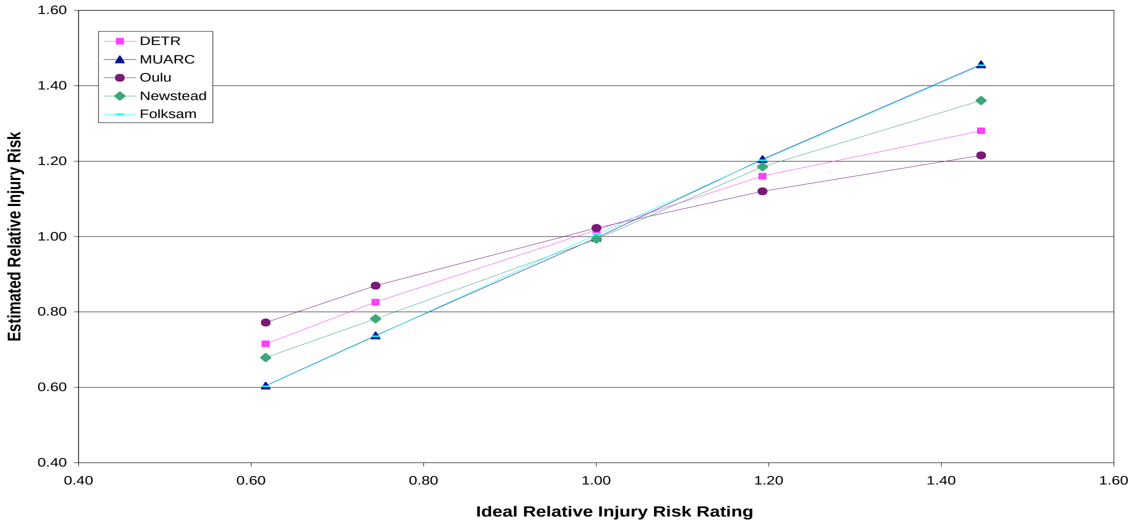
SCENARIO DEVELOPMENT AND RESULTS

Base Scenario

Under the base scenario, each of the five hypothetical vehicle models has an identical impact severity distribution derived from real crash data in Australia (Fildes et al, 1991). The injury risk distribution differs across vehicles with the probability of injury increasing monotonically for each vehicle with increasing impact severity. The risk of injury is proportional across the five vehicle models (as shown in Figure 1 for the case of 2 vehicle models) with the risk of injury being greatest for vehicle model 1 and declining for each subsequent vehicle model. It is assumed that injury risk is unaffected by the impact partner.

After running the simulation model with the inputs described above, relative injury risk values were estimated using each of the rating methods. Figure 4 below plots these estimated measures of injury risk against the ideal injury risk rating defined in Equation 4. A well performing safety rating measure will show strong correspondence between the ideal safety rating and each of the estimated safety ratings.

Figure 4. Estimated relative injury risk measures plotted against the ideal relative injury risk measure for vehicles one to five under the base scenario.



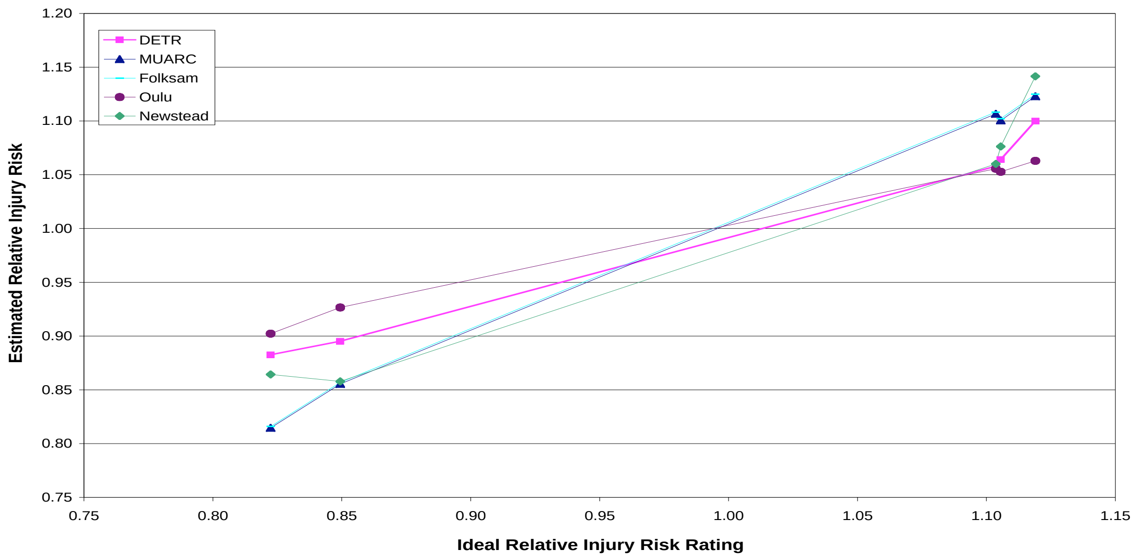
It is evident that all of the safety rating measures are able to differentiate the safety performance of each of the five vehicles. There is relatively strong correspondence between the ideal safety rating and each of the estimated ratings although slightly greater differentiation between vehicle models is achieved with the Folksam and MUARC methods.

Altered Injury Risk Scenario

Under this scenario five individual vehicle models were again generated, each representing an equal proportion of the total vehicle fleet. Each vehicle has an identical impact severity distribution as in the base scenario. The injury risk distribution differs across vehicles and the relationship between the five curves is not proportional (as shown in Figure 2 for the case of 2 vehicle models). This is the primary distinction between this scenario and the base scenario. It now becomes more difficult to assess which of the five vehicles has the best safety performance from visual inspection of the injury risk curves. None of the five vehicle models performs consistently better or worse across the range of impact severities than the other vehicle models. Therefore, the ultimate ordering of the five vehicle models in terms of safety will depend heavily on the distribution of impact severities at which the vehicles crash. Further, the effectiveness of the five safety rating measures will depend upon their ability to correctly capture the relationship between injury risk and impact severity.

Relative injury risk values were again estimated using each of the rating methods after running the simulation model. These estimated measures of injury risk are plotted against the ideal ratings and the results are shown in Figure 5 below. Each of the methods is able to differentiate between each of the five vehicle models. However, the ideal rating of vehicles one, two and three are very similar and these vehicles were not rated in the correct order by all of the rating methods. Given, the clustering of these three vehicle models, it is not unexpected that the estimated ratings deviate in a minor way from the ideal safety ratings. In practice, the confidence limits surrounding the point estimates of these three clustered vehicles ratings, derived from a logistic regression analysis, would be likely to overlap and thus no real distinction could be made.

Figure 5. Estimated relative injury risk measures plotted against the ideal relative injury risk measure for vehicles one to five under the altered risk scenario.

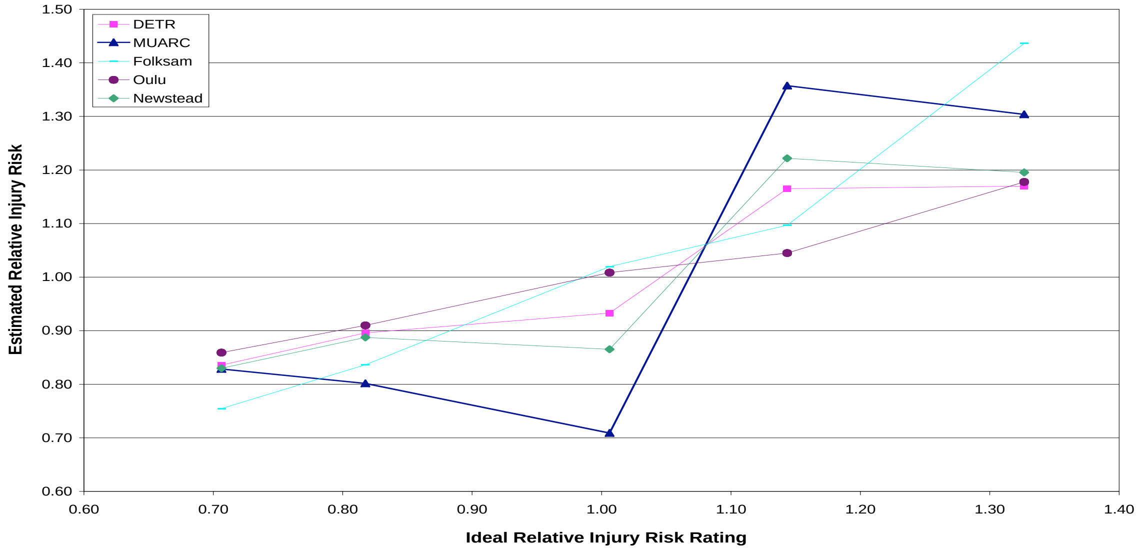


Varied Impact Severity Scenario

Under the previous two simulation scenarios the impact severity distribution of each of the five vehicle types was identical. However, in reality it is likely that vehicles of different types will be driven in different road environments and therefore the distribution of impact severities is unlikely to be identical across all vehicles. To model this, this simulation scenario applies different impact severity distributions for each of the five vehicle models. The injury risk distributions used are those used under the base scenario. That is, the relative performance of each of the five vehicles is consistent across the entire range of impact severities. Further, under this scenario the ideal safety rating is calculated as a weighted average of injury risk according to Equation 1. However, in contrast to previous scenarios, it was necessary to select a single impact severity distribution to use in the calculation. As each vehicle had its own impact severity curve, an average of the five impact severity distributions was used in the calculation of the ideal safety rating.

The relative injury risk values calculated from the simulation output are plotted against the ideal safety measure in Figure 6 below. The relative consistency between the estimated injury risk ratings and ideal safety ratings displayed in the previous two scenarios are not evident under this scenario. A possible exception to this is the rating estimated using the Folksam method, which appears to control well for variations in impact severity across vehicle models. It is noted, that the estimates calculated here are unadjusted for factors other than vehicle model, which may influence safety performance and act as proxies for impact severity. In practice such adjustments are made when these methods are used to estimate injury risk. Further simulations which account for these factors would need to be run in order to accurately assess the effectiveness of these measures as they operate in practice.

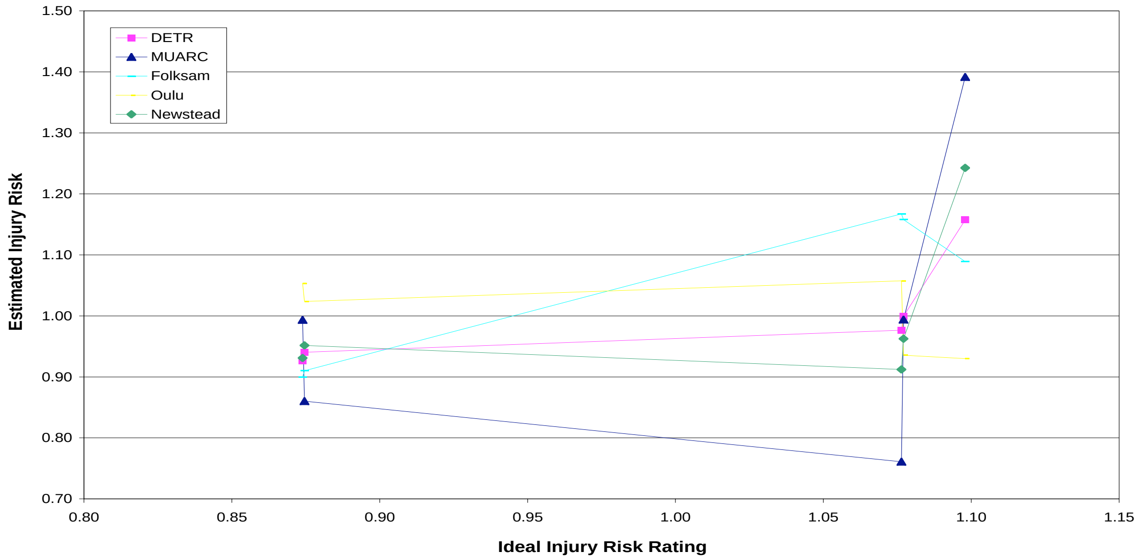
Figure 6. Estimated relative injury risk measures plotted against the ideal relative injury risk measure for vehicles one to five under the varied impact severity scenario.



Varied Impact Severity and Altered Risk Scenario

To this point either the injury risk or impact severity distributions of the five vehicle models have been adjusted away from the base scenario to assess the ability of the different rating methods to account for these complexities. Under this scenario both the impact severity and injury risk distributions are altered. This represents a complete departure from the base scenario. Further, in calculating the ideal safety rating under this scenario a single impact severity distribution, representing an average of the five impact severity distributions, was used given that each vehicle had a unique impact severity distribution.

Figure 7. Estimated relative injury risk measures plotted against the ideal relative injury risk measure for vehicles one to five under the varied impact severity and altered risk scenario.



The plot of the relative injury risk values calculated from the simulation output against the ideal safety measure (Figure 7) reveals that none of the methods employed to estimate injury risk accurately reflected the ideal safety rating when the impact severity distributions differ across vehicles and the injury risk curves for each of the vehicles are not proportional.

Aggressivity Base Scenario

The foundation for this scenario is the base scenario presented above. The impact severity distribution is identical across vehicles and there is a proportional relationship between the injury risk distributions of each of the five vehicle models. In addition however, each vehicle is assigned an aggressivity parameter intended to represent the relative risk posed by the focus vehicle to the driver of an opposing vehicle. This parameter acts as a multiplier on the injury risk curve of the vehicle with which the focus vehicle collides. The exact values of the aggressivity parameters are shown in Table 2 below.

Table 1. Aggressivity parameters for vehicles one to five.

Vehicle Type	Aggressivity Parameter
Car 1	0.7
Car 2	1.5
Car 3	0.95
Car 4	1.2
Car 5	0.83

An aggressivity parameter of 1 indicates neutral aggressivity whilst a value greater than one indicates increased aggressivity (i.e. the vehicle is more likely to cause injury to the occupant of the colliding vehicle) and a value less than one indicates a relatively less aggressive vehicle (i.e. the vehicle is less likely to cause injury to the occupant of the colliding vehicle). The inclusion of an aggressivity parameter does not affect the calculation of the ideal safety rating as a well performing safety rating will measure the performance of a vehicle independently of factors other than vehicle design including the collision partner. Therefore, the ideal safety ratings calculated under the aggressivity base scenario are identical to those calculated under the base scenario.

Figure 8. Estimated relative injury risk measures plotted against the ideal relative injury risk measure for vehicles one to five under the varied impact severity scenario.

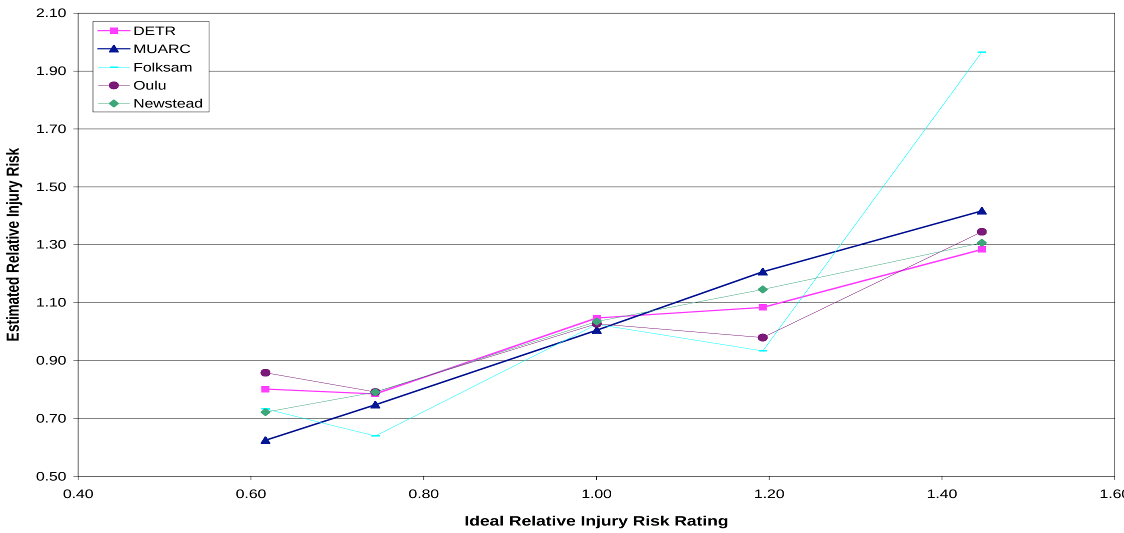


Figure 8 above plots the estimated measures of injury risk derived from the simulation output against the ideal ratings calculated from the input data. The effectiveness of the five safety ratings methods is varied under this scenario. Both the MUARC and Newstead methods are able to correctly differentiate between the five vehicle models and display the strongest correlation with the ideal relative injury risk ratings. However, there is some weakness in the other methods when the aggressivity of the vehicle is included in the scenario.

DISCUSSION AND CONCLUSION

This paper has developed a framework, based on simulation techniques, for objectively assessing the relative effectiveness of vehicle safety rating systems. The framework has been applied to some of the fundamental data generation mechanisms relevant to vehicle safety measures and provides additional information to that obtainable through theoretical comparisons alone. However, only a small number of scenarios have been considered here and the input data used is limited. Additional, more comprehensive analysis would provide additional insights into the effectiveness of vehicle safety rating systems. Nevertheless, some issues highlighted by the current analysis are worthy of discussion.

It is clear from the scenarios considered above that the effectiveness of the measures of vehicle crashworthiness in capturing the true crashworthiness of a vehicle is dependent upon the nature of the underlying injury data. The base scenario demonstrates the ability of each of the methods to accurately differentiate between five vehicles with known properties where the underlying injury data is uncomplicated. That is, where the probability of injury risk is proportionate across impact severities for each vehicle model and the impact severity distribution is identical across vehicle models. However, the five safety rating measures examined here are commonly applied to real world mass crash data sources where the underlying form of the input data is uncertain and unlikely to satisfy these conditions.

It is evident from the analysis that the effectiveness of each of the safety rating methods deteriorates when the above conditions are not met. Most particularly, when it is no longer assumed that the impact severity for each vehicle model is identical and variation is allowed across vehicle models, a number of the safety measures do not perform well. A possible exception to this is the Folksam rating method which appears to control well for variations in impact severity across vehicle models. The analysis highlights the need to adjust for variations in impact severity when applying the other methods in practice. However, further analysis would be required to assess the effectiveness of these adjustment procedures and the ability of factors other than vehicle model to successfully act as proxies for impact severity. Such adjustments may also aid in improving the ability of all safety ratings measures to accurately reflect the underlying injury data in circumstances where the injury risk curves are not proportional and impact severity distributions differ across vehicle models.

Finally, the analysis highlights the difficulty in separating the aggressivity of collision partners from the crashworthiness of the focus vehicle when using some methods. A well performing safety rating will measure the performance of a vehicle in protecting its occupants independently of factors other than vehicle design, including collision partner. The MUARC and Newstead methods perform well in this regard where the probability of injury risk is proportionate across impact severities for each vehicle model and the impact severity distribution is identical across vehicle models. Further investigation of the effectiveness of these measures under more complicated input conditions is warranted providing successful adjustment for variations in impact severity can be achieved.

REFERENCES

Cameron, M. H., Finch, C., & Le, T. (1994). *Vehicle crashworthiness ratings: Victoria and NSW crashes during 1987-92. Technical Report* (Vehicle safety No. Report No. 58). Melbourne, Australia: Monash University Accident Research Centre.

Craggs, A., & Wilding, P. (1995). *Cars: make and model: the risk of driver injury and car accident rates in Great Britain: 1993* (No. Transport Statistics Report). HMSO, London, U.K.: Department of Transport.

Fildes, B. N., Lane, J. C., Lenard, J., & Vulcan, A. P. (1991). *Passenger cars and occupant protection* (No. Report No. CR95). Canberra, Australia: Federal Office of Road Safety.

Gustafsson, H., Hagg, A., Krafft, M., Kullgren, A., Malmstedt, B., Nygren, A., et al. (1989). *Folksam car model safety rating 1989-90.*Stockholm, Sweden: Folksam Insurance.

Huttula, J., Pirtala, P., & Ernvall, T. (1997). *Car safety, aggressivity and accident involvement rates by car model 1997* (No. Publications 40). Oulu, Finland: University of Oulu.

Newstead, S., Delaney, A., Watson, L., Cameron, M., & Langwieder, K. (2005). *Injury Risk Assessment from Real World Injury Outcomes in European Crashes and Their Relationship to Euroncap Test Scores.* Paper presented at the Proceedings, 19th International Technical Conference on the Enhanced Safety of Vehicles (ESV), Washington, USA.

APPENDIX A

DETR Method: $$\frac{n_{ink} + n_{iik}}{n_{ink} + n_{iik} + n_{nik}}$$

Newstead Method: $$\frac{n_{iik}}{n_{iik} + n_{nik}}$$

Folksam Method: $$\frac{n_{ink} + n_{iik}}{n_{iik} + n_{nik}}$$

Oulu Method: $$\frac{n_{ink} + n_{iik}}{n_{ink} + 2n_{iik} + n_{nik}}$$

MUARC Method:

$$\frac{n_{ink} + n_{iik}}{n_{ink} + n_{iik} + n_{nik} + n_{nnk}}$$