THE ANALYSIS OF DUTCH NATIONAL DATA ON HEAVY TRUCK ACCIDENTS: THE NECESSARY EXTENSION OF TRADITIONAL FREQUENCY COUNTS WITH LOGISTIC REGRESSION ANALYSIS

Y.W.R. de Vries, H.G. Mooi,
TNO-Automotive
The Netherlands
Paper no. 153

ABSTRACT

For the 1993-1997 Dutch national accident data, logistic regression analysis was used to find the most important factors, that influenced the outcome of an accident with a truck involved. Frequency counts were used to identify factors that occurred most frequently. The combination of these two methods led to the most important factors influencing the number and severity of truck accidents. An important extension with respect to only frequency counts is that significance levels were taken into account to check whether differences are really distinguishable. It was concluded that the combination of frequency counts and logistic regression is a necessary extension to prevent the presentation of artefacts when basing conclusions solely on frequency counts and to find factors with high risk. Furthermore, it was found that national statistics are not detailed enough to find the real underlying causes and can therefore hardly be used to find points for improvement on vehicle and road design. Therefore, in-depth investigation of truck accidents is necessary to identify real causes.

INTRODUCTION

In 1995, heavy goods vehicles with a gross vehicle weight over 3500 kg (HGVs) were involved in about 15% of all serious accidents and in 30% of all fatal accidents in Europe[1]. In 1990 in the Netherlands, this was in about 25% of all fatal accidents[2]. Since then this number has decreased towards 16%, still involving 165 deaths[3]. HGVs make up about 12.5% of all registered vehicles in the European Community and their average mileage is between 3-5 times higher than that of cars[4]. Though the number of injuries per driven kilometre is smaller than that in passenger cars, the number of fatal accidents per driven kilometre is almost twice as high. Also the number of fatalities for the collision partner is 25 times higher than the number of fatalities for the HGV occupants. All this shows that HGVs are involved in an important percentage of the fatal accidents and it is worth looking at methods for improving this situation.
Comparison of international data[5,6,7,8] with the Dutch data to find if Dutch accident data differs from other countries, proves very difficult[3], due to the different categories used by the different countries. Also the collection methods, sampling and levels of registration differ greatly. Therefore accident information of other countries can not directly be used in the Netherlands and analysis of the Dutch accident data is necessary to find important variables for the causation and outcome of HGV accidents in the Netherlands.
Firstly the methods used for the analysis of the data are described. Next the results are presented, subdivided into frequency count results, regression analysis results and the combined analysis. Finally the discussion and conclusions are presented.

METHOD

For the analysis, 1993-1997 Dutch national accident data was used. This data is collected by the police and coded in the Traffic accident registration database (VOR: **V**erkeers**O**ngevallen **R**egistratie). Around 50 variables are coded, including environmental, personal, and technical information, in quite some detail
Only later in the study 1998 data became available. A short comparison of frequency distribution showed that no significant changes had occurred.

Statistical Methods

Two different kind of analysis were done: traditional frequency counts and a logistic regression analysis. Both analysis were done with Genstat©. For the detection of significant differences a 95% confidence interval is used.

Frequency counts were used to identify significant differences in occurrence. Frequency counts give important information about the number of occurrences. The frequency tables were analysed according to standard frequency table analysis methods. A Poisson distribution was assumed, to find significant differences in occurrence. The frequency counts were based on accidents with a killed (K) or hospitalised (Seriously Injured; SI) HGV or collision partner vehicle occupant, where the Gross Vehicle Weight (GVW) of the HGV was over 3500 kg. This resulted in 3121 accidents with a killed or seriously injured victim (KSI accident). The number of KSI accidents were taken into account, not the number of victims.
Trying to improve the situations where the number of severe accidents is highest will not necessarily

have a high impact on the total number of severe accidents. The high frequency situations also need to have influence on the outcome of an accident.

Logistic Regression can be used to predict the outcome of an accident based on values of a set of predictor variables. The influences of the predictor variables are corrected for each other, so that the *real* influence of a variable is found. Sometimes variables correlate very well (e.g. car mass and length). In that case only one of these variables is found to be of influence in the regression analysis. With this analysis, the categories which have the most influence on the outcome of an accident can be found. *The result of the logistic regression are categories which increase the risk of the HGV or collision partner occupant significantly (95% certainty)*. Thirty variables were included in the analysis, of which nine had significant influence on the probability to get killed or seriously injured. The other variables did not influence the outcome significantly. Due to memory problems only the five most influential variables were used to predict the outcome of an accident. Firstly the probability to get killed, with respect to the non-fatal accidents in the database, were analysed; secondly, the probability to get seriously injured (hospitalised) with respect to all non-fatal, non-hospitalised accidents. This was done separately for the truck occupants and the collision partner occupants. All vehicle types were included. It has to be noted that the presented values of the prediction of the probability are with respect to the cases in the database. Therefore the values can be rather high (see Table 4b: manoeuvre with a pedestrian, for which a probability of 57% to get seriously injured was found). This is because non-injury accidents were not taken into account. Absolute values can therefore only be used as an indication with respect to other values in that variable, but not as the absolute probability.

Statistically different values. The following example shows that even though a value is lower or higher than another value, it does not mean that this is really the case. In 1997, 1163 people were killed in a traffic accident. In 1996, this number was 1180. To find out if the number of killed people has decreased a Poisson distribution is assumed. Because of the large values the distribution may be approximated with a normally distributed curve with an expected value of μ= 1180 and a standard deviation of σ= $\sqrt{1180}$. The one sided 95% confidence interval equals to:

$$\mu_L = 1180 - 1.65 \cdot \sqrt{1180} = 1123 \qquad (1.)$$

This is lower than the number of killed in 1997. Therefore, it may be concluded that the number of killed has not reduced significantly. Not significant differences found in this study should be treated in the same way.

Under Representation

It is known that some under representation of accidents exists in the national accident data. The degree of representation is about 94% of the fatalities and 58-62% of the hospitalised victims[9]. The representation of non-hospitalised injured victims is even lower (approximately 16-18%). Per type of vehicle the representation is also different. The representation degree of seriously injured bicyclists is the lowest (around 36%), followed by pedestrians, HGVs and buses with approximately 60%. Mopeds and motorcycles are included in about 65-70% of the cases and about 80% of the cars and vans are registered. The data and outcomes are not corrected for these under representations.

Data Sanitation

For the analysis described in this paper, HGVs and their collision partners and single HGV accidents are looked at. In this database the selection for collision partners can only be done by, what is coded as, primary colliders (the two objects that had the first collision). Therefore only accidents with a HGV that was a primary collider are taken into account. This means that some HGVs with injured occupants are not included in the analysis. The number of injured occupants (killed, seriously injured and slightly injured) is therefore not equal to the total number of injured HGV occupants for the years looked at. Only accidents with an injured person were selected, adding up to 8848 accidents.

Due to the fact that in HGV-HGV accidents (316), both of the HGVs can be either a HGV or a collision partner, an accident between two HGVs is counted as two accidents. By this selection the total number of HGV accidents adds up to 9164. (8973 primary impacts and 191 single impacts).

From this selection a coupling with the licence plate registration (RDW-data; Vehicle Technology and Information Centre) was made for the HGV accidents with a killed or a hospitalised person (KSI accident). Of the 3596 requested HGVs, 3259 could be coupled (±91%). Of these HGVs, 237 did not qualify as such, because the GVW of the vehicle was under 3500 kg and therefore not falling into the definition of HGV used in this paper. This resulted in 3121 HGV accidents for which RDW data was available. Many of the 237 vehicles are (large) vans.

It was decided for the general regression analysis to use all the 9164 accidents and not to remove the accidents with the vehicles known to be non-HGVs. This would cause a bias, because for the non-KSI accidents this data is not available. For the

detailed analysis (only KSI accidents) these accidents were removed, because for all accidents it is known what type of HGV was involved. The detailed analysis includes extra information about the HGV; among others the GVW, HGV make and horse power of the involved HGV.
The total number of accident and injuries over the years 1993-1997 (clean database) that were used for the analysis are shown in Table 1. In Table 2 the number of involved HGVs, collision partner vehicles and the number of injured occupants are shown.

Table 1.
The number of accident used in the different analysis

Accident numbers		
	Number of accidents	Used in
Total	9164	- General regression
KSI	3596	-
KSI with detailed information	3121	- Frequency counts - Detailed regression

Table 2.
Number of accidents and injuries over the years 1993-1997 in the 'clean database'

	Maxium severity			
	Fatal	Serious	Light	Total
# Accidents	848	2748	5568	9164
# HGVs	64	346	968	1378
# CPs	788	2460	4890	8138
# HGV occupants	66	362	1079	1507
# CP occupants	853	2775	5863	9491
# Occupants	919	3137	6942	10998

Updating Impact Locations

In the VOR-database the collision point is coded for eight locations (three on front and rear and one on each side). It is however possible to extend this to twelve impact locations (three on each side) in a high percentage of cases (96%), because accidents are also coded as frontal, side, etc. In the case an accident is, for example, a side impact accident and one of the vehicles is hit in the front and another at the front left, this last vehicle must be hit in the front left side in order to make this a side impact accident. In the case where both vehicles were hit on a location shown in Figure 1, the exact impact location can not be found. In these cases the impact location will remain the same, with a remark that the updated location could not be extracted.

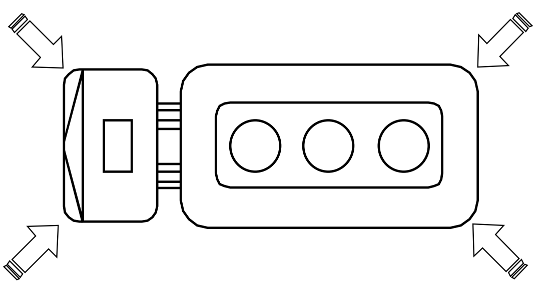

Figure 1 Impact location combinations, for which the exact location can not be found.

Year Of Admission On The Road

Although the used accident data is relatively new, the majority of the HGVs is from before 1994 (84%, see Figure 2). Therefore under run protections are hardly included in the accident set. Rear under run protections (RUPs) are required on new HGVs since 1996. Side under run protections (SUPs) are mandatory since 1994 on new HGVs, and front under run protections (FUPs) are not mandatory.

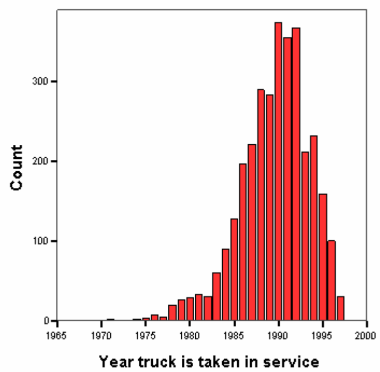

Figure 2 Number of HGVs in the database by year of admission (included are KSI accidents).

ANALYSIS RESULTS

In this study three main results were achieved. The results based solely on the two different analysis types and the results of the combined analysis. For the most important variables the results of the analysis are shown in the next paragraphs. For each variable the frequency counts are treated first, followed by the results of the regression analysis. Last, the combination of these two is presented.
The tables shown in the following paragraphs require some explanation. With typographical differences or crosses the significant groups are shown. Groups with the same typographical styles or with crosses in one column, can not be separated significantly. They therefore occur equally frequent or are equally dangerous, even though the numbers are not the same. For example: Table 5 should be read such, that a motorcyclist and a bicyclist are at equal risk, and significantly more dangerous than other vehicles.
In the heading is shown whether a variable was of influence on the probability to get killed (K) or seriously injured (SI).

Speed Limit (K)

In the Dutch national accident statistics, only the speed limit is coded. The collision speed, *delta v*,

or EES values are not included. Therefore, the variable closest to this is the maximum allowed speed. To reduce memory usage, this variable was included as a continues variable (not categorical) in the regression analysis.
An assumption is made that the actual collision speed is correlated with the speed limit.

Frequency Counts show a significant difference in accident frequency on the different road types (urban road = 0-70 km/h, secondary road =80-100 km/h and motorway = 120 km/h) and). Most KSI accidents occur on urban and secondary roads, followed by motorways. When looking at HGV occupants separately, it can be seen that most victims are found on secondary roads, followed by motorways.

Table 3.
The number of KSI accidents

Road types	#accidents with a KSI collision partner	#accidents with a KSI HGV occupant	# KSI accidents, with detailed information
motorway	*450*	*118*	*437*
secondary road	1312	212	1312
urban road	**1478**	**80**	1372
Total	3240	410	3121

Logistic regression showed that only for the probability to get killed for the collision partner occupant, a linear relationship could be found with the speed limit. For the number of seriously injured and for HGV occupants the speed limit was not a significant predictor.

Combined analysis shows that the number of killed collision partner occupants is significantly higher at lower speed limits, but that the speed limit is not an explanatory factor at low speeds for the number of killed victims. This means that other factors are involved in accidents at low speed limits, and that a high speed limit is an important predictor for the number of killed collision partner occupants. Even though HGV occupants are frequently involved in KSI accidents on motorways and secondary roads, this was not found to be an influencing variable on the outcome of an accident for the HGV occupant. An explanation might be that HGVs are more frequently found on secondary roads and motorways (exposure).

Manoeuvre (KSI)

In the variable *manoeuvre*, the movements of the vehicles with respect to each other are coded. Unfortunately in an accident with a pedestrian, not the movements are coded but that the accident was with a pedestrian. Because *pedestrian* is coded in more variables this leads to the problem of near multicollinearity.

Frequency counts show differences in frequency distribution on the three different road types. On urban roads, the manoeuvres *A: same road, same direction, with turning* (23%) and *B: two crossing roads, without turning* (22%) occur most frequent. On secondary roads, also the manoeuvre *B: two crossing roads, without turning* (22%) is most frequent. Motorways have a different distribution, with the highest occurrence of manoeuvre *D: same road, same direction, without turning* (58%).

Table 4.
Probabilities for the collision partner occupant to get a)killed or b)SI

a)	Manoeuvre	Prediction	Groups						
G:	Same road, opposing direction without turning	0.259	x						
H:	With pedestrian	0.240	x	x					
A:	Same road, same direction, with turning	0.204		x	x				
B:	Two crossing roads without turning	0.178			x	x			
C:	Two crossing roads with turning	0.141				x	x		
E:	Same road, opposing direction with turning	0.134					x	x	
I:	*With objects and animals*	*0.079*	-	-	-	-	-	-	-
F:	With parked vehicle	0.072						x	x
D:	Same road, same direction, without turning	0.050							x

b)	Manoeuvre	Prediction	Groups			
H:	With pedestrian	0.5695	x			
B:	Two crossing roads without turning	0.2221		x		
E:	Same road, opposing direction with turning	0.2187		x		
I:	*With objects and animals*	*0.2183*	-	-	-	-
G:	Same road, opposing direction without turning	0.2157		x		
C:	Two crossing roads with turning	0.1823			x	
A:	Same road, same direction, with turning	0.177			x	
F:	With parked vehicle	0.174			x	
D:	Same road, same direction, without turning	0.1275				x

Logistic regression results are shown in Table 4. X*'s* in one column of *Groups*, indicate that categories within that column are not significantly different. It can be seen that the manoeuvre *G: same road, opposing direction without turning* and *H: With pedestrian* have the highest predictive value for a collision partner to be killed, but can not be distinguished from each other. To get seriously injured the manoeuvre *H: with pedestrian* is the most dangerous. The manoeuvres that follow (*B*,*E*,*G*) do no differ significantly. Manoeuvre *I* is not significantly different from any manoeuvre due to the low number of cases. The manoeuvre that is least dangerous is *D: same road, same direction, without turning* (SI collision partner). For accidents with a killed collision partner this manoeuvre can not be separated significantly from *F: with a parked vehicle.*

Combined analysis shows that the two highest frequency manoeuvres (*A* and *B*) have relatively much predictive value for the probability to get killed, but for the probability to get seriously injured they do not differ significantly from other manoeuvres. The manoeuvre *H: with pedestrian* is highly predictive for the collision partner to get seriously injured or killed. The frequency of occurrence of KSI accident with pedestrians are thus very well explained by the severity of these accidents, and not necessarily by the high frequency of occurrence of this type of accident (including slight injuries).
Manoeuvre *D*, which is most frequent on motorways, has a very low predictive value to get killed or seriously injured. This manoeuvre is most

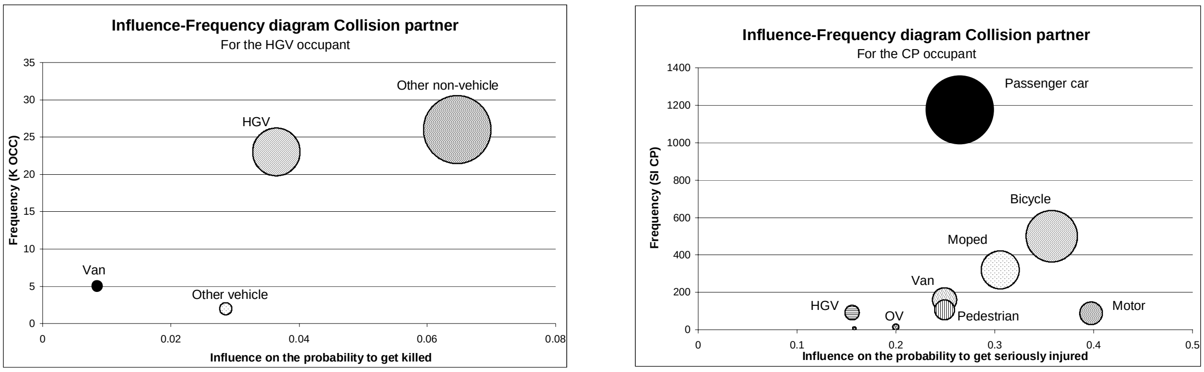

Figure 3 Influence-frequency diagram for the HGV occupant on the probability to get killed (left) and for the collision partner occupant on the probability to get seriously injured (right). The bubble size is equal to the frequency x probability.

likely better explained by the fact that it occurs on motorways, which have a high speed limit. However, this is not investigated further.

Type Of Collision Partner (KSI)

In this variable the types of collision partners are coded.

Frequency counts show a difference in distribution over the road types. On urban roads *bicycles* (38%) are included most frequently in KSI accidents with an HGV, followed by passenger cars (24%). On secondary roads and motorways passenger cars are involved most frequently, respectively 57% and 66%.

When looking only at the killed and seriously injured HGV occupants, the most frequent collision partner is an *other non-vehicle (e.g. house, pole, tree, etc)* (36%). Second most frequent is another HGV (28%) involved.

Table 5.
Probabilities for the collision partner occupant to get seriously injured by vehicle type

Type CP	Prediction	Group				
Motorcycle	**0.39751**	x				
Bicycle	**0.35774**	x				
Moped	*0.30551*		x			
Passenger car	0.26477			x		
Van	0.2493			x		
Pedestrians	0.2493			x		
Other vehicles	0.20004			x	x	
Bus	0.158			x	x	
HGV	0.15575				x	
Other non-vehicles	0.00114					x

Logistic regression shows that for both the HGV occupant and the collision partner occupant the type of collision partner is of influence on the probability to get seriously injured. For the HGV occupant it is also a predictive variable for the probability to get killed.

Motorcyclists and bicyclists have the highest probability to get seriously injured, followed by moped riders (see Table 5). Others are at equal risk (pedestrians do not have a high predictive value, but are already included in the manoeuvre), except for other HGV occupants which are at less risk.

For the HGV occupant *other non-vehicles* are the most dangerous, followed by another *HGV* as collision partner.

Combined analysis shows that the number of seriously injured vulnerable road users are well explained by the fact that they are vulnerable. Therefore accidents with these road users do not necessarily occur more frequent than accidents with other road users, but have a higher probability to end with a high severity. The high number of injured passenger car occupants however, is not explained by being in a car. A car is not significantly more dangerous than other vehicles. Other factors must be of influence for car occupants (see also Figure 3, right).

For the HGV occupant both the KSI accident frequency and the probability to get killed or seriously injured are highest for *other non-vehicles*, followed by another *HGV* as collision partner. The collision partner type was the *only* significant influential variable in the database for the HGV occupant. With respect to the average risk of all other variables a collision with an *non-vehicle (object)* is 7 times more than an accident with a *Van*; an accident with another *HGV* is 4 times more dangerous than an accident with a *Van* (see also Figure 3, left).

Collision Point On The HGV (K)

In Figure 4 the frequency of occurrence of KSI accidents is shown for each impact location on the HGV. The colours show the significance levels. Bars with the same colour do not differ significantly from each other. Bars with different colours do differ significantly from each other. Bars with a gradient colour are not statistically significant different from either of the two colour groups; e.g. on a secondary road, the left side of the HGV is hit equally frequent as (is not statistically different from) the left front, as well as the right

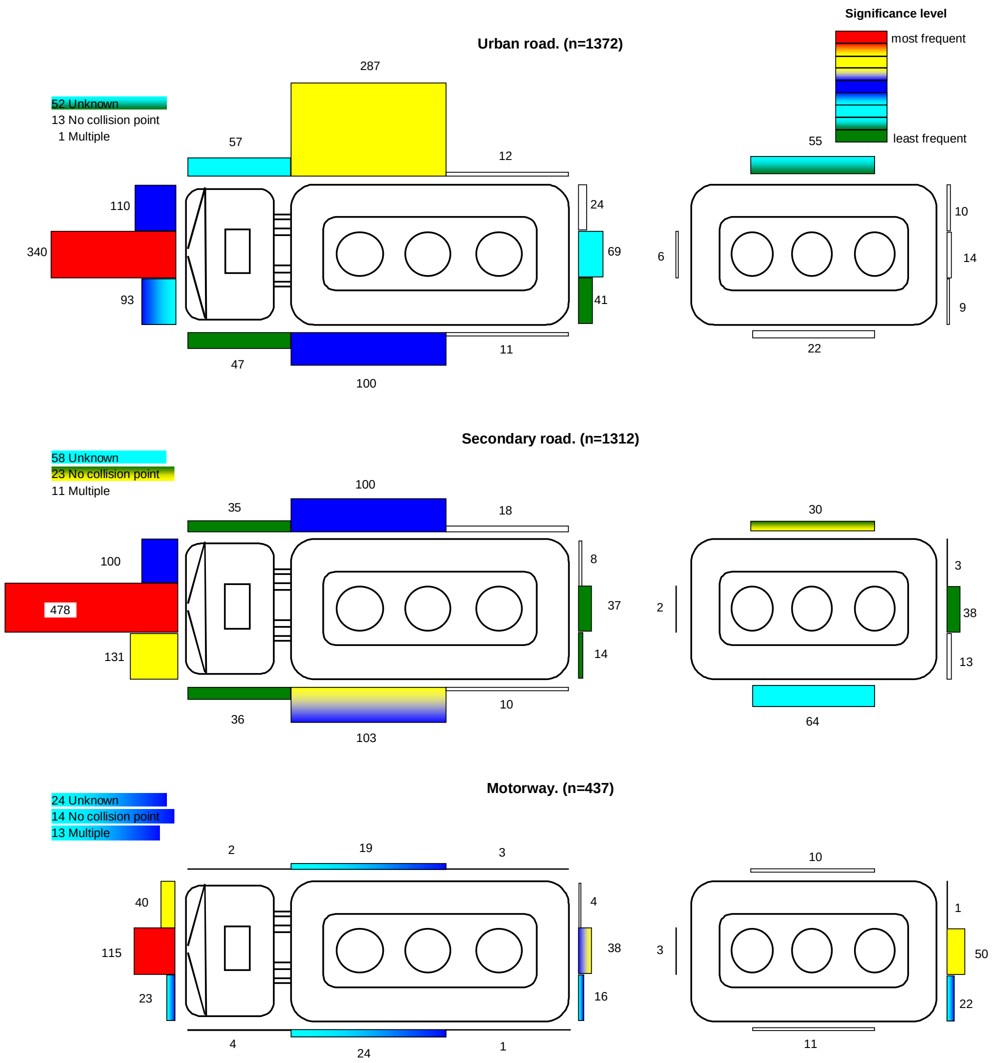

Figure 4 Frequency of KSI accidents per impact location. Locations with a different colour vary significantly. Gradient fills, do not differ significantly from either colours.

front. Not coloured impact locations are too complexly grouped together to provide any relevant information on significance.

On all road types, a collision on the *mid front* occurs significantly more than other collisions. An impact on the right side seems to be a typical urban impact location, because it is second most frequent and is less present on other road types.

Logistic regression shows that there is a statistical difference on the prediction to get killed with respect to the impact locations. The rear of the trailer is significantly more dangerous than any part of the truck/tractor. The *front of the trailer* has the second highest prediction value, but is not significantly different from many other impact locations, due to very low frequency of occurrence. The *front of the HGV* is therefore the second most dangerous impact location. The *right side* of the HGV is not found to be significantly more dangerous than many other impact locations. The *left side of the tractor/truck* is found to be the least dangerous.

Combined analysis shows that the accidents on the *trailer mid rear* and the high frequency of impacts on the *mid front of the HGV* are explained by the relative high danger with respect to other impact locations. The high frequency on the right side on urban roads are not explained by the danger of the right side of the HGV. Other variables are likely to explain this better. In the next paragraph will be shown that the main collision partner on the right side of the HGV is a bicycle, which is a vulnerable road user and more explanatory.

Collision Point On The Collision Partner (SI)

For the collision points on the collision partner no update of the collision points was done, so only the standard eight locations were used (see Updating Impact Locations).

Frequency counts show for urban and secondary roads the *mid front* (45%) as the significantly most frequent impact location on the collision partner, followed by the *left side* (21%)

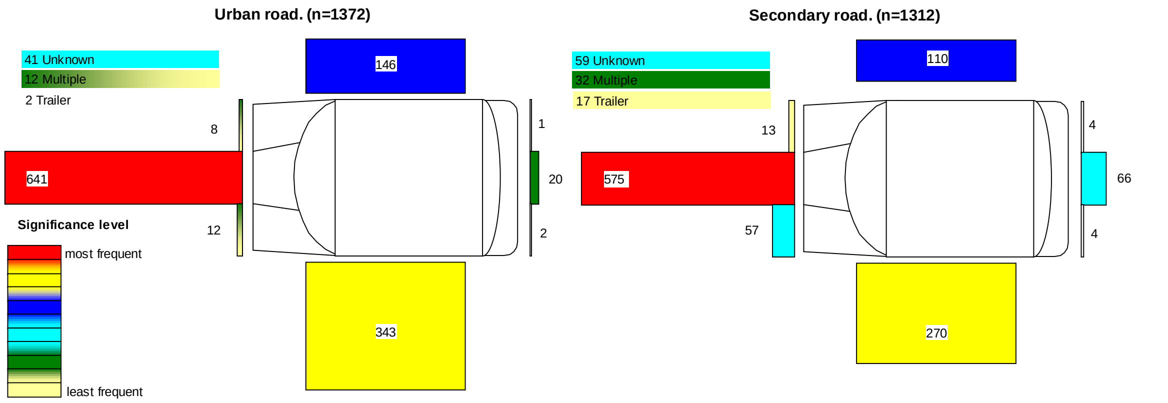

Figure 5 Frequency of KSI accidents by collision point on the collision partner. Locations with a different colour vary significantly. Gradient fills, do not differ significantly from either colours.

and thirdly the *right side* (9%) (see Figure 1). On motorways also the *mid front* is most frequent. The *right* and *left side* do not differ significantly. It is striking that the *left front* and *right front* are so much less pronounced than for the HGV front. This is most likely influenced by the involvement of two-wheelers.

Logistic regression shows a variation in the influence of the different categories in this variable for the seriously injured collision partner occupants. However, significant differences are difficult to present, because the categories do not differ from nearby categories. Due to the sliding scale, it can not be said that *mid front* is the most dangerous collision point, because it is not significantly different from the *left side*, *left front* and *right side*. But it is more dangerous than the rest of the categories.

Combined analysis shows that the three most frequent impact locations with SI accidents, also have the highest predictive value for these categories. The high frequencies are thus explained by the danger of these impact locations.

Combined Impact Locations

Unfortunately the combination of the impact locations on the HGV and the collision partner were not taken into account in the logistic regression. However, it is illustrative to present the frequency count analysis on this (without significant levels). For the most frequent collision interactions it is looked at the collision partners that were involved, the manoeuvre and the intended movements of the vehicles (see Figure 6). It can be seen that on urban roads, mainly bicyclists are involved in KSI accidents. The manoeuvres *Two crossing roads* and *same road, same direction, with turning* are very prominent. This last manoeuvre is highly prominent in KSI accidents with a right turning HGV and a bicycle going straight ahead. Both the involvement of a bicyclist and these two manoeuvres were found to have a high predictive value for the collision partner to get killed in the regression analysis. The *mid right side of the HGV* is not found to be significantly more dangerous than other impact locations. However, the *mid front* was found to be the second most predictive impact location on the HGV.

On secondary roads the *mid front of the HGV* is predominantly present, together with the also classified as dangerous manoeuvres *B: two crossing roads, without turning* and *G: same road, same direction without turning.* On motorways the highly predictive impact location *trailer mid rear* is prominently present. The involved manoeuvres are the least predictive found in the regression analysis. It has also to be noted that the *front of the HGV* is most frequently in contact with the *side* of a collision partner.

Age Of The Collision Partner (K)

The age groups are represented with the mean of the ten year interval.

Frequency counts show that there is a difference in distribution over the different road types (not shown here). On urban roads the age group 10-19 is most frequently involved in KSI accidents with a HGV. Overall the age group 20-39 is most frequently involved in KSI accidents.

Logistic regression shows that with increasing age the risk increases to get killed in an accident with an HGV. Between the ages 10-59 no significant difference could be found. Above and below this interval the risk increases.

Combined analysis shows a counteracting relationship between the logistic regression and the frequency counts (see Figure 7). The lower number of K accidents at higher ages are explained well by the fact that the victims are older. The high frequency of the ages 11-60, is not explanatory. This sounds reasonable, because it may be expected that younger people are less vulnerable, but likely more frequently involved due to exposure or other explanatory variables.

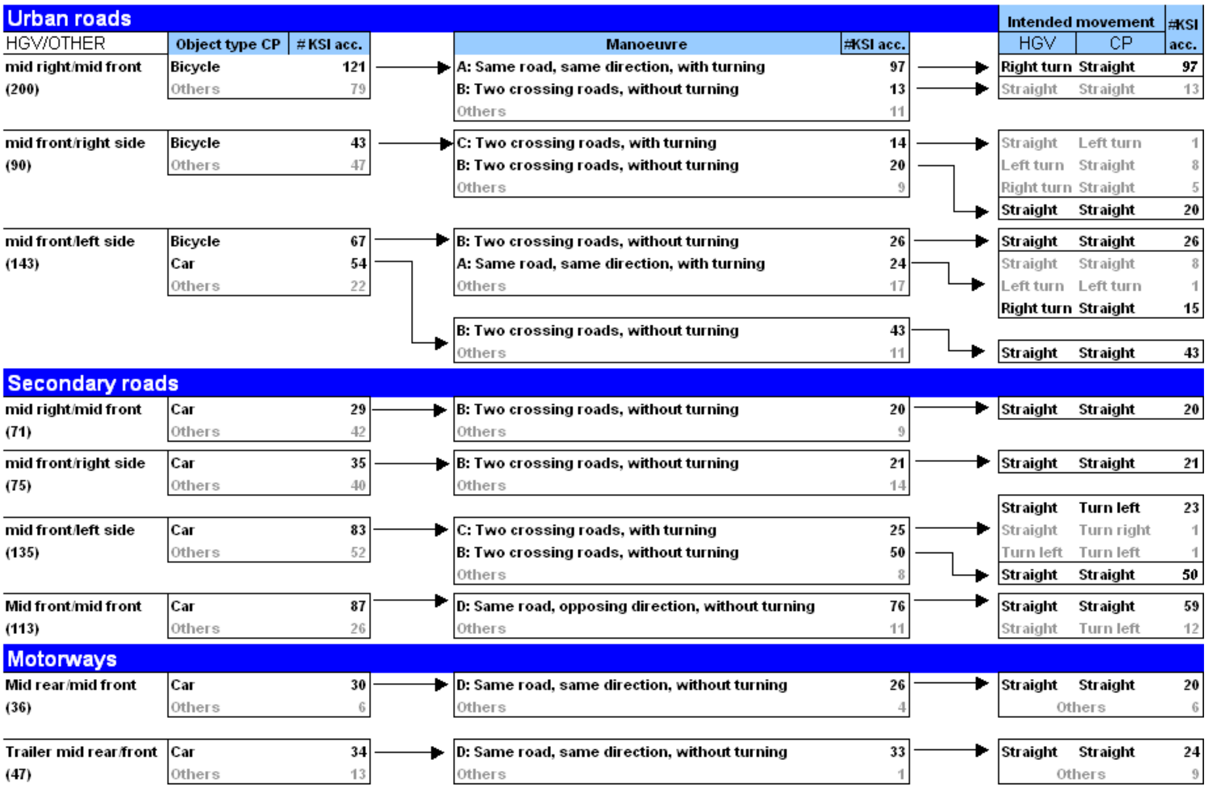

Figure 6 Tree diagram of the most frequent collision interactions, followed by the most frequent collision partners, manoeuvres and intended movements for KSI accidents.

Gross Vehicle Weight (K)

The Gross Vehicle Weight (GVW) was taken into account in the detailed analysis. See also Data Sanitation.

Frequency counts show that the largest number of KSI accidents occurs with HGVs with a GVW of 18 tons (22%) and 19 tons (19%). See also Figure 8. These categories are significantly more involved than other categories and also differ significantly from each other.

Logistic regression. In the regression analysis the maximum rear axle pressure proved to be a predictive variable for the collision partner to get killed. This variable correlates highly with the GVW (R^2=.92) and is considered by the authors to be a better predictive variable than the maximum allowable rear axle pressure. The probability to get killed for the collision partner increases with increasing GVW of the HGV. It must be taken into account that for the regression analysis a linear relationship was assumed, this is not necessarily the true relationship.

Combined analysis. Accident severity increase with increasing GVW. However, the frequency of involved vehicles in accidents with a killed collision partner occupant (N=695) shows

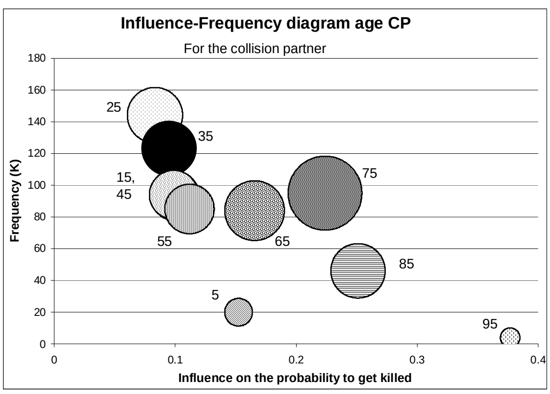

Figure 7 Influence-frequency diagram for the age of the collision partner occupant.

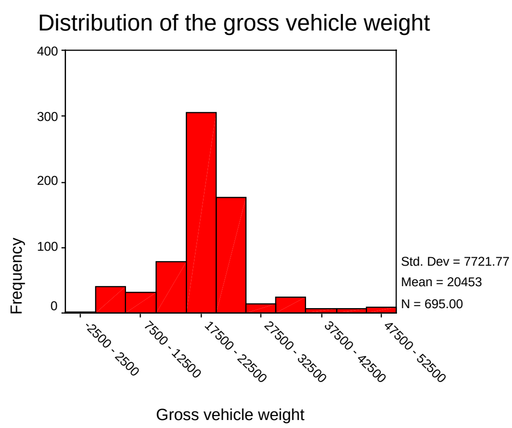

Figure 8 Distribution of the GVW of HGVs that were involved in KSI accidents.

maximum values at GVWs of 18 tons (23%) and 19 tons (21%). In this case the frequency counts show mainly the exposure rate.

DISCUSSION

Up till now accident statistics are often presented without taking significant levels into account. Neither statistical differences in occurrence nor logistic regression analysis have been used to find important variables. This paper shows this possibility and necessity.

With the presentation of solely the frequency counts, the influence of exposure or underlying variables that explain better why these frequencies are so high are left out of consideration.

Logistic regression analysis shows which categories have a high predictive value and are thus actually explaining the number of killed or seriously injured victims when that category is present in an accident. Some problems occur when using the regression technique due to the fact that some categories are coded in more than one variable (e.g. accident with a pedestrian). Also the actual collision speed is not coded, which possibly influences the predictive value of some variables (e.g. the type of manoeuvre and the impact location). In order to compensate for this problem it is recommended that in future exercises the combined impact locations are taken into account.

The results show that, especially in the case of high exposure (e.g. speed limit, age groups), high frequency categories are not necessarily the categories to look at to reduce the number of victims. Underlying variables may be the actual cause / explanation of these high frequencies. Therefore, measures to reduce the number of victims in categories which have both a high frequency and a high predictive value with respect to the other categories in this variable have the highest potential for success.

The risk exists that some of the presented results are interpreted as trivial, because it is already known that vulnerable road users are at high risk, and that high speeds result in more severe injuries. This is however the first time that is it shown with Dutch national accident data that these variables are, independently, of statistical significant importance for the increase in risk. It would have been strange if these variables were not found to be of influence.

Influential Variables For The HGV Occupant

Influential variable for the HGV occupant can hardly be found with the national statistics. The only influential variable was the type of collision partner. The categories *other non-vehicle* and *another HGV* had both the highest frequencies as well as the highest predictive values. The speed limit, collision point on the HGV and manoeuvre types are not significantly influencing the probability to get killed or seriously injured for the HGV occupant. Details on injury causation are necessary to better predict what can be done for HGV occupants (e.g. seatbelt usage, cab deformations, etc.). In-depth investigations are best suited to find these variables.

Influential Variables For The Collision Partner

The influential variables for the collision partner can be distinguished more clearly. The high frequencies of vulnerable road users are well explained by their vulnerability alone. Improvement of HGVs to be less aggressive towards pedestrians and other vulnerable road users has a high potential for reducing the number of killed and seriously injured. The mandatory fitting of side under run protections is in line with these results.

High speed limits have influence on the probability to get killed or seriously injured. When speed limits would be reduced, the number of KSI victims would reduce significantly. This is a trivial conclusion. Nevertheless this is an important influential factor for the number of KSI victims. Reduction of the speed limit seems not very realistic and therefore efforts should be made to prevent these accidents from happening or increase occupant safety by improving both HGVs and collision partner vehicles. Also the possibility exists that vehicles were going faster than the speed limit. In that case better enforcement would reduce the number of victims.

The manoeuvres *G: same road opposing direction without turning, A: same road same direction with turning* and *B: two crossing roads without turning* have relatively much influence on the high number of accidents with killed collision partner occupants. Manoeuvres *G* and *B* could be influenced to some extend by the collision speed which is not included as a separate variable. It could be the underlying explaining variable. Further in-depth research in this area is necessary. The relatively dangerous manoeuvre *A* is included frequently in accidents with a bicycle on urban roads and the relatively high risk collision point on the front on the bicycle (see Figure 6). This therefore seems to be a dangerous situation, which supports the current interest of the Dutch government for improving fields of view on HGVs.

The collision point on the *rear of the trailer* is the most significant predictive factor in the variable collision point. The number of seriously injured collision partners with this variable present is not extremely high. However, because easy modifications on the trailer rear are possible (lower under run protections) this seems an easy way to reduce the number of victims. The front of the

HGV is second most dangerous, and is also involved most frequent. Therefore, front under run protection devices seem to have a high priority. It has to be noted that the front of the HGV is most frequently in contact with the side of a collision partner. This may have implications on the design of especially energy absorbing under run protections. If large deformations are tolerated in the lower regions, a high level of intrusion could occur in the upper regions of the collision partner vehicle, thus endangering the collision partner occupants.
Higher GVW increases the risk for the collision partner. Fortunately, very heavy trucks are not involved very frequently, with respect to other HGVs. The mean GVW involved in accidents with a killed collision partner is around 20 tons. Therefore GVW reduction does not seem to be a first priority.
The impact locations on the collision partner do not differ greatly from each other. The *mid front* and *left side* occur most frequent and also have the highest predictive value. Improvements on these locations have the highest potential for reduction of victims.

CONCLUSIONS

Combining the regression analysis and frequency count results give a unique insight in potentials for improvements. Improvement of situations with high frequency and a high influence on the probability are most likely to decrease the number of killed and seriously injured. Seemingly important factors in the frequency counts or the regression analysis alone are not necessarily the right factors to improve. They might either have little influence on the outcome or do not occur frequently enough to have a high priority. Especially in the case where the frequency is highly influenced by the exposure (e.g. number of vehicles of that type on the road, age groups, type of manoeuvres made, etc.) wrong conclusions could be drawn.
Tree-diagrams as shown in Figure 6 provide good information about involved categories which had much influence on the frequency of these accidents.
Very often more details are needed to find the real underlying causes. In-depth research with the appropriate questions is necessary to find these causes. This will be performed in the near future in the Netherlands.
For the HGV occupant only the *type of collision partner* was of influence on the outcome of the accident. The collision partners *other non-vehicles (objects, poles, buildings, etc.)* and *HGVs* pose the most danger for the HGV occupant.
The number of KSI vulnerable road users are explained well by their vulnerability.
The manoeuvre *A: same road, same direction, with turning*, in combination with bicyclists and the relatively high risk collision point on the front of the bicycle is situation with high potential for improvements.

Recommendations

It is recommended in future analysis of the Dutch national data to include the collision interaction terms to compensate for the missing of the collision speed.
For several variables (collision speed, injury causation, underlying causes for the manoeuvres, etc.) not enough information is known. It is recommended to do further in-depth research to find the underlying causes.

ACKNOWLEDGEMENTS

This research was supported partly by DAF Trucks NV and Scania Holland in a co-financing arrangement with TNO-Automotive.

REFERENCES

1 Janssen, E.G., *et al.*, 1995, *Ontwikkelingen in de richting van duurzame voertuigveiligheid*, SWOV, Leidschendam, The Netherlands (Dutch)
2 Goudswaard, A.P., Janssen, E.G., 1990, *Passieve veiligheid bedrijfsvoertuigen; een literatuuronderzoek*, TNO-report 754080030, Delft, The Netherlands (Dutch)
3 Vries, Y.W.R. de, 2000, *Accident Analysis of Heavy Trucks*, TNO-Report 00.OR.BV.053.1 /YdV, Delft, The Netherlands (English)
4 Langwieder, K., Bäumler, H., 1995, *Energy-Absorbing Front Underrun Protection of Trucks*, Dep. of Automotive Engineering and Accident Research, Munich, Germany (English)
5 Grandel, J., *Interior safety of truck cabs. Results of an in depth study*, DEKRA AG, Stuttgart, Germany (English)
6 FAI, 1994, *Untersuchungen zur inneren Sicherheit van Lkw-Fahrerhausern*, FAI: Schriften Reihe nr. 115ISSN 0933-050 X, FAT, Frankfurt , Germany (German)
7 FARS database, internet: http://www-fars.nhtsa.dot.gov/www/query.html
8 Knight, I., 2000, *Accidents involving heavy goods vehicles 1994 to 1996*, TRL, ImechE Conference Transactions: Vehicle safety 2000, C567/015/2000, ISBN: 1 86058 271 0, Chippenham, Wiltshire, UK. (English)
9 Polak, P.H., Blokpoel, A., 1998, *Schatting van de werkelijke omvang van de verkeeronveiligheid 1997*, R-98-55, SWOV, Leidschendam, The Netherlands (Dutch)