PERGAMON

Accident Analysis and Prevention 35 (2003) 59–69

ACCIDENT ANALYSIS & PREVENTION

www.elsevier.com/locate/aap

The use of multilevel models for the prediction of road accident outcomes

Andrew P. Jones [a,*], Stig H. Jørgensen [b]

[a] School of Environmental Sciences, University of East Anglia, Norwich, Norfolk, NR4 7TJ, UK

[b] Department of Geography, The Norwegian University of Science and Technology (NTNU), N-7055 Dragvoll, Trondheim, Norway

Received 5 July 2000; received in revised form 16 July 2001; accepted 31 July 2001

Abstract

An important problem in road traffic accident research is the resolution of the magnitude by which individual accident characteristics affect the risk of fatality for each person involved. This article introduces the potential of a recently developed form of regression models, known as multilevel models, for quantifying the various influences on casualty outcomes. The application of multilevel models is illustrated by the analysis of the predictors of outcome amongst over 16,000 fatally and seriously injured casualties involved in accidents between 1985 and 1996 in Norway. Risk of fatality was found to be associated with casualty age and sex, as well as the type of vehicles involved, the characteristics of the impact, the attributes of the road section on which it took place, the time of day, and whether alcohol was suspected. After accounting for these factors, the multilevel analysis showed that 16% of unexplained variation in casualty outcomes was between accidents, whilst ~1% was associated with the area of Norway in which each incident occurred. The benefits of using multilevel models to analyse accident data are discussed along with the limitations of traditional regression modelling approaches.

Keywords: Road traffic accidents; Multilevel models; Logistic regression; Casualty survival; Epidemiology

1. Introduction

An important and frequently occurring problem in road traffic accident research is the resolution of the magnitude by which individual accident characteristics may affect the severity of the incident and hence the subsequent risk of fatality for each person involved. Attention is often focused on the relative contribution of human and vehicle related determinants of risk, along with the role that other elements such as highway design and environmental conditions may play. An understanding of such factors may be used to identify particular hazards (Jones and Bentham, 1995), assess the success or otherwise of safety measures (Murphy et al., 2000), and assist in elucidating the directions along which future research should be targeted (Noy, 1997).

From a methodological viewpoint, a wide variety of approaches have been employed to study fatality risk. For example, Kim et al. (1995) use data on accidents in Hawaii to illustrate the use of a categorical log linear model to examine personal and behavioural predictors of crash and injury severity, whilst Shankar et al. (1996) describe the development of a nested procedure for the analysis of accident severity on rural freeways. In the UK, Jones and Bentham (1995) calculated Odds ratios to determine probabilities of fatality amongst traffic accident causalities according to a complex matrix of explanatory variables. However, despite this apparent diversity of methodologies, common to most investigations is the requirement for a statistical model which will predict fatality risk for an individual based upon a range of explanatory variables. In line with this, the most frequently used technique is the generalised linear modelling (GLM) methodology of logistic regression (Menard, 1995).

The popularity of logistic regression models stems from their ease of interpretation, widespread acceptability, and the provision of suitable estimation routines in the majority of popular statistical packages. However, there can be problems with their application in certain areas of road accident outcome prediction. One of the most fundamental of these arise when the factors influencing the probability of survival for any individual casualty are seen to be operating at a variety of scales, with these scales comprising successive levels of a hierarchy. These may be associated with the personal characteristics of the casualty at the lowest level of the hierarchy, the features of the vehicle within which they are located or the distinguishing events of the incident in which they are involved. At the highest levels of the hierarchy, they may be extended to the properties of the road section upon which the incident took place, or even

* Corresponding author. Tel.: +44-1603-593127; fax: +44-1603-507719.
E-mail address: a.p.jones@uea.ac.uk (A.P. Jones).

0001-4575/02/$ – see front matter
PII: S0001-4575(01)00086-0

the attributes of the geographical region or country where it occurred.

The possible existence of hierarchical structures within accident data is commonly ignored. However, disregarding hierarchies, where they are present, can lead to the production of models giving unreliable estimates of prevision, incorrect standard errors, confidence limits, and tests (Skinner et al., 1989). Furthermore, the resultant model will present an over-simplistic picture of a complex reality, and hence may be open to mis-interpretation (Goldstein, 1995).

In recent years a new form of statistical modelling has been developed that allows hierarchical data structures to be easily specified and their influence to be eloquently and efficiently estimated (Duncan et al., 1998). Several terms have been used to describe this new development: multilevel models (Goldstein, 1995), random coefficient models (Longford, 1993), and hierarchical linear models (Bryk and Raudenbush, 1992). Hereinafter, only the term "multilevel models" is used.

This paper illustrates the utility of multilevel models for the analysis of road accident data, and specifically, the study of the factors affecting fatality risk for individual casualties. It begins with a brief review of traditional approaches, highlighting the problems associated with these models. It goes on to outline the theory of multilevel estimation strategies, and concludes with an illustration of practical usage by their application to a model of fatality risk amongst Norwegian casualties.

2. Limitations of traditional generalised linear modelling approaches to the prediction of fatality risk

Consider a simple situation where three vehicles, A, B and C are involved in an accident. Vehicle A bears three casualties, whilst both vehicles B and C are occupied by two casualties (Fig. 1). Two of the casualties by vehicle A are killed in the incident, whilst the others survive. To understand why some of the casualties survived whilst others did not, a GLM is required that will predict the proportion of casualties who are fatally injured or estimate the probability that the injuries of an individual casualty will be fatal. Normally, a logistic regression will be used where the log-Odds ratio (or logit) of an outcome variable y for each casualty i is estimated as a function of a matrix of explanatory variables.

Although logistic regression presents a convenient way of modelling a binary response, the basic tenet behind traditional models for casualty outcomes will be that each record entered into the estimation procedure corresponds to an individual causality, and that the residuals from the model exhibit independence. However, a consideration of the data structure illustrated by Fig. 1 suggests that the assumption of independence may often not hold true. For example, it is reasonable to assume that the characteristics of the vehicles within which the casualties are travelling will effect their probability of survival. If this is the case, then casualties within the same vehicle will tend to have more similar outcomes than casualties within different vehicles, and the assumption of residual independence will not be met. The

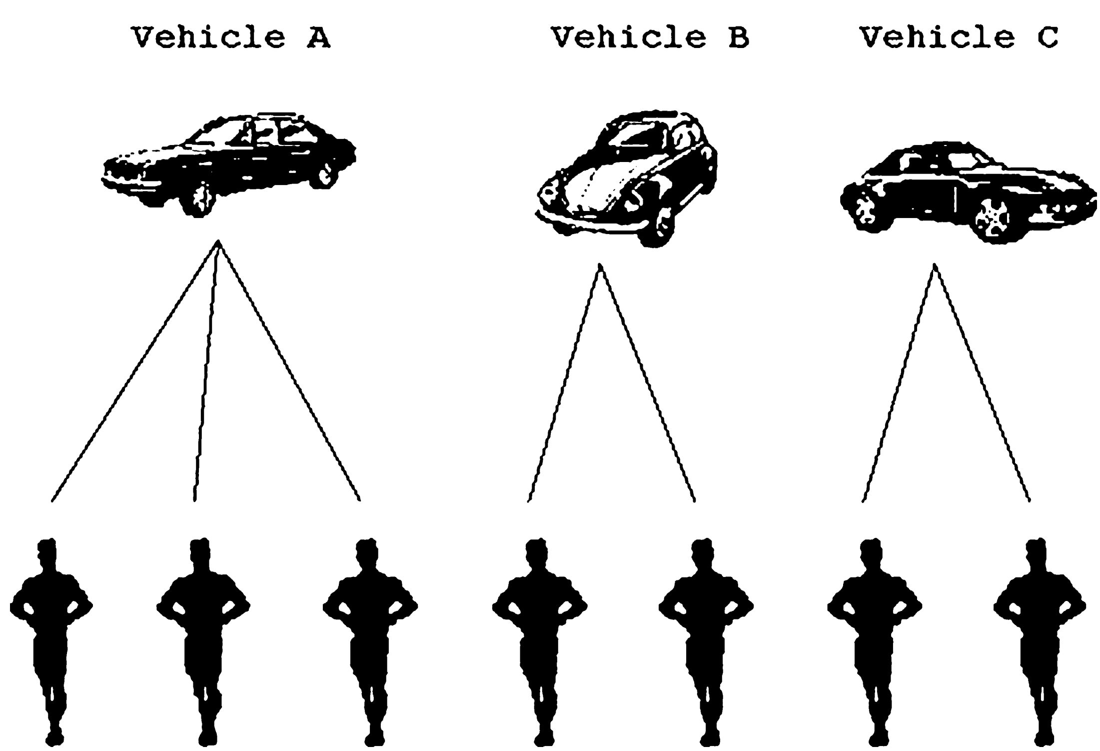

Fig. 1. A typical situation of casualties within vehicles.

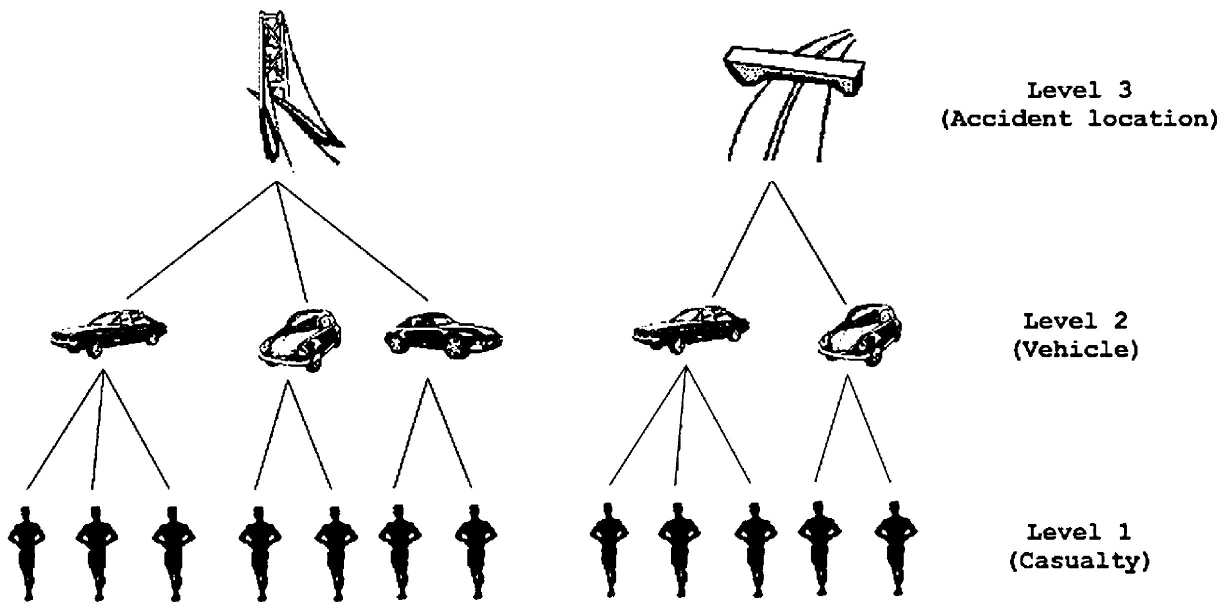

Fig. 2. This data set may be seen as exhibiting a hierarchy of casualties: at the lowest level, level 1; nested within vehicles, level 2; nested within accident locations, level 3.

same argument may be extended to encompass the effect of similarities between different incidents, road sections, or geographical regions. In this sense, the graphical depiction of the accident data set in Fig. 1 may be viewed as actually corresponding to a two-level hierarchy of casualties (level 1) within accidents (level 2). This hierarchy may be additionally extended with further levels representing, for example, the location of the road section upon which the accident took place (Fig. 2).

There are numerous problems in accommodating hierarchical data structures within the traditional generalised linear estimation framework. A technical limitation of traditional GLMs stems from the fact that they are likely to produce poorly estimated parameters and standard errors (Skinner et al., 1989). Problems with standard error estimation arise from the presence of intra-unit correlation where, in the manner described above, the assumption of independence in the model residuals is not met. If intra-unit correlation is small, then reasonably good estimates of standard errors may be expected (Goldstein, 1995). However, where intra-unit correlation is large then estimation strategies such as weighted least squares will underestimate standard errors, meaning that confidence intervals (CI) will be too short and significance tests will too often reject the null hypothesis.

A further limitation of the traditional GLM is that the values of vehicle or location related variables must be collapsed to the level of the individual casualty and simply replicated across all individuals sharing those characteristics. This is problematic in that it provides no information on, for example, the probability of occupants of the same vehicle, incident, or locality having similar outcomes; a factor that is of considerable interest in accident research. If the analyst is specifically interested in, for example, geographical variations in the outcome being studied, the limitation may be circumvented by fitting regression models which include dummy variables to indicate the region within which accidents occurred. However, this solution in itself presents difficulties as models estimated using dummies can quickly become large and complex if the data set contains observations for many regions.

An alternative to the use of dummy variables to model hierarchical data structures would be to fit a separate model to each incident, road length, or region. However, it is difficult to draw broad conclusions on the respect roles of contextual and compositional factors from the output of this strategy, as the significant variables may differ and unreliable results may be produced if some models are based on small numbers of observations.

3. The multilevel model

An alternative strategy, which circumvents all of the problems with traditional GLM techniques outlined above, is the fitting of true multilevel models. For algebraic simplicity, a two level hierarchy of i casualties (at level 1) involved in j accidents (at level 2) are considered here. As with a traditional logistic regression model, the observed responses y_{ij} are dichotomous (0, 1), indicating whether each casualty died or survived with serious injury, and the standard assumption that they are binomially distributed holds.

In the multilevel case, the accidents included in the model are regarded as a random sample of events from a population, and hence a regression relation is assumed for each.

Considering a situation with just one explanatory variable, casualty 'age', being tested the model may be written as

$$y_{ij} = \beta_{0j} + \beta_1 \mathrm{age}_{ij} + \in_{ij} \quad (1)$$

where the subscript *i* takes the value from 1 to the number of casualties in each accident, and the subscript *j* takes the value from 1 to the number of events in the sample. Using this notation, items with two subscripts *ij* vary from casualty to casualty, where an item that has a *j* subscript only varies across accidents but is constant for all the casualties within an event. If an item has neither of the subscripts, it is constant across all casualties and accidents.

As the events included in the analysis are treated as a random sample, Eq. (1) may be re-expressed as

$$\beta_{0j} = \beta_0 + \mu_j, \qquad \hat{y}_{ij} = \beta_0 + \beta_1 \mathrm{age}_{ij} + \mu_j \quad (2)$$

where β_0 is a constant and μ_j is the departure of the *j*th accident's intercept from the overall value, which means that it is an accident level (level 2) residual that is the same for all casualties within an event. Therefore, this term describes, after holding constant the effect of the explanatory variables within the model, the residual influence of the accident in determining the outcome for each individual casualty within it. It, hence, allows the identification to be made of which incidents carry a higher or lower risk in terms of outcome severity than would be expected from the values of the explanatory variables in the model.

The notations expressed in Eq. (2) can be combined. Introducing an explanatory variable 'cons', which forms a constant or intercept term, and associating every term with an explanatory variable, the model takes the form

$$y_{ij} = \beta_0 \mathrm{cons} + \beta_1 \mathrm{age}_{ij} + \mu_{0j} \mathrm{cons} + \in_{0ij} \mathrm{cons} \quad (3)$$

and finally, the coefficients can be collected together and written as

$$y_{ij} = \beta_{0ij} \mathrm{cons} + \beta_1 \mathrm{age}_{ij}, \qquad \beta_{0ij} = \beta_0 + \mu_{0j} + \in_{0ij} \quad (4)$$

In this equation, both μ_j (the level 2 or accident level residuals) and $\in_{ij}$ (the level 1 or casualty level residuals) are random quantities whose means are estimated to be equal to zero. It is assumed that, being at different levels, these variables are non-correlated. A comparison between the multilevel model and a traditional GLM illustrates the tenet of multilevel models. Traditionally the residual or error term of a model, $\in$, is seen as an annoyance and, the aim of the modelling process is to minimise its size. With multilevel models, $\in$ is of pivotal importance in model estimation. Rather than a single error term being estimated, it is stratified into a range of terms, each representing the residual variance present at each level of the hierarchy. Viewed in this sense, μ_j represents accident level effects, whilst $\in_{ij}$ represents those operating at the level of the individual.

The model presented in (4) is known as a variance components model, where the only random parameters are the intercept variances at each level (Lin, 1997). Remembering that, age_{ij} represents the age of each casualty, the variance components model depicts relationship between age and outcome as constant, but (provided $\mu_j \geq 0$) casualties are modelled as being at a higher risk of death in some accidents than others, as depicted graphically in Fig. 3. Each solid line in Fig. 3 represents the fitted relationship between Odds of death, $\hat{y}$, and casualty age (x_age) across all the casualties within each accident. In this case, there are 10 accidents in the data set, and hence 10 solid lines. The dotted line represents, the mean relationship across all accidents, and hence is similar to the result that would be obtained for a model that ignored the hierarchical structure of the data.

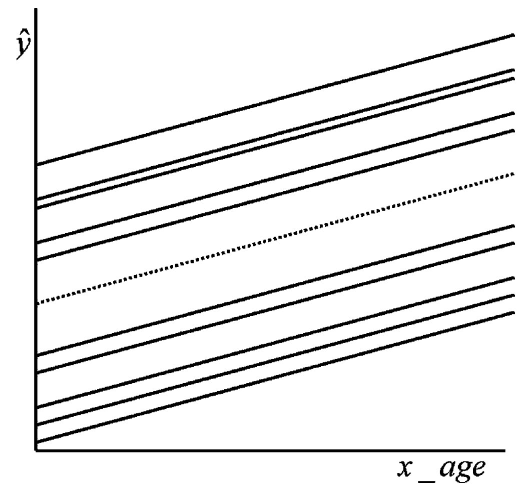

Fig. 3. The variance components model. Each solid line represents an accident, whilst the dotted line represents the mean relationship across all accidents.

The variance components model may be extended to allow the effects of the x_{ij} explanatory variables to differ at higher levels of the hierarchy. For example, it may be that casualty age is more important determining outcome in certain accidents than others. To facilitate this, the explana-

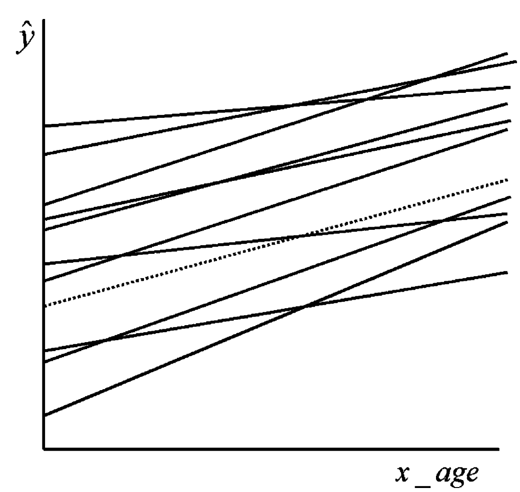

Fig. 4. The fully random model. Each solid line represents an accident, whilst the dotted line represents the mean relationship across all accidents.

tory variable age needs to be allowed to vary at level 2 (the accident level) of the model, and hence the notation needs a new term adding so that

$$y_{ij} = \beta_{0ij}\text{cons} + \beta_{1j}\text{age}_{ij}, \qquad \beta_{0ij} = \beta_0 + \mu_{0j} + \in_{0ij},$$
$$\beta_{1j} = \beta_1 + \mu_{1j}\text{age}_{ij} \tag{5}$$

This has the effect of estimating the term u_{1j} for each accident which is added to the coefficient to age, hence allowing the slope of the regression line to vary for each accident. This is depicted in Fig. 4, which may be compared with the parallel fitted lines of the variance components model in Fig. 3.

4. Parameter estimation for multilevel models

There are various methods available for parameter estimation in multilevel models. The use of full Bayesian treatment has recently become computationally feasible with the development of 'Markov Chain Monte Carlo' (MCMC) methods, especially, Gibbs Sampling (Gilks et al., 1996). These methods have the advantage, in small samples, that they account for the uncertainty associated with the estimates of the higher-level parameters, and can provide exact measures of uncertainty. Hoijtink (1998) also describes an alternative 'Bootstrap' procedure for taking account of this uncertainty which involves random re-sampling with replacement from the model to overcome the assumption of normality amongst the higher level variables. A less computationally intensive approach involves the use of Iterative generalised least squares (IGLS). IGLS is adequate for situations where there is a large sample of responses at each level of the hierarchy, for example many casualties nested within many accidents. It was the estimation method used in the following section.

The theory of IGLS is described in detail by Goldstein (1995). Briefly, initial estimates of the fixed (explanatory variable) terms are derived by ordinary least squares ignoring the higher-level random (residual variance) terms. The squared residuals from this initial fit are then regressed on a set of variables defining the stricture of the random part to provide initial estimates of the variances/covariances. These estimates are, then, used in a generalised least squares analysis to provide revised estimates of the fixed part, which is in turn used to revise the estimates of the random part, and so on until convergence. Crucially, a difficult estimation problem is decomposed into a sequence of linear regressions that can be solved efficiently and effectively. However, a limitation of IGLS for models with a binomial or Poisson distributed response variable is that it requires assumption of normally distributed variance above level one of the hierarchy.

It is important to note that the slopes and intercepts (as illustrated in Figs. 3 and 4) that are estimated for units within level 2 and above of the hierarchy will not be the same as those that would be obtained from an OLS solution; they are, in fact, shrunken residuals as they have, to a greater or lesser extent, been shrunken towards the overall mean relationship. Taking an example of a level 2 model at the accident level, if $\sigma^2_{e0} = \text{var}(\in_{0ij})$ and $\sigma^2_{u0} = \text{var}(u_{0j})$, then each accident level residual is estimated using

$$\hat{u}_j = \frac{n_j\sigma^2_{u0}}{n_j\sigma^2_{u0} + \sigma^2_{e0}}\tilde{y}_j \tag{6}$$

From this, it can be seen that if n_j is large and there are many casualties within the accident, then the predicted level-2 residuals will be closer in value to the raw OLS residual, than when n_j is small. If n_j is small, then the residual will be shrunken towards the mean. Similarly, if σ^2_{e0} is large and there is a lot of variability in outcomes of the casualties within an accident, then the predicated residual will also be shrunken. In this sense, multilevel estimates can be seen as conservative estimates of variability at different levels of the hierarchy, where units based on a small sample or a very variable outcome are considered to provide little information, and are shrunken towards a mean.

5. A multilevel model of accident outcomes for vehicle occupants in Norway

In this section, the application of multilevel modelling is illustrated by its use to model outcomes amongst vehicle occupants involved in road accidents in Norway.

Whilst, levels of mortality per kilometre travelled in Norway are one of the lowest in the world (WHO, 1997), there are large geographical variations in proportional survival amongst casualties (Fig. 5). To highlight the utility of multilevel models, this analysis was undertaken to determine whether these variations could be explained by risk factors for which information is commonly available from police accident records (such as the characteristics of casualties and vehicles and the circumstance of the incidents) or whether some other unidentified factors were apparent.

Data on all vehicle users who were injured in police recorded road traffic accidents on Norwegian public roads between 1985 and 1996, the most recent period for which data is available, was provided by the Norwegian Public Roads Administration. Pedestrians were excluded but pedal cycles and motor cycles were included. In these records, the severity of injury for each casualty was coded as 'slight', 'serious', 'dangerous', or 'fatal'. Casualties with only slight injuries were not used here as these injuries are not life-threatening, and there is usually significant under-reporting amongst this group (Simpson, 1996). The final data set comprised 16,332 casualty records.

For each casualty, the variables outlined in Table 1 were provided. These can be broadly grouped into those associated with the individual, their vehicle, the accident, and it's surroundings. Whilst, the intuitive hierarchy for this data is one of casualties (level 1), within vehicles (L2), within accidents (L3), within localities (L4); this structure

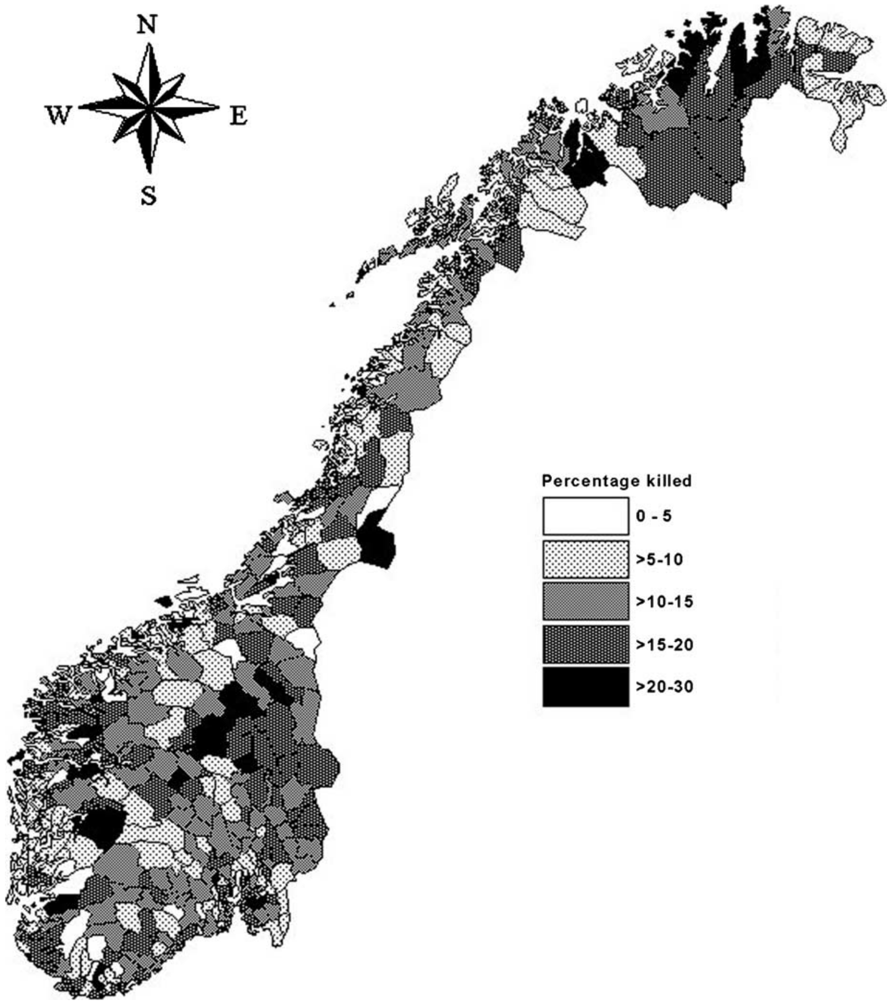

Fig. 5. Percentage of vehicle occupant casualties killed in serious road traffic accidents in Norwegian communes 1985–1996: numerator—numbers of deaths, denominator—numbers of casualties with at least serious injuries.

Table 1
Variables included in the analysis

Variable	Description
Casualty outcome	0, Survived with serious injuries; 1, died
Casualty sex	0, Female; 1, male
Casualty age group	1, 0–15 years; 2, 16–64 years; 3, >65 years
Casualty type	1, Cyclist, 2; motor cyclist, 3; car occupant; 4, bus/coach occupant, 4; lorry occupant; 5, other
Casualty suspected of drinking	0, No; 1, yes
Casualty wearing seat-belt	0, No; 1, yes
Collision type	1, Rear impact; 2, head on; 3, side impact
Daylight conditions	1, Night; 2, day; 3, dusk
Season	1, Spring; 2, summer, 3, autumn; 4, winter
Road condition	1, Dry; 2, wet; 3, icy; 4, slippery
Accident at junction	0, No; 1, yes
Accident in tunnel	0, No; 1, yes
Accident environment	0, Urban; 1, rural
Road type	1, European grade; 2, national grade; 3, provincial grade; 4, local grade
Weather	1, Fine; 2, raining; 3, snowing; 4, foggy
Poor weather visibility	0, No visibility problems, 1, visibility problems
Population density	Population density of municipality (persons/km^2)
Distance from municipality to nearest district hospital	Straight-line distance to hospital (km)

presents a problem as the majority of vehicles contained, only a single occupant with at least serious injuries. Hence, little information is available to differentiate levels 1 and 2. Therefore, the data was analysed using a hierarchy of casualties (level 1, $n = 16,332$ of which 3024 were fatally injured), within accidents (level 2, $n = 12,943$), within municipalities (level 3, $n = 436$). Municipalities are the Norwegian administrative districts for which accident data is collated, and were included as a third level to represent the broad geography of the country. If the interest here had been in especially determining the effect of vehicles, then it would have been possible to select a subset of the data containing only vehicles with more than one person receiving fatal or serious injuries. However, this sample would not then have been representative of the at-risk population.

The units at each level of the hierarchy were identified by the provision of a unique identifying code. As vehicles were not included as a specific level, level 2 actually represents a combination of both vehicle and accident-level effects.

Although the accident rates depicted in Fig. 5 follow the Poisson distribution, the aim here was to the model the Odds of survival for each casualty, and hence binary logistic regression models were fitted. The data was modelled using the MlwiN package (Rasbash et al., 2000) developed by the Multilevel Models Project at the Institute of Education, London. An IGLS estimation strategy was employed to fit a logistic regression model where the response variable indicated whether each casualty survived with serious injuries (response = 0) or died (response = 1) as a result of the incident. For simplicity, only a variance components model is given here.

Table 2 shows the output of the final model produced when non-significant variables were excluded. Because corner point estimation was used, the parameter estimates indicate the effects of the variables relative to the intercept β_0. The intercept hence measures the effect of a 'base line' casualty. That is, a female cyclist aged between 0 and 15 years, involved in a head-on impact on an urban local grade (minor) road at night. It would be the only casualty in the accident, which would not be at a junction, and drinking would not be suspected.

Although technically different, the fixed part of the multilevel model in Table 2 can be interpreted in the same way as an ordinary logistic regression. The results show that there is a significantly elevated risk of death for females. There is also a trend of increasing risk with increasing casualty age. A comparison of the different casualty types in Table 2 shows the highest risk amongst occupants of non-specified vehicles and cyclists, and the lowest amongst bus and coach occupants. Head on impacts carried an elevated risk, as did those occurring at night, with rear impacts being less risky. Whilst no information was available on the actual speed of the vehicles in each accident, the road type and urban–rural nature of the location were modelled as surrogates for these factors. It is, hence, unsurprising that the highest risk was found on fast rural and provincial roads, with the lowest on urban and local roads. Accidents away from junctions were also found to carry an elevated risk, doubtless also indicating the effects of vehicle speed. Finally, casualties suspected of being under the influence of alcohol were more likely to be killed.

One of the key interests of fitting a multilevel model here was to determine if, after controlling for the variables in the fixed part of the model, there remained statistically significant variations in casualty outcomes between accidents and Norwegian municipalities. These random effects are shown in the lower part of Table 2. As a variance components model was fitted, this part of the model is relatively simple.

Table 2
Odds ratios and 95% CI of predictors found to be associated with outcome

Variable	Odds ratio	95% CI	P
Fixed effects			
Intercept	0.14	0.10–0.19	<0.001
Sex			
Male	0.75	0.66–0.83	<0.001
Casualty age group			
16–64 Years old	1.36	1.16–1.57	<0.001
65 Years old or above	3.53	2.93–4.30	<0.001
Casualty type			
Motor cyclist	0.62	0.49–0.76	<0.001
Car occupant	0.69	0.57–0.84	<0.001
Bus/coach occupant	0.28	0.15–0.53	<0.001
Lorry occupant	0.65	0.51–0.84	<0.001
Other vehicle	1.36	0.98–1.90	0.07
Collision type			
Rear impact	0.44	0.36–0.55	<0.001
Side impact	0.69	0.58–0.82	<0.001
Accident environment			
Rural	1.29	1.11–1.44	<0.001
Road type			
European road	2.21	1.69–2.48	<0.001
National road	1.59	1.33–1.91	<0.001
Provincial road	1.40	1.15–1.69	0.001
Daylight conditions			
Day	0.85	0.77–0.94	<0.01
Dusk	0.86	0.76–0.97	<0.001
Accident severity			
Two casualties	1.26	1.13–1.41	<0.001
More than two casualties	1.70	1.47–1.95	<0.001
Accident at junction			
Yes	0.82	0.69–0.96	0.01
Alcohol suspected			
Yes	1.14	1.00–1.32	0.05
Hierarchical effects			
Level 1 (casualty)			
Intercept variance σ^2_{e0}	1	–	–
Level 2 (accident)			
Intercept variance σ^2_{u0}	0.627	0.49–0.77	<0.001
Level 3 (municipality)			
Intercept variance σ^2_{v0}	0.050	0.017–0.083	0.003

Although the multilevel methodology involves estimating a separate intercept value for each accident and municipality, the variance between the three levels of the model may be neatly summarised by the three parameters σ^2_{e0}, σ^2_{u0}, $\sigma^2_{\nu0}$. These are the same parameters used in the calculation of the shrinkage factor illustrated in Eq. (6). They are known as variance parameters, as they measure the variance in the parameters $\in$ (at level 1), υ (at level 2), and ν (at level 3). In other words, they show the relative variability in the model residuals that may be attributed to the effects of casualties, accidents, and municipalities respectively. In this case the casualty level variance parameter was constrained to the value 1 to correspond to a binomially distributed response, although this constraint could be relaxed to allow for extra-binomial variation.

The random parameters in Table 2 show that there is considerable residual variance at both the accident and municipality levels. To estimate the proportion of overall residual variability that is associated with each level, it is normal to calculate the ratio of each of the three variance terms to total variance. The result is known as the intra-unit correlation coefficient, ρ and, taking the accident level (level 2) as an example, it can be calculated using the formula

$$\rho = \frac{\sigma^2_{u0}}{\sigma^2_{\nu0} + \sigma^2_{u0} + \sigma^2_{e0}} \tag{7}$$

The implication of constraining the level 1 variance parameter to be equal to 1 here, is that it is not possible to use the raw value of σ^2_{e0} in Eq. (7). This limitation may be overcome in one of the two ways; either the model is re-fitted with the response variable specified as being normally distributed, or the parameter σ^2_{e0} is multiplied by $\Pi^2/3$ to provide an approximation of the ratio of the between-casualty to total variances (Rasbash et al., 2000). The latter solution is preferable and its application here reveals that whilst the majority of variation in outcomes (83%) occurs at level 1 (that is between individuals), 16% at the accident level, and ~1% between municipalities. The suggests that unmeasured accident characteristics are an important influence, in addition to a relatively small geographical effect.

For a large sample, such as that used here, the significance of the variance parameters may be assessed by using a Wald test (Korn and Graubard, 1990). For a smaller sample (for example, if there had only been far fewer municipalities included in the analysis), the assumption of normality between higher levels of the hierarchy may not hold true, and higher level variances are better modelled using simulation methods (Browne and Draper, 2000). In this case, it is clear that there is statistically significant variance at both the accident and commune levels. This indicates the presence of residual contextual effects and suggests that the geography of outcomes mapped in Fig. 5 cannot be entirely explained by the variables for which the police collect data.

The next factor of interest is the determination of which of the units at each level of the hierarchy are associated with a particularly high or low level of risk of fatality for the casualties nested within them. In other words, taking level three as an example, within which municipalities are casualties most likely to die or survive with serious injuries after controlling for all the fixed effects in the model? This question requires the derivation of the separate intercept term and associated standard error for each hierarchical unit, as outlined in Eq. (4). These terms may then be used to compare units.

Fig. 6 shows the residual variance between municipalities in intercepts, drawn with comparative CI and ranked in order of magnitude. The figure is complex due to the large number of municipalities (436) included in the analysis. However, it is apparent that, whilst the confidence intervals

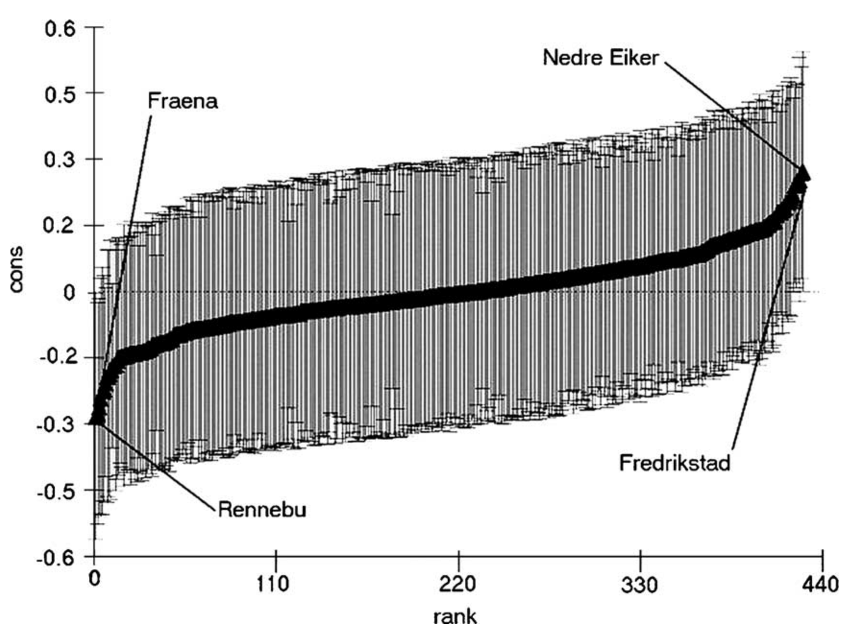

Fig. 6. Ranked municipality-level residuals and comparative 95% CI from multilevel analysis.

overlap for many of the municipalities, there are clearly two sub-populations at the extremes which are associated with a particularly high or low risk. The highest risks of casualty death were predicted to be in Nedre Eiker and Fredrikstad (both close to Oslo), whilst the lowest were in Rennebu and Fræna (both on the north coast).

To illustrate the difference between multilevel estimates and those obtained from a tradition GLM, the model was refitted as a single level model ignoring the multilevel structure and a residual and associated standard deviation was estimated for each municipality based on the distribution of casualty level residuals within it. These residuals and CI are given in Fig. 7. A comparison of Figs. 6 and 7 illustrates how the multilevel shrinkage algorithm detailed in Eq. (6) has acted to lessen apparent variations between municipalities; estimates for those with little information (i.e. few casualties or large variations in casualty outcomes) have been shrunken towards the mean. Hence, the distribution of values and CI is much smoother in Fig. 6 than in Fig. 7, providing a more conservative estimate of the contextual influences on the outcome being studied.

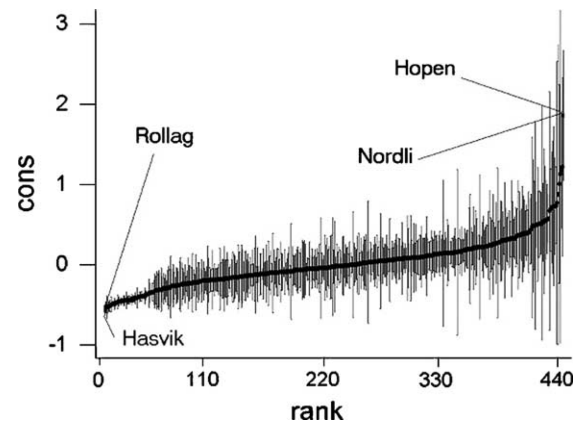

Fig. 7. Ranked municipality-level residuals and comparative 95% CI from non-level analysis.

Fig. 8 shows the data from Fig. 6 mapped onto municipality boundaries, and hence depicts the residual geography of fatality risk. In many respects, the map is rather similar to that of Fig. 5, illustrating the inadequacy of routinely

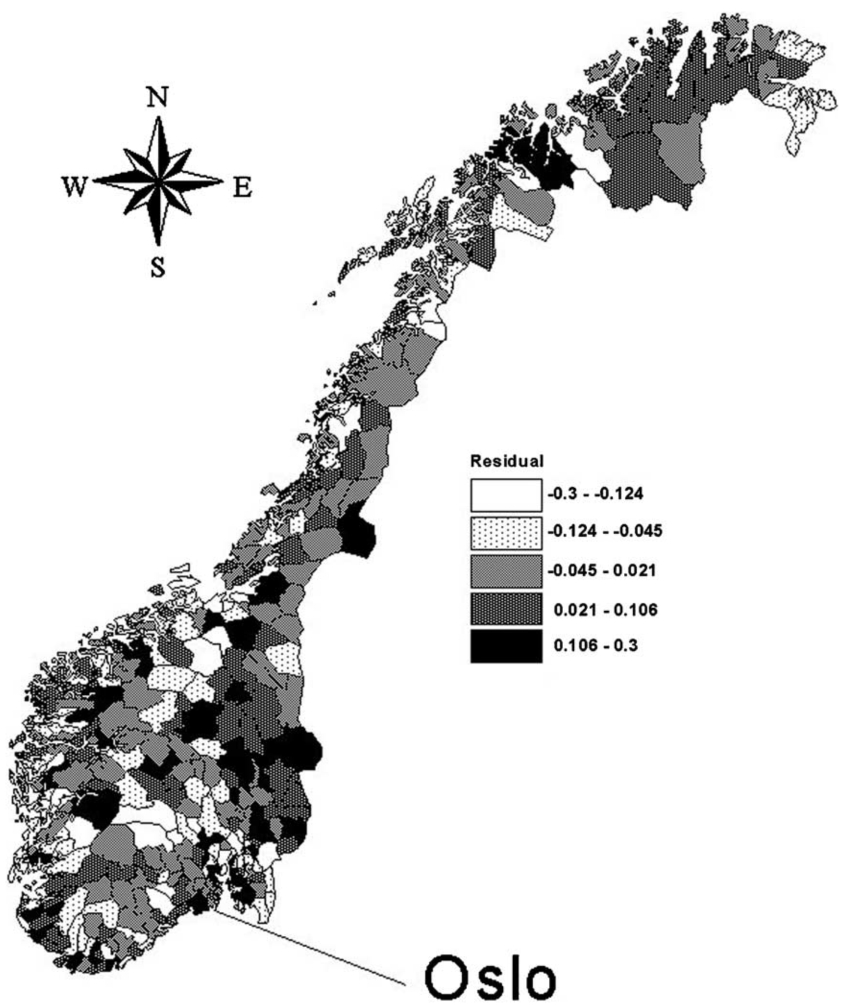

Fig. 8. A map of the municipality-level residuals.

collected data for explaining the geography of risk in this country. The distribution of risks depicted in Fig. 8 suggests that there may be some unidentified risk factors in municipalities close to Oslo, on the south coast, to the north-east of Oslo, and in the northernmost counties of Finnmark and Troms.

6. Discussion and conclusions

This article has introduced the use of multilevel models for the analysis of road traffic accident outcome data, and has illustrated the application of the methodology to a study of seriously and fatally injured casualties in Norway.

The analysis presented here found statistically significant residual variation in casualty outcomes between separate accidents and different geographical locations. This finding signifies the presence of intra-unit correlation in the data set whereby casualties involved in the same accidents, or in accidents in the same municipalities, are more likely to have similar outcomes than those drawn from the rest of the sample. By ignoring this correlation, the analyst runs the risk of producing mis-specified and poorly estimated models.

Whilst the ultimate aim of any analysis would to produce a model where the residual variation was close to zero, this is, in practice, unlikely to be achievable. Here, 17% of the residual variance, and hence non-quantified circumstances of each outcome, occurred at the accident and municipality levels. A multilevel model cannot explain the reasons for the presence of this intra-unit correlation. Nevertheless, where the technique is used to obtain conservative estimates of its magnitude and form, it can assist in further hypothesis formulation. For example, it is likely that a large proportion of the crude geographical disparities in outcomes between municipalities mapped in Fig. 5 are associated with random influences. However, a comparison of the multilevel and non-multilevel residual plots in Figs. 6 and 7 show how the multilevel residuals have been smoothed whereby values upon which little certainty can be placed are shrunken towards the mean. This has the effect of removing much of the unavoidable randomness in the data and allows the reasons for the relative ranking of units at each level of the hierarchy to be investigated with more confidence.

The exact nature of the remaining residual variation that remains in the models will differ between incidents, but is likely to be associated with factors such as the design and build quality of the vehicles involved, and the specific forces associated with the impact. Paramedical response times may also have a role to play, whilst broad variations in the accessibility of hospital emergency facilities might explain some of the observed variation between municipalities. Certainly, the presence of substantial residual variation illustrates some of the weaknesses in routinely collected data, and indicates the potential for further research to qualify these effects.

Although the essential ideas of multilevel models we developed over 20 years ago, it is only recently that improvements in computing power and advances in the understanding of effective model implementation have meant that their implementation has become a practical proposition (Bull et al., 1998). We are currently on a wave of innovation as use spreads from the original developers to the wider research community. Having said that, the multilevel approach retains some of the limitations of more traditional quantitative techniques, as well as introducing new ones.

In the research presented here, the influences on casualty outcomes are modelled more powerfully than traditional techniques allow, yet the random parameters can ultimately offer only limited insight into the reasons behind between-accident and between-place variations in outcome. In addition, technical limitations remain that are associated with the estimation of non-linear models using IGLS. In these cases, the level one variance is allowed to be distributional (in this case, possessing a binomial distribution), but the variance at higher levels is always assumed to be normal. This assumption can lead to the generation of biased estimates of higher level variances. A solution that is currently being developed may lie with application of computationally intensive simulation methods such as Gibbs and Metropolis Hastings Sampling and Bootstrapping.

As with any other analytical method, multilevel models are limited in their power by the shortcomings of the data used to construct them. In this instance, injury related information was crude, limiting a detailed consideration of injury severity. From a public health point of view, the comparison between uninjured individuals and those receiving injuries could be of interest, although this was not made here, as no information was available on uninjured residents involved in accidents. A further limitation of this analysis is associated with the fact that the probability of a trip ending in an accident is likely to be related the factors used to explain severity. Hence, the sample we have analysed is somewhat self-selected. It would have been preferable to model the outcome for each casualty conditionally on the accident dating place, although no information on general characteristics of each journey was available to us.

The ultimate health outcome for a casualty involved in a road traffic accident will always involve a complex interplay between a wide range of factors that are difficult to quantify and may be subject to random variation. This unpredictability will undoubtedly introduce a considerable amount of uncertainty into any model developed to identify and predict the important influences on casualty survival. Whilst multilevel models cannot remove this uncertainty, they can allow it to be quantified and accounted for, providing significant advances over more traditional techniques. Given the strongly hierarchical nature of most accident data sets, the use of these models should be seriously considered in future research.

Acknowledgements

We thank the Norwegian Public Roads Administration (Statens Vegvesen, Vegdirektoratet) for providing the accident data used in this study. This research collaboration was made possible by the provision of a travel grant from the British Council in Oslo.

References

Browne, W.J., Draper, D., 2000. Implementation and performance issues in the Bayesian and likelihood fitting of multilevel models. Computat. Statistics 15, 391–420.

Bryk, A.S., Raudenbush, S.W., 1992. Hierarchical Linear Models. Sage, Newbury Park.

Bull, J.M., Riley, G.D., Rasbash, J., Goldstein, H., 1998. Parallel Implementation of A Multilevel Modelling Package. University of Manchester, Manchester, UK.

Duncan, C., Jones, K., Moon, G., 1998. Context, composition, and heterogeneity: using multilevel models in health research. Soc. Sci. Med. 46 (1), 97–117.

Gilks, W.R., Richardson, S., Spiegelhalter, D.J., 1996. Markov Chain Monte Carlo in Practice. Chapman and Hall, London.

Goldstein, H., 1995. Multilevel Statistical Models, 2nd Edition. Edward Arnold, London.

Hoijtink, H., 1998. Bayesian and Bootstrap Approaches to Testing Homogeneity and Normality in a Simple Multilevel Model. Utrecht University, Utrecht.

Jones, A.P., Bentham, G., 1995. Emergency medical service accessibility and outcome from road traffic accidents. Public Health 109, 169–177.

Kim, K., Nitz, L., Richardson, J., Li, L., 1995. Personal and behavioral predictors of automobile crash and injury severity. Accident Anal. Prevent. 4, 469–481.

Korn, E.L., Graubard, B.I., 1990. Simultaneous testing of regression coefficients with complex survey data. Am. Statistician 44, 270–276.

Lin, X., 1997. Variance component testing in generalised linear models with random effects. Bimetrika 84, 309–325.

Longford, N.T., 1993. Random Coefficient Models. Clarendon Press, Oxford.

Menard, S., 1995. Applied Logistic Regression Analysis (Quantitative Applications in the Social Sciences, Vol. 106). Sage, New York.

Murphy Jr., R.X., Birmingham, K.L., Okunski, W.J., Wasser, T., 2000. The influence of airbag and restraining devices on the patterns of facial trauma in motor vehicle collisions. Plastic Reconstruct. Surg. 105 (2), 516–520.

Noy, Y.I., 1997. Human factors in modern traffic systems. Ergonomics 40 (10), 1016–1024.

Rasbash, J., Browne, W., Goldstein, H., Yang, M., Plewis, I., Healy, M., Woodhouse, G., Draper, D., Langford, I., Lewis, T., 2000. A Users Guide to MlwinN, Version 2.1. Insititute of Education, London.

Shankar, V., Mannering, F., Barfield, W., 1996. Statistical analysis of accident frequency of rural freeways. Accident Anal. Prevent. 28 (3), 391–401.

Simpson, H., 1996. Comparison of Hospital and Police Casualty Data: A National Study, TRL Report 173. Transport Research Laboratory, Crowthorne, UK.

Skinner, C.J., Holt, D., Smith, T.M.F., 1989. Analysis of Complex Surveys. Wiley, Chichester, UK.

World Health Organisation (WHO), 1997. Health and Environment in Sustainable Development: Five Years After The Earth Summit. WHO, Geneva.