Pergamon

Accid. Anal. and Prev., Vol. 30, No. 5, pp. 641–649, 1998

Printed in Great Britain
0001-4575/98 $19.00 + 0.00

PII: S0001-4575(98)00009-8

THE INFLUENCE OF TREND ON ESTIMATES OF ACCIDENTS AT JUNCTIONS

LINDA MOUNTAIN[1]*, MIKE MAHER[2] and BACHIR FAWAZ[1]

[1]Department of Civil Engineering, University of Liverpool, Liverpool, UK
and [2]Department of Civil and Transportation Engineering, Napier University, Napier, UK

(Received 17 September 1997; in revised form 18 January 1998)

Abstract—While reliable estimates of expected accidents can be achieved by combining observed accidents and accident model predictions using an empirical Bayes approach, there are a number of obstacles to the widespread adoption of the method. This paper concentrates on problems associated with the available predictive models. Of particular concern is the effect on model predictions of accident trends over time resulting from, for instance, traffic growth or national road safety programmes. Since accident models invariably include traffic flow as an explanatory variable, the effects of flow changes can be included provided that account is taken of the non-linear relationship between accidents and exposure. It is, however, common to assume that accident risk per unit of exposure is constant over time, whereas national data imply that accident risk is declining. In addition, there is a need, in practice, to rank and evaluate remedial sites in terms of the specific accident types or severities which might be targeted by treatment (for example, wet road accidents in the case of surface treatment). This then raises the question of whether the proportions of accidents of various types varies over time or with traffic flow and site characteristics. Generalized linear modelling was used to develop regression estimates of expected junction accidents (both in total and disaggregated by severity, road surface condition and lighting condition) which allow for the possibility of accident risk varying over time. Accident risk at the sample of some 500 junctions was shown to be declining annually by an average of 6%, with no significant difference in the value of trend between accident types. The factors which affected the proportions of accidents of various types included the method of junction control, speed limit and traffic flow. © 1998 Elsevier Science Ltd. All rights reserved

Keywords—Generalized linear models, Highway safety, Junction accident risk, Trend

INTRODUCTION

The prioritization of sites for accident remedial treatment, and the monitoring and evaluation of treatment effectiveness requires a knowledge of the accidents which would occur in the absence of treatment. Estimates based on observed accidents in the period prior to possible treatment are notoriously unreliable due to factors such as regression-to-mean and trend effects and it has been established that better estimates can be achieved using an empirical Bayes (EB) technique [see, for example, Persaud and Dzbik (1993), Kulmala (1994) and Mountain et al. (1995)]. In this the true underlying mean accident frequency (m) at a site is estimated as a weighted combination of the observed accidents (x_b) and a predictive model estimate of the expected accident frequency (μ). The inclusion of x_b allows some account to be taken of specific site characteristics not included in the prediction model while μ smoothes out random variation, controlling for regression-to-mean effects. The quality and usefulness of the EB estimates will depend on the availability and precision of suitable predictive models and the cost of the data required to apply them.

Over the years numerous models have been developed relating accidents to various measures of flow, site characteristics and geometry. While early models were based on the assumption of a normal error structure using classical least squares regression modelling, it is now accepted that the use of generalized linear models, with a quasi-Poisson or (preferably) negative binomial error structure, is more appropriate [see, for example, Joshua and Garber (1990), Miaou and Lum (1993) and Maher and Summersgill (1996)]. Thus more recent models have been developed using statistical programs specifically designed for fitting generalized linear models, such as GLIM (Francis et al., 1993) and GENSTAT (Lane

*Corresponding author. Tel.: 0151 794 2000; Fax: 0151 794 5218; e-mail: l.mountain@liv.ac.uk

et al., 1988). In spite of these advances, from the point of view of remedial site selection and monitoring, there remain a number of problems associated with the available predictive models.

Of particular concern is the effect on model predictions of accident trends over time due to traffic growth and the effects of local or national road safety policies and programmes. Since accident models invariably include traffic flow (Q) as an explanatory variable, the effects of flow changes over time can be included provided that account is taken of the non-linear relationship between accidents and exposure. Thus, for example, a simple model would have the form

$$\mu = \alpha Q^{\beta}$$

where α is the accident *risk* per unit of exposure (Q^{β}). Errors only arise where accidents are implicitly assumed to be proportional to traffic flow (i.e. $\beta = 1$) and accident *rates* per million-vehicles or million-vehicle-kilometres used. Although accident rates are commonly quoted most models indicate a sub-linear relationship between accidents and flow [see, for example, Hauer (1995), Summersgill et al. (1996) and Mountain and Fawaz (1996)].

More of a problem is the effect of road safety policies and programmes on accident risk since, if these are effective, it might be hoped that α would actually decline over time. It is, however, universal practice in accident modelling to assume that α is constant. If, however, risk changes from year to year by a factor γ, better estimates might be obtained for the expected number of accidents in year t, μ_t, using a model of the form

$$\mu_t = \alpha_0 \gamma^t Q_t^{\beta}$$

where α_0 is the risk in year 0 and Q_t is the flow in year t (Maher and Summersgill, 1996). Such a model for the trend effect assumes that the 'general safety development' due to safety programmes is the same from year to year. Clearly such a model is inevitably an oversimplification, as there could be larger effects in some years due perhaps to a change in legislation or reporting practise. Furthermore, it is primarily intended for application to circumstances where it is expected that the risk is reducing over time (that is $\gamma < 1$); in cases where $\gamma > 1$, the model will lead to an 'explosion' of the accident frequency. Also, alternative forms for the trend could be assumed instead of the simple geometric: for example, a linear function of time, or one in which the risk tended towards a non-zero lower bound. However, it was felt that the geometric form was the most simple and convenient and that it was most unlikely that the data would support the fitting at this stage of any more complex form.

To obtain an estimate of the average national trend, a model of this form was fitted to the U.K. national accident data for the period, 1975–1995. In this μ_t was the total annual accidents in the U.K. in year t ($t = 0$ for, 1975) and Q_t was the total national annual traffic volume in year t (measured in vehicle-kilometres). Using GLIM with a quasi-Poisson error structure and log link function, γ was estimated to be 0.980 with *SE* 0.005 implying that, for the period, 1975–1995 accident risk in the U.K. declined annually by an average of 2%, accumulating into a reduction of 33% over the whole period. Although the annual rate of decline is small it is statistically significant ($p < 0.01$) and in a relatively short period of time has a cumulative effect which is of practical significance. Thus the national data imply that accident models which do not allow for this effect will rapidly become outdated, in the long term resulting in substantial over-estimates of accident frequencies. In addition, where the models are to be used in an EB approach to evaluate the effects of remedial treatment, the inclusion of γ would allow the trend in accident risk to be separated from the effects of treatment.

A second problem is the need to be able to rank and evaluate remedial sites in terms of the specific accident types or severities which might be targeted by treatment. So, for example, in the case of surface treatment, it is common to select sites on the basis of the observed wet road accidents and to evaluate treatment effectiveness in terms of the reduction in wet road accidents. A similar argument can be made for models of accidents disaggregated by severity and lighting conditions. This then raises the question of whether the proportions of accidents of various types varies with site characteristics, traffic flow or over time.

This paper presents models for accidents at junctions (for both aggregate accident totals and accidents disaggregated by severity, road surface condition and lighting condition) which allow for the possibility of accident risk varying over time. In developing the models the objective was to improve on the quality of estimates based on observed accidents alone while keeping additional data requirements (and hence costs) to a minimum. The models are designed to give as reliable as possible predictions of expected accident frequencies based on data that could be reasonably expected to be readily available for all candidate sites at the site selection stage. The combination of these estimates with the observed accidents at specific sites using the EB method allows some account to be taken of factors not included in the prediction models. The models demonstrate the importance of allowing for the effects of trend in accident models and give an indication of the factors which affect the proportions of accidents of various types.

THE DATA

The data used in this study comprised details of highway and junction characteristics, personal injury accidents (PIAs) and traffic flows on networks of A- and B-roads in six U.K. counties for periods of between 5 and 15 years between 1980 and 1994. Major junctions were defined as junctions between two A-roads or an A- and B-road, and other junctions as minor junctions. The networks included some 3400 km of highway, with 501 major junctions and some 5000 minor junctions. Models for accidents on links and at minor junctions are dealt with in a separate paper (Mountain et al., 1997). Of the 501 major junctions 300 were priority junctions, 105 were roundabouts and 96 were traffic signals. Table 1 shows the distribution of these junctions by traffic volume: the road carrying the largest total entry flow being defined as the *major* road. It will be noted that priority junctions are reasonably well distributed across the range of major road entry flows but tend to have low minor road entry flows. Traffic signals and roundabouts on the other hand tend to have high major and minor road entry flows.

For the purposes of this study, the relevant information concerning each accident was its date, location, severity, road surface condition and lighting condition. Accidents occurring within 20 m of the extended kerblines of a junction were classified as junction accidents. Data were obtained concerning a total of some 25,000 PIAs of which some 3200 occurred at major junctions.

In developing the models, the aim was to restrict the explanatory variables and factors to quantities which would be routinely available to the road safety engineer so that few or no additional data costs would be involved beyond those incurred in calculating accident rates. The junction characteristics included were thus the number of arms and method of control, together with the major road carriageway type and speed limit. The flow data comprised the total entry flows via the major and minor road approaches. Records of all traffic counts carried out on the approaches to each junction during the study period were obtained and converted to annual average daily traffic flows (AADT) using appropriate factors (Department of Transport, 1991). For some junctions, traffic counts were not available for every year of the study period. Flows for years without counts were obtained by linear interpolation. Overall interpolated flows accounted for $<25\%$ of the flow data. The data were stored in a database, using the relational database package FOXPRO (Lupton et al., 1996).

It should be stressed that the data sample included only junctions located outside the central areas of major conurbations and thus the models developed from the data may only be applicable to such junctions. A recent series of U.K. junction accident studies which included central area junctions in fact found only a London factor, with the risk of junction accidents in London between 1.3 and 1.4 times higher than elsewhere (Layfield et al., 1996; Summersgill et al., 1996; Taylor et al., 1996).

Table 1. Distribution of junction sample by traffic volume

Method of control	Major road inflow [AADT (vehicle day^{-1})]	Number of junctions with minor road inflow [AADT (vehicle day^{-1})]		
		<2000	2000–4000	>4000
All types	<6000	86	21	6
	6000–12,000	82	64	40
	$>12,000$	46	45	111
Priority	<6000	79	15	4
	6000–12,000	71	42	21
	$>12,000$	24	22	22
Traffic signals	<6000	2	2	2
	6000–12,000	6	10	9
	$>12,000$	11	12	42
Roundabouts	<6000	5	4	0
	6000–12,000	5	12	10
	$>12,000$	11	11	47

AGGREGATE JUNCTION MODELS WITHOUT TREND

Initially the models were fitted without trend using a generalized linear modelling approach with a negative binomial error structure employing the GLIM statistical package. On the basis of the findings of an earlier study (Mountain and Fawaz, 1996), the fitted models were of the form

$$\mu = \alpha Q^{\beta_1} q^{\beta_2}$$

where μ is the expected number of accidents per year, Q is the major road inflow (AADT in 1000 vehicles day^{-1}), q is the minor road inflow (AADT in 1000 vehicles day^{-1}) and α, β_1 and β_2 are parameters to be estimated.

Since the annual accident frequency is not a true Poisson variate, the models were fitted using the total number of accidents occurring during the whole t year study period at each junction as the unit of data.

[The term $\log(T)$ was used as an *offset variable* in the fitting process so that its coefficient was constrained to 1 and consequently the resulting models gave estimates of annual accident frequencies.]

In order to investigate the effects of the main junction features on accident frequencies it was necessary to group the data into mutually exclusive subsets. This is achieved in GLIM by using *factors* or dummy variables which take a sequence of integer values or *levels* for each subset of data. The factors used for the junction models are summarized in Table 2. In fitting the models it was found that there was no significant difference in the parameter estimates for traffic signals and roundabouts. Thus, the three level factor, JTYP, was replaced by the two level factor, JTYP2.

Details of the fitted models are shown in columns 3 and 4 of Table 3. It will be noted that the only factor which has a significant effect on the estimates is the method of junction control and that the estimates of β_1 and β_2 are <1. For all methods of control the estimated value of β_1 is 0.546 implying that a doubling of the major road entry flow would increase accidents by 46%. The estimated values of β_2 imply that doubling the minor road entry flow would increase accidents by some 13% at priority junctions and by 33% at signals and roundabouts. The use of accidents *rates* at junctions (i.e. accidents per million vehicles) implies a doubling of accidents in response to a doubling of the total vehicle inflow (i.e. $Q+q$). The models suggest that the accident increase would, in practice, depend on the distribution of any flow increases between the major and minor arms and the relative magnitudes of the flows. If both major and minor road inflows doubled, the models would predict an increase in accident frequency of 65% at priority junctions and of 92% at traffic signals and roundabouts.

Table 2. Factors tested for junction models

Name	Description (number of sites)	
CWY	Carriageway type	1 = Single (429) 2 = Dual (72)
ARMS	Number of arms	1 = Three-arm junction (379) 2 = Four- or more arm junction (122)
URBR	Speed limit	1 = ≤40 mph (urban) (225) 2 = >40 mph (rural) (276)
JTYP	Method of junction control	1 = Major–minor priority (300) 2 = Traffic signals (96) 3 = Roundabouts (105)
JTYP2	Method of junction control (two levels only)	1 = Major–minor priority (300) 2 = Traffic signals and roundabouts (201)

The estimated values of β_2 would seem to imply that, for similar levels of major and minor entry flow [for minor road entry flows >1000 vehicles per day (i.e. $q>1$)], priority junctions are safer than traffic signals and roundabouts, which is contrary to expectations. It must be stressed that, given the limited range of factors and variables included, the relationships represented by the models are likely to be associative rather than causal. Thus, in the case of priority junctions, it is probable that this method of control is not used at inherently dangerous junctions such as those with large turning flows, high levels of pedestrian activity, high approach speeds or restricted visibility. Such junctions are more likely to be controlled by traffic signals or a roundabout. To be able to draw conclusions about the underlying causal factors it would be necessary to include in the models all the site characteristics and variables thought likely to influence safety. Only then would the effect of method of control be compared at junctions which are similar in all other respects and a causal relationship be established. While such models would be essential for assessing the safety implications of particular junction features this is not the purpose of the models presented here.

AGGREGATE JUNCTION MODELS WITH TREND

In modelling accidents without trend, the models were fitted using the total number of accidents occur-

Table 3. Parameter estimates for aggregate junction accident models

		Without trend		*With* trend	
Model parameter (1)	Method of control (2)	Parameter value (3)	Standard error (4)	Parameter value (5)	Standard error (6)
α	All	**0.167**	*0.030*	**0.280**	*0.012*
β_1	All	**0.546**	*0.086*	**0.536**	*0.084*
β_2	Priority	**0.182**	*0.069*	**0.183**	*0.067*
	Signals and roundabouts	**0.406**	*0.060*	**0.407**	*0.058*
γ	All	—	—	**0.938**	*0.026*
Shape parameter, k	All	**1.523**		**1.551**	

ring during the whole t year study period at each junction as the unit of data. In considering trend in accidents, it would perhaps seem more natural to disaggregate the data so that accidents occurring in each year at each junction represent a unit of data. The problem with using this type of disaggregated data is that the error structure becomes a mixture of random between site errors and highly correlated within site errors. Maher and Summersgill (1996) suggest that it is preferable to use the aggregated form of the data. This, however, complicates the problem of incorporating trend into the model. Let the expected number of accidents in year t, μ_t, be given by

$$\mu_t = \alpha_0 \gamma^t Q_t^{\beta_1} q_t^{\beta_2}$$

where Q_t and q_t are the major and minor entry flows in year t (so that the risk—accidents per unit of exposure—decreases geometrically over time). The total expected accidents, $E(A)$, at a junction over the whole of its study period (between years t_1 and t_2) is then

$$E(A) = \sum \mu_t = \alpha_0 \left(\sum \gamma^t Q_t^{\beta_1} q_t^{\beta_2} \right)$$

with the summations over the study period (i.e. $t = t_1$ to $t = t_2$, where t_1 and t_2 may vary from junction to junction). Before fitting, the model must be transformed to an additive, linear form. In this case, however, the standard $\log_e$ transformation gives

$$\eta = \log[E(A)] = \log(\alpha_0) + \log(\sum S_t)$$

where

$$S_t = \gamma^t Q_t^{\beta_1} q_t^{\beta_2}$$

which is not in standard, additive linear form (as far as the parameters γ, β_1 and β_2 are concerned) and hence cannot be fitted using the standard approach in GLIM. However, if the values of γ, β_1 and β_2 were known, the S_t could be calculated and hence α_0 could be estimated by regression. Building on this idea, Maher and Summersgill (1996) suggest an iterative approach using *constructed variables* [see Atkinson (1985)]. In this, initial estimates of γ, β_1 and β_2 (say $\gamma = \beta_1 = \beta_2 = 1$) are adjusted by means of a series expansion of the expression above for η in the following two-stage process:

(1) Using the current estimates of γ, β_1 and β_2, calculate the S_t and fit the model above using $\log(\Sigma S_t)$ as an offset variable. The fitted value is $\hat{A}$.

(2) Constructed variables X_1, X_2 and X_3 are then calculated as

$$X_1 = \frac{[\Sigma(\log Q_t) S_t]}{\Sigma S_t}$$

$$X_2 = \frac{[\Sigma(\log q_t) S_t]}{\Sigma S_t}$$

$$X_3 = \frac{(\Sigma t S_t)}{\Sigma S_t}$$

and GLIM is used to fit accidents (A), with X_1, X_2 and X_3 as explanatory variables and $\log(\hat{A})$ as an offset variable. The linear predictor is of the form

$$\eta = \log[E(A)] =$$

$$\log(\hat{A}) + \Delta \log(\alpha_0) + \Delta\beta_1 X_1 + \Delta\beta_2 X_2 + \left(\frac{\Delta\gamma}{\gamma}\right) X_3$$

and the output gives estimates of $\Delta \log(\alpha_0)$, $\Delta\beta_1$, $\Delta\beta_2$ and $\Delta\gamma/\gamma$ which are the (small) changes in the values of $\log(\alpha_0)$, β_1, β_2 and γ from their current values. Using the updated values of β_1, β_2 and γ, return to (1) and repeat until convergence is achieved (generally in two or three iterations). (For the purpose of completeness, it should be noted that in the cases where the exponent of minor road flow was allowed to be different for the different junction types, so that an interaction term was included in the model, a further constructed variable was required in addition to those above.)

Details of the fitted model are given in Table 3, columns 5 and 6 (which can be compared with the fit without trend in columns 3 and 4). The only parameter estimates significantly affected by the inclusion of trend are those of α. It can be seen that the estimated value of γ is significantly <1 ($p < 0.05$), and that the shape parameter, k, is larger (implying that the degree of overdispersion is reduced).

It will be noted that, at 0.938, the estimated value of γ is somewhat smaller than the estimated national average value (0.980). This may be attributed to the differences in the basic characteristics of the data sets. For example, the sample data used to develop the models summarized in Table 3 relates only to junctions outside major conurbations in six counties and these models relate accidents to specific junction characteristics and entry flows, rather than to overall national network volumes. In addition, in Table 3, the values of the flow exponents β_1 and β_2 are *primarily* found by modelling the between-site variation in stage 1 and the value of γ (which models the between-years effect) is found in stage 2 (that this is the case can be seen from a comparison in Table 3 of the estimates of β_1 and β_2 for the without trend and with trend cases; the values are virtually identical indicating that the estimation of β_1 and β_2 and of γ is almost uncoupled). In the case of the national data, both parameters are found from the between-years variation. The estimate of trend at junctions is in fact not significantly different from the value of 0.948 estimated for highway links (Mountain et al., 1997).

Although, at 6% per year, the annual rate of decline in accident risk is not particularly large, it has a considerable effect on estimated accidents over the study period: the model without trend under-estimates accidents in 1980 ($t=0$) by about 40% and over-estimates accidents in 1995 ($t=15$) by about 55%. Since it is inevitable that accident prediction models are based on historical data, the inclusion of trend is vital if the models are to be used to predict current or future accidents. Even a model based on the most recent five full years data (1992–1996) would over-estimate accidents in 1997 by some 17% if an annual trend of 0.938 was ignored.

DISAGGREGATE JUNCTION MODELS WITH TREND

The next stage of the analysis was to develop disaggregate predictive models. Inspection of the data revealed that only some 2% of major junction accidents were classified as fatal and thus fatal and serious accidents were modelled together giving two categories for severity. Some 20% of the junction accidents were fatal or serious. In the case of road surface condition, again only some 1% of major junction accidents were classified as occurring in ice, snow or flood and thus these accidents were classified with wet accidents to give two categories: wet etc. and dry. Some 36% of the junction accidents were classified as wet etc. Two categories were used for lighting condition: daylight and night (dark).

In modelling accidents disaggregated by lighting condition an important explanatory variable for night accidents would seem likely to be the presence or absence of street lighting and ideally this should be included as a factor in the model. This information is included on the STATS19 form (the standard injury accident report form used in the U.K.) and hence is available for junctions where accidents have occurred. There will, however, inevitably be a number of junctions with no recorded injury accidents over the study period. In the case of the sample data, junctions with no recorded injury accidents accounted for 43% of the total sample. Street lighting inventories were sought from local authorities but the information was not generally available in a form which could be readily related to the other network data held on the database. Clearly a visual inspection of the network would have been time consuming and would not necessarily represent the lighting situation throughout the study period. Accidents were thus modelled without a factor to reflect the availability of street lighting. It is, however, worth noting that street lighting was almost invariably available at traffic signals and roundabouts. At these junctions some 30% of accidents occurred at night but only 1% of accidents were classified as night/unlit. This compares with priority junctions where some 24% of accidents occurred at night with 7% classified as night/unlit.

The model fitting process was repeated as before but with accidents of the relevant type replacing the total accidents as the dependent variable. Initially the models were fitted without trend but again the inclusion of trend improved the model fit, resulting in larger shape parameters. A summary of the estimated model parameters is given in Table 4 with a summary of the models in Table 5. It will be noted that none of the estimated values of γ were significantly ($p<0.05$) different from 0.938 (the value of γ estimated for aggregated accidents) implying that there was no significant difference in the value of trend between accident types.

PROPORTIONS OF JUNCTION ACCIDENTS OF VARIOUS TYPES

The parameter estimates in Table 4 indicate that there tends to be some variation in the proportion of accidents of various types with method of control and entry flow, although it is clear from the values of the standard errors that not all of these variations are significant. In order to obtain a clearer picture of which factors and variables did have a significant effect on the proportions of accidents of various types, the proportion data were modelled using GLIM with a binomial error structure and logit link function. So, for example, in the case of accidents disaggregated by severity, the proportion of fatal and serious accidents was $p_{f\&s}$, the proportion of slight accidents was then $(1-p_{f\&s})$ or p_{sl} and the pure binomial variance ($Ap_{f\&s}p_{sl}$), where A was the accident total at the junction over the period of observation as before. In practice there was some degree of overdispersion in the data. This was allowed for, using a similar procedure to that used in the quasi-Poisson model, namely by multiplying the binomial variance by a factor, K^2, calculated as (χ^2/degrees of

Table 4. Parameter estimates for disaggregate junction accident models with trend

Model parameter	Junction type	Accidents disaggregated by: Severity — Fatal and serious: Parameter value	Fatal and serious: Standard error	Severity — Slight: Parameter value	Slight: Standard error	Road surface condition — Dry: Parameter value	Dry: Standard error	Road surface condition — Wet, snow etc.: Parameter value	Wet, snow etc.: Standard error	Lighting condition — Day: Parameter value	Day: Standard error	Lighting condition — Night: Parameter value	Night: Standard error
α	All	**0.043**	*0.002*	**0.237**	*0.011*	**0.152**	*0.007*	**0.107**	*0.005*	**0.190**	*0.009*	**0.061**	*0.003*
β_1	All	**0.642**	*0.116*	**0.502**	*0.089*	**0.617**	*0.092*	**0.482**	*0.099*	**0.593**	*0.088*	**0.557**	*0.114*
β_2	Priority	**0.225**	*0.071*	**0.179**	*0.072*	**0.198**	*0.074*	**0.308**	*0.062*	**0.297**	*0.055*	**0.196**	*0.092*
	Signals and roundabouts	**0.225**	*0.071*	**0.461**	*0.062*	**0.433**	*0.062*	**0.308**	*0.062*	**0.297**	*0.055*	**0.522**	*0.074*
γ	All	**0.956**	*0.036*	**0.935**	*0.027*	**0.928**	*0.027*	**0.954**	*0.032*	**0.938**	*0.027*	**0.940**	*0.033*
Shape parameter k	All	**1.256**		**1.408**		**1.604**		**1.534**		**1.523**		**1.152**	

freedom). This does not affect the parameter estimates but increases their standard errors by a factor K. With the logit link function, the linear predictor is not for $p_{\mathrm{f\&s}}$ but for the logit transformation of $p_{\mathrm{f\&s}}$, $\log_e(p_{\mathrm{f\&s}}/p_{\mathrm{sl}})$. Thus the linear predictor is

$$\eta = \log_e(p_{\mathrm{f\&s}}/p_{\mathrm{sl}}) = \log(\alpha_0) + \beta_1 \log(Q) + \beta_2 \log(q)$$

and the models summarized in Table 6 are for the odds ratios, $p_{\mathrm{f\&s}}/p_{\mathrm{sl}}$, not $p_{\mathrm{f\&s}}$. The value of $p_{\mathrm{f\&s}}$ can, however, be determined straightforwardly using

$$p_{\mathrm{f\&s}} = (1 + 1/\text{odds ratio})^{-1}$$

For accidents disaggregated by road surface condition p_{wet} and p_{dry} denote the proportions of wet etc. and dry accidents respectively while, for accidents disaggregated by lighting condition, the proportions of dark and light accidents are p_{dark} and p_{light}.

From Table 6 it can be seen that, for accidents disaggregated by severity, the ratio of fatal and severe to slight accidents varies only with the method of control, with a higher value at priority junctions (by a factor of 1.6). In the case of accidents disaggregated by road surface condition, the ratio of wet to dry accidents varies with speed limit and the major road entry flow: the ratio is higher on rural roads (by a factor of about 1.2) and is slightly higher for smaller major road entry flows (perhaps reflecting the effects of higher approach speeds at lower flows). The ratio of night to daytime accidents is the same for all methods of control and speed limit but increases slightly with increases in the minor road entry flow.

Clearly, as an alternative to modelling disaggregated accidents separately, it is possible to use the estimates of the proportions of accidents of various types together with estimates of accident totals obtained using the aggregate models summarized in Table 5. Since a suitable independent data set was not available, the data used for developing the models was used to compare the errors in the estimates obtained using these alternatives. The results are summarized in Table 7. It will be noted that there appears to be little to choose between the two approaches, but a comparison using an independent data set would be worthwhile.

CONCLUSIONS

Junction accident models have been developed to predict both aggregate accident totals and accidents disaggregated by severity, lighting and road surface condition. The main findings can be summarized as follows:

(1) Effective road safety policies and programmes will result in a decline in accident risk from year

Table 5. Summary of best fit models with trend

<table>
<tr><th></th><th></th><th colspan="2">Method of control</th></tr>
<tr><th>Accident class</th><th></th><th>Major–minor priority</th><th>Traffic signals and roundabouts</th></tr>
<tr><td>All</td><td></td><td>$\mu_t\ (0.938)^t 0.280 Q^{0.536} q^{0.183}$</td><td>$\mu_t = (0.938)^t 0.280 Q^{0.536} q^{0.407}$</td></tr>
<tr><td>By severity</td><td>Fatal and serious</td><td colspan="2">$\mu_t = (0.956)^t 0.043 Q^{0.642} q^{0.225}$</td></tr>
<tr><td></td><td>Slight</td><td>$\mu_t = (0.935)^t 0.237 Q^{0.502} q^{0.179}$</td><td>$\mu_t = (0.935)^t 0.237 Q^{0.502} q^{0.461}$</td></tr>
<tr><td>By road surface condition</td><td>Dry</td><td>$\mu_t = (0.928)^t\ 0.152 Q^{0.617} q^{0.198}$</td><td>$\mu_t = (0.928)^t 0.152 Q^{0.617} q^{0.433}$</td></tr>
<tr><td></td><td>Wet, snow, ice etc.</td><td colspan="2">$\mu_t = (0.954)^t 0.107 Q^{0.482} q^{0.308}$</td></tr>
<tr><td>By lighting condition</td><td>Day</td><td colspan="2">$\mu_t = (0.938)^t 0.190 Q^{0.593} q^{0.297}$</td></tr>
<tr><td></td><td>Night</td><td>$\mu_t = (0.940)^t 0.061 Q^{0.557} q^{0.196}$</td><td>$\mu_t = (0.940)^t 0.061 Q^{0.557} q^{0.522}$</td></tr>
</table>

Table 6. Summary of best fit models for odds ratios

Accident disaggregated by:	Junction type	Odds ratio
Severity	Priority	$p_{\text{f\&s}}/p_{\text{sl}} = 0.318$
	Signals and roundabouts	$p_{\text{f\&s}}/p_{\text{sl}} = 0.199$
Road surface condition	Urban	$p_{\text{wet}}/p_{\text{dry}} = 0.800 Q^{-0.170}$
	Rural	$p_{\text{wet}}/p_{\text{dry}} = 0.983 Q^{-0.170}$
Lighting condition	All	$p_{\text{dark}}/p_{\text{light}} = 0.301 q^{\ 0.143}$

to year. This trend effect should be included in accident prediction models both to ensure that the models do not become rapidly outdated and (where the models are to be used in an EB approach to evaluate the effects of remedial treatment) to allow trend and treatment effects to be separated. Based on the data available for this study, it was estimated that accident risk at junctions is declining annually by an average of 6%, with no significant variation in trend with accident type.

(2) The use of accident *rates* at junctions (i.e. accidents per million vehicles) implies a doubling of accidents in response to a doubling of the total vehicle inflow. The models suggest that, in practice, the accident increases would generally be rather less than this.
(3) The ratio of fatal and severe to slight accidents depends on the method of junction control. The ratio of wet to dry accidents depends on the speed limit and the major road entry flow. The ratio of dark to night accidents depends on the minor road entry flow.
(4) Disaggregate accidents can be estimated either using separate models for each accident type or using estimates of proportions together with estimates of total accidents. There appears to be little to choose between these two approaches.

Acknowledgements—The authors gratefully acknowledge the financial support of EPSRC and the assistance of the staff of the county councils who supplied data for this project, namely Cheshire, Cornwall, Cumbria, Devon, Lancashire and Warwickshire.

Table 7. Comparison of model errors

	Accidents disaggregated by								
	Severity			Road surface condition			Lighting condition		
	Total	Fatal and serious	Slight	Total	Dry	Night	Total	Dry	Night
Observed mean annual accidents	0.927	0.180	0.747	0.927	0.591	0.337	0.927	0.678	0.250
Estimated mean annual accidents using:									
(a) Proportions	0.914	0.179	0.736	0.914	0.582	0.333	0.914	0.668	0.247
(b) Separate models	0.908	0.179	0.730	0.905	0.579	0.326	0.906	0.664	0.243
Mean error in estimates using:									
(a) Proportions	−0.013	−0.001	−0.011	−0.013	−0.009	−0.004	−0.013	−0.010	−0.003
(b) Separate models	−0.019	−0.002	−0.017	−0.022	−0.012	−0.011	−0.021	−0.014	−550.007
Root mean squared error using:									
(a) Proportions	1.003	0.267	0.849	1.003	0.683	0.402	1.003	0.702	0.427
(b) Separate models	1.003	0.269	0.847	1.007	0.682	0.409	1.015	0.717	0.425

REFERENCES

Atkinson, A. C. (1985) *Plots, Transformations and Regression: an Introduction to Graphical Methods of Diagnostic Regression Analysis*. Clarendon Press, Oxford.

Department of Transport (1991) *Traffic Appraisal Manual—the Guide to Good Practice*. HMSO, London.

Francis, B., Green, M. and Payne, C. (eds.) (1993) *The GLIM System: Release 4 Manual*. Clarendon Press, Oxford.

Hauer, E. (1995) On exposure and accident rate. *Traffic Engineering and Control* **36**, 3, 134–138.

Joshua, S. C. and Garber, N. J. (1990) Estimating truck accident rate and involvements using linear and Poisson regression models. *Transport Planning Technology* **15**, 41–58.

Kulmala, R. (1994) Measuring the safety effects of road measures at junctions. *Accident Analysis and Prevention* **26**, 6, 781–794.

Lane, P. W., Galwey, N. W. and Alvey, N. G. (1988) *GENSTAT 5: an Introduction*. Oxford University Press, Oxford.

Layfield, R. E., Summersgill, I., Hall, R. D. and Chatterjee, K. (1996) Accidents at urban priority cross-roads and staggered junctions on urban single carriageway roads. (Department of Transport, Report TRL 185): Transport Research Laboratory, Crowthorne, U.K.

Lupton, K., Fawaz, B., Jarrett, D., Mountain, L. and Wright, C. C. (1996) A database for assessing the safety effects of highway improvements. *Traffic Engineering and Control* **37**, 5, 321–326.

Maher, M. J. and Summersgill, I. (1996) A comprehensive methodology for the fitting of predictive accident models. *Accident Analysis and Prevention* **28**, 3, 281–296.

Miaou, S-P. and Lum, H. (1993) Modelling vehicle accidents and highway geometric design relationships. *Accident Analysis and Prevention* **25**, 6, 689–709.

Mountain, L. and Fawaz, B. (1996) Estimating accidents at junctions using routinely available input data. *Traffic Engineering and Control* **37**, 11, 624–628.

Mountain, L., Fawaz, B. and Jarrett, D. (1995) The influence of minor junctions on link accident frequencies. *Traffic Engineering and Control* **36**, 12, 673–679.

Mountain, L., Maher, M. J. and Fawaz, B. (1997) The effects of trend over time on accident model predictions. In *Proceedings of the 25th European Transport Forum*, Brunel University, September, 1997 PTRC, London.

Persaud, B. and Dzbik, L. (1993) Accident prediction models for freeways. *Transportation Research Record* **1401**, 55–60.

Summersgill, I., Kennedy, J. and Baynes, D. (1996) Accidents at three-arm priority junctions on urban single carriageway roads. (Department of Transport, Report TRL 184): Transport Research Laboratory, Crowthorne, U.K.

Taylor, M., Hall, R. D. and Chatterjee, K. (1996) Accidents at 3-arm traffic signals on urban single-carriageway roads. (Department of Transport, Report TRL 135): Transport Research Laboratory, Crowthorne, U.K.