

# A Probabilistic Quantitative Risk Assessment Model for Long-term Work Zone Crashes

Qiang Meng\*, Jinxian Weng, Xiaobo Qu

*Department of Civil Engineering, National University of Singapore, 117576, Singapore*

## ABSTRACT

Work zones especially long-term work zones increase traffic conflicts and cause safety problems. Proper casualty risk assessment in work zone is of importance for both traffic safety engineers and travelers. This paper develops a probabilistic quantitative risk assessment (QRA) model to evaluate the casualty risk that combines frequency and consequence of all accident scenarios in long-term work zone crash. The casualty risk is expressed by individual risk and societal risk. The individual risk is interpreted as the frequency of a driver/passenger being killed or injured and the societal risk describes the relationship between frequency and the number of casualties in work zone crash. The event tree diagram and consequence model are applied to determine the frequency and consequence of accident scenario, respectively. Seven intermediate events comprising age (*A*), crash unit (*CU*), vehicle type (*VT*), alcohol (*AL*), light condition (*LC*), crash type (*CT*) and severity (*S*) are considered in the event tree analysis. Since the probability of intermediate event may have a large variation, the uncertainty associated with the probability can be expressed by probability distribution function. The consequence model takes into account the combination effects of speed and Emergency Medical Service Response Time (*ERT*) on the consequence of work zone crash. Finally, a numerical example that utilizes the Southeast Michigan work zone crash data is carried out in this paper. The result shows that the presented model has the capability of evaluating the casualty risk in long-term work zone crash. There would be a 62% decrease of individual fatality risk and 44% reduction of individual injury risk if the mean speed were slowed down by 20%, while only be a 5% reduction of individual fatality risk and 0.05% reduction of individual injury risk if *ERT* were reduced by 20%. It suggests that slowing down speed is more effective than reducing *ERT* in casualty risk mitigation.

**Keywords:** Work zone, Crash, Quantitative risk assessment, Parameter uncertainty

## 1. Introduction

Work zone safety problem has become a high-priority issue for traffic engineer professionals. The presence of work zone increases traffic conflicts and causes severe traffic safety problems. As deemed by many researchers ((Ha and Nemeth, 1995; Roupail et al., 1988; Ullman et al., 2006), the occurrence likelihood of severe crash in work zone is higher than that in non-work zone and severe crashes would lead to casualties. Hereafter, casualties refer to fatalities or injuries among the drivers and passengers traveling across work zone in this study. As reported by (Daniel et al., 2000), most casualties are caused by vehicle crashes occurred in long-term work zone. Long-term work zone is defined as the work zone where work occupies more than 3 days (*MUTCD*, 2003). Therefore, there is a critical need for the assessment of

casualty risk in long-term work zone crash.

Although various models have been developed for the crash risk assessment by previous studies, most of these models emphasize on estimating the occurrence likelihood or frequency of work zone crash. However, in practice, traffic safety engineers seem to pay more attention on the casualty risk, i.e., the likelihood of a driver/passenger being killed or injured in work zone and the relationship between the frequency and consequence of work zone crash. So far, few work zone risk assessment studies have focused on the evaluation of casualty risk by simultaneously taking into account the frequency and consequence of work zone crash.

Fortunately, the quantitative risk assessment (QRA) model has been proven as an effective methodology quantitatively to assess the overall safety level of hazards (Meng et al. 2009; NUREG, 1975). It allows a quantitative assessment of a facility's risks rendered by a broad range of accidents from frequent-minor to rare-major accidents. This study therefore develops a QRA model in an attempt to evaluate the casualty risk combining frequencies and consequences of long-term work zone crashes. The individual risk and societal risk are used to express the casualty risk. The individual risk is interpreted as the frequency of a driver/passenger being killed or injured and the societal risk describes the relationship between frequency and the total number of casualties in work zone crash.

### *1.1. Literature review*

Two types of approaches are currently being used for the risk analysis of vehicle crash. One is the collective approach, characterized by crash frequency modeling (Abdel-Aty and Keller, 2005). Frequency of crashes is aggregated over specific time periods and locations. Some studies (Agresti, 2002; Miaou, 1994; Milton and Mannering, 1998) attempted to explore the relationship between crash frequency and explanatory variables, such as roadway and traffic conditions, by using Poisson or Negative Binomial (NB) regression models. These studies were based on the ability of the model formulation to capture the underlying distribution of the crash count data. Recently, some researchers (Abdel-Aty and Keller, 2005; Chang and Chen, 2005) proposed the distribution-free methodologies that are essentially driven by the observed data. These methodologies do not need inherent assumptions about the distribution of crash data.

The other is the individual approach, characterized by each individual crash case. This approach focuses on investigating the relationship between the occurrence probability of specified severity categorization of vehicle crash and the contributing factors of driver, vehicle, roadway, and environment characteristics. Various models including logistic regression model (Dissanayake and Lu, 2002; Li and Bai, 2006), logit model (Ouyang et al., 2002), ordered probit model (Khattak and Cassidy, 1999) and artificial neural network model (Mohamed and Hassan, 2004) have been applied to estimate the occurrence probability of vehicle crash. More advanced logit-based approaches, such as nested logit model (Shankar et al., 1996) and mixed logit model (Milton et al., 2008) were also employed to analyze the same issue. For the individual approach, the consequence of vehicle crash is implicitly evaluated according to the severity categorizations (i.e., possible injury, disabling injury or fatality). The severity categorization cannot quantitatively describe crash consequence in terms of the number of casualties.

In practice, the frequency of a driver/passenger being killed or injured and the frequency of major work zone crash has become a matter of great concern of traffic

safety engineers. Without taking into account crash frequency, Nilsson (1981) developed the Power Model to formulate the number of casualties as a function of mean speed. Evanco (1996) addressed the relation between the number of fatalities and the Emergency Medical Service Response Time (*ERT*) using the logistic regression method. As mentioned above, both collective and individual approaches are only available to evaluate crash risk. Fortunately, the QRA model provides a numerical evaluation of accident consequences, frequencies, and their combination into an overall risk measure (Borysiewicz et al., 2006). It is capable of assessing the casualty risk in work zone crashes.

A number of accident scenarios may be possible in work zone crash because a variety of contributing factors may affect the likelihood and the consequence. Previous studies (Dissanayake and Lu, 2002; Li and Bai, 2008; Ouyang et al., 2002) provide valuable evidences that the occurrence likelihood of work zone crash varies with the contributing factors comprising age, crash unit, vehicle type, alcohol, light condition, crash type and crash severity. These seven contributing factors can be regarded as intermediate events to estimate the occurrence likelihood of a particular accident scenario. The reliability of the occurrence likelihood is dependent upon the probabilities of these intermediate events. However, these probabilities may have large variations due to impacts of external factors. In this case, the occurrence probability may not be known with perfect confidence (Borysiewicz et al., 2006). A second probability concept uses a distribution pattern to reflect the uncertainty (variation) associated with the probability of intermediate event (Vose, 2000). The propagation of uncertainty associated with these intermediate event probabilities can be performed by using the Monte Carlo sampling approach (Kalos and Whitlock, 1986).

### *1.2. Objectives and contributions*

The objective of this paper is to develop a QRA model to assess the casualty risk that combines the occurrence frequency and consequence of crash in long-term work zone. To achieve this objective, it is required to estimate the frequencies and consequences of crash accident scenarios. Work zone crash frequency function is formulated by the conventional statistical method. The event tree diagram and consequence model are applied to determine occurrence frequencies and consequences of accident scenarios, respectively. In the event tree, the probability of intermediate event may have large variation/uncertainty. Therefore, we propose a new idea that the probabilities of these events are formulated as random variables that have probability distribution patterns. Finally, two risk expressions are utilized to express the casualty risk caused by long-term work zone crash.

The contribution of this paper comprises the following aspects:

- 1). This paper fills the research void in the assessment of casualty risk which simultaneously takes into account the frequency and consequence of crash in long-term work zone.
- 2). It presents an effective method to formulate work zone crash frequency and solve the problem that occurrence probabilities of events may have large variations.

## **2. Probabilistic Quantitative Risk Assessment Model Formulation**

Occurrence of a work zone crash may lead to various consequences including fatalities or injuries. Hence, there would be a number of possible accident scenarios

with distinct consequences for the work zone crash. These possible accident scenarios can be logically illustrated by an event tree diagram in which any path from the top event (i.e. the work zone crash) to an end node reflects one accident scenario with some intermediate events triggered by the work zone crash. The occurrence frequency of a particular accident scenario hence equals to the product of frequency of the work zone crash and the occurrence probability/likelihood of this scenario. The consequence of this accident scenario can be estimated by the consequence model. Therefore, the major task in model formulation is to estimate work zone crash frequency, build an event tree diagram and consequence model. The flowchart of QRA model formulation is depicted in Fig.1.

### 2.1. Frequency of work zone crash

Fig.2 depicts a typical long-term work zone configuration according to *MUTCD* (2003). Total work zone length  $L$  is defined as the sum of  $L_1$ ,  $l$  and  $L_2$ , where  $L_1$  is the length of advance warning and transition areas,  $l$  is the length of activity area and  $L_2$  is the length of termination area. The work zone crash frequency obviously increases with work zone length ( $L$ ) and work zone duration ( $D$ ). In addition, traffic volume ( $Q$ ) and road type ( $U$ ) are also two factors influencing work zone crash frequency. From a statistical viewpoint, the lower traffic volume on a work zones, the lesser effect on work zone crash it results in (Rouphail, 1988). Work zones located in rural areas have a higher crash potential than those in urban areas (Harb et al., 2008).

Khattak et al. (2002) pointed out that work zone length ( $L$ ), work zone duration ( $D$ ), traffic volume ( $Q$ ) and road type ( $U$ ) are the important determinants of crash frequency in work zone. To determine the relative importance of each factor, we can use the methods provided by SAS and random forest (Breiman, 2001). The random forest performs well in handling a very large number of variables with a large number of data sets. In this paper, we use SAS to determine the relative variable importance because there are only 14 available sets of Ohio long-term work zone crash data in 2002, listed in Table 1. For easier interpretation, the relative variable importance can be standardized, with the most important variable being assigned a relative variable importance of 100. With these 14 sets of data, SAS outputs the relative importance of 100 for work zone duration, 88.8 for work zone length and 80.0 for traffic volume. It suggests that these three factors are the most relevant predictors. Although road type ( $U$ ) is somewhat less influential, it also has roughly half of the relevance of traffic volume. Therefore, these four contributing factors are the determinants of crash frequency in work zone.

By using the conventional statistical approach, the frequency of work zone crash can be modeled as a function of these four explanatory variables. In this paper, road type is formulated by a binary variable  $U$  that is equivalent to one if it is urban road; zero if it is rural road. For estimation purpose, the function of work zone crash frequency can be expressed as the following form:

$$\ln(f) = \alpha_0 + \sum_{i=1}^4 (\alpha_i \times x_i) + \varepsilon \quad (1)$$

where  $f$  is work zone crash frequency,  $\alpha_0, \alpha_i, i=1, \dots, 4$  are coefficients to be estimated,  $\varepsilon$  is the random error term,  $x_i$  is the  $i^{\text{th}}$  explanatory variable (i.e., work zone length).

The explanatory variables of work zone length, traffic volume and work zone

duration can each have three forms. For example, the explanatory variable of work zone length can be expressed in terms of origin form ( $L$ ), logarithm form ( $\ln(L)$ ) and reciprocal form ( $1/L$ ). There are totally  $3 \times 3 \times 3 = 27$  combinations of forms for explanatory variables in Eq. (1). A trial and error method is applied to determine the best combination form. First, we enumerate all possible combination forms and then choose all feasible combination forms. The feasible combination form is defined as that can yield significant coefficients (all  $p$ -values less than 0.05). For these feasible combination forms, we select the form that has the best model performance (largest adjusted  $R^2$ ) as the optimal combination form.

Using the above trial and error method, we can determine the best functional form of Eq. (1) and the values of coefficients with the field data. For example, with work zone crash data listed in Table 1, we can obtain the trial and error results of the combination forms, shown in Table 2. It can be found that there are four feasible combination forms. Among these feasible combination forms, the best combination form of the four explanatory variables is found to be  $(\ln(L), 1/D, 1/Q, U)$  because it yields the largest adjusted  $R^2$  (95.4%). Therefore, the crash frequency function can be expressed by the following form:

$$\ln(f_i) = \alpha_0 + \alpha_1 \times \ln(L_i) + \frac{\alpha_2}{D_i} + \frac{\alpha_3}{Q_i} + \alpha_4 \times U_i + \varepsilon_i \quad (2)$$

The results of the coefficients  $\alpha_0, \alpha_1, \alpha_2, \alpha_3$  and  $\alpha_4$  are shown as follows:

Parameter	Estimate	Std Err	$p$ -value
$\alpha_0$	6.12	0.24	<.01
$\alpha_1$	0.429	0.13	<0.01
$\alpha_2$	-215	37.8	<0.01
$\alpha_3$	-66468	11278	<0.01
$\alpha_4$	-0.235	0.10	0.04

Therefore, the crash frequency function in long-term work zone can be expressed by:

$$f(L, D, Q, U) = e^{6.12 - 215/D - 66468/Q - 0.235U} \times L^{0.429} \quad (3)$$

Note that the work zone crash frequency function shown by Eq.(3) is only valid for  $1.17 \leq L \leq 5.9$ ,  $41,192 \leq Q \leq 113,108$  and  $122 \leq D \leq 365$  since the observed data is only within this range.

## 2.2. Event tree building and occurrence probability calculation with parameter uncertainty for an accident scenario

In work zone crash, the occurrence probability of accident scenario will depend on the driver, vehicle, roadway, and environment conditions. The event tree is applied to determine all possible accident scenarios and their corresponding occurrence probabilities.

Since the occurrence likelihood of an accident scenario varies based on the age, crash unit, vehicle type, alcohol, light condition, crash type and crash severity, these seven contributing factors are regarded as intermediate events for the event tree. In general, the event tree is investigated from left to right. Thus, the event tree is started at the "Vehicle Crash" (top event) column and terminated at the "Severity" column, as shown in Fig.3. Due to space limits, the event tree is decomposed into three sub-event trees, namely sub-event tree (a), sub-event tree (b) and sub-event tree (c). Sub-event

tree (b) will continue from the intermediate event of vehicle type of the sub-event tree (a). Similarly, sub-event tree (c) will be linked to the intermediate event of light condition of the sub-event tree (b).

In the “Age” column of Fig.3, the factor of age is categorized into two types: young and others. This is because the specific safety needs of young (<25) might need to be separately addressed (Harb et al., 2008). Since heavy vehicles involved work zone crashes are severer than non-heavy vehicle involved crashes (Pigman and Agent, 1999), the work zone crash is classified into two groups: i) heavy vehicle involved in crash and ii) non-heavy vehicle involved in crash. According to the number of vehicles involved, four types of crash units are listed in the “Crash Unit” column, which are respectively 1-unit, 2-unit, 3-unit and  $\geq 4$  unit. In addition, light condition is categorized into daylight, dark with lighting and dark without light because work zone crash is primarily occurred in these three types of light condition. According to the severity of crash, three types of work zone crashes comprising fatal crash, injury crash and property damage-only (PDO) crash are considered, as shown in Fig.3. In addition, all possible consequence of work zone crash can be described as: i) deaths and injuries in fatal crash; ii) injuries in injury crash and iii) property damage only in PDO crash, shown in the “Consequence Description” column.

A particular accident scenario depends on the path, which is represented by a specific sequence of intermediate events in the event tree. Therefore, the occurrence probability associated with a particular accident scenario can be simply determined by multiplying the probabilities of all intermediate events on that path. According to the structure of event tree in Fig.3, there are a total of  $2 \times 4 \times 2 \times 2 \times 3 \times (2+1) = 288$  possible accident scenarios. Let  $p_{seq_r}$ ,  $r=1,2,\dots,288$ , be the occurrence probability of the  $r^{\text{th}}$  accident scenario. Mathematically, it can be expressed by:

$$p_{seq_r} = \prod_{k=1}^7 P(E_k | E_{k-1}), \quad k = 1, \dots, 7 \quad (4)$$

where  $E_k (k=1, \dots, 7)$  is the intermediate event  $k$  along the path possessed by the  $r^{\text{th}}$  accident scenario,  $P(E_k | E_{k-1})$  is the probability that event  $E_k$  is triggered by event  $E_{k-1}$  along the path possessed by the  $r^{\text{th}}$  accident scenario.

From Eq.(4), it can be recognized that the accuracy of estimated occurrence probability associated with the accident scenario depends on the probability of intermediate event along the path. The probabilities of the intermediate events can be evaluated from historical data.

However, variation and uncertainty may exist in the probability of intermediate event along the path. For example, the probability that a vehicle collides with a fixed object, that is, the probability of “1-unit” event, can be estimated from historical crash data. However, this probability is not suitable to be regarded as a crisp value. This is because the value may vary due to statistical variation in external conditions, such as the likelihood of a fixed object appearing in work zone, or the variability the variable has. A second probability concept can be used to describe this variation/uncertainty as a probability distribution around a “point estimate” for the probability of “1-unit” event. Therefore, the probabilities of the intermediate events in the event tree are considered as random variables, which can be represented by means of probability distribution in this study. The Monte Carlo sampling method is a well-recognized method addressing the problem.

### 2.3. Consequence model

In general, the number of casualties depends on the number of crash units involved in a crash, vehicle type and crash severity. In addition, the speed and *ERT* also affect the consequence of work zone crash. In this section, the consequence model is developed to evaluate the number of casualties for each accident scenario.

According to the nature of PDO crash, it is reasonably assumed that there are no casualties in a PDO crash. Regarding the estimation of casualties in a fatal or injury crash, we separately analyze these two kinds of severe crashes. In fatal work zone crash, passengers or drivers in the involved vehicles have the risk of both being killed and injured while only the risk of being injured is considered in injury work zone crash. When the mean speed is  $V_0$  km/h, the number of fatalities and injuries in a vehicle of type  $i$ , respectively denoted by  $N_{F_i}$  and  $N_{I_i}$ , can be calculated by

$$N_{F_i} = \begin{cases} N_{FF_i}, & \text{in a fatal crash accident} \\ 0, & \text{in an injury crash accident} \end{cases} \quad (5)$$

$$N_{I_i} = \begin{cases} N_{IF_i}, & \text{in a fatal crash accident} \\ N_{II_i}, & \text{in an injury crash accident} \end{cases} \quad (6)$$

$$i = \begin{cases} 1, & \text{for light vehicle} \\ 2, & \text{for heavy vehicle} \end{cases} \quad (7)$$

where  $N_{FF_i}$  = the number of fatalities in a vehicle of type  $i$  in a fatal work zone crash;

$N_{IF_i}$  = the number of injuries in a vehicle of type  $i$  in a fatal work zone crash;

$N_{II_i}$  = the number of injuries in a vehicle of type  $i$  in an injury work zone crash.

Further, the number of fatalities and injuries should fulfill the following threshold limits:

*In fatal crash*

$$N_i = N_{F_i} + N_{II_i} \quad (8)$$

$$0 < N_{F_i} \leq N_i \quad (9)$$

*In injury crash*

$$0 < N_{II_i} \leq N_i \quad (10)$$

where parameters  $N_i$  is the average occupancy of a vehicle of type  $i$ . Constraints (8) and (9) imply that passengers or drivers have the risk of being killed in fatal work zone crash. Constraint (10) ensures that the number of injuries should not exceed the vehicle occupancy.

### 2.3.1. Effects of speed and ERT

Speed can affect the number of casualties in work zone crash. Intuitively, the greater the speed in a crash accident, the more casualties it will lead to. The Power Model, originally proposed by Nilsson (1981, 2004), first addressed the following quantitative relationship between the number of casualties and speed.

$$\frac{n_1}{n_0} = \left( \frac{V_1}{V_0} \right)^\alpha \quad (11)$$

where  $n_1$  is the number of casualties per involved vehicle at the mean speed of  $V_1$ ;  $n_0$  is the number of casualties per involved vehicle at the mean speed of  $V_0$ , the

exponent parameter is  $\alpha = 4$  in fatal accidents and  $\alpha = 2$  in injury accidents.

Elvik et al. (1997) demonstrated that the Power Model is in a very good agreement with the empirical experience for injury accidents and fatal accidents. To test the relative plausibility of the Power Model, other alternative models including a linear model and a logistic model were examined by Elvik et al. (2004). They pointed out that the linear model was highly implausible and the logistic model with exponent depending on initial speed did not perform better than the Power Model because the gain in precision was too small to be justified. In addition, European Commission (1999) explicitly pointed out that the relationship between speed and accident severity can be expressed by the Power Model. However, the Power Model with the exponents of four and two may not be the best. Elvik et al. (2004) assembled a large dataset from 97 published studies containing 460 results to reformulate the Power Model. They presented the best estimate of exponents and the corresponding confidence interval for the causalities in fatal accidents and injury accidents. Because of its plausibility and simplicity, the Power Model has been found a widespread use to quantify the effects of speed on road safety (Cameron and Elvik, 2008; Evans, 1991; Kallberg, 1998; Woolley, 2005).

In this study, the Power Model is also employed to estimate the effects of speed on the consequence of accident scenario. Suppose that the mean speed has the same effects on the risk of being killed or injured for driver/passenger in different vehicle types. Therefore, we can obtain:

$$N_{F_i}^{(1)} = \left( \frac{V_1}{V_0} \right)^{\alpha_1} \times N_{F_i} \quad (12)$$

where  $N_{F_i}^{(1)}$  is the number of fatalities in an involved vehicle at the mean speed of  $V_1$ ,  $N_{F_i}$  is the number of fatalities in a vehicle of type  $i$  at the mean speed of  $V_0$ ,  $\alpha_1$  is the best-estimated exponents in fatal crash.

Taking into account the threshold limit shown by Eq.(9), the number of fatalities in Eq. (12) should satisfy

$$N_{F_i}^{(1)} = \min \left\{ \left( \frac{V_1}{V_0} \right)^{\alpha_1} \times N_{F_i}, N_i \right\} \quad (13)$$

where  $N_i$  is the average occupancy of a vehicle of type  $i$ .

*ERT* is crucial for the timely delivery of emergency medical services (EMS) to accident casualties. It is expected to have an impact on the number of fatalities because little or no first aid to serious injurers may lead to more deaths in fatal crashes. Evanco (1996) addressed the linear relationship between accident *ERT* and the number of fatalities. After taking into account the effect of *ERT*, the number of fatalities in an involved vehicle can be evaluated by the following linear function:

$$N_{F_i}^{(2)} = \left( \beta + (1 - \beta) \times \frac{T_1}{T_0} \right) \times N_{F_i}^{(1)} \quad (14)$$

where  $N_{F_i}^{(2)}$  represents the number of fatalities taking into account the impact of *ERT*;  $T_0$  is the normal mean response time (min);  $T_1$  is the actual mean response time (min),  $\beta$  is the parameter to be estimated;  $N_{F_i}^{(1)}$  is the number of fatalities in an involved vehicle of type  $i$ .

Also considering the above threshold limit, the number of fatalities in Eq. (14) should satisfy

$$N_{F_i}^{(2)} = \min \left\{ \left( \beta + (1-\beta) \times \frac{T_1}{T_0} \right) \times N_{F_i}^{(1)}, N_i \right\} \quad (15)$$

where  $N_{F_i}^{(1)}$  is the number of fatalities in an involved vehicle;  $N_i$  is the average occupancy of a vehicle of type  $i$

An assumption is made that *ERT* has no impact on the number of injuries in injury crash accident. Hence, we can obtain the number of injuries after taking into account the effects of speed and *ERT*:

$$N_{I_i}^{(2)} = \begin{cases} N_i - N_{F_i}^{(2)} & , \text{ in a fatality crash accident} \\ \min \left\{ \left( \frac{V_1}{V_0} \right)^{\alpha_2} \times N_{I_i}, N_i \right\} & , \text{ in an injury crash accident} \end{cases} \quad (16)$$

where  $N_{I_i}^{(2)}$  is the number of injures in a vehicle of type  $i$  after taking into account the effects of speed and *ERT*;  $N_{I_i}$  is the number of injures in a vehicle of type  $i$ ;  $N_i$  is the average occupancy of a vehicle of type  $i$ ;  $\alpha_2$  is the best-estimated exponents in injury crash accident.

In this study, the total number of fatalities and injuries involved in a multi-unit vehicle crash can be obtained using the simple method of algebraic sum. Let  $N_{TF_r}$  and  $N_{TI_r}$  respectively represent the total number of fatalities and injuries in the  $r^{\text{th}}$  accident scenario,  $r = 1, 2, \dots, 288$ , and they can mathematically be calculated by:

$$N_{TF_r} = m_1 \cdot N_{F_1}^{(2)} + m_2 \cdot N_{F_2}^{(2)} \quad (17)$$

$$N_{TI_r} = m_1 \cdot N_{I_1}^{(2)} + m_2 \cdot N_{I_2}^{(2)} \quad (18)$$

where  $N_{TF_r}$  and  $N_{TI_r}$  are respectively the total number of fatalities and injuries in the  $r^{\text{th}}$  accident scenario.  $N_{F_1}^{(2)}$  and  $N_{I_1}^{(2)}$  are respectively the number of fatalities and injuries of a light vehicle.  $N_{F_2}^{(2)}$  and  $N_{I_2}^{(2)}$  are respectively the number of fatalities and injuries of a heavy vehicle.  $m_1$  and  $m_2$  represent the number of light vehicles and heavy vehicles in the  $r^{\text{th}}$  accident scenario, respectively.

#### 2.4. Quantitative casualty risk expressions

It is now widely recognized that the risk from an accident refers to a function of the likelihood of occurrence of possible undesired events and the magnitude of their associated consequences (Borysiewicz et al., 2006). The present study also adopts two commonly used ways to quantitatively express casualty risk in work zone crash. One is individual risk and the other is societal risk. The expressions of individual and societal risks incorporate the frequencies and consequences of all the accident scenarios that have been identified.

##### 2.4.1. Individual risk

Considine (1984) defines individual risk as the likelihood or frequency of the fatality/injury occurring to a person in the vicinity of a hazard. In this study, individual risk is defined as the frequency of a passenger or driver being killed or

injured when traveling through work zone. It includes two aspects: (i) individual fatality risk and (ii) individual injury risk. The individual fatality and injury risk from all accident scenarios, respectively denoted by  $IR_F$  and  $IR_I$ , can be calculated by:

$$IR_F = \frac{f_{\text{crash}}(Q, D, L, U) \cdot \sum_{r=1}^{288} (p_{\text{seq}_r} N_{TF_r})}{Q(p_1 N_1 + p_2 N_2)} \quad (19)$$

$$IR_I = \frac{f_{\text{crash}}(Q, D, L, U) \cdot \sum_{r=1}^{288} (p_{\text{seq}_r} N_{TI_r})}{Q(p_1 N_1 + p_2 N_2)} \quad (20)$$

where,  $IR_F$  and  $IR_I$  are respectively the individual fatality and injury risk from all accident scenarios;  $f_{\text{crash}}(Q, D, L, U)$  is work zone crash frequency;  $p_{\text{seq}_r}$  is the occurrence probability of the  $r^{\text{th}}$  accident scenario;  $N_{TF_r}$  and  $N_{TI_r}$  are respectively the total number of fatalities and injuries in the  $r^{\text{th}}$  accident scenario;  $N_1$  and  $N_2$  are respectively the average number of persons in a light vehicle and a heavy vehicle.  $p_1$  and  $p_2$  are percentages of light vehicles and heavy vehicles, respectively.  $Q(p_1 N_1 + p_2 N_2)$  represents the total number of persons traveling across work zone.

#### 2.4.2. Societal risk

The individual risk implies that the risk can be aggregated into a single number. Obviously, single number cannot address the relationship between the frequency and the total number of casualties caused by a work zone crash. Therefore, the societal risk is employed to complement the individual risk measure in this study. It is a measure of risk to a group of people.

The most common form of presentation of societal risk is the  $F/N$  curve, which illustrates the relationship between crash consequence and the corresponding occurrence frequency. The number of casualties  $x$  (fatalities or injuries) and the corresponding frequencies  $F(x)$  is shown on the abscissa and ordinate, respectively. Mathematically,  $F/N$  curve can be expressed by

$$F_N(x) = P(x \geq N) \quad (21)$$

where  $x$  is the number of casualties caused by a work zone crash, and  $N$  is a given value;  $P(x \geq N)$  can be described as the frequency that the number of casualties caused by work zone crash,  $x$ , is not less than the number  $N$ .

### 3. Numerical example

One numerical example is carried out to assess the developed QRA model in evaluating casualty risk expressed by the individual and societal risks caused by vehicle crash in long-term work zone. Considering a long-term work zone with the length of 3.5 kilometers (i.e.,  $l=3.5$  km) and the work zone duration of 130 days (i.e.,  $D=130$  d). The work zone is located at an urban road of the Southeast Michigan State. In addition, the length of  $L_1$  and  $L_2$  are respectively set as 0.5 km and 0.1 km. Therefore, the total work zone length  $L$  is 4.1 km ( $\approx 2.6$  mile).

Suppose that the work zone crash frequency satisfies the relationship shown by Eq. (3), the frequency of work zone crash is 23.66 crashes per work zone duration (130 d) if the average daily traffic volume  $Q$  is 45,000 vehicles per day in this example.

### 3.1. Data

The Southeast Michigan Traffic Crash Records Database for years from 1999 to 2008 is utilized and is originally obtained from the Southeast Michigan Council of Governments. The database contained 89 fatal work zone crashes involved fatalities and injuries, 10,142 work zone injury crashes involved injuries and 35,036 PDO work zone crashes. From this 10-year's work zone crash database, it can be found that the probabilities that more than one heavy vehicle are respectively involved in 2-Unit, 3-Unit, and  $\geq 4$ -Unit work zone crashes are extremely low. Therefore, only one heavy vehicle is considered in heavy vehicle-involved crashes. In addition, it does not take into account the crashes involving more than 4-unit vehicles because such crashes occupy extremely low percentage (less than  $1.0E-04$ ) among all recorded work zone crashes. Therefore, the number of light vehicles  $m_1$  and the number of heavy vehicles  $m_2$  in each accident scenario are shown in Fig.3.

Assuming that the value of  $V_0$  and  $T_0$  are respectively 60km/h and 5.2 min, we can obtain the average number of fatalities and injuries per vehicle in a fatal/injury crash accident using the linear regression method from the database, which are shown in Table 3. To apply the proposed QRA model, we take a set of values of the parameters  $V_1$ ,  $T_1$ ,  $\alpha_1$ ,  $\alpha_2$ ,  $\beta$ ,  $p_1$  and  $p_2$  (see in Table 3). Car and motorcycle are considered as the major two types of light vehicle and the heavy vehicle consists of van and truck. The average value of car and motorcycle occupancy is thus regarded as the light vehicle occupancy ( $N_1$ ). The heavy vehicle occupancy ( $N_2$ ) approximates the average occupancy of van and truck.

#### 3.1.1. Occurrence probabilities of intermediate events in the event tree

Table 4 reports the means and relative standard deviations of the probabilities of all intermediate events from the database. It can be seen that 2-unit crashes are the primary crashes among all work zone crashes because of high percentage of 0.7404 (for age less than 25) and 0.7351 (for age larger than 25). The relative standard deviations of the probabilities of intermediate events of vehicle type ( $VT$ ), crash type ( $CT$ ) and severity ( $S$ ) are large. It implies that the probabilities of these intermediate events have large variations (uncertainties). Therefore, the probabilities of intermediate events  $VT$ ,  $CT$  and  $S$  are described by probabilistic distributions in the context of event tree analysis. On the other hand, the occurrence probabilities of intermediate events of age ( $A$ ), crash unit ( $CU$ ), alcohol ( $AL$ ) and light condition ( $LC$ ) that have small relative standard deviations are assigned by sample means.

The @Risk software is allowed to fit probability distributions to the sample data. This software makes use of chi-squared fit statistic to measure how good the distribution fits the sample data. The smaller the value of chi-squared statistic is, the better the fit. Fig.4 presents the best-fitted probability density functions of occurrence probabilities of intermediate events of vehicle type ( $VT$ ), crash type ( $CT$ ) and severity ( $S$ ). In Fig.4,  $p(v_j)$ ,  $j=1, \dots, 8$  represent probability density functions for eight states of the event vehicle type ( $VT$ ).  $p(v_9)$  and  $p(v_{10})$  are respectively probability density functions in the events of crash type ( $CT$ ) and severity ( $S$ ). After probability density functions  $p(v_j)$ ,  $j=1, \dots, 10$  are determined, the propagation of the uncertainty of intermediate events  $VT$ ,  $CT$  and  $S$  can be performed by resorting to the Monte Carlo sampling method. In the  $i^{\text{th}}$  Monte Carlo simulation,  $i=1, 2, \dots, 10000$ ,

the occurrence probabilities  $(v_1^i, \dots, v_{10}^i)$  of event vector  $(VT_1, \dots, VT_8, CT, S)$  can be sampled from the corresponding probability density functions.

### 3.2. Results and discussions

Given all the parameters required by the proposed model, the individual and societal risks that combine work zone frequency and consequence of all accident scenarios could be calculated. For each realization of Monte Carlo simulation, the proposed model can yield one combination result, comprising the number of casualties and frequency, associated with a particular scenario. Percentile ranks are used to clarify the interpretation of all possible results. Fig.5 reports the percentile-based individual fatality and injury risks that a driver or passenger might experience during the work zone period. In Fig.5 (a), the 25<sup>th</sup>, 50<sup>th</sup> and 75<sup>th</sup> percentile of individual fatality risk are respectively  $1.07 \times 10^{-6} / D$ ,  $1.33 \times 10^{-6} / D$  and  $1.58 \times 10^{-6} / D$ . Here,  $D$  is work zone duration, which is equal to 130 days. It implies that a passenger has the probability of  $1.58 \times 10^{-6}$  of being killed when s/he is traveling through work zone at a 75% confidence level. By comparing Fig.5 (a) with Fig.5 (b), it is shown that the passenger has a higher injury risk than fatality risk. The result is consistent with the fact that injuries are a relatively high percentage of overall traffic accident casualties.

Fig.6 (a) and (b) depicts the relationship between the frequency and the number of casualties in a work zone crash. It can be clearly seen that the frequency decreases with the number of fatalities/injuries resulted from a work zone crash. It implies that the major crash accident (i.e. large number of fatalities or injuries) has a relatively low occurrence frequency than the minor accident in work zone.

From the practical point of view, the selection of  $F/N$  curve should be made based on the worst case. This is because the occurrence probabilities of intermediate events in the event tree have variations/uncertainties, which might lead to underestimated results. In view of highly safety requirement, the conservative attitude is recommended in assessing work zone risk. Therefore, 95<sup>th</sup> percentile of  $F/N$  curve could be considered as the societal risk caused by work zone crash (see in Fig.6 (a) and (b)).

#### 3.2.1. Impact of uncertainty associated with the probability of intermediate event

The larger uncertainties associated with the occurrence probabilities of intermediate events may cause higher degree of uncertainty associated with the output (i.e., individual fatality/injury risk). To evaluate the impact of the uncertainty with respect to the specified event (i.e.,  $S$ ), the corresponding occurrence probability is still represented by means of probability distribution. Simultaneously, the occurrence probabilities of other intermediate events (i.e.,  $VT$  and  $CT$ ) with large variations are assigned by the sample mean values. Fig.7 (a) and (b) respectively describes the impact of uncertainty of event occurrence on the degree of uncertainty associated with the individual fatality and injury risks. Intuitively, the uncertainty associated with the probability of intermediate event  $S$  can cause a significant degree of uncertainty associated with individual fatality risk. This is because the occurrence frequency of a driver being killed is greatly affected by the probability of intermediate event  $S$ . Similarly, the uncertainty of individual injury risk is significantly affected by the event  $CT$ .

In order quantitatively to determine the degree of uncertainty associated with the output of individual risk, the ratio of upper bound to lower bound is introduced in this section. Let  $\rho(e)$  denote the uncertainty ratio of upper bound to lower bound for the uncertainty of outcome caused by the uncertain event  $e$ , and it can then be expressed by:

$$\rho(e) = \frac{X_{0.95}(e)}{X_{0.05}(e)} \quad (22)$$

where  $X_{0.95}(e)$ =upper bound based on 95% percentile for event  $e$ ;  
 $X_{0.05}(e)$ =lower bound based on 5% percentile for event  $e$ .

The larger the uncertainty ratio is, the bigger uncertainty of the outcome will be resulted from the uncertainty event  $e$ .

Fig.7 (a) shows that the intermediate event  $S$  leads to the highest degree of uncertainty associated with individual fatality risk because the corresponding uncertainty ratio  $\rho(S)$  is the largest, up to 2.88. Instead of intermediate event  $S$ , the intermediate event  $CT$  has the biggest impact on the degree of uncertainty associated with individual injury risk, shown in Fig.7 (b). The degree of uncertainty associated with individual fatality/injury risk is slightly affected by the intermediate event  $VT$ . The major reason is that the relative standard deviation for the probabilities of intermediate event  $VT$  is relatively small, as compared with the intermediate events  $CT$  and  $S$ .

### 3.2.2. Impact of ERT and speed on individual risk

As mentioned above, both the  $ERT$  and mean speed have an impact on the consequence in work zone crash. In Fig.8 (a), it can be seen that the individual fatality risk decreases with the  $ERT$ , whereas Fig.8 (b) shows that the individual injury risk is insignificantly affected by  $ERT$ . This is because delayed  $ERT$  decreases the chance of survival of passengers or drivers who are seriously injured in fatal crash. It can be concluded that more deaths would be caused in work zone if the local emergency medical service were low-efficiency.

Fig.9 (a) and (b) demonstrates that the mean speed poses great influence on the individual fatality risk as well as injury risk. There would be a 62% decrease of individual fatality risk and a 44% reduction of individual injury risk if the mean speed were slowed down by 20%. However, there would only be a 5% reduction of individual fatality risk and a 0.05% reduction of individual injury risk if the  $ERT$  were also reduced by 20%. Therefore, we can conclude that slowing down speed is a more effective way to reduce the driver/passenger's fatality and injury risk.

## 4. Conclusions

In contrast to previous crash risk assessment models, the developed QRA model aims to evaluate the casualty risk that combines frequency and consequence of long-term work zone crash. Since a number of accident scenarios may be possible in a work zone crash, the present study utilizes an event tree diagram to identify all possible accident scenarios. Seven intermediate events comprising age ( $A$ ), crash unit ( $CU$ ), vehicle type ( $VT$ ), alcohol ( $AL$ ), light condition ( $LC$ ), crash type ( $CT$ ) and severity ( $S$ ) are considered in the event tree analysis. The reliability of the estimation of occurrence likelihood associated with a particular accident scenario is dependent upon the probability of each intermediate event, whereas they may have large

variations. Therefore, the probabilities of the intermediate events are formulated as random variables that have probability distribution patterns. In contrast to simply categorizing consequence in past studies, the QRA model provides a consequence model to estimate the quantitative consequence in terms of the number of casualties in each accident scenario. Two risk expressions are utilized to express the casualty risk. One expression is the individual risk that is used to express the risk of a driver/passenger being killed or injured in long-term work zone. The other casualty risk expression is the societal risk describing the relationship between the frequency and total number of casualties caused by vehicle crash.

A numerical example is carried out and the Southeast Michigan work zone crash data are utilized to calibrate the proposed QRA model. The probability of intermediate events *VT*, *CT* and *S* with large variations are represented by means of probability distribution. The proposed QRA model demonstrates the capability of reporting individual risk and societal risk in long-term work zone. It also presents that the occurrence probability of intermediate event *S* that has large statistical variation causes a big uncertainty on the individual fatality risk. Slowing down speed is found to be more effective than reducing the *ERT* in mitigating individual fatality/injury risk.

## References

- Abdel-Aty, M. and Keller, J., 2005. Exploring the overall and specific crash severity levels at signalized intersections. *Accident Analysis and Prevention*, 37(3), pp.417-425.
- Agresti, A., 2002. *Categorical data analysis*. New York, NY: John Wiley and Sons, Inc.
- Borysiewicz, M.J., Borysiewicz, M.A., Garanty, I. and Kozubal, A., 2006. *Quantitative Risk Assessment (QRA)*. CoE MANHAZ. Institute of Atomic Energy.
- Breiman, L., 2001. Random forests, *Machine Learning* 45, 5 - 32.
- Cameron, M.H. and Elvik, R., 2008. Nilsson's Power Model connecting speed and road trauma: Does it apply on urban roads? 2008 Australasian Road Safety Research, Policing and Education Conference. Adelaide, South Australia.
- Chang, L.Y. and Chen, W.C., 2005. Data mining of tree-based models to analyze freeway accident frequency. *Journal of safety research*, 36(4), 365-375.
- Considine, M., 1984. *The assessment of individual and societal risks*. SRD Report R-310, Safety and Reliability directorate. Warrington: UK Atomic Energy Authority.
- Daniel, J., Dixon, K., and Jared, D., 2000. Analysis of fatal crashes in Georgia work zones. *Transportation Research Record*, 1715, 18-23.
- Dissanayake, S., Lu, J., 2002. Analysis of severity of young driver crashes, sequential binary logistic regression modeling. *Transportation Research Record* 1784, 108-114.
- Elvik, R., Vaa, T., Mysen, A.,B., 1997. *Trafikksikkerhetshåndbok*, Transportøkonomisk institutt, Oslo. 313-136 (In Norwegian)
- Elvik, R., P. Christensen, and A. Amundsen., 2004. *Speed and road accidents: An evaluation of the Power Model*. 740/2004., Oslo: Institute of Transport Economics (TOI).
- European Commission. 1999. *MASTER: Managing speeds of traffic on European roads*. Final report. European Commission, Brussels.
- Evanco, W.M., 1996. *The Impact of Rapid Incident Detection on Freeway Accident Fatalities*, Mitretek Systems, Inc., WN96W0000071.

- Evans, L. 1991. Traffic safety and the Driver. ISBN 0-422-00163-0. pp. 153-154 and 322.
- Kallberg, V-P. 1998. The two effects of speed on accidents: Number and severity. Transportation Research Board, Paper No. 981191.
- Ha, T.J., Nemeth, Z.A., 1995. Detailed Study of Accident Experience in Construction and Maintenance Zones. Transportation Research Record 1509, 38–45.
- Harb, R., Radwan, E., Yan, X., Pande, A., and Abdel-Aty, M., 2008. Freeway Work-Zone Crash Analysis and Risk Identification Using Multiple and Conditional Logistic Regression. Journal of Transportation Engineering, 134(5), 203-214.
- Kalos, M.H and Whitlock, P. A., 1986. Monte Carlo Methods. Volume I Basics. New York: Wiley Interscience.
- Khattak, A. and Cassidy, G., 1999. Factors that influence multivehicle rear-end crashes: Analysis of crash propagation and injury severity. Southern Transportation Center, Univ. of Tennessee.
- Khattak, A.J., Khattak, A.J. and Council, M.F., 2002. Effects of work zone presence on injury and non-injury crashes. Accident Analysis and Prevention, 34(1), 19-29.
- Li, Y., Bai, Y., 2006. Investigating the characteristics of fatal crashes in the highway construction zones. In: CIB W99 International Conference on Global Unity for Safety & Health in Construction, Beijing, China, June 28–30, pp. 301–309.
- Li, Y., Bai, Y., 2008. Development of crash-severity-index models for the measurement of work zone risk levels. Journal of the Accident Analysis and Prevention, 40(5), 1724-1731.
- Meng, Q., Wang, X.C., Qu, X.B., Thong, K. M., Lee, S.P. and Wang, S.C. 2009. Quantitative risk assessment models-state of the art and their application to Singapore's road tunnel. In: Proceedings of the 2<sup>nd</sup> international tunnel safety forum for road and rail, Lyon, France, 20-22.
- Miaou, S.P., 1994. The relationship between truck accidents and geometric design of road sections: Poisson versus negative binomial regression. Accident Analysis and Prevention, 26(4), 471-482.
- Milton, J. and Mannering, F., 1998. The relationship among highway geometrics, traffic-related elements and motor-vehicle accident frequencies. Transportation, 25(4), pp.395-413.
- Milton, J.C., Shankar, V.N. and Mannering, F.L., 2008. Highway accident severities and the mixed logit model: An exploratory empirical analysis. Accident Analysis and Prevention, 40(1), pp.260-266.
- Mohamed, A. A. and Hassan T. A., 2004 Predicting Injury Severity Levels in Traffic Crashes: A Modeling Comparison. Journal of Transportation Engineering, 130(2), pp. 204-210
- MUTCD, 2003. Manual on Uniform Traffic Control Devices.
- Nilsson, G. 1981. Traffic safety in terms of accidents, injuries, risks and consequences- a multidimensional method for description of traffic safety situation. Road Design and Safety. IX:th IRF World Meeting Stockholm.
- Nilsson, G., 2004. Traffic safety dimensions and the power model to describe the effects of speed on safety. Bulletin 221. Lund Institute of Technology, Department of Technology and Society.
- NUREG, 1975. WASH 1400 (NUREG-75/014), Reactor Safety Study, Washington, DC.
- Ouyang, Y., Shankar, V., Yamamoto, T., 2002. Modeling the simultaneity in injury causation in multi-vehicle collisions. Transportation Research Record 1704, 143–152.

- Pigman, J.G., and K.R. Agent., 1990. Highway Accidents in Construction and Maintenance Work Zone. Transportation Research Record 1270, National Research Council, Washington D.C., 12-21.
- Rouphail, N.M., Yang, Z.S., Fazio, J., 1988. Comparative Study of Short- and Long-Term Urban Freeway Work Zones. Transportation Research Record 1163, Transportation Research Board, National Research Council, Washington, DC, 4-14.
- Shankar, V., Mannering, F., and Barfield, W., 1996. Statistical analysis of accident severity on rural freeways. Journal of Accident Analysis and Prevention, 28(3), pp.391-401.
- Ullman, G.L., Ullman, B.R., and Finley, M.D., 2006. Analysis of Crashes at Active Night Work Zones in Texas. In: Proceedings of the TRB 85<sup>th</sup> Annual Meeting, CDR-OM, National Research Council, Washington, DC.
- Vose, D., 2000. Risk Analysis-A Quantitative Guide. Chichester, UK: John Wiley & Sons.
- Woolley, J., 2005. Recent advantages of lower speed limits in Australia. Journal of the Eastern Asia Society for Transportation Studies, 6, pp. 3562-3573.

## List of Figures and Tables

**Fig.1.** A flowchart of QRA model formulation

**Fig.2.** Typical work zone configuration

**Fig.3.** Structure of event tree for work zone crash

**Fig.4.** Probability density functions for the probabilities of intermediate events in the event tree

**Fig.5.** Analysis of individual risk in work zone

**Fig.6.** Analysis of F/N curve for casualties in work zone

**Fig.7.** Impact of uncertainty associated with the probability of intermediate event

**Fig.8.** Influence of *ERT* on individual risk

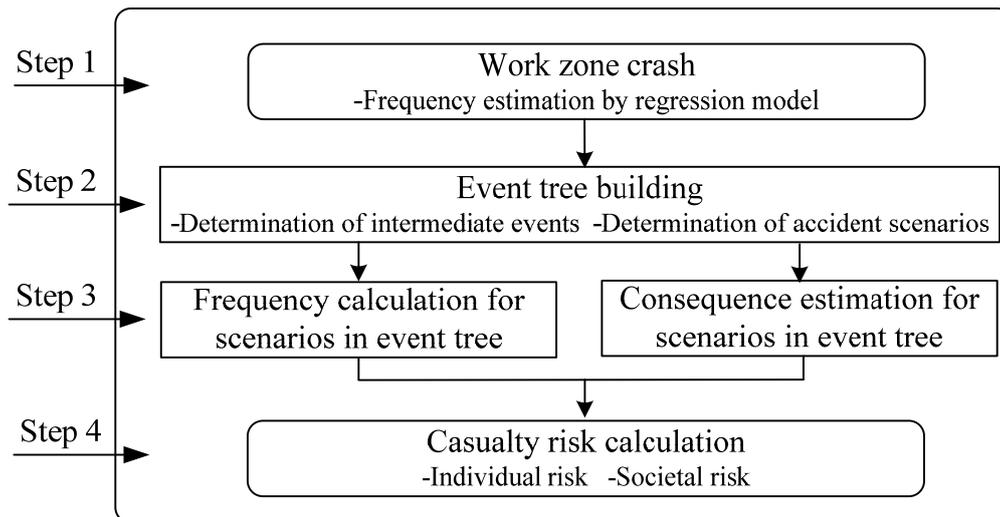
**Fig.9.** Influence of speed on individual risk

**Table 1** Ohio work zone accident data in 2002

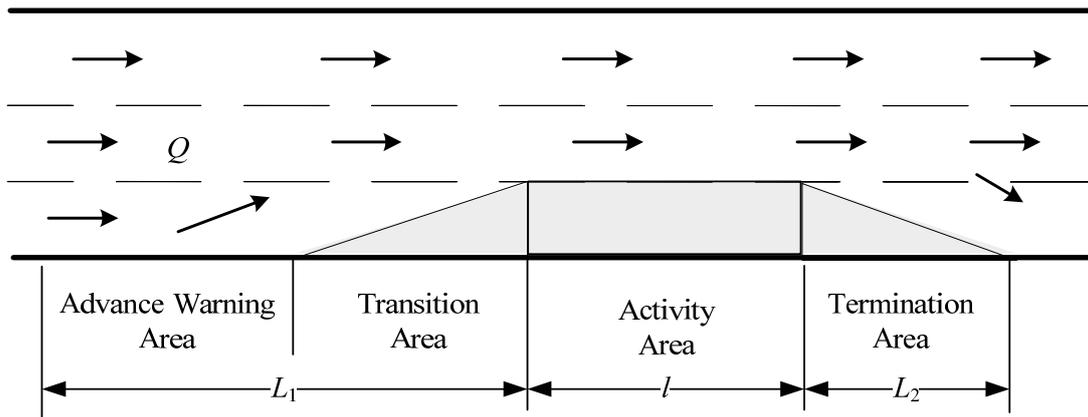
**Table 2** Trial and error results for work zone crash frequency

**Table 3** Input parameters for the numerical example

**Table 4** Estimated probabilities of intermediate events in the event tree

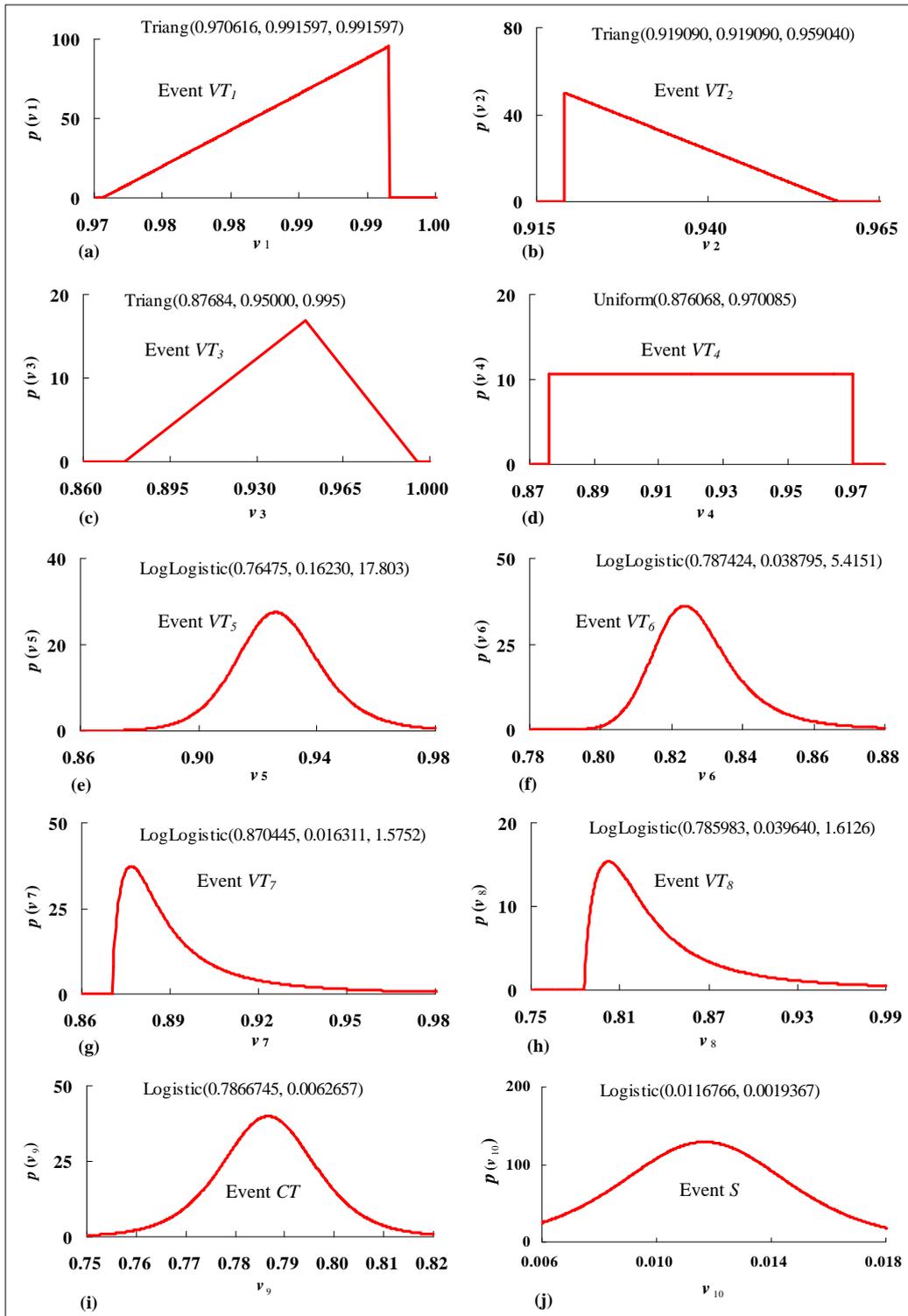


**Fig.1.** A flowchart of QRA model formulation

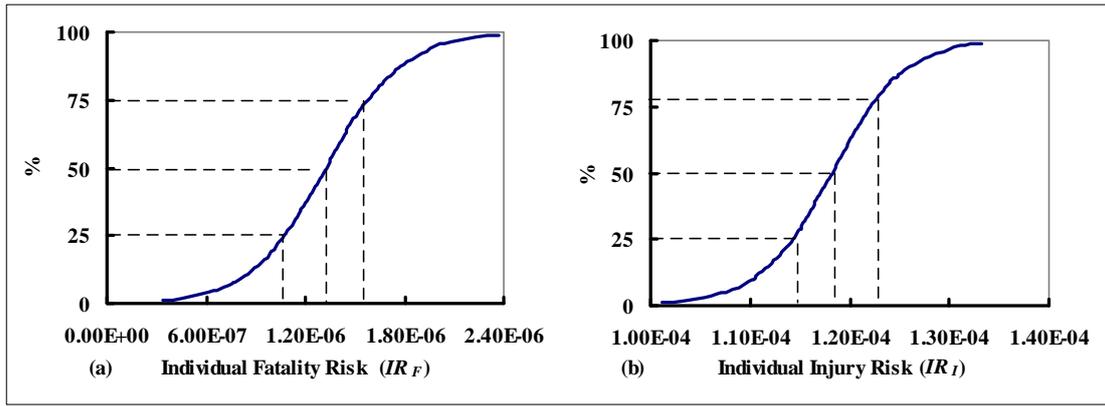


**Fig.2.** Typical work zone configuration

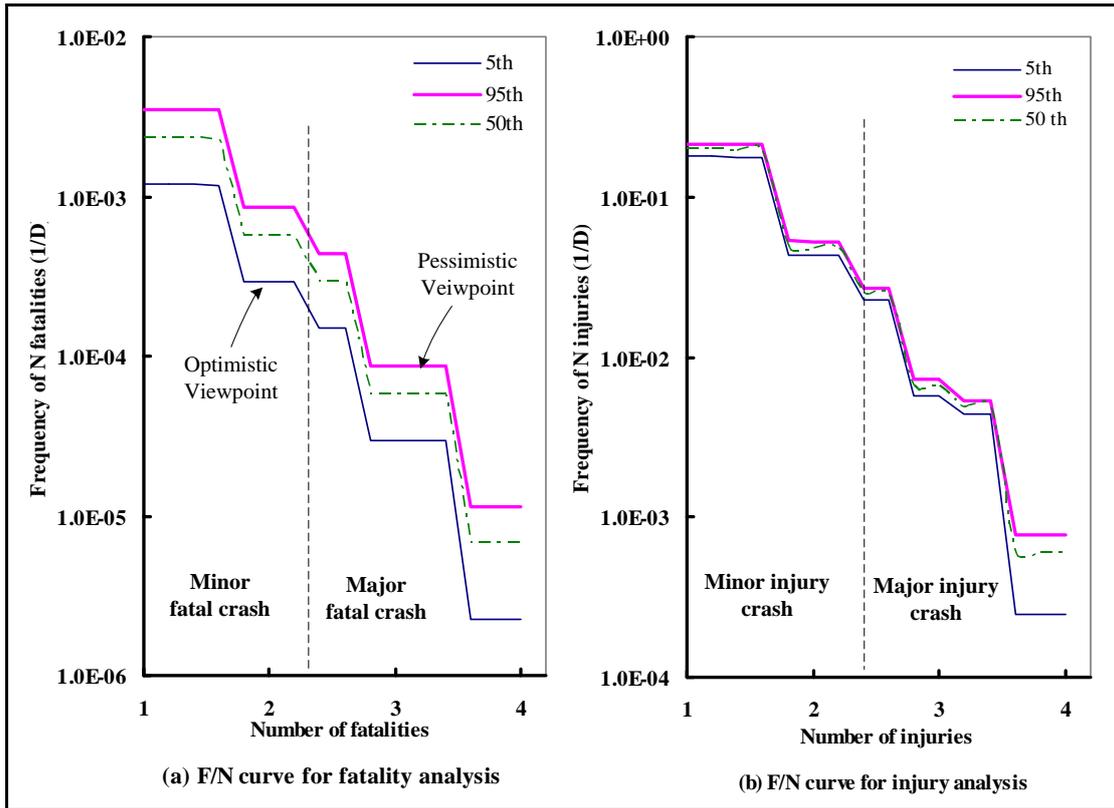




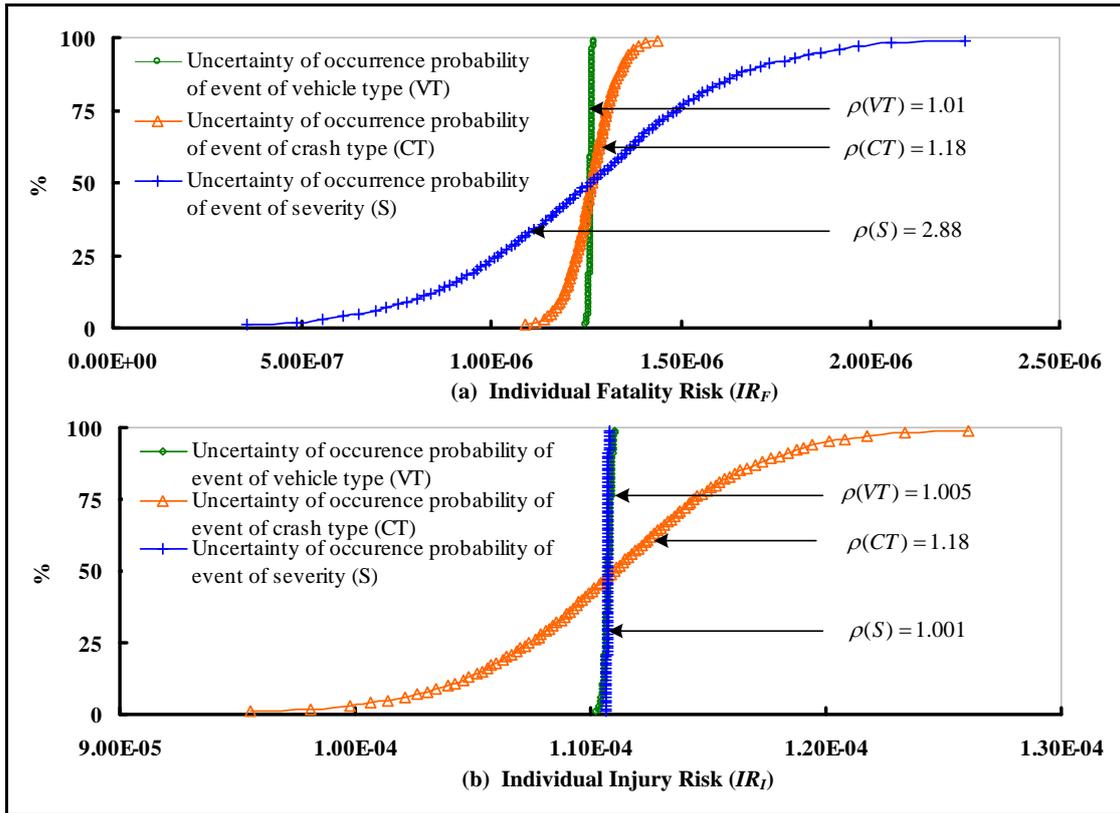
**Fig.4.** Probability density functions for the probabilities of intermediate events in the event tree



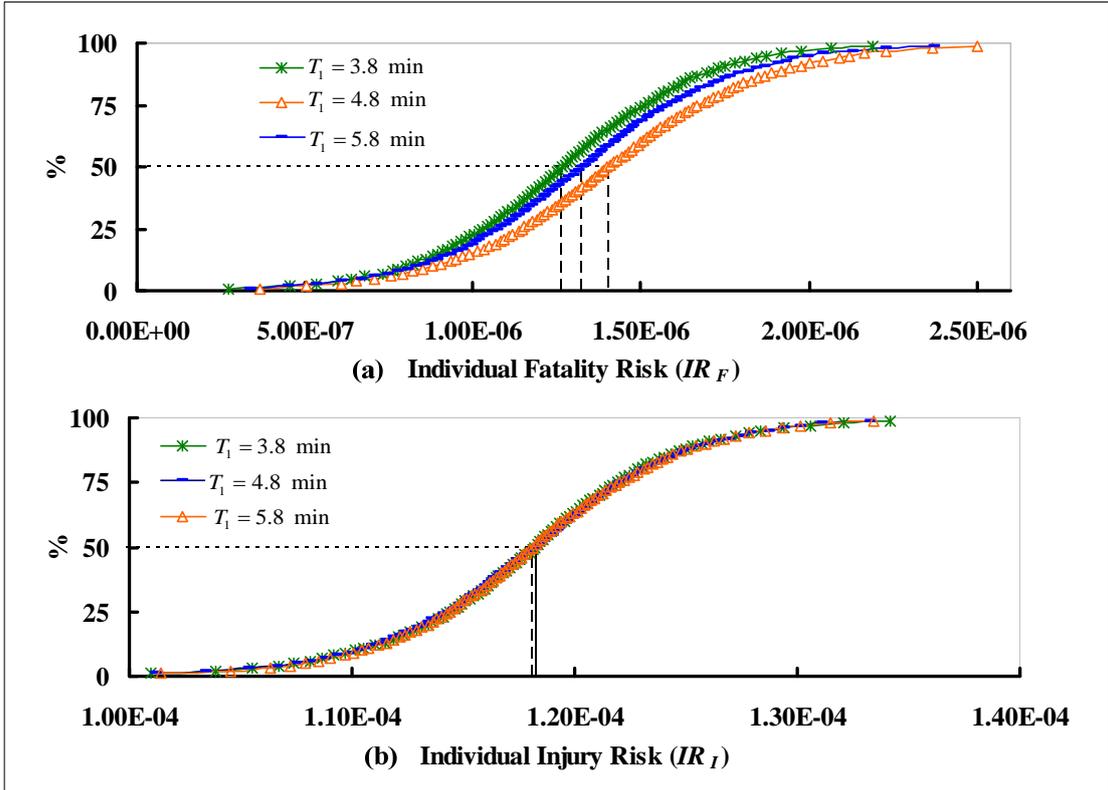
**Fig.5.** Analysis of individual risk in work zone



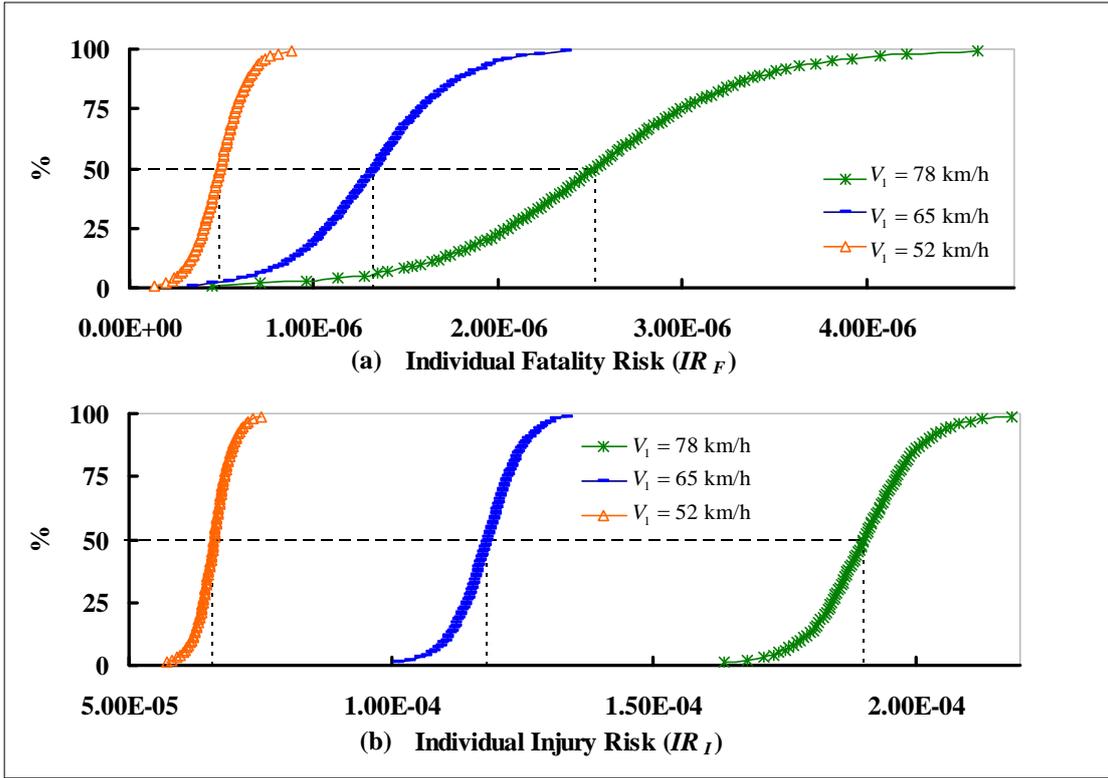
**Fig.6.** Analysis of F/N curve for casualties in work zone



**Fig. 7.** Impact of uncertainty associated with the probability of intermediate event



**Fig.8.** Influence of *ERT* on individual risk



**Fig.9.** Influence of speed on individual risk

**Table 1**

Ohio work zone accident data in 2002

Length (mile)	Annual average daily traffic (veh/day)	Work zone duration (days)	Road Type	Work zone crash frequency (crash per work zone duration)
4.12	113,108	304	1*	174
3.20	84,947	274	0	163
4.29	86,397	274	1	134
5.50	71,116	304	1	140
3.79	60,257	213	1	100
5.90	63,656	365	0	212
1.17	85,111	207	0	84
4.00	41,470	213	1	48
2.70	41,192	122	1	23
4.18	63,032	335	1	117
3.00	41,865	122	0	21
2.50	41,487	243	1	43
5.42	55,334	365	1	125
2.92	60,681	213	1	55

*Note:* \* 1 for urban road; 0 for rural road

**Table 2**  
Trial and error results for work zone crash frequency

Combination form ( $x_1, x_2, x_3, x_4$ )	$p$ -values for coefficients					Feasibility	Adjusted $R^2$
	$\alpha_0$	$\alpha_1$	$\alpha_2$	$\alpha_3$	$\alpha_4$		
( $L, D, Q, U$ )	<0.01	0.34	<0.01	<0.01	0.38		88.9%
( $L, D, 1/Q, U$ )	<0.01	0.22	<0.01	<0.01	0.46		92.9%
( $L, D, \ln(Q), U$ )	<0.01	0.24	<0.01	<0.01	0.42		91.7%
( $L, 1/D, Q, U$ )	<0.01	0.02	<0.01	<0.01	0.04	Feasible	93.4%
( $L, 1/D, 1/Q, U$ )	<0.01	<0.01	<0.01	<0.01	0.09		94.7%
( $L, 1/D, \ln(Q), U$ )	0.02	<0.01	<0.01	<0.01	0.07		94.1%
( $L, \ln(D), Q, U$ )	0.01	0.12	<0.01	<0.01	0.15		92.1%
( $L, \ln(D), 1/Q, U$ )	0.71	0.06	<0.01	<0.01	0.19		95.6%
( $L, \ln(D), \ln(Q), U$ )	<0.01	0.07	<0.01	<0.01	0.16		94.1%
( $\ln(L), D, Q, U$ )	<0.01	0.40	<0.01	<0.01	0.35		88.6%
( $\ln(L), D, 1/Q, U$ )	<0.01	0.22	<0.01	<0.01	0.38		92.9%
( $\ln(L), D, \ln(Q), U$ )	<0.01	0.27	<0.01	<0.01	0.36		91.5%
( $\ln(L), 1/D, Q, U$ )	<0.01	0.03	<0.01	<0.01	0.04	Feasible	92.7%
( $\ln(L), 1/D, 1/Q, U$ )	<0.01	<0.01	<0.01	<0.01	0.04	Feasible	95.4%
( $\ln(L), 1/D, \ln(Q), U$ )	0.03	0.01	<0.01	<0.01	0.04	Feasible	94.5%
( $\ln(L), \ln(D), Q, U$ )	<0.01	0.16	<0.01	<0.01	0.13		91.7%
( $\ln(L), \ln(D), 1/Q, U$ )	0.52	0.07	<0.01	<0.01	0.13		94.8%
( $\ln(L), \ln(D), \ln(Q), U$ )	<0.01	<0.01	<0.01	<0.01	0.13		93.8%
( $1/L, D, Q, U$ )	<0.01	0.52	<0.01	<0.01	0.37		88.2%
( $1/L, D, 1/Q, U$ )	<0.01	0.27	<0.01	<0.01	0.38		92.6%
( $1/L, D, \ln(Q), U$ )	<0.01	0.35	<0.01	<0.01	0.37		91.1%
( $1/L, 1/D, Q, U$ )	<0.01	0.07	<0.01	<0.01	0.06		91.3%
( $1/L, 1/D, 1/Q, U$ )	<0.01	0.03	<0.01	<0.01	0.06		94.4%
( $1/L, 1/D, \ln(Q), U$ )	0.08	0.04	<0.01	<0.01	0.06		93.3%
( $1/L, \ln(D), Q, U$ )	0.01	0.25	<0.01	<0.01	0.14		91.0%
( $1/L, \ln(D), 1/Q, U$ )	0.62	0.11	<0.01	<0.01	0.13		94.3%
( $1/L, \ln(D), \ln(Q), U$ )	<0.01	0.14	<0.01	<0.01	0.13		93.2%

Note: A combination of variable form is feasible only if all  $p$ -values are less than 0.05.

**Table 3**  
Input parameters for the numerical example

Parameters	Descriptions	Values
$p_1$	Percentage of light vehicles	83.0%
$p_2$	Percentage of heavy vehicles	17.0%
$N_1$	Average number of persons in a light vehicle	1.54
$N_2$	Average number of persons in a heavy vehicle	1.70
$N_{FF_1}$	Expected number of fatalities of a light vehicle involving a fatal crash accident	0.54
$N_{FF_2}$	Expected number of fatalities of a heavy vehicle involving a injury crash accident	0.62
$N_{II_1}$	Expected number of injuries of a light vehicle involving an injury crash accident	0.65
$N_{II_2}$	Expected number of injuries of a heavy vehicle involving an injury crash accident	0.90
$V_1$	Actual mean speed	65km/h
$T_1$	Actual mean value of EMS response time	4.8 min
$\alpha_1$	The best estimate exponents in fatal crash	4.5 <sup>a</sup>
$\alpha_2$	The best estimate exponents in injury crash	2.7 <sup>a</sup>
$\beta$	The parameter in Eq. (14)	0.73 <sup>b</sup>

Note: <sup>a</sup> source from Elvik et al. (2004), confidence interval:  $\alpha_1$  (4.1~4.9);  $\alpha_2$  (0.9-4.5) ;

<sup>b</sup> source from Evanco (1996)

**Table 4**  
Estimated probabilities of intermediate events in the event tree

Event	Acronym	Probability of occurrence	pdf	Mean	Relative standard deviation	Description
Age	$A$	$p(A)$	—	6.795E-01	1.10E-02	This event will happen as age>25
Crash Unit	$CU_1$	$p(CU_1)$	—	1.062E-01	1.02E-02	This event will happen as age≤25 and $CU=1$ -unit
	$CU_2$	$p(CU_2)$	—	7.404E-01	1.74E-03	This event will happen as age≤25 and $CU=2$ -unit
	$CU_3$	$p(CU_3)$	—	1.217E-01	9.78E-03	This event will happen as age≤25 and $CU=3$ -unit
	$CU_4$	$p(CU_4)$	—	3.170E-02	1.81E-02	This event will happen as age≤25 and $CU>=4$ -unit
	$CU_5$	$p(CU_5)$	—	1.791E-01	1.74E-02	This event will happen as age>25 and $CU=1$ -unit
	$CU_6$	$p(CU_6)$	—	7.351E-01	2.22E-02	This event will happen as age>25 and $CU=2$ -unit
	$CU_7$	$p(CU_7)$	—	6.950E-02	8.79E-03	This event will happen as age>25 and $CU=3$ -unit
	$CU_8$	$p(CU_8)$	—	1.630E-02	1.39E-02	This event will happen as age>25 and $CU>=4$ -unit
Vehicle Type	$VT_1$	$v_1$	$p(v_1)$	9.844E-01	5.29E-02	This event will happen as age≤25, $CU=1$ -unit, $VT=LV$
	$VT_2$	$v_2$	$p(v_2)$	9.335E-01	6.21E-02	This event will happen as age≤25, $CU=2$ -unit, $VT=LV$
	$VT_3$	$v_3$	$p(v_3)$	9.419E-01	6.05E-02	This event will happen as age≤25, $CU=3$ -unit, $VT=LV$
	$VT_4$	$v_4$	$p(v_4)$	9.220E-01	8.86E-02	This event will happen as age≤25, $CU=4$ -unit, $VT=LV$
	$VT_5$	$v_5$	$p(v_5)$	9.279E-01	7.73E-02	This event will happen as age>25, $CU=1$ -unit, $VT=LV$
	$VT_6$	$v_6$	$p(v_6)$	8.281E-01	8.86E-02	This event will happen as age>25, $CU=2$ -unit, $VT=LV$
	$VT_7$	$v_7$	$p(v_7)$	8.968E-01	7.04E-02	This event will happen as age>25, $CU=3$ -unit, $VT=LV$
	$VT_8$	$v_8$	$p(v_8)$	8.475E-01	6.89E-02	This event will happen as age>25, $CU=4$ -unit, $VT=LV$
Alcohol	$AL$	$p(AL)$	—	9.632E-01	3.54E-03	This event will happen as non-alcohol drivers involved
Light Condition	$LC_1$	$p(LC_1)$	—	7.653E-01	1.84E-02	This event will happen as $LC=$ daylight
	$LC_2$	$p(LC_2)$	—	1.655E-01	6.78E-03	This event will happen as $LC=$ dark with lighting
Crash Type	$CT$	$v_9$	$p(v_9)$	7.869E-01	1.37E-01	This event will happen as $CT=$ PDO crash
Severity	$S$	$v_{10}$	$p(v_{10})$	1.179E-02	2.92E-01	This event will happen as $S=$ fatal crash

Note: Relative standard deviation=Standard deviation/Mean;  $LV=$  light vehicle