401

VTT PUBLICATIONS

Matti Roine

Accident risks of car drivers in wintertime traffic

TECHNICAL RESEARCH CENTRE OF FINLAND ESPOO 1999

VTT PUBLICATIONS 401

Accident risks of car drivers in wintertime traffic

Matti Roine

VTT Communities and Infrastructure

Dissertation for the degree of Doctor of Science of Technology to be presented with due permission for public examination and debate in Auditorium R1 at Helsinki University of Technology, Department of Civil and Environmental Engineering (Rakentajanaukio 4 A, Espoo, Finland) on the 9th of December, 1999, at 12 o'clock noon.

TECHNICAL RESEARCH CENTRE OF FINLAND
ESPOO 1999

SBN 951–38–5546–5 (soft back ed.)
ISSN 1235–0621 (soft back ed.)

ISBN 951–38–5547–3 (URL: http://www.inf.vtt.fi/pdf/)
ISSN 1455–0849 (URL: http://www.inf.vtt.fi/pdf/)

JULKAISIJA – UTGIVARE – PUBLISHER

Valtion teknillinen tutkimuskeskus (VTT), Vuorimiehentie 5, PL 2000, 02044 VTT
puh. vaihde (09) 4561, faksi (09) 456 4374

Statens tekniska forskningscentral (VTT), Bergsmansvägen 5, PB 2000, 02044 VTT
tel. växel (09) 4561, fax (09) 456 4374

Technical Research Centre of Finland (VTT), Vuorimiehentie 5, P.O.Box 2000, FIN–02044 VTT, Finland
phone internat. + 358 9 4561, fax + 358 9 456 4374

VTT Yhdyskuntatekniikka, Liikenne, logistiikka ja yhdyskunnat,
Metallimiehenkuja 10, PL 1902, 02044 VTT,
puh. vaihde (09) 4561, faksi (09) 464 850

VTT Samhällsbyggnad och infrastruktur, Transport och samhällsplanering,
Metallmansgränden 10, PB 1902, 02044 VTT
tel. växel (09) 4561, fax (09) 464 850

VTT Communities and Infrastructure, Transport and Urban Planning,
Metallimiehenkuja 10, P.O.Box 1902, FIN–02044 VTT, Finland,
phone internat. + 358 9 4561, fax + 358 9 464 850

Technical editing Leena Ukskoski

Libella Painopalvelu Oy, Espoo 1999

Roine, Matti. Accident risks of car drivers in wintertime traffic. Espoo 1999. Technical Research Centre of Finland, VTT Publications 401. 137 p. + app. 20 p.

Keywords winter traffic, vehicular traffic, cold environments, automobiles, drivers, traffic accidents, risk factors, research, models

ABSTRACT

The wintertime accident risks of drivers and the factors affecting the risks were analysed using statistical accident models. The evaluation method was based on reliability theory and on survival modelling. The data consisted of two parts: responses to a postal questionnaire, addressed to 10,000 vehicle owners, about driving and accidents during wintertime in the years 1991–1993; detailed records of fatal accidents, in the years 1987–1991, generated by in-depth accident investigation teams and compiled by the Motor Insurers' Road Safety Committee (VALT). Replies to the postal questionnaires were received from 5,881 vehicle owners, giving a response rate of 59%. The replies included 296 self-reported accident involved drivers. The VALT data on fatal wintertime accidents contained 658 drivers involved in fatal accidents.

The analyses of the two data sources confirmed that driving conditions and kilometreage driven during the wintertime contribute to amount of accidents. The best explaining variables in both survival and risk models of wintertime driving were driver characteristics (age, driver behaviour, kilometreage driven, speed, and use of safety-belt), drivers' state (driving under the influence of alcohol) and vehicle characteristics (vehicle weight and condition of tyres).

Young and inexperienced on the one hand, or old (and experienced) drivers, on the other hand, had the highest fatal accident risks. There were no direct sex-related differences in accident risks between female and male drivers. Drivers using non-studded winter tyres had a somewhat greater accident risk than drivers using studded tyres, but the difference was not statistically significant.

Survival modelling with two distinct data sources, a postal survey data and an in-depth accident data, were used in the analysis. Modelling methods were also analysed by simulation and it was concluded that survival modelling promises to be a useful tool for safety analysis but the method needs further development.

FOREWORD

I have studied since the 1980's, mainly with Risto Kulmala and Veli-Pekka Kallberg, the use of accident prediction models for analysis of the safety of road links and later on for many other research purposes in the Technical Research Centre of Finland (VTT). During this, we learned the required statistical background and orientation to practical methodology, especially the principles of generalised linear models, from Professor Anders Ekholm from the University of Helsinki. In 1995, Risto Kulmala finalised his thesis on Safety at rural three- and four-arm junctions; Development and application of accident prediction models.

When working with accident prediction models, we realised that the development of transport modelling had since the 1970's been focused on disaggregated modelling based on data on individual behaviour. I started to investigate these models and their applications in the field of safety. Very soon I found out that Jovanis & Chang (1989) had been analysing highway accident occurrence by using disaggregated survival modelling. I found also some other articles but none of those dealt with modelling of the normal accident process. The topic in these analyses was usually the survival of certain known small population of identified drivers.

I started to study the theory and modelling methods, and made the first practical analyses during the beginning of the 1990's. I published my first review of this area in 1993, looking into disaggregated models and fatal accidents. This study was done with the support of VTT. On the basis of these analyses I concluded that survival modelling can be developed and used in explaining also the normal accident process based on survey data. I also noted that there still remained both theoretical as well as technical problems. Technical problems were mainly related to software. During the beginning of the 1990's, there were no proper tools available, but very soon e.g. SPSS released a version including proportional survival models.

I applied survival theory based modelling techniques during the Winter and Road Traffic Project commissioned by the Finnish National Road Administration (FinnRA) to explore the safety effects of studded tyres. The contact person in FinnRA, MSc. Anne Leppänen, gave me valuable support and advice on the client view of the required form and use of results.

During the accident risk analysis of car drivers in wintertime traffic, my closest colleagues were those invaluable supporters who gave me answers to most of the questions and problems I had. Risto Kulmala was my closest supporter during this long process from the beginning to the end. He also helped me with the first translations into English.

PhD. David Zaidel (VTT) gave his substantial support in the final phase of this work. He gave me advice, comments and critique on the whole research and helped me to produce

a decent research report. Arja Wuolijoki has all the time been involved in the reporting and also finalised this report with its layout.

Professor Matti Pursula supervised the final phase of my work. When starting the work my supervisor was Professor Sulevi Lyly, who retired during the process. During the preliminary examiner phase of the work PhD, Senior lecturer Pertti Laininen from the Helsinki University of Technology gave me valuable advice and suggestions to improve the work from the point of view of statistical modelling and analysis. I also want to thank the other preliminary examiner Professor Matti Syvänen from the University of Tampere for his great support and effort. I have also received support and advice on statistical modelling from Professor Anders Ekholm from the University of Helsinki.

I want to thank all colleagues who have contributed this study and supported me. I want especially to thank my family Helena, Milla and Henna who have endured this all during so many years.

SUMMARY

The study examined the effects of human factors, kilometreage driven, vehicle characteristics, and roadway factors on amount of accidents and risks of driving during wintertime months. The study was based on data from the project "Winter and Road Traffic" by the Finnish National Road Administration's (FinnRA) which was reported in 1995. The objectives of the study included both analysis and explanation of driver wintertime accident risk factors and investigation of the survival modelling method.

The theory of hierarchical systems served as the reference framework for the study. It assumes that traffic can be viewed as several systems at different levels. A general system, which covers all mobility, is at the highest level. The next level is that of the traffic system itself. The traffic system can be further divided into mode specific transport systems. Safety in the road transport mode, which is the focus of the present study, can be examined as a driver-vehicle-traffic environment system.

The study involved a postal questionnaire addressed to vehicle owners during the years 1991–93, combined with analysis of the data-base of fatal accidents examined by accident investigation teams during 1987–1991. This data-base, compiled by Motor Insurers' Road Safety Committee (VALT), included information on 658 drivers who had been involved in fatal accidents during the winter months. The postal questionnaire was addressed to a random sample of 10,000 vehicle owners. Replies to the postal questionnaires were received from 5,881 vehicle owners, giving a response rate of 59%. The final data-base contained information on 4,344 drivers, 296 of whom had reported to have been involved in an accident sometimes during the winter months of 1987 to 1991.

Drivers' wintertime accident risks and their dependence on various factors were examined using statistical accident models. Modelling survival is commonly applied in medicine to the study of serious diseases and treatment methods. An evaluation method for accident risks based on survival models has been previously tested (Roine 1993b). The present study further assessed the development needs of survival models in the area of traffic accident analysis.

Several factors have a bearing on the risk of wintertime accidents. Based on the questionnaire data, the probability of drivers being involved in wintertime accidents can best be explained by driver's age and sex, kilometreage driven in wintertime, the proportion of the kilometreage driven in built-up areas, the age of the vehicle and the interactive terms of the explanatory variables.

According to the VALT data, nearly all the same variables also proved to be important risk factors in fatal accidents. Other successful variables in the models based on the VALT data included driving under influence of alcohol, the use of the safety belt, the speed at the site of the accident, the weight of the vehicle and the tyre-tread depth.

More specifically, the combined analyses of the two data sources confirmed the following answers to the main questions of this study concerning the accident risk of winter driving:

- kilometreage driven during winter driving conditions determines wintertime accident probability as a unique risk factor – the more kilometres are driven during the winter, the higher the probability of being involved in an accident.
- wintertime traffic places the highest demands on young, inexperienced drivers as well as on old, experienced drivers.
- there does not appear to be clear sex-related differences between the accident risks of male and female drivers, when taking also account their driving experience in wintertime conditions
- some vehicle characteristics can be separated into own individual risk factors but they are strongly interrelated to many other risk factors.

Drivers of non-studded winter tyres had a somewhat larger accident risk than the users of studded tyres, but the obtained difference was not statistically significant.

The effects of driver's motives and attitudes were not directly modelled in this study, but some of their influence can be assumed to underlay the effects of driver age, sex, or experience, as well as determine their choosing to speed, drive under influence of alcohol, or not use the safety belt.

The survival modelling method was analysed by the models produced during the study and with the models based on simulated data. The analysis included model formulation, estimation, model validation and comparisons. The main survival model type used in the analysis was Cox type distribution-free proportional model. The explanatory power of the produced survival models and the consistency with data and background knowledge seemed to be as good as e.g. logit models used in the comparisons. The simulation analysis pointed out that some explanatory power is lost when the survival models are based on random time period. The results proved also that better exposure data e.g. time spent in traffic used in the analysis could improve the models. The application of survival models to the accident data appears to be a promising approach but needs still further development.

CONTENTS

APPENDICES

A: Questionnaire layout
B: Survival models compiled from the postal survey data
C: Survival models compiled from the VALT data
D: Calculation example of a survival model

List of Figures

List of Tables

1. INTRODUCTION

1.1 Objectives of the study

The objective of the study was to investigate the following issues:

- which factors influence the accident risk of drivers in wintertime traffic and especially how does accident risk depend on the type of tyres on the vehicle,
- how do principles of survival modelling apply to investigating accidents and accident risks,
- what are the data requirements in using survival models, and
- how can accident risks be evaluated by survival models when data include only basic information on the parties involved in an accident.

Drivers' wintertime accident risks and the factors influencing them were examined with statistical accident models. Survival modelling is commonly applied in medicine to the study of serious diseases and treatment methods (e.g. Kalbfleisch & Prentice 1980; Lee 1992). An evaluation method for accident risks based on survival models has been previously tested (Roine 1993b).

The data for modelling were derived from two sources. One was a postal survey by a questionnaire sent to 10,000 vehicle owners selected at random from the Motor Vehicles Registry. The survey (59% response rate) provided information about wintertime driving practices and experience in the years 1991–93.

The other source of data were the wintertime fatal accidents documented by accident investigation teams during 1987–91. A total of 658 drivers were on record.

The postal survey provided data on drivers who *were*, as well as on drivers who *were not* involved in traffic accidents during the reference period. However, the fatal accident data-base only includes drivers who *were* involved in accidents. Roine (1994) and Roine (1996) previously reported winter accident risks based on cohort and case-control methods as well as logit models. Deriving accident risks based on survival models from these two sets of data was, therefore, a special challenge. The present study further assessed the development needs of survival models in the area of accident analysis.

1.2 Theory of hierarchical systems

Drivers are faced with risky situations and potential accidents every time they are on the road. Counter-measures are actions taken by society to prevent accidents or moderate their consequences. Such measures are based on ideas regarding why and how dangerous situations and accidents evolve. These ideas can be interpreted as theories on the accident process, even though they might not have the form of formal theories.

Several theories on the occurrence of traffic accidents have been proposed. A basic classification of the theories has been suggested by a number of authors (Elvik 1991, Häkkinen 1978):

- human oriented theories,
- environmental or risk factor theories,
- causal chain theories,
- system theories, and
- theories of hierarchical systems.

This study is based on the theory of hierarchical systems. It assumes that traffic safety consists of several systems on different levels. A general system, which covers all mobility, is at the top level. At this level, traffic safety is determined by how much mobility can be decreased or its growth be restrained. This is due to the fact that accidents depend more on mobility, exposure, than on any other factor.

The next level in the hierarchy is the level of the traffic system itself, where it is vital to find the way to manage vehicle mileage as safely as possible. The traffic system can be further divided into mode-specific transport systems, in which case safety can be examined as a driver-vehicle-traffic environment system. Drivers' actions can be further examined within these sub-systems, e.g. as a detection-judgement-action-feedback system (Häkkinen 1978).

This research project will focus on the person-vehicle-traffic environment system. According to Häkkinen (1978), the successful functioning of the system requires that its parts – the person, the vehicle and the traffic environment, as well as the ”functions” which connect them – should remain within certain limits of variation. According to the theory, accidents happen when road users cannot adapt their actions to the varying demands of the traffic environment. Consequently, the risk of accidents can be lowered by improving road users' performance in traffic or by reducing system demands on road users (Elvik 1991). Put another way, humans inevitably make errors but by altering the circumstances in which they operate one can minimise the frequency of errors or moderate their consequences (Elvik & Vaa 1990, Björnskau et al. 1994).

Wintertime accidents have, for the most part, been examined from the standpoint of either traffic environment, winter conditions or the behaviour of individual drivers. Such studies have not sufficiently taken into consideration the impact of possible differences between driver groups or interactions between factors. The systems theory approach in the present study examines which characteristics connected with drivers and vehicles, together with the prevalent roadway conditions, influence the occurrence of wintertime accidents.

1.3 Traffic accident probability and risk

The occurrence of traffic accidents can be explained only to a limited extent by a deterministic causal relationship, in which the occurrence of certain conditions will always lead to accident consequences. Due to this, the occurrence of accidents is often explained with a probabilistic causal relationship, in which the occurrence of the cause will increase the probability of the occurrence of the effect (consequence) (Uusitalo 1974).

In a probabilistic model accidents are assumed to be produced by a stochastic process that can be described by a series of trials with certain outcomes (Elvik 1991). In the classic definition, probability is the ratio between the amount of positive outcomes and all possible outcomes. According to the frequency interpretation of probability, probability is defined as a long-term frequency, which is estimated by the relative frequency of a certain occurrence. The basis of the frequency interpretation is the empirical observation that in the case of random-nature mass phenomena, the relative frequency of a certain occurrence is reasonably stable (Niiniluoto 1983).

The frequency interpretation of accident probability means that it is the long-run number of negative outcomes (accidents) divided by the total number of trials (positive and negative outcomes). From this follows the basic equation:

$$\text{The expected number of accidents} = \text{Number of trials} \times \text{Accident probability} \quad (1)$$

In safety research number of trials is usually called “exposure” and accident probability is the “accident risk”. When applying the frequency interpretation, exposure is the number of traffic situations that can potentially result in accidents (number of trials) per a selected unit of time. The ratio between the number of accidents and the exposure is therefore the relative frequency. The probability of an accident (accident risk) is the limit value of the relative frequency.

In defining population accident risks based on actual accident statistics, data on accidents and exposure to accidents are usually gathered separately. Statistical accident risks are often represented by accident rates, which are formed as the ratio between the accidents that have occurred and aggregated vehicle mileage. Vehicle mileage is presumed to correspond with exposure to accidents.

It has been common to assume that accident rates calculated with vehicle mileage are comparable across varying circumstances, the way theoretical accident risks are. This, evidently, is not the case. As Hauer (1997) has demonstrated, the relationship between accidents and amount of traffic may not be linear and, therefore, one must consider a risk function relating accidents and mileage rather than a single rate.

Accident rates represent the expected number of accidents per vehicle mileage unit. Another estimate of accident probability can be derived from the ratio of accident involved drivers to non-involved drivers in similar situations (in which case the estimate receives a value between zero and one). When traffic conditions for the two groups are similar, the estimate for the risk of driving in a given condition may actually be a better estimate of accident probability than one based on general mileage exposure. However, such data are hard to come by (Elvik & Vaa 1990).

1.4 Traffic risk factors

Systems theory assumes that accidents are generated by numerous factors from different levels of the hierarchy. These factors often have interactions and confounding effects with each other. Such factors may be considered as accident causes if they either increase or decrease the probability of accident occurrence. Therefore, in order to prevent accidents, one must know which of the numerous traffic risk factors have a real strong influence on the number and probability of accidents.

The risk factors considered here represent human factors and mobility, vehicle factors and traffic environment factors. In the following sections, an overview of the risk characteristics of each sub-group of factors is provided. The next chapter describes in more detail previous findings about the risks of several of the factors mentioned here.

1.4.1 Human factors

People have driving objectives and are the most active actors in the system. While their actions might be influenced by subconscious motives and subliminal cues (Uusitalo 1974), they are also the most adaptive element in the traffic system. They can create risky situations as well as respond to ever changing new demands of the traffic environment.

According to studies in the 1970’s and 1980’s, factors associated with the road user were the direct cause of about 95% of accidents investigated. Factors associated with the traffic environment were the direct cause of 28–34%, and factors associated with the vehicle directly caused only 8–12% of accidents (Elvik 1991). The factors are overlapping which explains that the total exceeds 100%. All of the studies also acknowledged the contribution of background or indirect factors to the causation of accidents.

Regardless of which evaluation model was used to assess the causes of accidents, human factors have been given a prominent position in it. Human factors relevant to driving behaviour can be classified in different ways. One classification, suggested by Marek & Sten (Elvik & Vaa 1990), divides the factors into four groups (Table 1):

- permanent or slowly changing physiological characteristics,
- temporary physiological characteristics,
- permanent psychological characteristics,
- temporary psychological characteristics.

Table 1. Human psychological and physiological factors that have been studied as possible accident risk factors (Elvik & Vaa 1990).

Permanent or slowly changing physiological factors	**Permanent or slowly changing psychological factors**	**Temporary physiological factors**	**Temporary psychological factors**
Age	Intelligence	Fatigue	Biorhythm
Sex	Personality factors	Stress	Emotional state
Vision	Attitudes	Alcohol	Breakdowns in concentration
Hearing	View of own skills	Drugs	
Reaction time	Recognition of dangerous situations	Acute illnesses	
Physical disorders and deficiencies	Mental illness	The menstrual cycle	
A heart condition		Pregnancy	
Diabetes			
Epilepsy			

Elvik and Vaa (1990) summarised the research about the relative impact of the above human factors on accidents (Table 2). The impacts of driver's age, vision, reaction time, knowledge of traffic regulations, actual skills, fatigue and use of alcohol on traffic accidents have been proven by research. The impacts of driver's sex, personality, and use of drugs are less conclusive.

The models and theories of driver behaviour are constantly developing and their development has been speeded up by the application possibilities of information technology in traffic. The models can be classified as follows (Björnskau et al. 1994, Björnskau 1994):

- models which describe behaviour,
- perceptual and cognitive models,
- motivational models and theories.

The models that describe behaviour structure driver behaviour according to the basic driving tasks but do not deal with the causal factors or mechanisms of behaviour. The perceptual and cognitive models deal with the processing of information performed by a driver. According to the motivational, intentional, models, drivers' intentions (motives) explain their behaviour and actions in traffic.

Table 2. The causal nature, certainty and significance of human factors variables in traffic accidents (Elvik & Vaa 1990).

The variable or group of variables	The causal nature of the effect	The causal relationship according to the research results	The significance from the point of view of the accident risk and exposure
Age	Indirect	yes	great
Sex	Indirect	varying results	minor
Vision	probable	yes	minor
Hearing	probable	exists according to some reports	minor
Reaction time	certain	yes	not known
Physical injuries	probable	exists according to some reports	minor
A heart condition	improbable	exists according to some reports	minor
Diabetes	improbable	exists according to some reports	minor
Epilepsy	improbable	exists according to some reports	minor
Intelligence	Indirect	exists according to some reports	minor
Reading disorders	probable	no	minor
Knowledge of the regulation	probable	yes	minor
Personality	probable	varying results	minor
Attitudes	certain	no	not known
View of skills	certain	exists according to some reports	great
Actual skills	certain	yes	great
Mental illness	Indirect	some reports	minor
Fatigue	certain	yes	moderate
Physical stress	Indirect	exists according to some reports	not known
Alcohol	Indirect	yes	great
Drugs	Indirect	varying results	not known
Acute illnesses	Indirect	exists according to some reports	minor
Menstrual cycle	Indirect	exists according to some reports	minor
Pregnancy	Indirect	no	minor
Biorhythm	Indirect	yes	minor
Emotional state	probable	no	not known
Breakdowns in concentration	certain	no	not known

Motivational models have been developed as a reaction to system theory's strong emphasis on the traffic environment. For example, risk compensation model (Wilde & Murdoch 1982) presumes that drivers intend to maintain a certain level of risk when driving. It is presumed that drivers perceive a decreased risk level associated with road improvements and, therefore, compensate for the available safety benefit by altering their behaviour, e.g. by increasing their speed of driving.

Ultimately, human factors influence traffic accident risk through drivers' behaviour. Even the impact of non-human factors on accident can be modulated by human behaviour. The potential risk effect of a road condition or a vehicle factor can be dampened or enhanced by drivers' alertness, fatigue or compensatory behaviour. Conversely, in the background of many human errors often lie misleading and over-demanding roadway, vehicle, or environmental factors.

1.4.2 Traffic and roadway factors

The environment with its roadway network creates the framework for the behaviour of traffic and exposes those who are on the network to various accident risks. Traffic environment can support and promote safe behaviour, but it can also encourage or lead to risky behaviour. The possible accident risk and exposure factors associated with traffic and its environment are:

- amount of traffic,
- characteristics of the traffic,
- road networks and land use,
- roadway conditions,
- management and control of traffic.

The amount of traffic relates to the amount of mobility and, therefore, represents exposure to accidents. Accident risks are dependent on the amount of traffic. Various studies (e.g. Kulmala 1995b) have ascertained that accident type distribution depends on the amount of traffic. The amount of traffic affects other traffic characteristics such as speed, headway distribution and the amount of overtaking, all of which can influence accident risks.

Roadway networks are made up of various types of junctions and sections, which have different risks associated with them. Land-use characteristics influence mobility and access needs which determine the structure and density of a network, the modes of transport, and their potential accident risks.

Traffic control and management are needed to make sure that traffic is as fluent and safe as possible on the network. Control devices and management procedures regulate drivers' mobility and behaviour in traffic. Therefore, they directly influence exposure as well as accident risk.

1.4.3 Vehicle factors

Case studies of traffic accidents have found that in a small proportion of vehicle factors (such as mechanical failures) were direct and primary causes of the accidents. Sudden failures associated with vehicles were just 1% of the “key events” in accidents investigated by Finnish teams. This is compared to 70% to 80% of the events attributed to drivers’ errors in vehicle control and operation, anticipation, and judgement (Karttunen 1994).

The role of vehicle factors appears to be larger in fatal accidents. In fatal collisions investigated by the teams, 22–26% of the “key events” evaluated were associated with the vehicle, 15–28% with the traffic environment, 43–56% with human factors and about 1% with traffic regulations. The most cited vehicle factors were unsuitable or worn-out tyres (20–42%), misuse or malfunction of safety belts (11–27%), poor crashworthiness (27–43%), and defective steering (Karttunen 1994).

1.4.4 Environmental conditions factors

Up to 20% of all the risk events analysed by accident investigation teams in Finland (Karttunen 1994) were related to characteristics of the traffic environment – mainly road surface, weather and lighting conditions.

Accident risk is usually 1–2 times as great in the dark as it is in the light (Kulmala & Peltola 1985). Darkness is often associated with weather and road surface conditions and with the amount of traffic as these factors may vary with time of day. The accident rate on wet roads is usually higher than on dry roads. The “Winter and Road Traffic” study showed that injury accident rates on the main roads of southern Finland were 13.6–38.9 on slippery road surface conditions and 6.1–21.7 on dry road surfaces (injury accidents/million vehicle kilometres). The corresponding values for central and northern Finland were 15.2–21.3 and 4.7–23.0; 6.7–16.2, and 5.4–23.8, respectively (Alppivuori et al. 1995).

These results agree with previous Finnish and Nordic studies. Polvinen (1984) found that, compared to dry road surfaces with sufficient friction, accident rates on icy or slushy road surfaces were 12 times and 7.4 times greater, respectively. Salting significantly reduced the accident rate. The Swedish Road and Traffic Research Institute (VTI) found that accident rates on icy and snowy roads were 2–3 times greater in northern, 3–6 times greater in central, and 7–10 times greater in southern Sweden, compared to the corresponding rates on dry road surface conditions (Valtonen 1986).

Pavement's physical condition is also a risk factor. Studded tyres have been found to wear out pavements and cause ruts on them. The impact of rutting on safety has been researched in Finland and elsewhere, especially in the 1980’s (Kallberg 1983, Hemdorff et al. 1989). While in dry conditions accident rate was 7% lower on rutted pavements than on undamaged ones, the reverse trend was observed in wet surface conditions.

2. DETAILED REVIEW OF PREVIOUS RESEARCH RESULTS

2.1 Driver characteristics

2.1.1 Driver age and driving experience

Driver age and sex are naturally not causal factors of accidents. They sum up well the effects of the other factors which actually influence the risk, such as mobility, driving experience, reaction time, motives and other factors (Elvik & Vaa 1990).

The relationship between accident risk and age is usually U-shaped. The youngest and the oldest drivers, both male and female, have a clearly higher than average accident risk (Elvik & Vaa 1990).

For example, Broughton (1988), found that 17–20-year-old male drivers had an average risk rate of 346 (per/100 million vehicle kilometres). The rate decreased with age to 79 in the 39–43-year-old age group and it rose again all the way to 164 in the age group of 74+ years old.

The injury accident risk by age and sex varied considerably at different times and on different days. The young male driver injury rate was at its highest at about 10–12 p.m., the under-65-year-old driver injury rate peaked at about 4–6 p.m. and that of the over-65-year-olds at about 12 a.m. –4 p.m. The under-25-year-old male driver injury rate was at its highest on Saturdays, but with most other drivers, on Fridays.

A study in the USA found that the accident rate in the 16–19-year-olds and the 75+ year-olds is considerably higher than the average rate. The fatality rate (number of fatalities/100 million miles) was 9.2 in the 16–19-year-olds, 1.8 in the 40–44 year-olds, and 11.5 in the 75+ year old age group (Massie et al. 1994).

The older drivers had the highest fatality rates, while the youngest drivers had the highest accident involvement rates. Male drivers fatality rates were higher than those of female drivers. However, female drivers had higher injury rates and accident involvement rates than male drivers (Massie et al. 1994).

Similar results concerning drivers' accident rates and age were reported in Finland (Ernvall & Pirtala 1992). Vehicle kilometres data were gathered in October 1990 at vehicle inspection stations, and accident data were taken from insurance companies' records for the years 1987–89. The average accident rate in the entire population was 2.3 accidents/million vehicle kilometres. It was 18 in the 18–19-year-old group going down to 2 in the 30–54-year-olds group and increasing again in the older age groups.

Young drivers' accident involvement has been clarified in more detail in a Finnish study of young servicemen (Hatakka et al. 1995) which addressed their choice of vehicle, driving habits and attitudes. Risk taking attitude was related to driver's traffic offences and accidents, choice of vehicle and the amount and nature of driving. However, young drivers can not be considered a homogenous group, as only 17% of the young drivers were considered risky drivers.

Similarly, a European study (Lynam & Twisk 1995) listed the special factors that may underlie the association between young age and accidents. They include: psycho-biological immaturity, a limited recognition of danger, the acceptance of risk, an excessive belief in one's own abilities, lack of experience, driving culture, and lifestyle induced risky type of exposure such as night driving. It is typical for young drivers' accidents that they

- occur during the week-end and at night,
- are often single accidents,
- are associated with speeding as major risk-factor,
- tend to be serious and
- involve mainly male drivers.

With young drivers, the direct effect of age, as such, on accidents is small, while the impact of other, age-associated, factors, such as alcohol, speeding, inexperience and motivation, is relatively great (Rajalin et al. 1989). As drivers become older, the relative impact of age on accident risk increases in comparison with other factors. Ageing begins to clearly affect the number of accident fatalities at the age of 60, when vision and hearing begin to weaken, reactions slow down, muscles weaken and there are problems with maintaining and developing knowledge and skills. The injuries sustained by older road users in an accident are more serious, in comparison to younger ones, due to general frailty that advances with age (Mäkinen 1985). This can contribute also the accident risk estimates.

The effects of the age and driving experience are difficult to distinguish from one another, because age and driving experience are highly correlated with each other. From the point of view of accident risks, driving experience is accumulated slowly. It has been stated that the development of driving skills with experience continues until drivers are middle-aged or at least until a driver has driven about 50,000 kilometres (Näätänen 1988). It has been stated as another estimate that the optimum level of driving experience is reached after driving for 7–8 years or about 100,000 kilometres (Häkkinen et al. 1996).

With novice drivers, it has been estimated that a decrease in accident risk is due more to the accumulation of driving experience than that of driver's age, while with old drivers it is equally due to both factors.

It is evident that in addition to the mileage driven, the conditions in which the mileage has been accumulated also influence driving experience. Therefore, quite detailed information would be needed to reliably estimate the impact of driving experience (Keskinen et al. 1994).

2.1.2 Driver sex

Systematic differences have usually been found in the fatal accident rates, degrees of injury and accident rates of male and female drivers. Evans (1987) concluded that according to various comparison criteria (fatal accidents/ vehicle kilometres, participation in accidents leading to the death of vulnerable road users, etc.), male drivers' level of traffic safety was lower than that of the female drivers.

According to a more recent study in the USA, male drivers' fatality rate was 55% higher than that of the female drivers (Massie et al. 1994). Female drivers' injury rate was 25% lower than of males, but the total accident rate was 16% higher than that of male drivers. These findings are very similar to results obtained in GB Broughton (1988).

Massie et al. (1994) discussed the possible causes for the differences between the accident rates of male and female drivers. It was suggested that the differences between the accident rates were possibly due to smaller exposure by female drivers, differences in the specific traffic environments experienced by the two groups, differential reaction times and errors of judgement. Difference in accident reporting by males and females was also suggested as an explanation for the disparity.

The differences found in accident risks between female and male drivers were analysed in more detail, with the help of survival models, at the University of Washington in 1990 (Mannering 1993). The study was based on a survey conducted on 200 students. Statistically significant differences in the relative accident risks between male and female drivers as well as in the background factors were found. Accident probability increased with mileage for both male and female drivers. Married drivers were found to have smaller risks than unmarried drivers, and drivers who had a good income, were involved less in accidents than drivers who had a low income. Suggested explanatory variables for the differences between these driver groups were exposure, risk behaviour, and the safety of the vehicles used by the groups.

According to a Finnish study (Ernvall & Pirtala 1992) based on insurance companies' accident data, female drivers' accident rate was about 10% lower than that of male drivers, in built-up areas. Young male drivers were prone to be involved in rear-end accidents. Female drivers were more likely than men to be involved in accidents while reversing.

Finnish in-depth accident investigation teams reported that vehicle control skills and understanding of traffic situations were common background causes in accidents of drivers who had only recently obtained their driving licenses. These were more often noted in the case of females than males. For example, the accident risk of novice female drivers, during wet or slippery road conditions, increased more than that of males.

Female drivers, more often than male drivers, were the principal guilty parties in the accidents investigated by the Finnish teams. The difference was largest for the experienced drivers involved in the accidents (Laapotti 1991).

Other risk factors, such as speeding, use of alcohol, carrying large number of passengers, and not using a safety belt, were more common with male than with female drivers.

It should be pointed out that other studies, (e.g. Spolander 1983 as reported by Rajalin et al. 1989) found little difference between male and female drivers' average accident risks once the type and nature of exposure were taken into account.

2.1.3 Condition of drivers

The effects of alcohol use on accident risk have been examined extensively. Elvik and Vaa (1990) state as a summary of many studies that no other single human factor affects the occurrence of accidents as dramatically as alcohol. The risk of a fatal accident is over 100 times as high for drivers under the influence of alcohol, as it is for sober drivers. Drivers' accident risk grows exponentially as blood alcohol level increases.

Evans (1991) has estimated that in 1982 about 53% of traffic accident fatalities in the USA had alcohol in their blood. The total elimination of alcohol use would have decreased the amount of fatalities by over 50%.

Zador (1991) estimated that a driver with a blood alcohol level of 0.5–0.9 g/l had at least nine times the risk of being involved in an accident compared to the average sober driver. Alcohol increased the accident risk of young drivers more than that of old drivers and that of female drivers more than that of male drivers.

According to epidemiological and experimental research, the use of sedatives increases driver's accident risk. Accident risk is also increased by drug abuse.

Sudden strokes, seizures and other medical emergencies while driving have been the primary causes of accidents in 1% of accident cases.

Several studies confirmed that prolonged, continuous driving leads to fatigue and increases driver's accident risk. Fatigue slows drivers' reactions and reduces their level of attention in traffic (Elvik & Vaa 1990).

Drivers who get stressed in difficult driving situations tend to have more accidents than drivers who get less stressed in similar situations (Elvik & Vaa 1990).

To sum up, driver factors associated with higher accident risks are young age, being male, inexperience, delinquent socio-economic background, risky exposure, and driving behaviour associated with several lifestyle, attitudes and personality factors.

2.2 Mobility of drivers

According to the Finnish National Road Administration’s passenger travel survey (FinnRA 1993) the annual average kilometreage of a car in Finland was 21,250 km. This was an increase of 9% compared to the 1984 survey. Privately owned cars were used less than cars owned by companies, 19,700 km and 29,100 km, respectively. As household income grows, so does the number of kilometres driven annually.

Pirtala and Ernvall (1992) calculated annual average kilometreage of cars that were in active use and were brought in for a vehicle inspection. The average car use in 1990 was 20,200 km. Cars driven by young, 20–30 year-old male drivers, accumulated 26,000 km annually. Vehicle kilometreage slowly decreased with driver age up to the 55-year-old age group and then decreased more rapidly. Female drivers drove less than men, in every age group. The average annual kilometreage was divided between built-up areas (8,500 km) and outside the built-up areas (11,700 km). Young drivers and female drivers drove proportionally more in built-up areas, compared to other drivers. An average of 7,900 km was driven during the winter, with winter tyres that were used for 4.8 months. Essentially similar results were reported by Mäkelä et al. (1993).

Vehicle kilometreage also depended on size and performance of cars. Cars over 1,200 kg were driven 87% more kilometres than small cars weighing up to 750 kg. The kilometreage of cars in the 80 kW + power class was double that of cars in the 35 kW power class or less.

While drivers who drive a lot experience, on the average, more accidents than those who drive little, their accident rate is actually lower. This finding holds also for exposure based on time spent in traffic (Elvik & Vaa 1990).

2.3 Vehicles and their equipment

2.3.1 Characteristics of vehicles

A car’s mass has been found to be the central factor influencing passengers’ injury level in accidents. If cars that weigh 900 kg and 1,800 kg collide, the risk of dying in the lighter car is 13 times higher than in the heavier car. In a single car accident, the fatality risk in a 900 kg car is 2.4 times that of a 1,800 kg car. Light cars cause others to have lower fatality risks, while heavy cars decrease one’s own fatality risks (Evans 1991 and Evans 1993).

Finnish studies demonstrated other car characteristics that affect accidents. Small mass but high capacity cars had the highest risks of fatal accidents. Rear-end collisions were common to large and powerful cars. Rear-wheel-drive cars were more risky in both the very small and the large car classes. The risk of getting into a fatal off-road accident increased as the vehicle’s mass increased and as the power-to-weight ratio increased. The off-road accident risk of rear-wheel-drive cars was 1.6 times that of front-wheel-drive cars. Cars involved in accidents at junctions were usually smaller than average. (Huttula & Ernvall 1994, Tapio et al. 1994).

The dependence between accident rate, vehicle's mass, and driver factors was studied in the USA in the early 1980's. Within same driver age groups, users of small cars had a lower accident rate than users of heavy cars. This was interpreted to reflect differences in drivers' risk behaviour (Evans 1983). Wasielewski & Evans (1983) observed that drivers of small vehicles drove slower, kept larger head-ways and were more likely to use safety belts.

A Finnish study (Ernvall & Pirtala 1992) based on insurance records found that relatively new, medium-sized family cars (1,300–1,600 cc) had a lower than average accident risk. So did cars equipped with diesel engines. A common characteristic among these vehicle groups was that their owners were predominantly middle-aged men, whose traffic accident risk is usually low. Old, inexpensive and predominantly rear-wheel-drive car models, favoured by young drivers, were less safe than other car models. Rear-wheel-drive cars driven by inexperienced drivers had a particularly high accident risk.

Young drivers who used rear-wheel-drive cars, reported more kilometres annually than drivers using front-wheel-drive cars, and did more driving in slippery road surface conditions and in the dark (Keskinen et al. 1994).

Hatakka et al. (1995) examined the relationship between young male drivers' risk attitudes, accidents, and the choice of car model. High-risk attitude drivers had more traffic accidents than other drivers of the same car model. Young drivers use relatively old, small cars, but also cars that have powerful engines. An Austrian study (mentioned by Hatakka et al. 1995) confirms popular knowledge about differences between men and women in car choice. Women, more often than men, selected their cars based on price and running costs, while men selected cars based on the accessories, performance and comfort. Clearly, a prerequisite for the comparison of vehicles' accident risks are better data on the amount and nature of exposure, the motives of driving and drivers' lifestyles.

2.3.2 Tyre type and condition

Studded tyres were developed in order to improve drivers' mobility and safety during slippery road surface conditions. The level of use of studded tyres in Finland during the winter exceeded 90% already in the mid-1970's. When it was found that studded tyres wear out the pavement considerably more than non-studded tyres, a research and development program was initiated to improved both tyres and road pavements. Regulations were put in place concerning the period of the year when studded tyres are allowed, as well as the number, force and protrusion of the studs (Valtonen 1986).

Saastamoinen (1994) reported that during the winter of 1993–94 almost 100% of cars and vans used some type of winter suited tyres. Of these, 93.8 were studded, 5.9% were non-studded type (including previously studded old tyres), and 0.3 % used summer tyres. Altogether, 21.7% of all tyres were new, while 1.8% were worn-out.

The significance of studded tyres for safety was investigated extensively in Europe and North America (Valtonen 1986). Studies in the 1970's lead to significant restrictions on use or a total abandonment of studded tyres. Studies in Quebec and Ontario, indicated that the use of studded tyres could be abandoned without jeopardising traffic safety.

According to North American studies, the safety margins created by the use of studded tyres are offset by a decrease in drivers' caution and the more careless manner of driving (Lundell 1993, Valtonen 1986).

The safety impacts of studded tyres as well as the problems caused by their use have also been studied in the Nordic countries. A study by Roosmark et al. (1976) was significant in showing that, when weather and surface conditions are considered, drivers who had used studded tyres had fewer accidents than drivers who had used ordinary tyres. The study estimated that a ban on studs would have increased the amount of accidents by 6%–13%. On the other hand, it was estimated that making studded tyres compulsory during the winter period (1.10–31.4) would have decreased accidents by 7%–15%. Another study, based on VALT data, estimated that the number of fatal accidents during 1979–1983 would have increased by about 10% if a ban on studs had been in force (Valtonen 1986).

Similar conclusions were reached in Norwegian studies although the positive impacts of studded tyres there were slightly lower (Huhtala & Kallberg 1978). The proposed mechanism for the positive effect of studs was that most drivers, with and without studded tyres, do not reduce their speed sufficiently when driving on slippery road surface; the studs increase the effective friction.

In Sweden, the difference in accident risks between studded tyres and summer tyres on slippery roads was investigated in a driver survey. The accident risk of cars with studded tyres was about 36–50% lower than that of cars with summer tyres (Öberg et al. 1993). More recent study (Carlsson & Öberg 1995) concluded that abandoning studded tyres would increase the number of wintertime accidents by 10–20%. The study further estimated that studded tyres reduce winter type accidents by about 25% in rural areas and by about 20% in built-up areas.

In the beginning of the 1990's, the impact of tyre type and the condition of the tyres on accidents in Norway were investigated with extensive interviews, surveys and tyre condition inspections (Ingebrigtsen & Fosser 1991). A main conclusion of the studies was that tyres with good friction capability reduced accidents in snowy and icy road surface conditions. Tyres' tread depth was particularly important. The authors state that drivers offset some of the safety benefits, possible with the use of proper tyres, with their risky behaviour.

Recent Norwegian study (Fosser & Saetermo 1995) found that users of studded winter tyres are different from users of non-studded winter tyres. The users of studded tyres are usually men, they have newer cars and they drive a lot in comparison with the users of non-studded tyres. When the differences between driver characteristics, driving conditions and vehicle types were all considered, no statistically significant differences in accident probabilities were found between users of studded and users of non-studded tyres.

Lundell (1993) used the 1987–90 VALT data (Finnish in-depth accident investigation teams). He compared the tyres of the "main guilty party" involved in the fatal accidents with those of the "other involved parties". Tyres' tread depth and stud protrusion were found to be in worse condition in the "guilty party" group. The innocent parties had somewhat newer and more expensive cars, which were in somewhat better condition than the cars of the main guilty parties. Driver characteristics were also found important. The main guilty parties included relatively more males, had less driving experience, and more of then were under the influence of alcohol, compared to the innocent parties.

2.4 Winter conditions

The various accident statistics records in Finland give quite similar figures of the share of wintertime accidents and their distribution. Central Bureau of Statistics records all road accidents reported by the police. Accident statistics of the Finnish National Road Administration include mainly accidents on non-built up areas, excluding also accidents in cities. Statistics of insurance companies (VALT) are based on reports by insurance holders.

The Central Bureau of Statistics (1992) reported 1,922 injury accidents during wintertime (1.11.–31.3), 1991 in Finland. National Road Administration's accident statistics put the number at 1,515 accidents (FinnRA 1992). According to the Central Bureau of Statistics, 35% of accidents in 1987–1991 took place in wintertime. National Road Administration's statistics place the number at about 37%.

When road condition at the time of the accident is considered, 74% of all wintertime injury accidents and about 71% of the fatal accidents, occurred when the road surface was snowy, icy or slushy.

The insurance companies' statistics for 1993 include 75,814 accidents, compensated by vehicle insurance, 37.2% of which had occurred on snowy or icy road surfaces (VALT 1994).

The number and distribution of winter accidents in Finland has changed little during 1989–1992, according to Central Bureau of Statistics records. In 1989: 30,521, in 1990: 29,277, in 1991: 30,594 and in 1992: 26,554 accidents. Male drivers were involved in 77–78% of the accidents and 24–25% of the accidents had occurred in the dark. 50%

of the accidents occurred on streets, 20% on public roads and the rest on other roads and areas. About 38% of these accidents happened at junctions.

When distinguishing accidents by vehicle direction, the largest groups were driving in the same direction (21–24%), intersecting driving directions (21–23%), opposite driving directions (about 13%) and off-road accidents (3–4%).

In recent years special attention has been given to frontal accidents due to their serious nature (Karttunen 1994). Frontal accidents have often occurred during wintertime conditions. According to VALT accident investigation teams, the most significant driver risk factors in frontal accidents are "speed too high for the situation" and vehicle control and handling errors. Roadway geometry characteristics were risk factors in only a small number of cases.

2.5 Driver behaviour

In addition to drivers' general abilities and skills, driving in winter time conditions may require certain special skills on the one hand, while amplifying the negative impacts of skill deficiencies, errors in judgement, or risk inclined motivations. Driver' perception of the objective hazards and of their own abilities to cope with them are often discussed as explanatory concepts for drivers' behaviour as well as accident involvement, (e.g. Elvik & Vaa 1990).

Näätänen and Summala (1976) argued that, generally, drivers do not base their decisions on safety considerations. The threshold of subjective risk is so high that road users do not have a sense of danger, although accident probability may have increased significantly.

Behaviour adaptation (OECD 1990) is another explanation for the failure of drivers to avoid accidents. It addresses the common finding that safety impacts of specific measures which were meant to improve safety, such as ABS-brakes and reflector posts, have proven to be smaller than expected. The explanation is that drivers have altered their behaviour (in an unsafe manner) and thus offset some of the potential benefits of the measures.

Studded tyres improve friction on otherwise slippery roads and the question is if drivers fully benefit from that advantage. OECD summary of Swedish and German studies in the 1980's on the impact of studded tyres (OECD 1990) concluded that drivers did not completely offset the better friction by increasing their speed. In fact, German studies reported that drivers who used studded tyres drove slower on icy motorways compared to drivers with summer tyres.

In the Finnish "Winter and Road Traffic" project of 1992–1994 it was found that driving speeds decreased by only 4 km/h in snowy road surface conditions in comparison with speeds in good road surface conditions (Saastamoinen 1993). One out of four drivers who travelled in platoons kept a headway distance of less than 1.5 seconds. Studded tyres were usually in better condition than the non-studded winter tyres, (2/3 of which were previously studded tyres) but no differences in driving speeds were found between vehicles equipped with different types of winter tyres (Heinijoki 1994). An analysis of speeds in curves and platoons suggested that drivers who used studded tyres had slightly greater safety margins than drivers who used non-studded winter tyres. However, these differences were not statistically significant (Roine 1993a).

The Finnish study included a driver survey. Drivers were underestimating the slipperiness of the road. When road surface conditions were slippery, over half of the drivers estimated that the road surface conditions were good, and only 14% thought road surface conditions to be slippery. The more slippery the roads were, the larger the discrepancy between actual and judged conditions was. Drivers slowed only slightly in slippery road surface conditions, and drivers who thought that the condition of their tyres was bad did not drive slower than the others.

Recent studies in Finland confirmed earlier results. Mean speeds on slippery roads are usually lower by only 2–4 km/h compared to good road surface conditions (Estlander 1995 and Rämä et al. 1996).

An experiment in the “Winter and Road Traffic” project compared behaviour of drivers, who were asked to drive either studded or friction tyres (Mäkinen 1994). Tyre type did not influence drivers’ mobility or amount of travel. Drivers in the friction tyres group drove at slightly higher speeds compare to drivers who used studded tyres. This was explained by the better comfort, the low noise level and good directional stability of the friction tyres.

The application of telematics in traffic is currently being examined in Finland for its potential to intensify traffic management and improve driver behaviour and safety in winter conditions. Preliminary results are promising. Changeable warning signs on slippery roads have been found to reduce driving speeds by about 2 km/h (Rämä et al. 1996).

Wintertime driving and safety has been also addressed by setting special wintertime speed limits, providing skill training on wet practice tracks, providing road surface weather information to drivers, and optimally scheduling winter road maintenance (Kallberg et al. 1991, Alppivuori et al. 1995).

3. MAIN QUESTIONS OF THE STUDY

Wintertime accident rates (or accident risks) are significantly higher during slippery road surface conditions compared to dry condition.

How can this fact be explained and what can be done to improve the safety of driving during wintertime conditions? Systems theory suggests that traffic accidents occur when road users can not sufficiently adapt their own performance level to the varying demands of traffic situations. However, the theory also suggests that various environmental and vehicle factors interact with driver factors and increase or decrease the probability of their making errors.

The study used new empirical data, based on survey questionnaire, to identify the main factors and to quantify their contribution to the probability of wintertime accidents. Specifically, the following variables were considered:

- driver age and sex,
- amount and nature of driver's mobility,
- characteristics and accessories of the vehicle,
- use of studded tyres.

The data collected by accident investigation teams (VALT data) were analysed in conjunction with the survey data, as well as their own, to gain additional insight into the following factors, which may influence drivers' accident probability and risk:

- socio-economic characteristics,
- driving experience,
- driver condition, especially the use of alcohol,
- factors which describe driver mobility,
- vehicle characteristics, accessories and ownership,
- speeding behaviour.

The main problems were formulated according to the objectives of this study (chapter 1.1, p. 17) emphasising the accident risk factors. They are based on previous research results, showing that traffic accident risks vary with driver characteristics, vehicle attributes, and the time and place of driving.

The main questions are:

1. Can driver involvement in wintertime accidents be examined on the basis of exposure over time; specifically, the amount and nature of wintertime mobility?

2. Do age and sex capture differences in accident risk at a general level, as background factors?

3. Do vehicle characteristics contribute to accident risk beyond their interaction with exposure and speed?

4. Does the use of studded tyres decreases driver wintertime accident risk?

The methodological objectives were not formulated as main problems. They are discussed in the context of data analysis and interpretation.

Statistical hypothesis testing will be used as a tool when the main problems are solved in the analysis.

In this study, it was possible to estimate accident probability (accident risk) of drivers using the direct frequency interpretation. The accident probability of drivers in wintertime traffic was calculated from a random sample (the postal survey) of drivers in wintertime and from data on accident drivers in fatal wintertime accidents in Finland (VALT data).

Driver wintertime accident probability (accident risk) according to the postal survey is estimated as:

$$\text{Probability of accident} = \text{Number of accident drivers} / (\text{Number of accident drivers} + \text{Number of non-accident drivers}) \qquad (2)$$

This is also how survival models can be interpreted. They produce the share of non-accident drivers (1 – share of accident drivers) as a survival function of time.

In the fatal accident data, we use similar definition but because all of the drivers in this data-base have been involved in accidents we consider “first parties” as accident drivers and “other parties” as non-accident drivers.

Driver, vehicle, and other variables in the accident data can be analysed as risk factors that may have an influence on basic driver accident risk. These risk factors either increase or decrease the basic level of calculated accident probability or accident risk.

4. DATA

4.1 Data Collection Approach

The methodological approach of the study required that data on accident involved and accident free drivers should come from a random sample of the driver population. This was accomplished through a postal survey. Even in a large survey the expected number of accident involved drivers could not be high. Furthermore, the number of questions in a postal survey must be limited. Therefore, we used a complementary set of data that included detailed information on a relatively large number of accident-involved drivers. This data-base was originally created by the VALT in-depth investigation teams.

4.2 Postal survey data

4.2.1 Determining sample size

A pilot sample was used in order to test the questionnaire and to estimate the required sample size. The sample size depends largely on the proportion of drivers who used non-studded winter tyres. An earlier study found this proportion very small. The sample size was calculated using both the Mantel–Haenszel method (McPherson 1990) and the case-control method (Björnskau 1994).

In the Mantel-Haenszel method, the sample size was estimated on the basis of the expected relative difference in risks and its 95% confidence interval. The 2 x 2 cross tabulation showed accident drivers and drivers in the control group who used either studded or non-studded tyres. For estimating sample size requirements , it was assumed that 8% of the users of studded tyres (TVH 1982) and 10% of the users of non-studded winter tyres were wintertime accident drivers, that the relative risk of non-studded tyres was about 1.30 times that of others (Huhtala & Kallberg 1978) and that the response rate would be about 60%. The required amount of observations was calculated so that the difference between the relative risks of drivers who used studded tyres and those who used non-studded winter tyres could be estimated with the probability of 95%. In practice, this condition meant that the lower limit of the 95% confidence interval calculated for the odds-ratio had to be > 1,00 (chapter 5). According to the calculations, the number of observations in the 2 x 2 cross tabulation were therefore 26,000 and after the expected response rate was taken into consideration, the required sample size was about 43,300 vehicle owners.

The reliability of the differences of the relative risks at different sample sizes can also be estimated based on the proportions of involved drivers. Therefore the normally distributed test score z (Snedecor & Cochran 1980) is:

$$z = \frac{(p_1 - p_2)}{\sqrt{p_1 \times q_1 / n_1 + p_2 \times q_2 / n_2}} \tag{3}$$

where

- z is the normally distributed random variable
- p_1 is the share of accident involved drivers who used non-studded winter tyres, out of all the drivers who used non-studded tyres; here 0.10
- p_2 is the share of accident involved drivers who used studded tyres, out of all the drivers who used studded tyres; here 0.08
- q_1 is $1 - p_1$
- q_2 is $1 - p_2$
- n_1 is the number of involved drivers who used non-studded winter tyres
- n_2 is the number of involved drivers who used studded tyres.

When p(z) reaches the needed 1.96 (normal distribution, probability of 95%), n_1 is about 633 and n_2 about 24,667 giving the needed number of observations as 25,300. Taking into account the response rate 60 % this gives the required sample size as 42,200.

This sample size would provide a reliable estimate of relative risk (within 95% confidence limit). A sample size of only 10,000 drivers would provide an estimate with a confidence interval of 83%.

4.2.2 Survey procedure

Postal Survey 1991–1993 of vehicle owners

Because of financial limitations, it was not possible to survey the required sample size of 42,000 vehicle owners. A random sample of only 10,000 vehicle owners could be taken from the Motor Vehicles Registry in the spring of 1993. These vehicle owners were sent a questionnaire. It inquired about the owners' personal background, car ownership, car use, opinions about winter tyres ,the tyre type of the car mainly used by the owner, and information on accident involvement during 1991–1993 (Appendix A). As an incentive to return the questionnaire, owners were told that three sets of tyres donated by Nokia Oy would be raffled amongst drivers who would reply.

Answers were received from 4,168 vehicle owners, in the first survey wave. A second form was sent to vehicle owners who had not yet returned the questionnaire. A further 2,604 answers were received after this second wave, so that the total number of replies was 6,772. Of these, 5,881 were usable replies that contained an adequate amount of data. The final response rate was 59%.

Postal survey of accident involved drivers during wintertime 1992–1993

All 3,000 drivers who had been involved in accidents during winter 1992–1993 were sent the same postal questionnaire. Of the 898 returned forms 763 were usable, an effective return rate of about 30%. These data were only used during the preliminary comparisons of the surveys.

Fine tuning the sample for analyses

5,881 completed forms were available for analysis. (Non useable returns included forms lacking personal data, those stating they have no car or a driving license). After removing respondents who reported not driving at all, a total of 5,423 valid owners, primary users of their cars, remained for detailed analysis. Of this group, 660 (12.2%) reported having an accident during 1991–1993.

Further cleaning of the database included the following steps.

1. Drivers, who had not driven at all during 1991–1993 or during winter months, were eliminated from further analysis; 5,101 drivers remained.
2. Drivers whose reported accident data were lacking, inconsistent, or out of the boundaries of the defined time period were removed; 4,518 drivers remained, 346 of whom were involved in wintertime accidents.
3. Three distinctly different vehicle groups were present in the sample: 1) ordinary passenger cars 2) four-wheel-drive cars and all-terrain vehicles 3) vans (Table 3). In order to analyse a uniform class of vehicles, only drivers who drove ordinary passenger cars were kept in the sample; 4,387 drivers remained.
4. While examining the data a small group of drivers whose wintertime driving was in the hundreds of thousands of kilometres range stood out. They are certainly not typical drivers and, therefore, it was decided to remove them from the sample. The criterion for removal was wintertime driving of over 100,000 km per season.
5. The final sample size of the postal survey used in the analyses contained data on 4,352 drivers, 296 of whom had been involved in 327 wintertime accidents (Table 4) during 1991–1993 (1993 data included only January, February and March.)

4.2.3 Wintertime accident experience reported by drivers in the survey

The following tables show the makeup of the sample and the distributions of driving experience and accidents. The various totals may vary slightly because of missing observations on some variables.

Table 3 shows the type of cars and the type of tyres of vehicles involved in wintertime accidents during 1991–1993, as reported by drivers.

Table 3. Vehicle and tyre type of drivers involved in wintertime accidents during 1991–1993.
(There were altogether 21 missing observations.)

Tyre type		Ordinary cars		Four-wheel drives		Cars that resemble vans	
		in accidents	not in accidents	in accidents	not in accidents	in accidents	not in accidents
Studs	amount	312	3966	10	65	2	5
	(%)	94.6	97.5	90.9	73.0	100.0	62.5
Non-studded winter tyres	amount	11	98	1	24	-	3
	(%)	3.4	2.3	9.1	27.0	-	37.5
Altogether	amount	323	4064	11	89	2	8
	(%)	100.0	100.0	100.0	100.0	100.0	100.0

Table 4 shows reported wintertime accidents and injuries by three age groups. Most involved drivers were mature, 26–50 years old. A comparison of accident involved drivers with the non-involved drivers, shows a similar proportion of the middle age group (59% and 58%, respectively) but a higher involvement of the young group (19% vs. 8%) and, correspondingly, a lower involvement of the older age group 22% vs. 35%).

Table 4. Reported wintertime accidents by severity and driver age group.
(22 drivers had some missing data.)

Driver age	Number of drivers	Number of accidents	Fatalities	Injuries	Damaged cars
25 years at maximum	56	64	-	7	60
26–50 years	175	193	-	16	175
Over 50 years	66	70	-	8	62
Totals	296	327	-	31	297

Some drivers experienced more than one accident during the period, and 31 of them sustained injuries. (There was one fatality accident, to a passenger, which is not shown in the Table 4). Older drivers had a higher response rate in the survey and their calculated accident involvement was lower than that of other age groups. It would seem that compared to younger drivers, older drivers either under-reported their accidents or that responding older drivers tended to be those who did not experience accidents. Comparisons of reporting trends in the Finnish general travel survey (HLT) support the latter assumption.

In the postal survey, we have not received data from those cases with fatal outcome to the driver. This bias is small because of the small number of such cases in the population. However, all such cases are considered in detail when analysing the in-depth database of the present study.

Table 5 shows the kind of tyres drivers reported using in wintertime, in general, and on the vehicle involved in wintertime accident. It was ascertained already during the planning stages of the postal survey that only a very small number of drivers used the new non-studded winter tyres – friction tyres. Therefore, the survey did not ask for separate information on the use of friction tyres. However, five of the drivers mentioned that they had used friction tyres. In the analyses, these five drivers were added to the drivers who used other kinds of non-studded winter tyres.

Table 5. Reported tyre type used in winter in general, and in the accident car.

	Tyre type at the time of accident		
Winter tyre type in general	**Studs**	**Non-studded winter tyres**	**Summer tyres**
Studs	306	8	5
Non-studded winter tyres	3	3	-
Summer tyres	1	-	-
Total	310	11	5

Most of the vehicles (95.0%) were equipped with studded tyres, 3.4% were using non-studded winter tyres and 1.6% were using summer tyres. (The type of tyre one of the involved drivers could not be determined).

Altogether 30 drivers using studded tyres (out of 310) were injured in the accidents and a similar proportion (one of nine) were injured among users of non-studded tyres

Of the eleven accident involved drivers who had non-studded tyres at the time of accident, eight reported that they usually did use studded tyres during the winter. So had reported the five drivers who have had accidents with summer tyres on.

Table 6 shows the distribution of multiple accidents among drivers who reported one or more accidents for the 1991–1993 period. Most (92%) drivers were involved in only one accident, 7.4% in two accidents and the rest in three or four accident events. Age had no effect on the number of repeat accidents.

Table 6. Distribution of accident drivers by self reported number of accidents during the period and driver age group.

	Number of accidents during winter 1991–1993				
Driver age	**1**	**2**	**3**	**4**	**Total**
25 years at maximum	49	7			56
26–50 years	159	12	2	1	174
Over 50 years	63	3			66
Total	271	22	2	1	296

Table 7 shows the distribution of reported accidents by year and month. While data for 1991 and 1992 are similar, the figures for the first months of 1993 show a marked increase. It is likely that memory was better for the more recent season and perhaps the sampling procedures inadvertently favoured the last season; for example, if return rates were higher for more current cases.

Table 7. Reported wintertime accidents by month and year.

Year	Month					Total
	January	February	March	November	December	
1991	15	19	25	19	25	103
1992	17	27	26	29	23	123
1993	34	35	32	-	-	101
Total	66	81	84	48	48	327

The age distribution of drivers in the survey was compared with the age distribution of driver respondents in the general travel survey of 1992 (FinnRA 1993). The comparison sample consisted of drivers who had a car at their disposal and used it. Table 8 shows the age distributions for all drivers, male and women drivers, in the general travel survey (HLT) and the postal survey (PK).

Table 8. The age distributions of the car drivers in the 1992 passenger transport survey (HLT) and in the postal survey data (PK).

		All drivers		Male drivers		Female drivers	
Driver age		**HLT**	**PK**	**HLT**	**PK**	**HLT**	**PK**
≤ 25 v		9,1	8,5	9,0	7,9	9,4	10,8
26–35		21,3	20,7	19,8	19,4	24,4	25,0
36–45		26,4	24,9	24,8	24,4	30,1	26,6
46–55		21,8	20,7	21,8	20,6	21,6	21,0
56–65		15,7	15,2	17,7	16,1	11,5	12,2
> 65		5,7	10,0	6,9	11,7	3,0	4,5
Total	%	100,0	100,0	100,0	100,0	100,0	100,0
	number	(3971)	(4351)	(2710)	(3340)	(1261)	(1011)

It can be seen that the distributions are essentially similar except for the older males (> 65) age group, a point mentioned earlier. There, a higher proportion was present in the PK (wintertime) sample; 11.7% vs. 6.9%, significant by the Kolmogorov–Smirnov test (Siegel 1956). This supports the opinion that no critical non-sampling errors were included in the final data.

4.3 VALT data

4.3.1 Nature of the database and case selection

The Motor Insurers' Road Safety Committee (VALT) accident investigation teams have studied the most serious traffic accidents in Finland since the 1970's. VALT, in co-operation with the Ministry of Transport, road authorities, police and other organisations, manage the operation of the investigation teams The accident investigation teams gather extensive and detailed information from each accident they investigate. VALT compiles the data into the database and for further use in safety research and analysis. In recent years, the investigation teams have examined almost all of the fatal traffic accidents. The investigation teams have also worked on special projects, examining safety problems that are of current interest.

The whole VALT database available during the beginning of this study contained information on 2,700 fatal accidents in 1987–1991. The database includes information on accident time, location, conditions, consequences, traffic, and specific information about each "party" to the accident-driver personal background and actions, and vehicle characteristics.

Of interest to this study were fatal accidents that occurred during wintertime. VALT data included 386 fatal wintertime accidents (November–March) in which 658 car drivers were involved.

4.3.2 Driver and accident characteristics

Table 9 presents the age of the involved drivers and the consequences to drivers (every accident had at least one fatality who could be a driver or another occupant). Out of 658 drivers, 110 (17%) had survived without injury, 315 (48%) had died and 233 (35%) had been injured. Most (52%) drivers were 26–50-year-olds, 170 (26%) were younger, and 144 (22%) older. Only 150 (23%) were female drivers. The share of female drivers was at its lowest (8%) in the older age group.

Table 10 presents the yearly and monthly (in the winter months) distribution of the accidents. There are clear fluctuations in the number of cases over years and months which reflect both changes in number of accidents and shifting investigation case-loads and selection.

Table 11 shows tyre types on vehicles used by "Primary" and "Other party" involved drivers in the database. About 90% of the drivers drove cars with studded tyres at the time of the accident, 6% had non-studded winter tyres and 4% had summer tyres on.

Table 9. Distribution of drivers in VALT 1987–1991 winter accidents database by injury severity and age.
(Six drivers gave inadequate information.)

Driver age	Severity of injury Total	No injuries	Fatalities	Injured
25 years at maximum	170	28	84	58
26–50 years	344	70	143	131
Over 50 years	144	12	88	44
Total	658	110	315	233

Table 10. The number of drivers in 1987–1991 VALT database by year and month.

Year	Month January	February	March	November	December	Total
1987	29	19	12	32	34	126
1988	22	29	32	40	56	179
1989	27	29	25	19	1	141
1990	14	21	15	17	28	95
1991	29	13	32	12	31	117
Total (%)	121 (18.4%)	111 (16.9%)	116 (17.6%)	120 (18.2%)	190 (28.9%)	658 (100.0%)

Table 11. The type of tyre used by "primary" and "other" involved drivers in the VALT database.
(35 drivers had given inadequate information.)

Tyre type	Primary Number	Primary %	Other Number	Other %	Total Number	Total %
Studded tyres	321	87.0	237	93.3	558	89.6
Non-studded winter tyres	27	7.3	11	4.3	38	6.1
Summer tyres	21	5.7	6	2.4	27	4.3
Total	369	100.0	254	100.0	623	100.0

Somewhat higher proportion of primary involved drivers has used non-studded tyres, compared to "other involved parties" (13% vs. 7.7%).

A tyre with a tread depth of 4 mm or less is considered in bad condition for winter driving. Studded tyres were clearly in **better** condition than other types of tyres. When the worst tyre was considered, 20% of the vehicles with studded tyres, 41% of the vehicles with non-studded winter tyres, and 79% of the vehicles with summer tyres, had tyres with tread depth of 4 mm or less.

Table 12 compares the age distribution of drivers in the VALT database with that of the postal survey (PK) and the general travel survey (HLT). Clearly, the VALT cases are not a random sample of all winter accidents or of all car drivers. Nevertheless, the age distributions of responding self-reported-accident drivers in the postal survey and in the VALT accident database did not differ significantly from one another (Kolmogorov–Smirnov test).

Table 12. Driver age distributions in the 1992 general travel survey (HLT), the postal survey and in the VALT database.

Driver age		HLT	Postal survey		VALT data
			All drivers	Accident drivers	
≤25		9.1	8.5	19.6	25.8
26–35		21.3	20.7	23.9	21.3
36–45		26.4	24.9	23.9	21.3
46–55		21.8	20.7	18.0	15.0
56–65		15.7	15.2	8.9	8.1
>65		5.7	10.0	5.8	8.6
Total	%	100.0	100.0	100.0	100.0
	amount	(3971)	(4351)	(327)	(658)

5. MODELLING APPROACH

5.1 Overview of the approach

The basic approach was to derive hypotheses from earlier research, and test them with new data from complementary sources. Each data source by itself was not sufficiently reliable or robust but the combination provided better possibilities for meaningful conclusions.

The postal survey data contained basic biographical information on drivers, their mobility and trips, their vehicles and tyres, and a description of their accidents during a reference period 1991–1993. The survey thus provided estimate of exposure to wintertime driving risks. The unit of analysis could be a driver, a vehicle, or an accident.

The VALT data contained, by definition, only drivers involved in accidents. However, the “involved drivers” could be divided into those drivers whose vehicles or behaviour have been largely responsible for the crash (primary or main guilty party) and those drivers who were taken as “other involved parties.” This distinction enables a calculation of an exposure estimate based on the concept of “induced exposure”. This concept was published by J. D. Thorpe in 1964 and later on also by Carr and Haight (Cerrelli 1972). It assumes that the number of “other involved parties” is in a systematic relationship with exposure so that the ratio of “primary parties” to “all involved parties” is a measure of risk.

The main method in this study was survival modelling. The results were compared with the results from cohort and case-control analysis and logit modelling (Figure 1).

Methods based on survival analysis have been developed recently as a new tool for safety research (e.g. Hensher & Mannering 1993; Jovanis & Chang 1989). The method was tested at VTT in previous research when developing accident models (Roine 1993b; Kulmala 1995a). Survival models were used to explain the impact of different factors on driver’s conditional accident risk. The conditional accident risk refers to the accident probability at moment *t*, given that an accident has not occurred before that moment *t*. The accident time in the present study described the exposure to wintertime accidents in the models. Data on the kilometres travelled by drivers were used to make exposure estimates more precise.

Accident and non-accident drivers were compared by the cohort and case-control methods with respect to each of several potential risk factors. When number of cases permitted, intermediate factors were also considered (Collet 1991, McPherson 1990 and Schlesselman 1982). The application of the historical cohort method required a modification of the usual procedure in cohort studies. Here the test and control groups are based on sampling the phenomena of interest (accidents) **after** they have already occurred (Roine 1996).

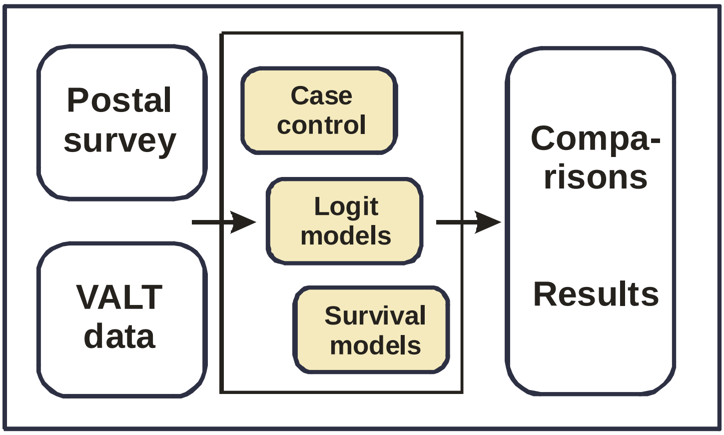

Figure 1. The data and the modelling of the study.

The probability of getting into an accident was modelled with logit models. The explained variable in the models was the share (proportion) of accident drivers, which was assumed to follow the binomial distribution. The use of this method was limited by the small number of observations, but modelling gave the opportunity to examine the effects and interactions of the various variables (e.g. Aitkin et al. 1989, Collet 1991). This method was especially suited for evaluating the relative risks of using studded or non-studded winter tyres (Roine 1996).

Data from the national accident statistics could only be used in the preliminary comparisons of the postal survey sample and the fatal accident data.

5.2 Case-control and cohort studies

Cohort and case-control study methods have been developed for research concerning epidemiological diseases. Case-control studies are used to identify the factors that influence the probability of a disease in a group of people. The group under examination is compared to appropriate control groups. The studies are usually carried out as longitudinal studies. The people under examination and the factors, which possibly affect the disease or expose people to it, are observed for several years. The significance of the factors is determined by comparing the sick and healthy groups after the examination period has ended (Collet 1991, Lee 1992). The relative risk is estimated based on the odds-ratio (Schlesselman 1982).

In a cohort study, a group of people is chosen to be followed and their health status is examined for a certain period. The people in the study are chosen based on interesting factors and are arranged into subgroups for comparisons. Cohort studies can also be carried out as historical studies, in which case the phenomenon under examination, such as being taken ill, has happened in the past.

Both methods of observation compare samples picked from two different base populations and evaluate whether being sick (or dying) is distributed the same way in the base populations. The comparison can be carried out in different ways and in the simplest cases by 2×2 cross tabulations and the χ^2-test.

In the case-control method, two comparable groups of people are chosen to be the objects of the research. The first group will include a sample of those individuals who have a characteristic that is being studied, and the second group will include a sample of individuals who do not have this characteristic. The research is often carried out in retrospect, when the phenomenon under examination, such as a disease, has already appeared in the people in the case group. The control group is chosen such that it will correspond with the case group as well as possible in every way, except for the factors under examination. Comparisons may be strengthened by matching pairs, one in each group (e.g. Elvik & Vaa 1990, Freedman et al. 1991).

If the impacts of the phenomenon under examination are evaluated with the odds-ratio method, the case-control groups can be small. In that case, the general occurrence of a disease in the base population can not be estimated. However, by comparing the groups one can estimate the probability of getting a disease from the relative risk (Schlesselman 1982).

In this study the impacts of different factors on drivers' accident probability were estimated with the case-control method. Accident drivers make up the "case" test group and accident-free drivers make up the "control" group. The distributions of the relevant factors in the two groups can then be compared. The impacts of the comparable factors were estimated with the odds-ratio (Ψ), which is approximately the same size as the relative risk (R), when rare phenomena are compared.

A further advantage of the odds-ratio is that it can be modelled with a linear logit model.

The odds-ratio and its standard error can be calculated from cell frequencies, as shown in Table 13 for a 2×2 table. The control group sets a starting level, in relation to which the relative risk of the test group is calculated.

Table 13. The data needed to calculate the odds-ratio.

Groups	Accident drivers	Non-accident drivers
Test group	a	b
Control group	c	d

On the basis of the cell frequencies presented in Table 13 one gets (Collet 1991, Schlesselman 1982):

$$\Psi = ad / bc \tag{4}$$

One can also calculate the standard errors of the odds-ratio (Ψ). The calculation is based on the value of log (Ψ) and the standard error (s.e.) of Ψ is obtained:

$$\text{s.e. } \Psi \approx \Psi \text{ s.e. } (\ln (\Psi)) \tag{5}$$

where

$$\text{s.e.}(\ln \Psi) = \sqrt{(\frac{1}{a} + \frac{1}{b} + \frac{1}{c} + \frac{1}{d})} \tag{6}$$

Because the odds-ratio approximates the expected value of the relative risk, it is important to calculate the confidence interval of the expected value with which one can estimate the reachable accuracy. Expected values that are, in practice, close to 1.0, indicate that the factor under examination does not significantly influence drivers' accident probability. When the value of the relative risk 1.0 is included in the calculated confidence interval, the effect of the factor studied is not statistically significant. The 95% confidence interval is usually used in estimation and its approximation is calculated as follows:

Lower limit

$$\ln(\Psi) = \ln(\Psi) - Z_a \times \text{s.e.}(\ln\Psi) \tag{7}$$

Upper limit

$$\ln(\Psi) = \ln(\Psi) + Z_a \times \text{s.e.}(\ln\Psi) \tag{8}$$

where Z_a is the chosen percentile point of the normal distribution, and for 95% confidence interval $Z_a = 1.96$.

A cohort study usually advances from cause to effect as a prospective, or longitudinal, study. The entities which are chosen to be examined, can often be chosen as a random sample, in which case the phenomenon's frequency can be estimated also in the base population. A cohort study can also be conducted on historical data (Schlesselman 1982).

In the cohort method the relative risk is derived from cell frequencies in the cross tabulation, as shown in Table 13 (Schlesselman 1982).

$$R = \frac{a/(a+b)}{c/(c+d)} = \frac{p_1}{p_2} \tag{9}$$

$$\Psi = ad / bc \tag{10}$$

The confidence interval of the relative risks are calculated as follows (Schlesselman 1982):

$$R \times \exp\left(-Z_a \times \sqrt{v}\right),\ R \times \exp\left(Z_a \times \sqrt{v}\right) \tag{11}$$

where

R is relative risk

Z_a is the percentile point of the normal distribution

$v = \operatorname{var}\ln(R) \approx \dfrac{b}{a(a+b)} + \dfrac{c}{c(c+d)}$

$p_1 = a / (a + b)$

$p_2 = c / (c + d)$.

Case-control and cohort methods were used in the preliminary study mainly when analysing the effect of tyre type on relative accident risk (chapter 7). The main problem with these methods seemed to be the restricted possibilities to take account various confounding factors. They had also restricted capabilities to reveal the relationships and dependencies between various variables.

Linear logit models

The impacts of different factors on drivers' accident probability can be examined with linear logit models. The probability of accidents is first calculated from the proportion of accident drivers in the sample data.

The linear logit model presumed that the share of "accident drivers" is composed of n_i binomially distributed observations y_i (drivers), where y_i has the values 0 = non-accident driver and 1 = accident driver. The parameters of the binomial distribution are n_i and p_i. In this case the expected number of accident drivers is, in accordance with the binomial distribution, $E(y_i) = n_i \times p_i$, where p_i is the probability of driver i getting into an accident. The linear logit model is defined as follows (Aitkin 1989, Collet 1991 and Chatfield 1994):

$$\text{Logit}\ (p_i) = \log\ [p_i\ /(1 - p_i)] = b_0 + b_1\ x_1 + b_2\ x_2 + \ldots + b_k\ x_k \tag{12}$$

where

p_i is the probability of getting into an accident

b_i is the coefficients of the model's variables

x_i is the model's variables

The probability of an accident can thus be calculated with the logit model as follows:

$$\text{Prob}(\mathit{accident}) = \frac{\exp(XB)}{1 + \exp(XB)} \tag{13}$$

where

$\exp(XB) = \exp(b_0 + b_1\ x_1 + b_2\ x_2 + \ldots + b_n\ x_n)$

and XB defines the matrice.

In logit models, the impacts of different variables are estimated by calculating the share of impacts for the class of variables chosen as the reference level. The reference level impact is set to 1,0. When examining variables with several classes, such as driver age, the lowest level of the youngest drivers is usually the reference level.

The logit models of the VALT data can be interpreted as the modelling of a case-control study. The test group (case) is composed of the primary involved parties in the accidents and the control group consists of those drivers involved in the accidents as other drivers. Following the method of induced exposure, the accident probability is derived by comparing the number of the primary involved parties with that of the other involved parties in the same accidents (Cerrelli 1972; Stamatiadis & Deacon 1995).

Logit models are generally a suitable tool when analysing this kind of driver related data (Roine 1993b). They lack time dependency as the survival models with their hazard functions. Here logit models are used in comparisons with survival models in theoretical analysis and also when comparing the estimated effects of studded tyres on accident risk. (chapter 6 and 7). The preliminary analysis (Roine 1993b) pointed out using VALT data and logit models that important explanatory variables for wintertime accident risks are e.g. use of alcohol, annual vehicle kilometreage, familiarity of route, driver age, type of tyre and use of safety belt.

6. SURVIVAL MODELLING

6.1 Defining survival

Survival research methods are typically used in the field of medicine to investigate the development of serious diseases and the effectiveness of medication. These same methods are also applied in the area of technology when examining the safe life-span and durability of technical systems and their parts (Crowder et al. 1991, Lee 1992).

An important issue in survival research is how to deal with cases whose survival can not be followed during the entire research period. Such individuals are called censored individuals (observations). Censoring can happen in the following three ways (Bunday 1991, Lee 1992):

- Type I: The duration of the study is fixed to a chosen period. The study includes cases that are monitored from a set starting point for as long as the phenomenon under examination occurs, or until the individuals are lost to the monitoring, or the entire monitoring period has passed. Individuals whose monitoring does not provide information about the occurrence of the phenomenon under examination until they drop out, or at the end of the study period, are censored observations.
- Type II: The length of the monitoring period depends on the desired number or proportion of uncensored observations. The length of the period is the same as the survival of the individual with the longest life span. Individuals, who are removed from the study for various reasons or survive less than the monitoring period, are censored observations.
- Type III: The duration of monitoring is fixed. However, individuals may enter the study at different starting points. Censored observations are the ones whose survival can not be defined, those who are lost from monitoring, or whose individual survival period continues after the overall monitoring period has ended.

Survival studies can be divided into the following two groups (Hensher & Mannering 1993):

- Monitoring censored to the right: The investigation is begun at a certain selected moment, when the individuals entering the examination are exposed to the phenomenon under investigation, e.g. a medicine or treatment. The investigation is continued from that moment on for a certain length of time.
- Monitoring censored to the left: The investigation is begun at a certain selected moment, but includes individuals whose exposure to the phenomenon under investigation has begun before the moment the examination period started, (as is in the present study).

Survival analyses can be based either on an assumption about survival following a certain distribution or on direct observation based on the actual data. Both procedures require dealing with censored and uncensored observations. The most commonly used survival distributions are the negative exponential distribution, the Weibull distribution, the Gumbel distribution, the logarithmic normal distribution and their combinations (Aitkin et al. 1989; Bunday 1991; Crowder et al. 1991; Hensher & Mannering 1993). Which type of function is best at describing the survival distribution is mainly dependent on the characteristics of the phenomenon under examination. Direct observation based on the data can be carried out with the Kaplan–Meier method (Bunday 1991).

6.1.1 Principles of survival modelling

In modelling survival is represented by a random variable (T), the probability density function of which is $f(t)$. The important functions in survival analysis, besides the density function, are the survival and hazard functions. The survival function $S(t)$ is a cumulative function, which depicts the share of "the living" as a function of time. The hazard function $h(t)$, on the other hand, represents the probability of an occurrence to end survival at point t on the condition that the individual or other object of examination has been "alive" until point t. The following relationship exist between these functions:

$$h(t) = f(t) \,/\, S(t) \tag{14}$$

$$H(t) = \quad h(t)dt \tag{15}$$

$$S(t) = \exp\,[-H(t)] \tag{16}$$

where

$h(t)$ is hazard function
$H(t)$ is cumulative hazard function
$S(t)$ is survival function.

This study applied distributionfree Cox's proportional survival models, which involve direct estimation based on data. Proportional survival models assume that survival t has its density, hazard and survival functions. There are no special starting assumptions made on the form of the density function of survival t.

The form of the model is:

$$h(t,X) = h_0(t) \times \exp(XB) \tag{17}$$

where

$h(t,X)$ is hazard function
$h_0(t)$ is the base level of the hazard function
$\exp(XB)$ is linear function formed by the variables and their parameters.

The model can be further written for calculations into the form:

$$S(t, X) = S_0(t)^{\exp(XB)} \tag{18}$$

where in this study

$S(t,X)$ is estimate of the share of drivers not involved in accidents until time t

$S_0(t)$ is base level of the survival function formed on the basis of data (19)

$$\exp(XB) = \exp\,(b_1\, x_1 + b_2\, x_2 + \ldots + b_n\, x_n) \tag{20}$$

b_i is the coefficients of the variables x_i (parameters).

When drawing up proportional survival or hazards models, stratification can also be used as an examination method. The different basic hazards ($h_0(t)$) are estimated with the model for the different values of the stratifying variable. The survival model explains the effect of the variables on these different base levels of the hazard function.

When using the negative exponential distribution or the Weinbull distribution as the survival distribution, the hazard and survival function are written:

$$h(t) = \lambda\beta(\lambda t)^{\beta-1} \tag{21}$$

$$\mathrm{S}(t) = \exp\,[-\,(\lambda t)^{\beta}] \tag{22}$$

In the model, λ is the so called rate parameter and β is the shape parameter. When the value of β is 1, it is the negative exponential distribution ($h(t) = \lambda$). When $\beta > 1$ or $\beta < 1$, the distribution is the Weibull-distribution. It is characteristic for the Weibull distribution that if $\beta > 1$, the hazard function is increasing and if $\beta < 1$, the hazard function is decreasing in relation to time (e.g. Bunday 1991).

6.1.2 Estimating models and choosing variables for the accident data

In this study, the survival, or accident time, has been defined as the number of winter days counted from the start of the analysis period. Accident time (time from the starting time until the moment an accident occurs) in the postal survey was calculated from a starting point of 1.1.1991. The starting point for the VALT data was 1.1.1987.

Only wintertime period was considered. Wintertime data include days between November to March for the period 1.1.1991–3.3.1993 (in the survey data,) and 1.1.1987–31.12.1991 (for the VALT data). The “survival” of drivers who got into accidents was the number of winter days from the beginning of the examination period until the day a driver was involved in an accident. If the driver had been in more than one accident, the winter days between the two accidents were considered survival days. Some of the postal survey drivers had only given the month the accident occurred. For these drivers, the middle of the month, the 15th, was defined as the exact date of the accident. In the case of drivers who obtained their driving license during the analysis period, starting time was defined as the date of obtaining the license. This procedure of calculations was successfully tested before (Roine 1993b).

As Figure 2 illustrates, there are two distributions. The first describes driver's exposure time until the beginning of the examination period (t) and the other is a distribution of exposures within the examination period ($T–t$). This is Type I according to the classification above.

It can theoretically be shown that if the survival distribution (*time*) obeys the negative exponential distribution, the examination period's ($T–t$) survival distribution is also a negative exponential distribution. If the survival distribution is of some other form, e.g. the Weibull distribution, the examination period's ($T–t$) survival distribution has to be separately derived based on the Weibull distribution (Kalbfleish & Prentice 1980). However, in the case of the proportional survival model, distribution assumptions need not be made and the estimation can be carried out without such assumptions.

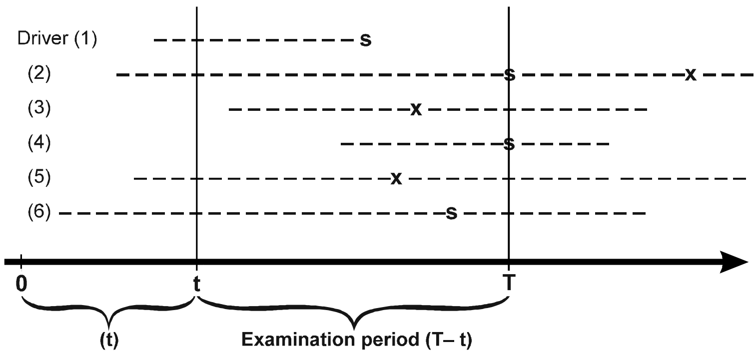

Figure 2. Survival time in the accident data (x = accident, s = censored observations).

While estimating the models, simplifications had to be made because no detailed information about exposure to accident risks was available. Differences between drivers in exposure to accident risks were assumed to be in proportion to amount of driving in wintertime. It was further assumed that drivers' wintertime kilometreage accumulated evenly during the entire winter period.

When estimating the proportional hazard model, the hazard's base level h_0 is first derived, followed by estimate of the final model with its variables and parameters. The basic assumption of the proportional survival model is that driver's accident risk can vary in its entirety as a function of time. The other variables are usually assumed to affect the accident risk in the same way no matter what the time was (e.g. Hensher & Mannering 1993). The model is multiplicative, meaning that the mean accident risk (hazard) is calculated as the product of the influences of the different factors.

The models were estimated with the help of the SPSS software package (Norusis 1993). The GLIM software package (Aitkin et al. 1989), which uses iterations when estimating a model, was used in estimating the Weibull models:

$$\log h(t) = \log \beta + (–1)\log t + \log\lambda \qquad (23)$$

where

β is unknown shape parameter

$\log \lambda = XB = b_1 x_1 + b_2 x_2 + ... + b_n x_n$.

The general principles of normal selective regression analysis can be applied, largely, when building survival models. Variables are accepted into the model if their addition to the model changes the likelihood-ratio significantly. Statistical significance is tested by Wald test. The test score used is the ratio between the square of the added variable's estimated likelihood-ratio coefficient and the coefficient's standard error. The test score is χ^2-distributed (Greene 1990).

When composing the models here, both forward selection and backwards variable elimination procedures were used. In forward selection, a model can have a fixed group of variables and other variables are selected into the model by adding variables and choosing those with the best explanatory power. In backward elimination , the initial model has the maximum number of variables, out of which the variables that have the lowest explanatory power are removed. Part of the model can be fixed also in this method.

Errors in proportional survival models are due to the models' fundamental assumption that all systematic variation is described with the use of the selected variables. However, other influential factors can be working in the background of time-dependent processes. The so-called heterogeneous problem (Greene 1990) implies the possibility of different hazard functions underlying the data. Another problem is connected with the state-dependence, which here can be understood as the influence of driving experience earlier in the season on later driving. These problems do not have acceptable standard solutions. The heterogeneous problem might be removed from the models by a heterogeneous factor, which takes the impact of variables not included in the model into consideration (Hensher & Mannering 1993). Because these correction methods are still being developed, the heterogeneous and state-dependence problems were not handled in the present modelling but their impact was considered when examining the results.

Individual potential variables were selected with the help of Kaplan–Meier estimates, directly from data. The survival functions' Kaplan–Meier estimates are calculated as follows (Bunday 1991; Crowder et al. 1991; Lee 1992):

$$S(t_j) = S(t_j - 1)\,[(n_j - f_j) / n_j] \tag{24}$$

$$\mathrm{var}\,[S(t_j)] = [S(t_j)]^2\,[f_j / n_j(n_j - f_j)] \tag{25}$$

where:

$S(t_j)$ is the value of the survival function's estimate at the moment t

$\mathrm{var}\,[S(t_j)]$ is the variance of the survival estimate at the moment t

n_j is drivers exposed to risk until the moment t

f_j is accident drivers until the moment t.

Next, the similarity of the survival functions obtained in this way was tested with the Log Rank test (Kalbfleish & Prentice 1980; Lee 1992).

The Log Rank test score L used in the comparisons of the survival functions is asymptotically normally distributed and is calculated as follows (Lee 1992):

$$LR = L / \sqrt{\mathrm{var}(L)} \tag{26}$$

$$\mathrm{var}(L) = \left[(n_1 \times n_2) \quad w_i^2\right] \Big/ \left[(n_1 + n_2)(n_1 + n_2 - 1)\right]$$

where

LR is Log Rank test score
w_i is the value of the combined points of the observation
L is the sum of the points given, or Σw_i
n_1 is the number of accident drivers at the moment $t = 0$,
n_2 is the number of non-accident drivers at the moment $t = 0$.

When there were distinct differences between the survival functions, the variable's significance in the model was tested against the rest of the variables.

6.1.3 Evaluation of survival models

The quality of survival models is determined by

- the scores of the models' quality (log-likelihood, the Pseudo R-squared index),
- the residuals (the Martingale and Cox–Snell residuals),
- the partial residuals,
- the impacts of removing observations that are deviant or have leverage (DfBeta) and
- the removal of variables.

The evaluation of the proportional models' explanatory power and statistical significance was based on the maximum likelihood method. The impact of the addition of each variable on the model's log-likelihood value (here –2×log-likelihood) was first calculated. Then the statistical significance of the model and the influence of adding variables were estimated, with the χ^2-distribution (e.g. Kalbfleish & Prentice 1980; Lee 1992).

An index value, similar to R-squared in normal regression, can be calculated when assessing the goodness of fit of the models. This index value is called Pseudo R-Squared and it is calculated as follows (Kalbfleish & Prentice 1980):

$$D = C / (n - p + C)$$

where

D is the Pseudo R-squared
n is the number of uncensored observations
p is the model's degrees of freedom
C is the χ^2 test score.

According to the survival model theory, the Cox–Snell residuals correspond with the values of the cumulative hazard function ($H(t)$). The Martingale residuals are calculated on the basis of the Cox–Snell residuals, and correspond with the estimated values of the cumulative hazard function. The Martingale residual was calculated separately for the uncensored observations (accident drivers) as the negative value of the Cox–Snell residual and for the censored (non-accident drivers) observations as (1 – Cox–Snell residual).

The general fit of the proportional survival and hazards model to the data is estimated by the partial residuals that are calculated for each variable in the model. They represent the difference between the estimated values and the values that are consistent with the variable's observations arranged in time order. When testing the suitability of the proportional hazard model, the partial residuals are examined as a function of time. If the pattern formed by the partial residuals is balanced and random, without having distinctive trends, the assumption concerning proportionality can be accepted.

DfBeta-coefficients are used while solving the impacts of observations with leverage. They help estimating how the coefficients of the model's variables change when a given observation is added or omitted from model estimation. Observations that receive deviant DfBeta-values are considered to have leverage and are further scrutinised.

Because the variables in survival models can strongly correlate with each other, their impact on the characteristics of the most important models composed were examined by alternately removing variables from the models.

6.2 Kaplan–Meier estimates of the variables in the survey

The candidate variables for the survival models were first evaluated by calculating survival or accident time functions for each variable with the Kaplan–Meier method (see section 6.1.2). The statistical significance of each variable and its effect (share of accident drivers, keeping the values of the other variables constant) were calculated. Table 14 presents for each variable the number of drivers, the share of accident drivers and the statistical significance of the effect on survival function based on the Kaplan–Meier estimation.

Table 14. Number of all drivers, accident drivers, share of accident drivers and the statistical significance (p-values) of the effect based on Kaplan–Meier estimates in the survey data.

Variable	All drivers	Accident drivers	Share of accident drivers	p-value
Driver age	**4352**	**327**	**7.5**	**0.000**
25 years at maximum	373	64	17.2	
26–50 years	2521	193	7.7	
over 50 years	1458	70	4.8	
Sex	**4374**	**328**	**7.5**	**0.106**
female	1018	88	8.6	
male	3356	240	7.2	
Employment situation	**4350**	**327**	**7.5**	**0.000**
working	2762	214	7.8	
not working	249	28	11.2	
unemployed	466	49	10.5	
retired	864	36	4.2	
Type of driving licence	**4366**	**328**	**7.5**	**0.147**
no	10	2	20.0	
motorcycle, tractor	4	0	0.0	
car	2733	218	8.0	
truck or bus	1619	108	6.7	
Wintertime kilometreage	**4375**	**328**	**7.5**	**0.000**
5,000 km/a at maximum	2011	122	6.1	
5,001–10,000 km/a	1481	112	7.6	
over 10,000 km/a	883	94	10.7	
Annual vehicle kilometreage	**4375**	**328**	**7.5**	**0.000**
less than 5,000 km/a	267	13	4.9	
5,001–15,000 km/a	1662	106	6.4	
15,001–25,000 km/a	1395	98	7.0	
25,001–40,000 km/a	776	79	10.2	
over 40,000 km/a	275	32	11.6	
Share of driving in built-up areas	**4304**	**327**	**7.6**	**0.266**
30% at maximum	2025	141	7.0	
31–60%	1375	108	7.9	
over 60%	906	78	8.6	
Vehicle total kilometreage	**4277**	**319**	**7.6**	**0.323**
50,000 km at maximum	831	53	6.4	
50,001–150,000 km	2131	159	7.5	
over 150,000 km	1315	107	8.1	
Car engine volume	**4138**	**305**	**7.4**	**0.736**
1,100 cm3 at maximum	249	22	8.8	
1,101–1,300 cm3	3679	276	7.5	
over 1,300 cm3	142	11	7.8	
Vehicle age	**4138**	**305**	**7.4**	**0.481**
5 years at maximum	1835	129	7.0	
6–10 years	1486	108	7.5	
over 10 years	817	68	7.8	
Company car	**4375**	**328**	**7.5**	**0.677**
yes	4106	306	7.5	
no	269	22	8.2	
Tyre type	**4356**	**321**	**7.5**	**0.260**
studded tyre	4247	310	7.3	
non-studded winter tyre	109	11	10.1	
Tyre type	**4356**	**321**	**10.1**	**0.272**
studded tyre ≤10,000 km/winter period	2824	178	6.30	
>10,000 km/winter period	1423	132	9.28	
non-studded tyre ≤10,000 km/winter period	70	7	10.00	
>10,000 km/winter period	39	4	10.26	

Driver age

Driver age was a significant survival variable (Log Rank = 69.10, df = 2, p = 0.000). With wintertime kilometreage kept constant, driver age still had a significant effect on the cumulative survival distribution (Log Rank = 62.66 df = 2, p = 0.000). The young drivers category accounted for the largest, 17.2%, share of accident-involved drivers. The oldest age group (over 50 years) had the smallest, 4.8%, share. Unsurprisingly, the share went up with exposure (both total and wintertime kilometreage).

Young drivers drove older and more used cars than other drivers ($\chi^2 = 129.6$, df = 10, p = 0.000). Keeping vehicle age constant, there was still a significant effect of age (Log Rank = 58.04, df = 2, p = 0.000). Young drivers had a higher share of accidents in all vehicle-age categories. Young drivers who used the oldest vehicles had the highest share (23.1%). Vehicle age did not have a major effect on the share of accident drivers in the older driver groups.

Driver sex

Table 14 shows a small, statistically non-significant difference (8.6% vs. 7.2%) in the accident share values of female and male drivers, respectively. It remained insignificant also when controlling for age. Controlling for wintertime kilometreage, however, showed that the effect was then statistically significant. Female drivers in the high (over 8,000 km) exposure category had a higher share of accident drivers than male in the same category, 14.1% vs. 8.7% (Log Rank = 22.50, df = 2, p = 0.000).

Female and male accident shares were compared with the χ^2-test in order to support the Kaplan–Meier estimates. Both the driver age and the wintertime kilometreage were kept constant. The comparison indicated that 26–50-year-old and over 50-year-old female drivers, who had a high wintertime kilometreage, accounted for a significantly higher share of accident drivers than male drivers of those age classes did.

Driver employment situation

Driver employment situation had a statistically significant effect on the cumulative accident time distribution. Drivers who reported that they were not working or that they were unemployed were involved in accidents more often than the others were (Log Rank = 25.88, df = 4, p = 0.000).

However, employment situation had strong correlation with other variables and especially with driver age. When driver age was controlled, employment status no longer had significant effect on the share of accident drivers. The young and the old age groups had a higher proportion of unemployed and not working (respectively) compared to the middle age group.

Wintertime kilometreage

The number of kilometres driven by a driver during wintertime had a significant effect on survival time (Log Rank = 18.69, df = 2, p = 0.000). The share of accident drivers increased with wintertime vehicle kilometreage.

Share of driving in built-up areas

The share of the kilometres driven in built-up areas did not have statistically significant effect on the cumulative accident time distribution. However, when the effect of wintertime kilometreage was kept constant the share of the kilometres driven in built-up areas was significant (Log Rank = 6.75, df = 2, p = 0.032); drivers who drove a larger share of their wintertime kilometreage in built-up areas had a higher share of accidents.

Annual vehicle kilometreage

The annual vehicle kilometreage had a statistically significant effect on survival time (Log Rank = 21.1, df = 4, p = 0.000).

Vehicle total kilometreage, car engine volume and company car

These variables proved to be statistically insignificant.

Vehicle age

According to the Kaplan–Meier estimates, vehicle age, in itself, did not have a clear effect on accident time. It was related to driver age and the wintertime kilometreage. Young drivers usually drove older vehicles. On the other hand, newer vehicles did the most kilometres during the winter months.

Type of tyre on the vehicle

Only 11 accident drivers reported not using studded tyres. The share of accident drivers was 10% in this group. The majority of drivers used studded tyres and their share of accident drivers was 7,3%. The effect of type of tyre was not significant (Log Rank = 1.27, df = 1, p = 0.260).

The users of non-studded tyres reported higher wintertime kilometreage than the users of studded tyres; 9,055 km vs. 7,832 km, respectively (F-value = 3.06 and p = 0.08).

Type of driving license

Ten (0.2 %) drivers reported not having a driving license. Two of them were involved in wintertime accidents. There were no significant differences in the accident share of car license holders compared to truck or bus licenses, especially after age and exposure were controlled.

Summary of the effects of the variables

The Kaplan–Meier estimates indicate that the main variables having an effect on survival time of drivers were driver age and the amount of exposure (wintertime kilometreage.) Young drivers with old cars driving a lot in wintertime have the highest relative share of accidents. Driving experience and driver's sex had an effect as well, interacting with age. For example, middle aged and old female drivers with high wintertime kilometreage had a higher share of accident involved drivers than the corresponding male groups. The practical significance of such an effect is small due to the small number of such female drivers in the population.

The nature of exposure (e.g. road surface conditions) could not be examined in the postal survey, but according to earlier studies, it can be assumed to affect the accident probability.

The type of tyre had a small, statistically non-significant effect. Only 109 drivers reported using non-studded tyres and only 10% of them reported having accidents.

6.3 Kaplan–Meier estimates of the VALT data variables

The VALT based accident data consistent of the fatal accidents studied in-depth by the investigation teams. Accident drivers were defined as those who were the primary involved (or so called guilty) parties. The other involved parties were a control group, other parties. Single accidents were not included in this database.

Table 15 shows the analysed variables and their effect on survival based on Kaplan–Meier estimates.

Driver age

Young and older drivers had a higher share of accidents than the middle age group (Log Rank = 28.3, df = 2, p = 0.000). When the impact of the annual kilometreage was kept constant, the impact of driver age remained statistically significant (Log Rank = 15.28, df = 2, p = 0.000). The effect of exposure here was in opposite direction to what was found with reported accidents in the postal survey; drivers with small annual exposure had a higher share of the accidents.

Driver sex

Female drivers had a higher share of accident drivers, primary parties, than male drivers. When the effect of age was kept constant, the difference in the survival distributions between female and male drivers remained (Log Rank = 15.1, df = 2, p = 0.000). However, controlling for annual vehicle kilometreage removed most of the difference (Log Rank = 2.61, df = 2, p = 0.106).

Driving experience (years since licensing)

Drivers with less than 5 years experience (license duration) had a share of 69.3% of accident drivers compared to 54.4% for more experienced drivers. The effect stayed even after controlling for driver age, (Log Rank = 13.3, degrees of freedom = 2, $p = 0.001$).

Use of alcohol

Driving under the influence of alcohol had a clear effect on the accident time distributions. Fully 96% of drivers who were assessed to have had more than 0.5 g/l of alcohol in their blood were the primary involved parties in accidents, compared to 56% of those whose alcohol blood level was below this limit.

About 8.5% of male drivers and 2.8% of female drivers drove under the influence of alcohol. Drivers who drove under the influence used older vehicles compared to the no alcohol group ($\chi^2 = 26.12$, df = 2, $p = 0.000$). The share of vehicles older than 10 years was 52% for those drivers who drove under the influence of alcohol and 20% for alcohol free drivers.

Drivers who drove under the influence of alcohol were less likely to use safety belts compared to alcohol free drivers – 25%/15%. This effect was statistically significant ($\chi^2 = 26.12$, df = 2, $p = 0.000$).

Use of safety belt

Drivers who had used safety belts had a lower share of primary parties than those who did not (Log Rank = 48.11, df = 1, $p = 0.000$). The effect remained when controlling for the effect of either driver age or driver annual vehicle kilometreage.

Annual vehicle kilometreage

As there was no information on driver wintertime kilometreage the reported annual vehicle kilometreage was used as the only exposure measure.

Drivers' annual vehicle kilometreage was negatively related to the share of accidents (Log Rank = 29.94, df = 2, $p = 0.000$), suggesting that drivers who had a high annual kilometreage had also more driving experience and lower involvement in wintertime accidents. However, lack of wintertime exposure data prevents determination whether winter driving experience or general exposure was the relevant factor.

Road section

Accident locations were separated into links and junctions. Junctions were further classified into the major or minor legs. There were no differences in drivers' survival distributions between the links and junctions (Log Rank = 0.28, degrees of freedom = 1, $p = 0.60$).

Table 15. Number of all drivers, accident drivers, share of accident drivers and the statistical significance (p-values) of the effect based on Kaplan–Meier estimates in the VALT data.

Variable	All drivers	Accident drivers	Share of accident drivers	p-value
Driver age	**658**	**386**	**58.7**	**0.000**
25 years at maximum	170	114	67.1	
25–50 years	344	174	50.6	
over 50 years	144	98	68.1	
Sex	**658**	**386**	**58.7**	**0.001**
female	150	104	69.3	
male	508	282	55.5	
Accidents	**494**	**278**	**56.3**	**0.087**
no	331	194	58.6	
1	125	69	55.2	
more than 1	38	15	39.5	
Offences	**565**	**334**	**59.1**	**0.480**
1 at maximum	421	248	58.9	
1–4	104	60	57.7	
more than 4	40	26	65.0	
Amount of alcohol	**626**	**369**	**58.9**	**0.000**
0,05% at maximum	581	326	56.1	
over 0,05%	45	43	95.6	
Use of safety belt	**634**	**369**	**59.2**	**0.000**
no	132	108	81.8	
yes	502	261	52.0	
Age of driving licence	**609**	**352**	**57.8**	**0.000**
under 5 years	140	97	69.3	
at least 5 years	469	255	54.4	
Familiarity of route	**588**	**345**	**58.7**	**0.012**
at least 3/month	423	235	55.6	
more seldom	165	110	66.7	
Annual vehicle kilometreage	**527**	**301**	**57.1**	**0.000**
under 5 000 km/a	80	63	78.8	
5 000–14 999 km/a	272	149	54.8	
at least 15 000 km/a	175	89	50.9	
Vehicle total kilometreage	**560**	**321**	**57.3**	**0.673**
50 000 km at maximum	172	98	57.0	
50 000–99 999 km	154	95	61.7	
100 000–199 999 km	171	95	55.6	
at least 200 000 km	63	33	52.4	
Age of vehicle	**659**	**383**	**58.7**	**0.053**
5 years at maximum	310	170	54.8	
6–10 years	198	116	58.6	
over 10 years	145	97	66.9	
Weight of vehicle	**658**	**386**	**58.7**	**0.005**
900 kg at maximum	99	71	71.7	
901–1 200 kg	350	206	58.9	
over 1 200 kg	209	109	52.2	

continued

Table 15. Continued.

Variable	All drivers	Accident drivers	Share of accident drivers	p-value
Own vehicle	**344**	**202**	**58.7**	**0.146**
no	56	29	51.8	
yes	288	173	60.1	
Tread depth of tyre	**617**	**365**	**59.2**	**0.002**
4 mm at maximum	147	102	69.4	
over 4 mm	470	263	56.0	
Tyre type	**623**	**369**	**59.2**	**0.012**
studded tyre	558	321	57.5	
non-studded winter tyre	38	27	71.1	
summer tyre	27	21	77.8	
Tread depth of tyre	**593**	**352**	**59.4**	**0.002**
Not winter road surface conditions				
4 mm at maximum	41	25	61.0	
over 4 mm	137	79	57.8	
Winter road surface conditions				
4 mm at maximum	103	75	72.8	
over 4 mm	312	173	55.4	
Tread depth of tyre	**579**	**337**	**582**	**0.010**
Non-studded winter tyre				
4 mm at maximum	22	15	68.2	
over 4 mm	15	12	80.0	
Studded tyre				
4 mm at maximum	432	236	54.6	
over 4 mm	110	74	67.3	
Speed limit	**655**	**383**	**58,4**	**0.001**
60 km/h at maximum (built-up area)	105	77	71,9	
70 and 80 km/h	322	187	58.1	
over 80 km/h	228	119	52.2	
Road section	**657**	**385**	**59.1**	**0.280**
junction	121	68	56.2	
link	536	317	59.1	

Speed limit at the scene of the accident

Speed limit was categorised into three classes: up to 60 km/h, 70 and 80 km/h, and over 80 km/h. The speed limit value was the lowest limit at the scene of the accident; at junctions it usually was the minor leg's speed limit. The share of primary parties, accident drivers, decreased with the speed limit – it was lowest in accidents that happened at high (>80 km/h) speed limits – (Log Rank = 14.35, degrees of freedom = 2, p = 0.001). A possible interpretation for this effect could be that in junction accidents, an involved party who comes from a minor road (with a lower speed limit) is likely to be the primary party. Therefore, the impact of the speed limit was examined separately for links and junctions. Speed limit remained a statistically significant variable, (Log Rank = 17.8, df = 2, p = 0.001). Another possible reason for the unexpected result could be the fact that the number of accident involved drivers tended to increase at high speed limits, resulting in smaller primary / all involved parties ratio. However, controlling for number of involved parties did not remove the effect.

Vehicle total kilometreage and vehicle age

Vehicle total kilometreage did not have an impact on the accident time distribution (Log Rank = 1.54, df = 3, p = 0.673), but vehicle's age did. This effect was co-dependent on driver age and on drivers' annual vehicle kilometreage, however. Young drivers tended to use older cars.

When the effect of driver age was kept constant, vehicle age no longer had a statistically significant effect on the survival distributions (Log Rank = 13.2, df = 2, p = 0.200). This was also the result when driver annual vehicle kilometreage was controlled (Log Rank = 0.47, df = 2, p = 0.789).

Weight of the vehicle

As vehicle weight increased the share of accident drivers, primary parties, became smaller. The effect remained when driver age was kept constant (Log Rank = 9.3, df = 2, p = 0.010), but became marginal when driver annual vehicle kilometreage was kept constant (Log Rank = 4.6, df = 2, p = 0.099).

The effect of weight also remained when the effect of vehicle's traction (front- or rear-drive) was kept constant (Log Rank = 10.81, df = 2, p = 0.005). There was a minor interaction between vehicle weight and traction. For the light vehicles class, the share of primary parties was the highest with front-drive vehicles, while in the heavy vehicles class, the front-drive vehicles had a lower share of primary parties.

Type of tyre and tread depth of tyre

Tyre type had a clear effect on accident share. About 78% of the drivers who used summer tyres, 71% of the drivers who used non-studded winter tyres and 58% of the drivers who used studded tyres, were primary involved parties. However, a separate test for the drivers using studded and non-studded winter tyres showed no difference in the accident time distributions and the shares of primary parties (Log Rank = 0.92, df = 1, p = 0.338).

Worn out tyres have a negative effect on safety. The vehicles with tyre tread depth below 4 mm had a 69% share of primary parties compared with 56% when tread depth was at or over 4 mm (Log Rank = 9.60, df = 1, p = 0.002).

The effect of tread depth was similar for both studded and non-studded winter tyres, and remained statistically significant also after controlling for tyre type (Log Rank = 6.57, df = 1, p = 0.010).

Route familiarity

Drivers, who had driven past the scene of the accident more often than three times a month, were less likely to be the accidents' primary involved party than drivers who had driven past the place less often.

Previous accidents and traffic violations

Drivers who had been in an accident more than once during the last five years were less likely to have been classified as main guilty parties. However, this was mainly correlated with exposure. When drivers' annual vehicle kilometreage was kept constant the effect of accident history was insignificant (Log Rank = 2.68, df = 2, p = 0.262).

Recorded traffic offences during the past five years did not have a significant effect on drivers' accident time distributions (Log Rank = 1.47, degrees of freedom = 2, p = 0.48).

Own vehicle

The ownership of the vehicle (own vehicle/not own vehicle) did not have a significant effect on drivers' accident time distributions. Own vehicles tended to be older, with tyres in poorer condition, their drivers either in the young or in the old age groups, and accumulated fewer kilometres in a year. It is likely that many "non own" cars were newer company cars or rented cars with high annual kilometreage and more accessible to middle aged drivers.

Summary of the effects of the variables in VALT data

Kaplan–Meier estimates identified variables that had major effects on accident time distribution. Driver age and sex, use of alcohol, annual vehicle kilometreage, speed limit at the scene of the accident, use of safety belt and tyre's tread depth. Type of tyre did not have a statistically significant effect on survival.

There were several inter-correlations between the variables, which reduced the size and statistical significance of the effect of individual variables. It should be pointed out that the process of defining who is a primary party involves the human judgement of an accident's investigating team. Therefore, it is possible that some of the main effects as well as interactions reflect the "accident models" of the investigators and not only the unbiased impact of an variable. For example, it is possible that the presence of alcohol in a driver's blood predisposes him or her as the guilty/ primary party.

6.4 Models for the postal survey data

6.4.1 Models and their compilation principles

The estimation method produces both proportional hazards and survival models because hazard function can be derived from the survival function (chapter 6.1.1, p. 53). We express the models here mainly as Cox proportional survival models.

Uniform principles were followed during the compilation of the hazards and survival models:

- First, the best individual variables were found by the method of elimination, so that the first model included all potential variables.
- The composition of the models continued from a maximal model, which included all of the best variables and all of their second-degree interactions.
- The basic models were composed by eliminating all the statistically insignificant variables and interactions from the maximal models.
- The basic models were compared to alternative models, which were composed by adding variables. The models were built stepwise, as potential variables were added one by one and tested if they improved the model.
- The statistically significant variables and interactions that could possibly be added to the basic models were then solved by adding and subtracting potential variables and interactions to and from the basic model one by one.
- Tyre type was added into the model as a variable only after the final basic model had been compiled, because the variable did not have a statistically significant independent explanatory power in the basic models.
- The significant correlations between the variables in the models were identified and analysed later when evaluating the models.
- The models were usually reported in a structure in which the first category of the classified variables was a reference category and the impact of various categories were relative to this category (simple contrast). Due to interactions, some variables were coded so that the reference category was the overall effect (deviation contrast).

The best survival models compiled from the survey data included the following variables:

- driver age,
- wintertime kilometreage,
- annual vehicle kilometreage,
- driver sex,
- share of kilometreage in built-up areas,
- vehicle age,
- type of tyre,
- interactions of the different variables.

Table 16 lists variable codes and their values.

Table 16. The most important variables used in the proportional survival models compiled from the postal survey data.

Variable code	Variable	Categories and variable values
CKIKA2	Driver age	1 ≤ 25 2 26–50 3 > 50
CKIKA	Driver age	1 ≤ 25 2 26–35 3 36–45 4 46–55 5 56–65 6 > 65
CTAKM	Wintertime kilometreage	1 ≤ 5 000 km 2 5 001–10 000 km 3 > 10 000 km
LNAKM	Ln(Wintertime kilometreage / 1000)	Continuous variable
CAJOKM	Annual vehicle kilometreage	1 15 000 km/a at maximum 2 15 000–25 000 km/a 3 Over 25 000 km/a
AUTONI	Vehicle age	1 10 years at maximum 2 Over 10 years
CTAJA	Share of driving in built-up areas	1 ≤ 50% 2 > 50%
RENGAS	Tyre type	1 Studded tyres 2 Non-studded winter tyres
SPUOLI	Driver sex	1 Male 2 Female

6.4.2 Characteristics of the models

The different models did not vary much in their Log-likelihood values and their explanatory power was approximately the same. A test of the models' Log-likelihood values against the χ^2-distribution showed them to be significant, meaning that any of the models can be used when explaining drivers' wintertime accident risks. A summary of the models is shown in Appendix B. Variables in the models were intercorrelated. When the correlations were examined in a model (model 5), the largest correlations between variables were found to be:

- driver age and wintertime kilometreage (0.22),
- driver age and vehicle age (0.19),
- driver age and sex (0.15),
- wintertime kilometreage and drivers sex (0.25),
- wintertime kilometreage and driving in built-up areas (0.19) and
- wintertime kilometreage and vehicle age (0.18).

According to the basic models, driver age had the greatest explanatory power in the models ($p < 0.000$), then wintertime kilometreage ($p < 0.001$) and finally the other variables and their interactions. According to basic models 1 and 2, the effect of driver sex and vehicle age depended strongly on driver age. Including the share of driving in built-up areas into the model also brought with it the interactions with wintertime kilometreage and driver age ($p > 0.03$).

The structure of the models and the estimates of parameters did not significantly depend on whether wintertime kilometreage was a category or a continuous variable (appendix B).

The main models are presented in Table 17 (models 6–8). In model 6, wintertime kilometreage has three categories and the model includes the interactions between driver age and sex and driver age and vehicle age. In model 7, wintertime kilometreage is a continuous variable. In model 8, wintertime kilometreage is in three categories. In addition it includes interactions between driver age and share of driving in built-up areas, and wintertime kilometreage and share of driving in built-up areas.

Models 6–8, include in the model driver sex, vehicle age and the effect of share of driving in built-up areas, through interactions ($p < 0.02$). When using this model type (deviation contrasts) the interactions also include the individual impacts of the variables.

In the interaction of driver age and sex the only significant coefficient was with driver age class of 26–50 years. The relative risk of female drivers was distinctly greater than that of male drivers in this age group.

The interaction between driver age and vehicle age was statistically significant in the age groups of young and old drivers (≤ 25, >50 years). Vehicle age had a strong correlation with driver age. In the youngest age group, about 31% of all drivers and 47% of the accident drivers drove vehicles that were over 10 years old. In the oldest age group, about 16% of all the drivers and 12% of the accident drivers drove vehicles that were over 10 years old.

In model 4, driver age was coded into six categories in order to estimate its effect more accurately. The interactions between driver age and vehicle age and between driver age and sex were used in the model (Appendix B, model 4). The interactions with the share of the driving in built-up areas were included in the model as well.

As already mentioned, tyre type was not a statistically significant variable in any of the models ($p > 0.28$), but it was included in the final model in order to estimate its effect on survival. When tyre type was included into the models, also variable describing vehicle characteristics was usually kept in the models. Earlier analysis of the data showed that tyre condition correlated with vehicle's age and, therefore, both were entered into the models.

According to the pseudo R-squared model 8 gives the best fit. The situation is not so clear when comparing the fit by loglikelihoods. The model 8 also includes many interactions and becomes then very complicated by structure.

Table 17. Proportional survival models 6–8 estimated from postal survey data. Estimated parameters and their standard errors (in parenthesis.)

Variables Name	Value	Estimated parameters Model 6	Model 7	Model 8
CKIKA2 Driver age	≤ 25	0,7761 (0,1326)	0,7669 (0,1343)	0,7697 (0,1327)
	26–50	-0,0045 (0,1073)	-0,0059 (0,1079)	-0,0455 (0,1104)
	> 50	-0,7716 (0,1558)	-0,7610 (0,1576)	-0,7242 (0,1572)
CTAKM Wintertime kilometreage	≤ 5 000 km	0,000*		-0,3503 (0,1327)
	5 000–10 000 km	0,1988 (0,1401)		-0,0496 (0,0874)
	> 10 000 km	0,5631 (0,1480)		0,3999 (0,0985)
LNAKM	Ln(wintertime /1000)		0,2652 (0,0710)	
CTAJA Share in built-up areas	driving in built-up areas ≤ 50%			0,000 *
	driving in built-up areas > 50%			0,4669 (0,1404)
CKIKA2×SPUOLI Driver age & sex	CKIKA2(1)×SPUOLI	-0,2544 (0,2508)	-0,2850 (0,2536)	-0,2879 (0,2494)
	CKIKA2(2)×SPUOLI	0,4439 (0,1559)	0,4674 (0,1563)	0,4874 (0,1559)
	CKIKA2(3)×SPUOLI	-0,1895 (0,2573)	-0,1825 (0,2603)	-0,1995 (0,2558)
CKIKA2×AUTONI Driver age & vehicle age	CKIKA2(1)×AUTONI(2)	0,6528 (0,2316)	0,6292 (0,2320)	0,6979 (0,2360)
	CKIKA2(2)×AUTONI(2)	-0,0451 (0,1828)	-0,0391 (0,1832)	-0,0615 (0,1843)
	CKIKA2(3)×AUTONI(2)	-0,6077 (0,2618)	-0,5901 (0,2624)	-0,6364 (0,2648)
CKIKA2×CTAJA Driver age & share in built-up areas	CKIKA2(1)×CTAJA(2)			0,3081 (0,2067)
	CKIKA2(2)×CTAJA(2)			-0,4036 (0,1686)
	CKIKA2(3)×CTAJA(2)			0,0955 (0,2065)
CTAKM×CTAJA Wintertime kilometreage & share in built-up areas	CTAKM(1)×CTAJA(2)			-0,4504 (0,1731)
	CTAKM(2)×CTAJA(2)			0,1830 (0,1745)
	CTAKM(3)×CTAJA(2)			0,2674 (0,1962)
RENGAS Tyre type	Studded tyre	0,000 *	0,000 *	0,000*
	Non-studded winter-tyre	0,3395 (0,3227)	0,3179 (0,3227)	0,3463 (0,3230)
Initial situation's -2log-likelihood		4952,4	4952,4	4943,4
Model's -2log-likelihood		4875,0	4874,4	4850,5
Model's degrees of freedom		9	8	14
Number of observations		4099	4099	4040
Pseudo R-Squared		0,24	0,24	0,30

*Reference level

6.4.3 Relative risks to drivers

The coefficients of the variables in a model reflect not only the direct effect of each variable but also the effect of other variables either in the model or background variables not in the model – which are correlated with them. The coefficients can be used in calculating relative risks. This will be illustrated with model 6.

Model 6 includes the variable values: CKIKA2 25–50 a, CTAKM ≤ 5000 km, SEX: female, AUTONI >10 a, RENGAS: non-studded tyres):

$$S(t,X) = S(t_0, X)^{\exp(-0.0045+0.000+0.4439-0.0451+0.3395)}$$

where

$S(t,X)$ is mean share of non-accident involved drivers at time t

$S(t,X)$ is baseline survival function calculated during the estimation.

Relative risk is calculated as:

$$RR = \exp(b_i) \tag{27}$$

where

RR is relative risk

b_i is estimated parameter of the variable in the survival model.

A 95% confidence interval for the relative risk can be obtained by using the confidence interval for *b*i as (e.g. Lee 1992):

$$C_I = b_i \; +/- \; Z_a \times (\text{estimated standard error of } b_i)$$

where Z_a is the chosen percentile point of the normal distribution, and for 95% confidence interval $Z_a = 1.96$.

The interpretation is that the quantified effect describes a relative risk factor having an effect on accident probability.

Table 18 shows the relative risks, their 95% confidence intervals and p-values for model 6. Table 19 shows the same for model 8. Figures 3–7 show the effects of each variable independently in the models. The total effect is multiplicative, the product of all effects of various variables.

Table 18. Relative risks and their 95% confidence intervals and p-values derived from survival model 6.

Model 6, Variables **Name**	**The relative risks and their 95% confidence intervals**			**p-value**
Value	Exp(B)	Lower limit	Upper limit	
CKIKA2 Driver age				**≤ 0.000**
≤ 25	2.173	1.676	2.818	
26–50	1.000	0.807	1.229	
> 50	0.462	0.341	0.627	
The average impact in the material as the reference level = 1.000				
CTAKM Wintertime kilometreage				**< 0.006**
≤ 5 000 km	1.000 *			
5 000–10 000 km	1.220	0.927	1.605	
> 10 000 km	1.756	1.314	2.347	
The average impact in the material as the reference level = 1.000				
RENGAS Tyre type				**< 0.293**
Studded tyres	1.000 *			
Non-studded winter-tyres	1.404	0.746	2.643	
The studded tyres as the reference level = 1.000				
CKIKA2×AUTONI Driver & vehicle age				**< 0.018**
Driver age ≤ 25 a and old/new vehicle	1.902	1.220	3.024	
Driver age 25–50 a and old/new vehicle	0.956	0.668	1.368	
Driver age > 50 a and old/new vehicle	0.545	0.326	0.910	
The average impact in the material as the reference level = 1.000				
CKIKA2×SPUOLI Driver & and sex				**< 0.017**
Driver age ≤ 25 a and woman/man	0.775	0.474	1.268	
Driver age 25–50 a and woman/man	1.559	1.148	2.116	
Driver age > 50 a and woman/man	0.827	0.500	1.370	
The average impact in the material as the reference level = 1.000				

*Reference level

Table 19. Relative risks and their 95% confidence intervals and p-values derived from survival model 8.

Model 8, Variables Name Value	Relative risks and their 95% confidence level Exp(B)	 Lower limit	 Upper limit	p-value
CKIKA2 Driver age				**≤ 0,000**
≤ 25	2.159	1.665	2.800	
26–50	0.956	0.770	1.186	
> 50	0.485	0.356	0.660	
The average impact in the material as the reference level = 1.000				
CTAKM Wintertime kilometreage				**≤ 0.000**
≤ 5 000 km	0.705	0.593	0.837	
5 000–10 000 km	0.952	0.802	1.129	
> 10 000 km	1.492	1.129	1.809	
CTAJA Share of driving in built-up areas				**< 0.001**
≤ 50%	1.000*			
> 50%	1.595	1.211	2.100	
The share of driving in built-up areas < 50% as the reference level = 1.000				
RENGAS Tyre type				**< 0.284**
Studded tyres	1.000*			
Non-studded winter tyres	1.414	0.751	2.663	
The studded tyres as the reference level = 1.000				
CKIKA2×AUTONI Driver & vehicle ages				**< 0.012**
Driver age ≤ 25 and old/new vehicle	2.001	1.265	3.192	
Driver age 25–50 and old/new vehicle	0.940	0.655	1.349	
Driver age > 50 and old/new vehicle	0.529	0.315	0.889	
The average impact in the material as the reference level = 1.000				
CKIKA2×SPUOLI Driver age & sex				**< 0.007**
Driver age ≤ 25 and female/male	0.750	0.460	1.223	
Driver's age 25–50 and female/male	1.628	1.199	2.210	
Driver's age > 50 and female/male	0.819	0.496	1.352	
The average impact in the material as the reference level = 1.000				
CTAKM×CTAJA Wintertime kilometreage & share of driving in built-up areas				**< 0.034**
5 000 km and maximum, and over 50%	0.637	0.454	0.895	
5 000–1 000 km, and over 50%	1.201	0.853	1.691	
Over 1 000 km, and over 50%	1.307	0.890	1.919	
The average impact in the material as the reference level = 1.000				
CKIKA2×CTAJA Driver age & share of driving in built-up areas				**< 0.049**
Driver age ≤ 25 and over 50%	1.361	0.908	2.041	
Driver age 25–50 and over 50%	0.668	0.480	0.930	
Driver age > 50 and over 50%	1.100	0.734	1.649	
The average impact in the material as the reference level = 1.000				

*Reference level

Driver age

Driver age was classified in the survival models either into three or six classes. Driver age had a highly significant effect ($p<0.000$) on driver accident probability (risk) and the share of accident drivers. The relative risk of drivers in the young age group was 4–5 time higher in comparison with oldest group – over 50 years old (Figure 3).

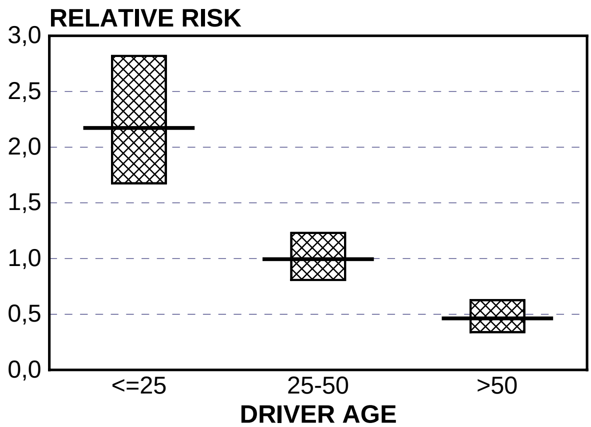

Figure 3. The relative risks of driver age and their 95% confidence intervals according to model 6 (survey data).

The effect of driver age was more refined in model 4, where age was represented by six categories. Driver's relative risk was lowest in the 56–65-year-old group and increased again in the 65+ age group (Appendix B, model 4).

Driver age is usually confounded with driving experience. However, the postal survey did not obtain data on experience. It is likely that the confounding is more significant in the young age group and not very important in the older groups, where the variables are almost perfectly correlated.

Driver sex

Driver sex was included in all of the models, usually as a statistically significant variable. It interacted with age. Middle aged (25–50) female drivers had a 1.6 times higher relative risk than male drivers of the same age group (Figure 4). The reference level = 1.000 is male drivers' risk.

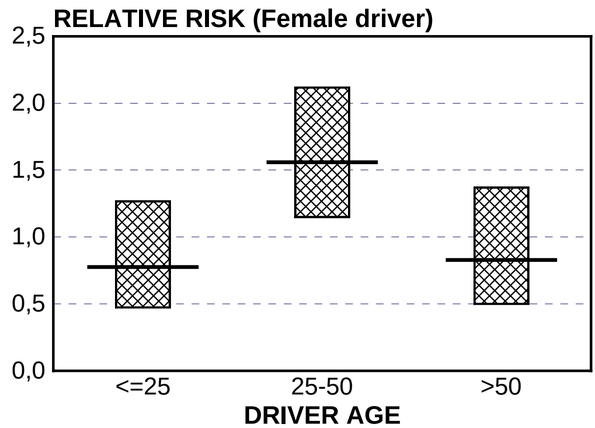

Figure 4. The relative risks of female drivers by age groups, and their 95% confidence intervals, in model 6 (survey data).

Wintertime kilometreage

Wintertime kilometreage represented the effect of exposure on accident probability and it is significant ($p < 0.000$) in all the models. Figure 5 shows how the relative risk in the model increased with exposure. The relative risk of drivers with the highest wintertime exposure was about 1.8 times that of drivers with the least exposure.

The wintertime kilometreage 5,000 km as the reference level = 1.000.

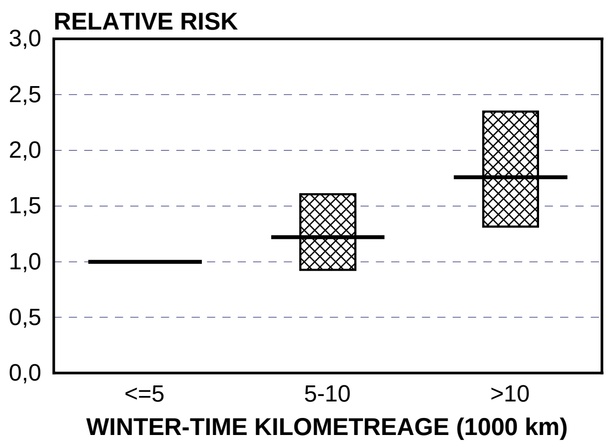

Figure 5. The relative risks of the driver's wintertime kilometreage and their 95% confidence intervals, in model 6 (survey data).

Share of driving in built-up areas

Drivers who drove over 50% of their annual kilometreage in built-up areas, had 1.6 times the relative risk of drivers who drove less than 50% of their kilometreage in built-up areas. The effect of this variable was stronger for young drivers. In addition, they drove in built-up areas more than other age groups.

The interaction between "share of driving in built-up areas" and "wintertime kilometreage" was included in model 8 (Table 17, page 71). This suggests that driving in urban areas may generate an added risk because of more demanding traffic conditions, compared to driving outside built-up areas.

Vehicle characteristics

Vehicle age was included in most of the models as a category variable with two values – up to 10 years and over 10 years. As pointed out earlier, this procedure was adopted because of the importance of vehicle characteristics and because it was assumed to take into account at least some of the variation in tyre condition.

The effect was dependent on driver age as well. Figure 6 shows how relative risk in model 6 changes with vehicle and driver age. Young drivers who used old vehicles had 1.9 times the risk of same age drivers using newer vehicles. Vehicle age did not have a significant effect on risk in the middle age group. Older drivers who use old vehicles had lower risk than old drivers using newer vehicles.

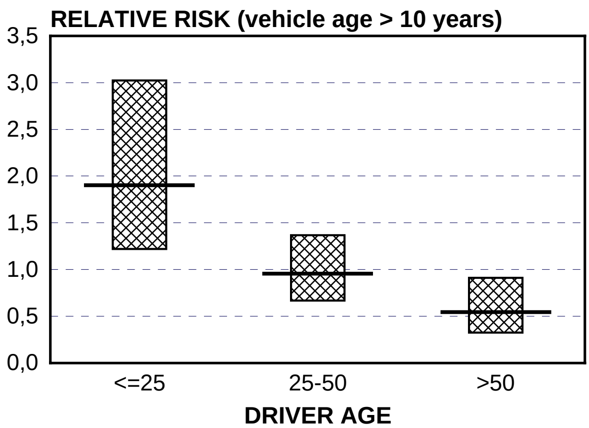

Figure 6. The relative risks of vehicle age and their 95% confidence intervals by driver age, in model 6 (survey data).

Type of tyre

Figure 7 shows that the risk of drivers who had used non-studded winter tyres was about 1.4 times that of the drivers who used studded tyres (as the reference level = 1.000.) However, the 95% confidence interval of the relative risk was 0.8–2.6, implying, that the difference could be entirely coincidental. It should be pointed out that the few cases of drivers who reported using the new kind of friction tyres were included in the non-studded category. This fact may have reduced the risk difference between users of studded and non-studded tyres.

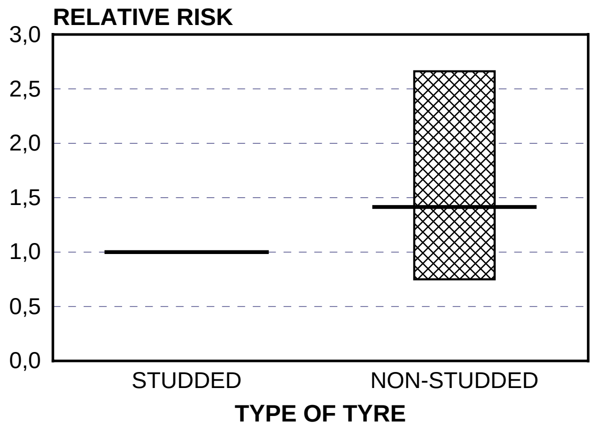

Figure 7. The relative risk of tyre type and its 95% confidence interval, in model 6 (survey data).

Joint impact of the relative risks

A cumulative hazard function (section 6.1.1) is a way of presenting jointly the relative risks associated with several variables. The function shows the expected number of accidents accumulated until moment *t*. Figure 8 shows hazard functions generated by model 6, for three age groups.

A hazard function can be transformed into a survival function (Figure 9), which shows how the effects of the model variables either increases or reduces the share of drivers involved in accidents during the examination period (accident probability). The effect of driver age is calculated in relation to the base level of the cumulative survival function (see section 6.1.1):

$$S(t) = [S_0(t)]^{\exp(XB)}$$

where

$S(t)$ is cumulative survival function (share of accident drivers at moment t)

$S_0(t)$ is calculated base level

$p = \exp(XB)$
$\exp(XB) = \exp\,(b_1 x_1 + b_2 x_2 + \ldots + b_n\, x_n)$

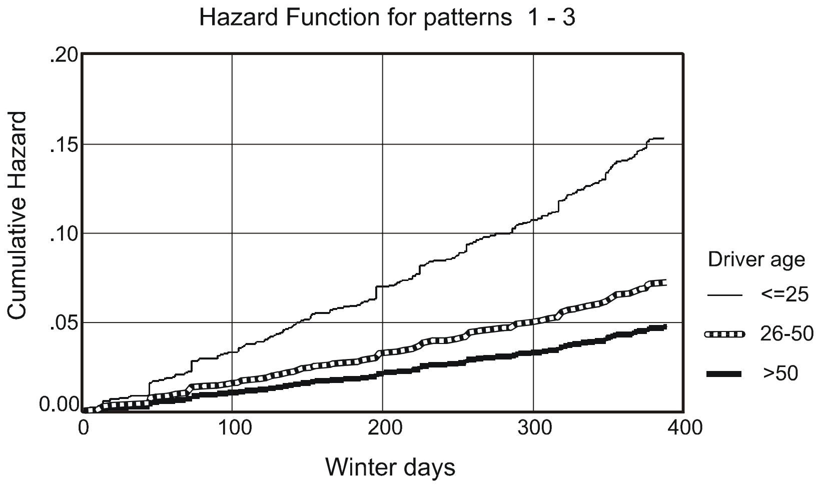

Figure 8. Cumulative hazard functions for drivers in different age groups, by model 6 (survey data).

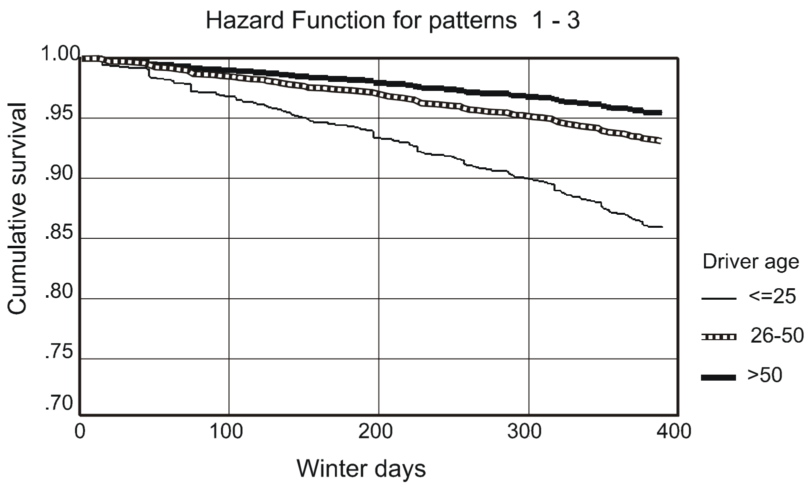

Figure 9. Survival functions for drivers in different age groups, by model 6 (survey data).

Changes in the number of accidents and in the share of accident drivers can be estimated with the coefficients of the variables in the function and their standard errors. This allows an estimate of relative risks of variables. An example with tyre type is given below.

The cumulative survival function, based on model 6, estimates the share of accident drivers as 6.6% (271, out of 4,099 drivers). Had all drivers used studded tyres, the share of accident drivers would have been 6.4% and the number of accident involved drivers would have been 262. On the other hand, had all drivers used non-studded winter tyres, the share of accident involved drivers would have been 8.8% (4.7–16.7%, 95% confidence interval) and the number of accident involved drivers would have been 361. The ratio between the shares is 1.38, which is interpreted as the relative risk of driving with non-studded tyres.

6.5 Models for the VALT data

6.5.1 Models and their compilation principles

To the extent possible, survival models for the VALT data were derived in a similar way as those compiled for the self-reported accident data in the postal survey. The same basic categories of driver, vehicle and exposure characteristics were used in the models. However, VALT data did not include wintertime exposure but it included variables which were not available in the postal survey. These were vehicle weight, route familiarity, speed limit, relative speed (difference between speed at accident time and posted speed limit), and road surface conditions.

The Kaplan–Meier estimates of variables' effects suggested which variables should be included in the models. Some variables represented characteristics associated with the driver and the vehicle, as in the survey data. The other variables referred specifically to the accident situation. The correlations between the variables were also examined.

The variables were first divided into two groups: 1) the most important variables from the point of view of comparing the VALT based and survey based models and 2) other interesting variables. The first group included: driver age and sex, annual vehicle kilometreage, use of alcohol, vehicle age and weigh, tyre type, tyre tread depth, and use of safety belt. The second group included route familiarity, the speed limit, the relative speed, and road surface conditions.

The compiling of the models followed the same principles as with the postal survey data (section 6.4.1). First all variables were included, and those that did not pass the statistical test were eliminated. Next, interactions were added to the model and then the statistically insignificant ones were eliminated. The models of the first variable group included all important interactions. When the second variable group was entered in the estimation, interactions with "use of alcohol" could not be integrated into the models because they distorted the structure of the models.

Finally, the models compiled by the elimination method were compared with the models that were compiled by adding variables. Since similar models resulted, the maximal first model was accepted. Next, vehicle age and weight, tyre types, and tyre tread depth were added to the models. The relative speed was added to models 1 and 2.

In preliminary analyses driver sex had strong effect in the models. This variable had also strong correlations with other variables, especially with annual vehicle kilometreage. The variable ‘sex’ was integrated into the models, even though Kaplan–Meier analysis did not support the integration. This was because the detailed analysis of the share of accident drivers pointed to middle-age females as having a large share of wintertime accidents. This finding held even after controlling for exposure. The same conclusion was reached with the postal survey data.

A summary of the models is presented in Appendix C.

The models included the following variables:

- *driver’s annual vehicle kilometreage,
- *driver age,
- use of alcohol,
- *driver sex,
- use of safety belt,
- route familiarity,
- vehicle weight,
- *type of tyres used in the vehicle,
- worst tyre’s tread depth,
- the speed limit,
- relative speed.

The variables marked with * are common to both VALT based and survey based models.

Table 20 lists the most important variables in the survival models based on VALT data.

Table 20. The most important variables, categories and values of proportional survival models for VALT data.

Variables Name	**Content**	**Categorisation**	
VAKM	Annual vehicle kilometreage	1	< 10 000 km
		2	10 000–29 999 km
		3	≥ 30 000 km
CKIKA2	Driver age	1	≤ 25
		2	26–50
		3	> 50
CKIKA	Driver age	1	≤ 25
		2	26–35
		3	36–45
		4	46–55
		5	56–65 a
		6	> 65
ALKO	Useof alcohol	0	≤ 0,5 ‰
		1	> 0,5 ‰
KEVYT	Vehicle weight	1	≤ 1 000 kg
		2	> 1 000 kg
TKELI	Winter road surface conditions	1	Dry or wet
		1	Snowy, icy or slushy
NRAJA	Speed limit	0	≤ 60 km/h
		1	> 60 km/h
AUTONI	Vehicle age	1	10 at maximum
		2	Over 10
AUTONI2	Vehicle age	1	5 years at maximum
		2	5–10 years
		3	Over 10 years
RENGAS	Tyre type	0	Non-studded winter tyres
		1	Studded tyres
RENGAS2	Tyre type	1	Summer tyres
		2	Non-studded winter tyres
		3	Studs in good condition
		4	Studs in bad condition
RENURA	Worst tyre's tread depth	0	4 mm at maximum
		1	Over 4 mm
NTONTKE	Non-studded tyres and winter road surface conditions	0	No
		1	Yes
NASTKE	Studded tyres and winter road surface conditions	0	No
		1	Yes
NTONPKE	Non-studded tyres and other road surface conditions	0	No
		1	Yes
SNOPE	Relative speed	0	< 0 km/h
		1	≥ 0 km/h
SPUOLI	Sex	1	Male
		2	Female
TVYO	Safety belt	0	Not in use
		1	In use
TUTTU	Familiarity of route	1	Scene of accident passed more often than once in a month
		0	Scene of accident passed more seldom than once in a month

6.5.2 Characteristics of the models

Log-likelihood values and test scores based on the χ^2-distribution proved that all models included statistically significant parameter estimates. There were no significant differences between basic models 1 or 2 in explanatory power.

There were several significant correlations between the variables of the first group of variables (model 1):

- sex and annual vehicle kilometreage (0.38),
- age and sex (0.28),
- use of alcohol and use of safety belt (0.41),
- tyre type and vehicle age (0.26),
- annual vehicle kilometreage and vehicle age (0.15),
- driver's sex and vehicle weight (0.15).

The added variables in model 2 had significant correlations between:

- driver age and route familiarity, especially in the oldest age group (0.22),
- speed limit and use of safety belt (0.15), and route familiarity (0.15),
- relative speed and driver age (0.22), and the use of alcohol (0.17).

Table 21 shows survival models 1–4. Model 1 is the basic model without the variables of the second group. Model 2 includes all of the most important variables. Model 3 includes all other important variables but leaves out the interaction between speed limit and relative speed. Model 4 includes the main variables, leaves out speed limit and relative speed, but integrates the interaction between vehicle age and tyre type.

Table 22 shows the relative risks, their 95% confidence intervals and their statistical significance, according to model 1.

Table 23 shows the relative risks, their 95% confidence intervals and their statistical significance, according to model 3.

Table 21. Proportional survival models 1–4 for VALT data; estimated parameters and their standard errors (in parenthesis.)

Variables		**Estimated parameters**			
Name	**Value**	**Model 1**	**Model 2**	**Model 3**	**Model 4**
VAKM	Annual vehicle kilometreage				
	< 10 000 km	0.000*	0.000*	0.000*	0.000*
	10 000–29 999 km	-0.6209 (0.1790)	-0.5974 (0.1973)	-0.6297 (0.1947)	-0.6425 (0.1804)
	≥ 30 000 km	-0.7231 (0.2039)	-0.6521 (0.2331)	-0.7086 (0.2262)	-0.7342 (0.2042)
CKIKA2	Driver age				
	Driver ≤ 25	0.000 *	0.000 *	0.000 *	0.000 *
	Driver 26–50	-0.4336 (0.1499)	-0.3286 (0.1630)	-0.3287 (0.1631)	-0.4353 (0.1501)
	Driver > 50	-0.0182 (0.1860)	0.1484 (0.2073)	0.1501 (0.2073)	-0.0167 (0.1862)
SPUOLI	Driver sex	0.3370 (0.1622)	0.3542 (0.1842)	0.3186 (0.1806)	0.3348 (0.1625)
ALKO	Use of alcohol	0.7317 (0.2450)	0.6575 (0.2756)	0.6575 (0.2754)	0.7246 (0.2441)
TVYO	Safety belt in use	-0.4932 (0.1682)	-0.4539 (0.1833)	-0.4586 (0.1830)	-0.4967 (0.1673)
KEVYT	Vehicle weight ≤ 1 000 kg	0.1828 (0.1359)	0.2122 (0.1455)	0.2238 (0.1449)	0.1896 (0.1364)
RENURA	Tyre's tread depth (mm)	-0.0597 (0.0304)	-0.0638 (0.0340)	-0.0633 (0.0339)	-0.0564 (0.0305)
TUTTU	Familiarity of route		-0.4178 (0.1537)	-0.4162 (0.1537)	-0.4081 (0.1456)
NRAJA	Speed limit > 70 km/h		-0.3667 (0.1724)	-0.3842 (0.1706)	
SNOPE	Relative speed > 0 km/h		0.0910 (0.1787)	0.1940 (0.1466)	
NRAJA×SNOPE	Relative speed >0 km/h & speed limit >60 km/h		0.3501 (0.3456)		
AUTONI	Vehicle age > 10 v	-0.2620 (0.1650)	-0.2737 (0.1763)	-0.2728 (0.1760)	-0.0220 (0.2849)
RENGAS	Tyre type	0.2373 (0.2344)	0.2244 (0.2806)	0.2151 (0.2804)	0.4012 (0.2741)
AUTONI×RENGAS	Vehicle age > 10v & non-studded winter tyres				0.5589 (0.5532)
Initial situation's -2log-likelihood		2881.7	2524.9	2524.9	2881.8
Model's -2log-likelihood		2809.8	2452.7	2453.8	2808.8
Number of observations		446	407	407	446
Pseudo R-Squared		0.25	0.28	0.27	0.26

*Reference level

Table 22. Estimated relative risks and 95% confidence intervals and p-values for variables in model 1 of VALT data.

Variables		**Relative risks and their 95% confidence interval**			**p-value**
Name	**Value**	**Exp(B)**	**Lower limit**	**Upper limit**	
VAKM	**Annual vehicle kilometreage**				0.001
	≤ 10 000 km	1.000*			
	10 000–29 999 km	0.537	0.378	0.763	
	≥ 30 000 km	0.485	0.326	0.723	
CKIKA2	**Driver age**				0.004
	≤ 25	1.000*			
	26–50	0.648	0.483	0.870	
	> 50	0.982	0.682	1.414	
SUKUPUOLI	**Sex**				0.038
	Male	1.000*			
	Female	1.401	1.019	1.925	
ALKOHOLI	**Use of alcohol**				0.003
	No	1.000*			
	Yes	2.079	1.286	3.360	
TVYÖ	**Use of safety belt**				0.003
	No	1.000*			
	Yes	0.611	0.439	0.849	
AUTON I	**Vehicle age**				0.112
	10 years at maximum	1.000*			
	Over 10 years	0.770	0.557	1.063	
KEVYT	**Vehicle weight**				0.179
	> 1 000 kg	1.000*			
	≤ 1 000 kg	1.201	0.920	1.567	
RENURA	**Tyre's tread depth**				0.049
	e.g. for 1 mm	0.942	0.888	1.000	
RENGAS	**Tyre type**				0.311
	Studded tyre	1.000*			
	Non-studded winter tyre	1.268	0.801	2.001	

*Reference level

Table 23. Estimated relative risks and their 95% confidence intervals and p-values for variables in model 3 of VALT data.

Variables		**Relative risks and their 95% confidence interval**			**p-value**
Name	**Value**	**Exp(B)**	**Lower limit**	**Upper limit**	
VAKM	**Annual vehicle kilometreage**				**< 0.007**
	≤ 10 000 km	1.000 *			
	10 000–29 999 km	0.550	0.374	0.810	
	≥ 30 000 km	0.521	0.330	0.823	
CKIKA2	**Driver age**				**< 0.015**
	≤ 25	1.000 *			
	25–50	0.720	0.523	0.991	
	> 50	1.160	0.773	1.741	
SUKUPUOLI	**Female driver**				**< 0.055**
	Man	1.000 *			
	Woman	1.425	0.993	2.045	
ALKOHOLI	**Use of alcohol**				**< 0.017**
	No	1.000 *			
	Yes	1.930	1.125	3.313	
TVYÖ	**Use of safety belt**				**< 0.013**
	No	1.000 *			
	Yes	0.635	0.443	0.910	
AUTONI	**Vehicle age**				**< 0.121**
	10 years at maximum	1.000 *			
	Over 10 years	0.761	0.538	1.075	
KEVYT	**Vehicle weight**				**< 0.145**
	> 1 000 kg	1.000 *			
	≤ 1 000 kg	1.236	0.930	1.644	
NRAJA	**Speed limit**				**< 0.033**
	60 km/h at maximum	1.000 *			
	Over 60 km/h	0.693	0.494	0.972	
SNOPE	**Relative speed**				**< 0.611**
	60 km/h at maximum	1.095	0.772	2.169	
	Over 60 km/h	1.419	0.721	2.794	
TUTTU	**Familiarity of route**				**< 0.007**
	Seldom	1.000 *			
	At least once a month	0.659	0.487	0.890	
RENURA	**Tyre's tread depth**				**< 0.060**
	for 1 mm	0.938	0.878	1.003	
RENGAS	**Non-studded winter tyre**				**< 0.424**
	Studded tyre	1.000*			
	Non-studded winter tyre	1.252	0.722	2.169	

*Reference level

Driver’s annual vehicle kilometreage (VAKMI, $p < 0.001$), use of alcohol (ALKOHOLI, $p < 0.003$) and driver age (CKIKA, $p < 0.004$) had the largest explanatory power in the models (Tables 22 and 23 as well as appendix C). Driver sex had a moderate influence. However, sex had a strong correlation with kilometreage suggesting that its effect reflected differences associated with exposure.

The effect of sex was further analysed by calculating a stratified model (model 7, appendix C). In this procedure, different base-level hazard functions were first calculated for female and male drivers and then the effect of each variable was introduced. The main difference between the two models (model 1 and model 7) was that the stratified model also included the interaction between driver age and vehicle age. There were no other major differences between the coefficients and their statistical significance. These results support the inclusion of sex as a separate factor in the models.

The use of alcohol had a strong correlation with the non-use of safety belt. Drivers who drove under the influence of alcohol were more likely to be classified as the accidents’ main guilty parties.

Tyre type did not enter in the models as significant variable, but tyre quality (tread depth) did, ($p < 0.05$). Decrease in the tyre’s tread depth increased driver’s accident risk.

Some risk factors, when combined, may have a synergetic effect. For example, 27% of primary accident drivers who were sober, also had tyres in bad condition, while in the non-sober group 42% had worn tyres.

6.5.3 Relative risks to drivers

Tables 22 and 23 present the relative risks along with their confidence intervals. The effect of each independent main variable is presented in the Figures 10–17.

Driver age

Figure 10 shows that the relative risk was high in both the youngest and the oldest driver age groups (The age group ≤ 25 serves as the reference level = 1.000.) Figure 11 is a model with more detailed age groups. It shows that the relative risk is lowest with the 36–45-, 46–55- and 56–65-year-old drivers and then increased again in the 65+ age group. This result is different from that in the postal survey data based models, where only the young drivers had a higher risk than all others.

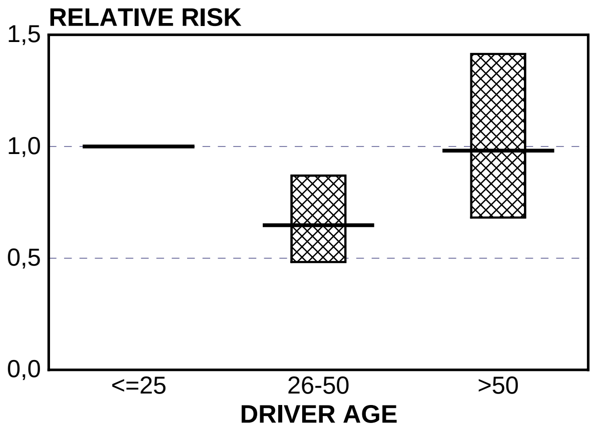

Figure 10. The relative risks of driver age their 95% confidence intervals, in model 1 (VALT data).

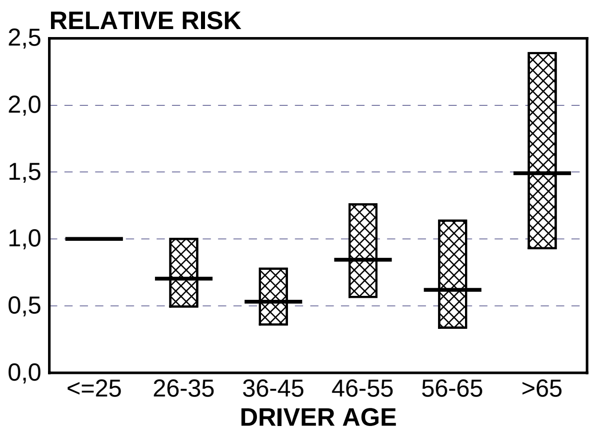

Figure 11. The relative risks of driver age and their 95% confidence intervals, in model 8 (VALT data).

Driver sex

According to the Kaplan–Meier estimates, the effect of sex would disappear when annual kilometreage is taken into account. The share of male and female drivers in the “guilty party” group differed significantly only in the 26–50 year age group.

However, in the models, sex was a significant variable in several models. Figure 12, of model 1, shows that female drivers had 1.4 times the risk (1.02–1.93 with 95%

confidence interval) of male drivers. According to the estimated stratified model, female drivers had a distinctly greater risk than male drivers (model 7, Appendix C).

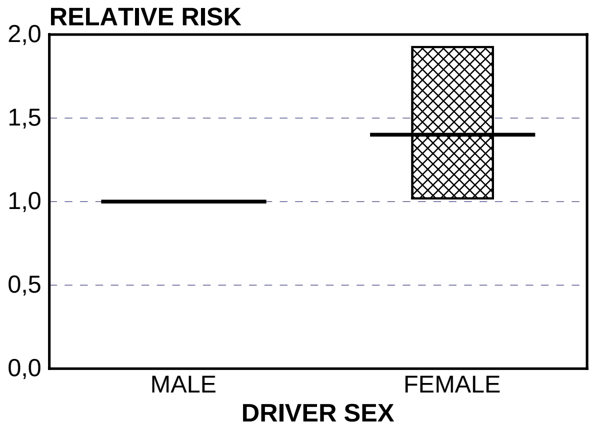

Figure 12. The relative risk of female drivers and its 95% confidence interval, in model 1 (VALT data).

Driver's use of alcohol

Driving under the influence of alcohol significantly increased driver's wintertime accident risk. According to model 1, the relative risk of drivers under the influence of alcohol, was 2.1 times (1.29–3.36 with 95% confidence interval) that of sober drivers (Figure 13).

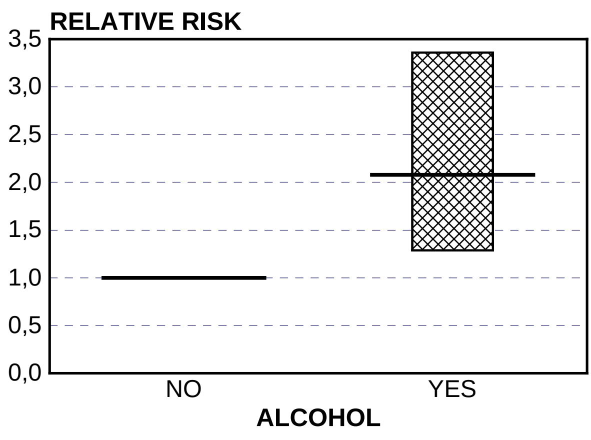

Figure 13. The relative risk of driving under the influence of alcohol and its 95% confidence interval, in model 1 (VALT data).

Use of safety belt

The relative risk of drivers who had used safety belts, was 39% lower (15–56% with 95% confidence interval) than the risk of drivers who had not used a safety belt (Figure 14). The correlations between non-use of belt, use of alcohol, and speeding, suggest that the hard core of non-users are different from users also in other aspects of behaviour that are associated with higher accident risks. They are also more likely of being identified as the primary involved party.

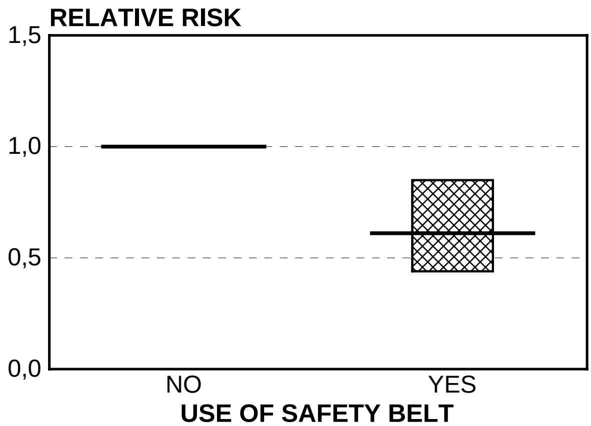

Figure 14. The relative risk of using of safety belt and its 95% confidence interval, in model 1 (VALT data).

Driver's previous accident history

Accident record did not have a separate effect once annual vehicle kilometreage was taken into account. The same conclusion was suggested by Kaplan–Meier estimates.

Annual vehicle kilometreage

Drivers' accident risk decreased as annual kilometreage increased. Figure 15 shows that drivers in the exposure category of less than 15,000 km/a had a relative risk about 50% higher compared to drivers in the larger exposure categories. The "guilty parties", as a group, had a lower annual vehicle kilometreage than the "other involved parties" had.

This result is in the opposite direction to what was found with the postal survey data where higher exposure was associated with higher risk. However, the VALT data, in contrast to the postal survey data, included only drivers involved in fatal accidents. In this population the annual vehicle kilometreage reflects, perhaps, specific wintertime related (beneficial) driving experience whereas in the postal survey data, which includes drivers who avoided accidents, the kilometreage reflects more the accumulated exposure to risk.

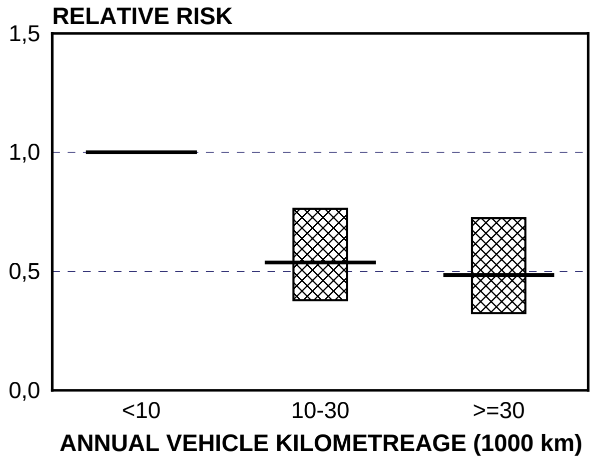

Figure 15. The relative risks of the driver's annual vehicle kilometreage and their 95% confidence intervals, in model 1 (VALT data).

Route familiarity

The variable was categorised to having driven at the site of the accident no more than once a month or more frequent than that. It had good explanatory power in the models ($p < 0.010$). According to model 2, the accident risk was about 34% (95% confidence interval: 11–52%) lower for drivers who were familiar with the site of the accident compared to other drivers.

Vehicle weight

Vehicle weight proved to be a discriminating factor although its statistical significance in the models was weak ($p > 0.15$). Weight was usually categorised as up to 1000 kg or over it. According to model 1, drivers using light vehicles had 1.2 times the relative risk (the 95% confidence interval: 0.92–1.57) of drivers using the larger weight vehicles.

Vehicle weight is typically associated with other vehicle characteristics. The earlier Kaplan–Meier estimates pointed out that vehicle's traction (front-wheel and rear-wheel drives) was one of such variables.

Vehicle age

Vehicle age had no significant independent effect on accident risk, in the models.

Tyre type and condition

Figure 16 shows that tyre type had a small, yet not significant, effect on driver's accident risk ($p > 0.3$). The relative risk of drivers, who used non-studded winter tyres, was about

1.3 times that of drivers who used studded tyres (Table 21, model 1). When the model included drivers who used summer tyres (model 9, Appendix C) the statistical significance of this variable improved (p-value 0.15). The relative risk of drivers, who used summer tyres, was about 1.7 times that of users of studded tyres

The effect of tyre type depended on the inclusion of tyre condition, use of alcohol and use of safety belt in the model. When these were omitted from the model the effect of tyre type on relative risk diminished.

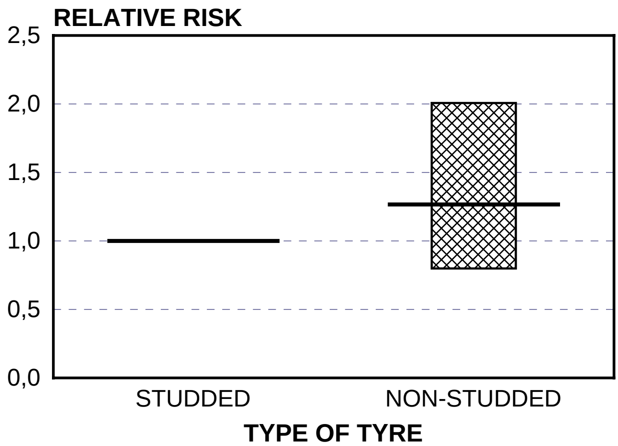

Figure 16. The relative risk of using non-studded tyres and its 95% confidence interval, in model 1 (VALT data).

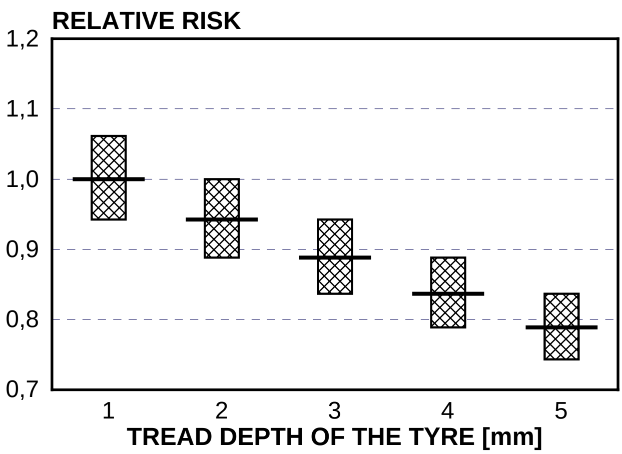

Figure 17. The relative risks of tyre tread depth and their 95% confidence interval, in model 1 (VALT data).

Tyre condition in itself, defined by tread depth of the worst tyre, has a systematic effect on relative risk. Figure 17 shows how risk changes with increasing tread wear.

Tyre type and condition interact. The relative risk of driving with non-studded and worn tyres is 1.9 times that of driving with studded tyres in good condition. Using non-studded tyres in good condition is only 1.1 times riskier.

Tyre condition, alcohol use and safety belt non-use were correlated. Drivers under the influence of alcohol drove old vehicles (52% vs. 20% in the sober group), had tyres in bad condition (40% vs. 23% in the sober group), and did not use safety belts (19% use rate compared to 84% in the sober group).

Speed limit and relative speed

According to model 2, drivers' relative risk was 32% lower on roads with speed limits over 60 km/h than on roads with low speed limits (95% confidence interval: 5–51%). Low speed limits are typical to built-up areas, with frequent junctions, bicycles and pedestrians. The higher relative risk on such roads on may reflect these factors as well.

Relative speed, the difference between the vehicle's estimated speed and the posted speed limit, had no significant effect in the models ($p > 0.6$).

Road surface conditions

The variable of specific road surface conditions at the time of the accident (whether the road had snow, slush, or ice on it vs. other condition) proved not to be statistically important variable. It was eliminated in the basic model specification. It did improve the models at a later stage, but not significantly (1.14 increase in risk, the 95% confidence interval: 0.84–1.55).

6.6 Evaluation of the models and methods

6.6.1 Evaluation of the postal survey models

The more successful survival (models 6 and 8 for the survey data; models 1 and 2 for the VALT accident data), were evaluated according to the following criteria:

1) size of residuals (Martingale and Cox–Snell residuals)
2) size of partial residuals
3) impact of eliminating observations that were deviant and had leverage (DfBeta)
4) impact of alternate elimination of the variables.

Model 6 contained seven observations which had high residual values (Cox–Snell); model 8 contained 13 such observations.

Partial residuals indicated that the proportional hazard models fits the data well, as no significant deviations were found.

Observations that were and had leverage were identified by DfBeta-coefficients (estimates the influence of each observation on the resultant coefficients of model variables.) Most of the deviant observations were found in driver age and its interactions. After deviant observations were removed, the models were re-estimated.

The elimination of seven observations with leverage affected the coefficients of driver age (CKIKA2), the interaction of driver age with vehicle age (CKIKA2 × AUTONI) and tyre type (RENGAS). Table 24 shows how model coefficients change as a result of removing the deviant observations. The elimination of the seven deviant observations from model 6 emphasised the relative risk of the youngest age group and diminished the relative risk of the oldest age group. It also reduced the coefficient of tyre type so that its relative risk decreased from 1.40 to 1.25. The changes in the correlation between driver and vehicle ages increased the relative risk of young drivers in old vehicles and lowered the risk of old drivers.

The elimination of the 13 deviant observations with leverage from model 8 affected the coefficients of driver age (CKIKA2), share of driving in built-up areas (CTAJA), tyre type (RENGAS) and the interactions of driver age and sex (CKIKA2 × SPUOLI), driver and vehicle ages (CKIKA2 × AUTONI) and driver age and the share of driving in built-up areas (CKIKA2 × CTAJA). The largest changes occurred in the coefficients of driver age and the interaction of driver age and the share of driving in built-up areas.

Elimination of the deviant observations and re-estimation of the models has changed the absolute value of some coefficients but did not affect the nature of effects. Minor changes occurred in the relative risks but the estimated values remained well within the original 95% confidence interval.

Table 25 shows the changes in the calculated accident risks, based on model 5, after removing variables. The reference model in this estimation did not include interactions, but was based on the same 4,040 observations as in the basic model. The percentages in Table 25 show the deviation of risk levels from the original reference model (5.1.)

The removal of driver age increased the coefficients of wintertime kilometreage (CTAKM), vehicle age (AUTONI) and driver sex (SPUOLI) in the model. The removal of wintertime kilometreage changed the coefficients of driver age (CKIKA2) and sex (SPUOLI). These most significant changes in the coefficients were in the range of –4 % to +17 %.

The removal of other variables from the model had only minor influence (–6% to +3%). The effect of tyre type remained very stable during removal of other variables from the model.

Table 24. The effects of leverage on models 6 and 8 of the postal survey data, estimated parameters and their standard errors (in parenthesis.)

The left column is the original estimation, and DfBeta-models are with leverage observations removed. The dark background indicates coefficients with the largest changes.

Variables		**Estimated parameters**			
Name	**Value**	**Model 6**	**DfB 6**	**Model 8**	**DfB 8**
CKIKA2 Driver age	≤ 25	0.7761 (0.1326)	0.8689 *	0.7697 * (0.1327)	0.8635 *
	26–50	-0.0045 (0.1073)	**0.1369 (0.1114)**	-0.0455 (0.1104)	**0.1490 (0.1160)**
	> 50	-0.7716 (0.1558)	**-1.0058 (0.1649)**	-0.7242 (0.1572)	**-1.013 (01760)**
CTAKM Wintertime kilometreage	≤ 5 000 km	1.000 *	1.000 *	-0.3503 (0.1327)	-0.3888 *
	5 000–10 000 km	0.1988 (0.1401)	0.2551 (0.1420)	-0.0496 (0.0874)	-0.0532 (0.0900)
	>10 000 km	0.5651 (0.1480)	0.6187 (0.1505)	0.3999 (0.0985)	0.4420 (0.1013)
CTAJA Share in built-up areas	Driving in built-up areas ≤ 50%			1.000*	1.000*
	Driving in built-up areas > 50%			0.4669 (0.1404)	0.3757 (0.1523)
CKIKA2× AUTONI Driver age & vehicle age	CKIKA2(1)×AUTONI(2)	0.6528 (0.2316)	**0.9646 ***	0.6979 (0.2360)	**0.8801**
	CKIKA2(2)×AUTONI(2)	-0.0451 (0.1828)	**0.1258 (0.1797)**	-0.0615 (0.1843)	0.0895 (0.1816)
	CKIKA2(3)×AUTONI(2)	-0.6077 (0.2618)	**-1.0904 (0.2829)**	-0.6364 (0.2648)	**-0.9695 (0.2873**)
CKIKA2× SPUOLI Driver age & sex	CKIKA2(1)×SPUOLI	-0.2543 (0.2508)	-0.2330 *	-0.2879 (0.2494)	-0.2466
	CKIKA2(2)×SPUOLI	0.4439 (0.1559)	0.4550 (0.1568)	0.4874 (0.1559)	0.5733 (0.1562)
	CKIKA2(3)×SPUOLI	-0.1896 (0.2573)	-0.2220 (0.2667)	-0.1995 (0.2558)	**-0.3267 (0.2766)**
CKIKA2× CTAJA Driver age & share in built-up areas	CKIKA2(1)×CTAJA(2)			0.3081 (0.2067)	**0.5240**
	CKIKA2(2)×SPUOLI			-0.4036 (0.1686)	-0.3166 (0.1786)
	CKIKA2(3)×CTAJA(2)			0.0955 (0.2065)	**-0.2074 (0.2346)**
CTAKM× CTAJA Wintertime kilometreage & share in built-up areas	CTAKM(1)×CTAJA(2)			-0.4504 (0.1731)	-0.4069 (0.1794)
	CTAKM(2)×CTAJA(2)			0.1830 (0.1745)	0.1594 (0.1797)
	CTAKM(3)×CTAJA(2)			0.2674 (0.1962)	0.2736 (0.2013)
RENGAS Tyre type	Studded tyre	1.000 *	1.000 *	1.000 *	1.000 *
	Non-studded winter tyre	0.3395 (0.3227)	**0.2236 (0.3400)**	0.3463 (0.3230)	**0.2571 (0.3403)**
Initial situation's -2log-likelihood		4952.4	4835.9	4943.4	4727.5
Model's -2log-likelihood		4875.0	4741.1	4850.5	4622.3
Model's degrees of freedom		9	9	14	14
Number of observations		4099	4092	4040	4027

*Reference level

Table 25. The relative risks and their 95% confidence intervals when removing different variables from model 5 of the postal survey data (Appendix B, model 5).

The dark background indicates the greatest changes in the relative risks.

Variables	Model 5.1 (Original)	Model 5.2 (No CKIKA2)	Model 5.3 (No CTAKM)	Model 5.4 (No RENGAS)	Model 5.6 (No SPUOLI)	Model 5.7 (No CTAJA)	Model 5.5 (No AUTONI)
CKIKA2(1) Driver age	1.000 *	-	1.000 *	1.000 *	1.000 *	1.000 *	1.000 *
CKIKA2(2) Driver age	0.448 0.332–0.606	-	0.434 (-3.1%) 0.321–0.586	0.454 (+1.3%) 0.332–0.609	0.442 (–1.3%) 0.327–0.597	0.441 (–1.6%) 0.326–0.595	0.442 (–1.3%) 0.328–0.597
CKIKA2(3) Driver age	0.322 0.223–0.465	-	**0.278** **(-13.7%)** 0.194–0.399	0.323 (+0.3%) 0.223–0.467	**0.305** **(–5.3%)** 0.212–0.440	0.313 (–2.8%) 0.217–0.452	0.315 (–2.2%) 0.218–0.453
CTAKM(1) Wintertime kilometreage	1.000 *	1.000 *	-	1.000*	1.000 *	1.000 *	1.000 *
CTAKM(2) Wintertime kilometreage	1.266 0.956–1.675	**1.397** **(+10.3%)** 1.060–1.842	- -	1.267 (+0.1%) 0.957–1.677	**1.203** **(–5.0%)** 0.914–1.585	1.230 (–2.8%) 0.931–1.625	1.250 (–1.3%) 0.946–1.651
CTAKM(3) Wintertime kilometreage	1.863 1.374–2.526	**2.181** **(+17.1%)** 1.621–2.935	- -	1.867 (+0.2%) 1.377–2.532	1.734 (–6.9%) 1.291–2.329	**1.759** **(–5.6%)** 1.305–2.371	1.828 (–1.9%) 1.353–2.471
AUTONI Vehicle age	1.126 0.851–1.489	**1.255** **(+11.5%)** 0.952–1.655	1.039 (-7.7%) 0.789–1.370	1.139 (+1.2%) 0.863-1.504	1.108 (–1.6%) 0.838–1.465	1.136 (+0.9%) 0.859–1.502	– –
SPUOLI Driver sex	1.311 1.004–1.710	**1.448** **(+10.5%)** 1.113-1.884	**1.160** **(-11.5%)** 0.897–1.501	1.309 (-0.2%) 1.003–1.708	– –	1.331 (+1.5%) 1.021–1.736	1.302 (–0.7%) 0.968–1.698
CTAJA Share in built-up areas	1.289 1.006–1.652	1.355 (+5.1%) 1.058–1.735	1.177 (-8.7%) 0.923-1.501	1.291 (+0.2%) 1.007–1.654	1.308 (+1.5%) 1.021–1.675	– –	1.293 (+0.3%) 1.009–1.657
RENGAS Tyre type	1.330 0.705–2.508	1.269 (-4.6%) 0.673–2.395	1.358 (+2.1%) 0.721–2.559	- -	1.321 (–0.7%) 0.701–2.491	1.341 (+0.8%) 0.711–2.529	1.368 (+2.9%) 0.728–2.571

In summary, the evaluation based on the partial residuals indicated that the models' structural supposition, proportionality, was acceptable. The elimination of deviant or leverage observations from the models caused rather minor changes to the coefficients of variables in the models. The calculated relative risks remained within the 95% confidence intervals of the basic models. The practical conclusion, therefore, is that deviant observations have not distorted the models and there was no need to eliminate observations or change the structure of the models.

6.6.2 Evaluation of the VALT data based accident models

The four evaluation criteria (6.6.1) were applied to the survival models 1 and 2 of VALT data pointing out:

a) There were three observations in the data with a high residual value (over ± 2.0), according to the Martingale residuals.
b) The partial residuals indicated that the proportional hazard model (2) fits the data well, since no significant deviations were found.
c) According to the DfBeta-coefficient criterion, model 1 had 8 deviant observations regarding driver's annual vehicle kilometreage, 9 regarding sex, 7 regarding tyre's tread depth, 5 regarding vehicle age, and 2 regarding tyre type. In model 2 there were 6 deviant observations related to route familiarity, 2 regarding relative speed, and 4 related to the speed limit.

The elimination of the deviant observations affected the coefficients of the following variables: use of alcohol (ALKOHOLI), use of the safety belt (TVYO), vehicle age (AUTONI) and driver age (CKIKA). The variables were highly correlated and their estimated coefficients were sensitive to changes in number of observation.

The elimination of the deviant observations with leverage increased the effect of average vehicle kilometreage, and reduced the coefficients of other variables, especially of use of alcohol and tyre type .

The elimination of the nine observations with leverage related to sex increased the variable's coefficient and the relative risk from 1.4 to 1.6. It also reduced the coefficient of tyre type. Elimination of deviant observations related to alcohol use caused a minor effect in the coefficients of the variable itself, but increased the coefficients of safety belt use and vehicle age.

Elimination of five deviant observations related to vehicle age reduced the variable's effect on driver's relative risk and further influenced the coefficients of alcohol use and the use of safety belt.

When the two observations with the most leverage related to tyre type were omitted from the estimation of the model the variable's coefficient increased and the relative risk increased from 1.27 to 1.43. The coefficient of tyre types was also affected by the elimination of observations with leverage related to annual vehicle kilometreage and driver sex.

Four observations with leverage related to speed limit (NRAJA) were eliminated from the estimation. This elimination increased the impact of speed limit on driver's relative risk and at the same time, it affected the coefficients of driver's annual vehicle kilometreage.

Table 26 shows the effects of removing variables on the coefficients and on relative risks. The same 446 observations were used in the comparisons. The main results are as follows.

Removal of annual vehicle kilometreage (VAKMI), driver sex (SPUOLI), use of alcohol (ALKOHOLI) and use of safety belt (TVYO) from the model caused significant changes, in the range of –15% to +34%, in the coefficients of the other variables. The removal of annual vehicle kilometreage (VAKMI) increased the coefficients of driver sex (SPUOLI) and vehicle age (AUTONI).

The removal of annual vehicle kilometreage resulted in an increased in female's relative risk in comparison to that of the male drivers, and the risk of old vehicles got closer to that of newer vehicles. The removal driver sex from the model effectively reduced the effect of annual kilometreage on relative risk.

The removal of alcohol use from the model decreased the relative risk of safety belt users by about 15%. When use of safety belt was removed from the model, the relative risk of use of alcohol increased by about 34%, while the changes in the other relative risks were small (–3% ...+4%). When both variables were omitted from the model, the coefficients of vehicle age and driver sex have changed.

The effects of tyre type and condition on the relative risk remained very stable when the variables were omitted. The effect of tyre type depended on the presence of at least one of the following variables: tyre's condition, use of alcohol, or use of safety belt. Tyre type had a greater effect on driver's relative risk when these variables were included in the model. The change was not just a reflection of the other variables' effects, but also a result of the model's improved explanatory power.

In summary, the alternate removal of variables from survival models based on VALT data resulted in changes in the coefficients of the following variables: use of alcohol, use of safety belt, driver sex, and vehicle age. Since the confidence intervals of the original coefficients were generally wide, the new coefficients stayed within the 95% confidence intervals of the full model. The coefficients of the same variables were also sensitive to elimination of deviant observations as determined by DfBeta-examinations.

Table 26. The relative risks and their 95% confidence intervals when eliminating different variables from model 1 of the VALT data.

The dark background indicates the greatest changes in the relative risks.

Variables	Model 1.1	Model 1.2 (No VAKMI)	Model 1.3 (No CKIKA2)	Model 1.4 (No KEVYTP)	Model 1.5 (No RENURA)	Model 1.6 (No SPUOLI)
VAKM(1) Annual vehicle kilometreage	1.000 *	–	1.000 *	1.000 *	1.000 *	1.000 *
VAKM(2) Annual vehicle kilometreage	0.537 0.378–0.763	–	0.502 (–6.5%) 0.354–0.712	0.547 (+1.9%) 0.386–0.776	0.553 (+3.0%) 0.390–0.786	**0.475** **(–11.5%)** 0.341–0.662
VAKM(3) Annual vehicle kilometreage	0.485 0.326–0.723	–	0.446 (–8.0%) 0.300–0.662	0.475 (–2.1%) 0.320–0.707	0.497 (+2.5) 0.333–0.742	**0.415** **(–14.4%)** 0.287–0.601
CKIKA2(1) Driver age	1.000 *	1.000 *	–	1.000 *	1.000 *	1.000 *
CKIKA2(2) Driver age	0.648 0.483–0.870	0.637 (–1.7%) 0.475–0.855	–	0.645 (–0.5%) 0.481–0.865	0.636 (–1.9%) 0.475–0.852	0.642 (–0.9%) 0.479–0.860
CKIKA2(3) Driver age	0.982 0.682–1.414	1.068 (+8.8%) 0.745–1.531	–	0.977 (–0.5%) 0.680–1.406	0.967 (–1.5%) 0.672–1.391	0.885 (–9.9%) 0.624–1.255
KEVYT Vehicle weight	1.201 0.920–1.567	1.214 (+1.1%) 0.935–1.576	1.208 (+0.6%) 0.927–1.575	–	1.225 (+2.0%) 0.938–1.599	1.253 (+4.3%) 0.964–1.630
RENURA Tyre's tread depth	0.942 0.888–1.000	0.949 (+0.7%) 0.894–1.008	0.938 (–0.4%) 0.884–0.995	0.940 (–0.2%) 0.886–0.997	–	0.939 (–0.3%) 0.885–0.996
SPUOLI Driver sex	1.401 1.019–1.925	**1.725** **(+23.1%)** 1.289–2.308	1.312 (–6.4%) 0.968–1.777	1.448 (+3.4%) 1.059–1.980	1.430 (+2.1%) 1.041–1.966	–
RENGAS Tyre type	1.268 0.801–2.007	1.272 (+0.3%) 0.803–2.015	1.278 (+0.8%) 0.807–2.023	1.244 (–1.9%) 0.787–1.968	1.274 (+.5%) 0.804–2.019	1.275 (+0.6%) 0.805–2.020
AUTONI Vehicle age	0.770 0.557–1.063	**0.847** **(+10.0%)** 0.614–1.171	0.780 (+1.3%) 0.565–1.077	0.774 (+0.5%) 0.561–1.069	0.838 (+8.8%) 0.614–1.146	0.757 (–1.7%) 0.549–1.043
ALKOHOLI Use of alcohol	2.079 1.286–3.360	1.973 (–5.1%) 1.222–3.185	1.989 (–4.3%) 1.236–3.201	2.002 (–3.7%) 1.241–3.231	2.193 (+5.5%) 1.360–3.535	2.010 (–3.3%) 1.248–3.237
TVYO Use of safety belt	0.611 0.439–0.849	0.614 (+0.5%) 0.443–0.853	0.600 (–1.8%) 0.432–0.832	0.594 (–2.8%) 0.428–0.825	0.607 (–0.7%) 0.436–0.844	0.627 (+2.6%) 0.452–0.869

*reference level

continued

Table 26. Continued.

Variables	Model 1.1	Model 1.2 (No RENGAS)	Model 1.3 (No AUTOY)	Model 1.4 (No ALKOHOLI)	Model 1.5 (No TVYO)	Model 1.6 (No ALKOHOLI. TVYO)
VAKM(1) Annual vehicle kilometreage	1.000 *	1.000 *	1.000 *	1.000 *	1.000 *	1.000 *
VAKM(2) Annual vehicle kilometreage	0.537 0.378–0.763	0.536 (–0.2%) 0.377–0.762	0.563 (+4.8%) 0.398–0.797	0.559 (+4.1%) 0.394–0.793	0.529 (–1.5%) 0.372–0.753	0.555 (+3.4%) 0.390–0.790
VAKM(3) Annual vehicle kilometreage	0.485 0.326–0.723	0.483 (–0.4%) 0.324–0.722	0.510 (+5.2%) 0.343–0.759	0.494 (+1.9%) 0.331–0.737	0.492 (+1.4%) 0.330–0.736	0.509 (+4.9%) 0.341–0.762
CKIKA2(1) Driver age	1.000 *	1.000 *	1.000 *	1.000 *	1.000 *	1.000 *
CKIKA2(2) Driver age	0.648 0.483–0.870	0.645 (–0.5%) 0.481–0.866	0.658 (+1.5%) 0.491–0.882	0.660 (+1.9%) 0.492–0.885	0.635 (–2.0%) 0.473–0.852	0.647 (–0.2%) 0.483–0.868
CKIKA2(3) Driver age	0.982 0.682–1.414	0.974 (–0.8%) 0.677–1.402	1.006 (+2.4%) 0.699–1.448	0.983 (+0.1%) 0.685–1.414	0.965 (–1.7%) 0.670–1.389	0.955 (–2.7%) 0.664–1.373
KEVYT Vehicle weight	1.201 0.920–1.567	1.192 (–0.7%) 0.913–1.555	1.195 (–0.5%) 0.915–1.560	1.153 (–4.0%) 0.883–1.505	1.249 (+4.0%) 0.959–1.626	1.224 (+1.9%) 0.939–1.596
RENURA Tyre's tread depth	0.942 0.888–1.000	0.942(+0.0%) 0.887–1.000	0.955 (+1.4%) 0.901–1.011	0.934 (–0.8%) 0.880–0.991	0.941 (–0.1%) 0.887–0.998	0.928 (–1.5) 0.874–0.985
SPUOLI Driver sex	1.401 1.019–1.925	1.405 (+0.3%) 1.021–1.932	1.424 (+1.6%) 1.036–1.958	1.364 (–2.6%) 0.994–1.873	1.363 (–2.7%) 0.999–1.876	1.281 (–8.6%) 0.932–1.760
RENGAS Tyre type	1.268 0.801–2.007	–	1.285 (+1.3%) 0.812–2.034	1.238 (–2.4%) 0.782–1.959	1.250 (–1.4%) 0.790–1.978	1.235 (–2.6%) 0.780–1.954
AUTONI Vehicle age	0.770 0.557–1.063	0.766 (–0.5%) 0.555–1.058	–	0.804 (+4.4%) 0.584–1.108	0.800 (+3.9%) 0.580–1.105	**0.879** **(+14.2%)** 0.641–1.206
ALKOHOLI Use of alcohol	2.079 1.286–3.360	2.062 (–0.8%) 1.275–3.335	1.993 (–4.1%) 1.237–3.211	–	**2.791** **(+34.2%)** 1.798–4.330	–
TVYO Use of safety belt	0.611 0.439–0.849	0.612 (+0.2%) 0.439–0.852	0.628 (+2.8%) 0.453–0.870	**0.517** **(–15.4%)** 0.383–0.697	–	–

*reference level

The correlations between the variables, a fact noted early in the process of compiling the models, may cause changes in coefficients when observations or variables are removed. Greene (1990) points out that even small changes in data can cause great changes in the coefficients of a model variables; the signs may be illogical or the values of the variables' coefficients can prove to be impossible. In the present evaluation, the variables behaved logically, continued to be in accordance with the expectations in their sign and value, and the standard errors of the variables remained moderate in size. Therefore, it was decided that all of the variables be included in the models. The analysis of the partial residuals confirmed that the models followed the requirement of proportionality.

In conclusion, the sensitivity testing of the of both postal survey and VALT data models, suggested that it was not necessary to eliminate observations or to alter the structure of the models. The models' actual problem turned out to be the strong correlations between the variables and their reflection in the models and, particularly, the interpretation of the correlative models. The explanatory power of the models is not very high suggesting a need to further improve the modelling methods.

6.6.3 Methodological issues in developing the accident survival models

The following issues had to be considered and resolved in the process of developing the survival models for the data in this study. Wherever possible, data-based testing and statistical analysis support the methodological choices.

- Has the chosen period (time slice) for the analysis (which was generally shorter than the exposure period of most drivers) caused any significant errors in estimations?
- Did the models consider adequately differences in exposure between drivers?
- Did survival models provide additional information that could not have been obtained with other methods, which do not depend on time?
- Could survival distributions be taken into account more reliably?
- How serious problem is the issue of multi-collinearity?
- How to deal with missing data?
- Can accidents of the same driver be assumed independent observations?
- How reliable and accurate were drivers' reports of accidents and exposure?
- What is the goodness of fit of the models?

Selecting the investigation period

A random sample data was generated so that the sample consisted of 4,350 drivers, 325 of whom were involved in winter time accidents. These data included randomly chosen values for the variables: accident time, driver age, driver sex, annual vehicle kilometreage, vehicle age and type of tyre. Random sample consisted of 57 % drivers in the first age class, 29 % in the second and 14 % in the last age class.

Drivers' accident times in the original records varied between 0–15 years or 0–5,400 days. In order to test the influence of the analysis period, three alternative models were built with different time-slices. Cox-type time models were built using the following variables: annual vehicle kilometreage driver age, sex, vehicle age and tyre type.

The time-slices were chosen so that they included the whole period, a period starting from the beginning of the whole period and a period somewhere during the whole long period.

Model 1: entire period, 15 years, beginning from moment $t = 0$; all 325 accident drivers are represented.

Model 2: shortened time period, 3,359 days, beginning moment $t = 0$; 200 accident drivers are represented in this time slice.

Model 3: 3,365 days, beginning from moment $t = 1,080$ days (from 0) and ending at $t = 4,445$.

Table 27 presents the comparison between the models.

The model for the entire period (model 1) gives driver age, vehicle age and driver sex as significant variables. In the shortened examination period (model 2), the same variables are indicated, in the same order, with somewhat different p values. In model 3, the significant variables are driver age, driver sex vehicle age.

It was concluded that in Cox-type survival models, a time-slice examination period selected carefully, could result in a model that is not significantly distorted. However, there are statistical differences between models especially when considering the effects of driver sex and vehicle age. The time slice model (model 3) does not capture these variables as statistically significant variables. This result supports the conclusion that explanatory power will be lost in survival models if random time slice data is used as a basis for modelling.

Impact of exposure time

Did the models consider adequately differences in exposure between drivers? This issue was examined by using different model types for survival time (Cox's model, Weibull-model and negative exponential distribution based model). The time dependent analysis should perhaps consider also time exposed in traffic and not only time in days. Modelling was based on the postal survey data.

Exposure time, (hours spent in traffic during winter time), and accident time (time in traffic up to the accident event), were calculated for each driver by dividing their individually reported, annual vehicle kilometreage (split between built-up areas and rural roads) by travel speed. Speed was assumed to be 35 km/h in built areas and 75 km/h on other roads. These calculations were repeated for each model. Table 28 presents a summary of the comparisons.

Table 27. Survival models with three time slices (models 1–3). Estimated parameters and their standard errors in (parenthesis.)

Variables Name	Value	Model 1	Model 2	Model 3
CKIKA2 Driver age	≤ 25	0.000 *	0.000 *	0.000 *
	26–50	-1.0500 (0.1270)	-1.2426 (0.1725)	-0.9958 (0.1630)
	> 50	-2.4066 (0.1648)	-2.2145 (0.1948)	-2.1860 (0.1985)
CTAKM Wintertime kilometreage	≤ 10 000 km	0.000 *	0.000 *	0.000 *
	10 000–20 000 km	0.2843 (0.1430)	0.3014 (0.1873)	0.0968 (0.1798)
	20 000–30 000 km	0.2190 (0.1440)	0.3544 (0.1845)	0.1212 (0.1778)
SPUOLI Driver sex	Male driver	0.000	0.000	0.000 *
	Female driver	-0.2014 (0.1114)	-0.3044 (0.1434)	-0.2753 (0.1422)
AUTONI Vehicle age	≤10 years	0.000 *	0.000 *	0.000 *
	>10 years	-0.2496 (0.1128)	-0.4282 (0.1473)	-0.2518 (0.1437)
RENGAS Tyre type	Studded tyre	0.000 *	0.000 *	0.000 *
	Non-studded winter tyre	-0.0701 (0.1113)	0.0651 (0.1419)	-0.1080 (0.1417)
Initial situation's -2log-likelihood		5420.8	3325.2	3358.5
Model's -2log-likelihood		5120.1	3141.5	3198.1
Model's degrees of freedom		7	7	7
Number of observations		4350	4350	4350

*Reference level

All the models were statistically significant. There were no significant differences in the coefficients of the variables in the models except for wintertime kilometreage. More specifically, the following was concluded from the comparisons of the models presented in Table 28:

- there were no great differences in the explanatory powers of Cox proportional survival model, models based on the time spent in traffic, and the logit model,
- the coefficient of wintertime kilometreage, especially in the high exposure category, was statistically significant in the model based on time in traffic (model 2),
- the explanatory power of the model based on the Weibull distribution was somewhat better than that of the model based on the negative exponential distribution and its shape parameter was $\beta = 1.573$, meaning that the hazard function was slowly increasing in relation to time.

Table 28. Different types of survival models, estimated parameters and standard errors (parenthesis) and goodness of fit.
Model 1: Cox proportional basic model, model 2: Cox proportional model of the time spent in traffic, model 3: the logit model, model 4: a survival model consistent with the Weibull distribution and model 5: a survival model based on the negative exponential distribution.

Variables	**Basic model (Cox model) (1)**	**Model of time spent in traffic (2)**	**Logit model (3)**	**Model of Weibull distribution (4)**	**Model of negative exponential distribution (5)**
CKIKA2 Driver age					
CKIKA2(1)	-0.0045 (0.1073)	-0.0012 (0.1074)	-0.0121 (0.1124)	-0.7379 (0.2239)	-0.7517 (0.2238)
CKIKA2(2)	-0.7716 0.1558	-0.7701 0.1554	-0.7994 0.1625	-0.9247 0.2482	-0.9385 0.2479
CTAKAM Wintertime kilometreage					
CTAKM(1)	0.1988 0.1401	0.1775 0.1401	0.2020 0.1463	0.2202 0.1410	0.2247 0.1408
CTAKM(2)	0.3395 0.3227	0.5100 0.1481	0.5875 0.1563	0.5976 0.1511	0.6086 0.1510
MARENG Tyre type	0.3395 0.3227	0.3448 0.3227	0.3602 0.3439	0.3222 0.3227	0.3214 0.3224
CKIKA2×SPUOLI Driver age & driver sex	0.0000	0.0000	0.0000	-0.1139 0.3096	-0.1210 0.3094
CKIKA2(1)×SPUOLI	0.4439 0.1559	0.4402 0.1558	0.4640 0.1632	0.4923 0.1647	0.5003 0.1646
CKIKA2(2)×SPUOLI	-0.1895 0.2573	-0.1931 0.2567	-0.1727 0.2676	0.0009 0.3398	0.0023 0.3389
CKIKA2×AUTONI Driver age & vehicle age	0.0000	0.0000	0.0000	0.7362 0.2671	0.7576 0.2670
CKIKA2(1)×AUTONI	-0.0451 0.1828	-0.0433 0.1828	-0.0516 0.1905	-0.0067 0.1945	-0.0060 0.1942
CKIKA2(2)×AUTONI	-0.6077 0.2618	-0.6017 0.2616	-0.6380 0.2781	-0.4489 0.3879	-0.4508 0.3871
Constant			-2.2235 0.1782	-6.473 0.2252	-8.899 0.2252
Initial situation's –2LL	4952.41	4952.19	2141.24	6087.3	60143.1
Final situation's –2LL	4875.04	4877.78	2064.97	5986.6	6061.5
Log-likelihood-change	77.36	74.40	76.27	100.7	81.6
Degrees of freedom	9	9	9	13	12
Significance	0.000	0.000	0.000	0.000	0.000
Number of observations	4099	4099	4099	4099	4099
Shape parameter β				1.53	1.00

The comparative analysis of the models suggests that basing the survival distribution on the negative exponential distribution is not necessarily the best solution. If another distribution is used, e.g. the Weibull distribution, methods need to be developed to account for two period distributions (time before the time-slice, and the time-slice). Cox proportional survival model is not necessarily more exact than the logistic model that does not depend on time. Each survival model would improve somewhat if better data on “time spent in traffic” were available.

The survival distribution does not necessarily follow the exponential distribution; a model based on the Weibull distribution was somewhat more exact than the model based on the negative exponential distribution. The hazard function derived with the negative exponential distribution was a constant but, according to the calculated Weibull model, it was slowly increasing. The data support an interpretation that drivers’ hazard function can be increasing during the reference period. However, there could be other reasons for this other than an increase in accident risk. For example, drivers’ accident recollection, which may have been better for the end (more recent) period than for it beginning.

Issue of multi-collinearity

The variables in the models were correlated with each other to various degrees. The so called multi-collinearity produced by correlations can cause problems in the estimation and interpretation of the models. The central problem of multi-collinearity is that parameters estimates or variable coefficients can not be estimated exactly, because the correlations increase the variables’ standard errors. The variables’ effects may be distorted and their coefficients may reflect the effects of other variables.

Some of the estimation problems observed in the study could be attributed, in part, to multi-collinearity. Small changes in the set of observations (following DfBeta-tests for deviant observations) caused some significant changes in parameter estimates. Many coefficients had a large standard error and minor statistical significance, although the model as a whole was statistically significant. However, none of the coefficients were of the wrong sign or unreasonably large.

How to deal with missing data

Missing data was a general problem. The amount of missing data was influenced by how drivers answered and how data were gathered. The amount of missing data varies from one variable to another but because of the small number of observations (particularly accident cases) its was not justified to remove all cases with some missing data. This resulted in different number of observations in each model. The models included all of the observations that did not have missing data concerning the variables included in the model.

Specific causes of missing data were drivers who had died in the accidents and could not be included in the postal survey. Their proportion in the driver registry, from which the sample was drawn, must be extremely small so that their obvious exclusion could not exert any significant bias. Some killed drivers were also omitted from the VALT data due to excessive missing data about them. Preliminary analyses showed that this case selection criterion reduced somewhat the number of primary involved parties in the database, and the absolute risk values. It did not affect the distributions of variables in the two groups of drivers (primary, and others).

Reliability and accuracy of reporting

Information given by drivers in a postal survey may have reliability problems such as, omission of answers, problems with recollection, answers that minimise driver's involvement or responsibility. Some analyses to logically check the reliability of the data were carried out.

The cumulative function of the number of wintertime accidents (accidents/month), calculated from the survey's data was compared to the random cumulative function of the number of accidents based on the Poisson process. The functions differed significantly (Kolmogorov–Smirnoff = 0.0775, $p<0.05$).

Figure 18 shows that the monthly number of reported wintertime accidents in the survey is increasing from the beginning of the study period.

A comparison between the theoretical and actual functions indicates that the reported function has a systematic growing trend: 59 involved drivers in January–March 1991, 70 in 1992 and 100 in 1993. Was this increase a true trend or, perhaps, a memory bias? In a previous roadside interview survey it was found that drivers tended to forget reporting accident data from the last 2–3 years (TVH 1980 and TVH 1982).

The observed trend of the survey was compared with the statistics compiled by the insurance companies. According to the insurance records (Figure 19), there was a slowly decreasing trend in wintertime accidents between 1991–1993. The same trend was also found in VALT data.

The accuracy of annual kilometreage reporting by drivers can be high. Joly et al. (1993) compared the previous week kilometreage reported the by 32 drivers in a phone survey, with the kilometreage calculated the following week on the basis of trip diaries. Drivers' estimates were, on the average, about 10% lower (and not statistically different) than the distance calculated with the help of trip diaries. The correlation coefficient between the estimated and measured kilometreage was 0.90. It is likely, however, that kilometreage estimates concerning long periods of time could be less accurate. To the extent that the inaccuracy is random, the average estimate for a sample may still be good.

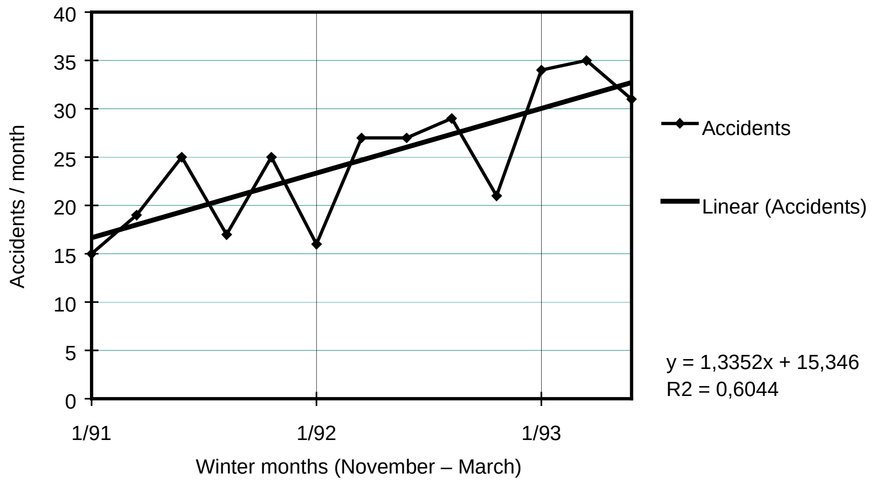

Figure 18. The monthly wintertime accidents in 1991–1993 reported in the postal survey.

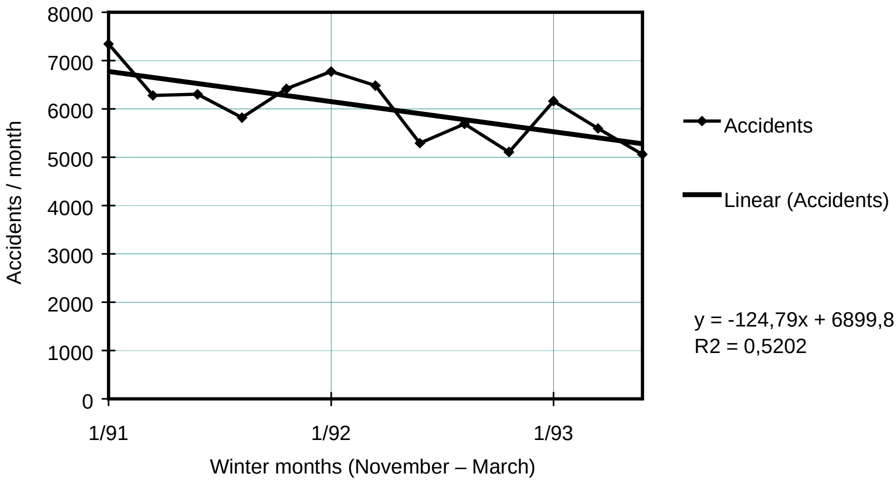

Figure 19. The monthly wintertime accidents in 1991–1993 according to insurance companies statistics.

Maycock reports that in postal survey a constant percentage memory loss effect from year to year is present. He concludes also that the older drivers are not having more difficulty than younger drivers in recalling the accidents as time passes, but the are not recalling the dates so well (Maycock et. al 1991).

Independence of observations

In calculating models, each driver's involvement in an accident was handled as a separate incident. It has been stated (e.g. Sheppard 1992) that an accident experience may affect drivers behaviour in some way, so that the probability of their next accident would not remain the same. On the other hand, Rajalin & Immonen (1993) concluded that for most drivers the personal impact of an accident is short-termed.

The postal survey sample was random including both accident and non-accident drivers. The models of VALT data were based on the concept of induced exposure. The VALT data was not random, included only drivers involved in fatal accidents (and who had not been killed themselves). Following the concept of "induced exposure", the accident group consisted of the "primary involved parties" and the exposure group consisted of "all other involved parties". The investigation teams' original decisions on the accidents' "primary involved parties" (main guilty party) determined, in part, the nature and fit of the models. It is possible that the effects of certain risk factors were emphasised too much, while the effects of other factors were not emphasised enough. An example of such factors could be the effect of alcohol and exceeding the speed limit.

Direct comparison of the accident records of the postal survey and VALT data files showed clear differences in the distributions of driver variables, mobility, vehicle characteristics and accident conditions. The survey data highlight ordinary drivers' common safety problems, such as those associated with driving in built-up areas, junctions and parking. The VALT data emphasised the factors associated with serious accidents on interurban roads and their background factors. In as strict sense, survival models of VALT data estimated the probability of being the "main guilty party" in a fatal accident and how various variables affect this probability. The actual parameters and values of the models may be less relevant to estimating drivers' general accident risks.

Goodness of fit of the models

Despite the differences in the basic sets of databases and the models' compilation principles, the essential outcomes of the models were quite similar. The effects of driver characteristics, such as the age and sex, were represented credibly and in accordance with earlier studies. An overall evaluation of the goodness of fit of the models certainly leaves room for improvement. The goodness of fit of these models can be evaluated by using maximum likelihood method but also by using a similar value as R-squared in normal regression, Pseudo R-squared index value. Pseudo R-squared values of the survival models varied between 0.24–0.30. For satisfactory modelling it is commonly considered that this index of quality should exceed 0.2.

7. SOLVING THE MAIN QUESTIONS

The approach for solving the four main questions that were put forward in chapter 3 (p. 35) included s the use a variety of research methods and data sources:

- Previous research results,
- Examinations of the postal survey results and VALT-bases accident data,
- Analysis with Kaplan–Meier method,
- The compiled accident models ,
- A separate case-control study and logit modelling that were done in association with this research. The results dealing with case-control study and logit modelling have been reported earlier (Roine 1996) and are included here only as controls

The four specified main questions (chapter 3) were:

1. Can driver involvement in wintertime accidents be examined on the basis of exposure over time; specifically, the amount and nature of wintertime mobility?

2. Do age and sex capture differences in accident risk at a general level, as background factors?

3. Do vehicle characteristics contribute to accident risk beyond their interaction with exposure and speed?

4. Does the use of studded tyres decrease driver wintertime accident risk, or does the use of non-studded tyres increase wintertime accident risk compared to use of studded tyres?

The methodological objectives were not formulated as main questions. They were discussed in the context of data analysis and interpretation.

Statistical hypothesis testing was used as a tool when defining the solutions to the four problems. The statistical decisions were based on survival models and the estimated parameters. The H_0 hypothesis is called null hypothesis and the other alternative H_1 is called alternative or research hypothesis. The purpose of the experiment is to decide whether the evidence tends to refute the null hypothesis. Since the research hypothesis is H_1, it is hoped that the evidence leads to reject H_0 and thereby accept H_1 (Milton & Arnold 1986).

When the values of the test statistics, here models and their estimated parameters, are in the critical region specified by α, the chosen level of significance, we reject H_0. Here we have specified α to be about 0.05 meaning that there is about 5 % risk to reject H_0 when it actually is true (Type I error). It is possible to make an other type of error (Type II error), to fail to reject H_0 when the alternative H_1 is true (Milton & Arnold 1986).

The next section first presents a summary of drivers' risk factors and relative risks based on the survival models. Then, each hypothesis is evaluated and finally, a summary is presented on the acceptance or rejection of the hypotheses.

7.1 Driver risk factors and relative risks

The probability of drivers to be involved in a wintertime accident was influenced by driver characteristics (age, gender, experience), their behaviour in traffic (speed, use of alcohol, use of safety belt), the amount and nature of exposure (annual and wintertime kilometreage, road surface conditions), and by vehicle and tyre characteristics (vehicle age, size, weight, tyre type and condition). The use of safety belt, which was only supposed to influence the seriousness of injuries, proved to be also a strong risk factor.

Survival models estimated the coefficients of the variables and their interactions which enabled calculation of relative risks associated with each factor. The time in the model referred to the number of winter-days until the (reported or investigated) accident.

The survival models of the survey data and of the VALT data gave quite similar views of many of the risk factors. Table 29 compares the interesting and important risk factors in the two sets of data. The largest differences were associated with the effects of driver age, annual vehicle kilometreage, and, to a less extent, sex.

The survival models derived from the two sources of data have, different conceptual interpretations, however. In the case of the survey data, the accident risk described by the modes is the probability of being involved in an accident (based on proportion of accident involved cases out of all exposed drivers). In the VALT-based accident data, survival models compare the "primary parties" and "other involved parties". The obtained results do not describe the accident risk, the probability of being involved in a fatal accident, but the probability of being an accident's primary involved party (main guilty party) on the condition that a fatal accident has occurred.

Table 29. The comparison of the interesting and important variables in the survival models of the survey and the VALT data.

Variable Postal survey	VALT data	Remarks
1 Driver age		
The relative risk was highest for the young and lowest for the old.	The relative risk was high for both the young and old.	
2 Sex		
Young males and females, or old males and females had similar risks. Middle aged had a higher relative risk than middle aged males.	Female drivers had a higher relative risk than male drivers.	Sex correlated strongly with exposure in the VALT data.
3 Use of alcohol		
No information in the data.	Significantly increased driver's relative risk.	Strong correlation with the use of safety belt.
4 Not using the safety belt		
No information in the data.	Significantly increased driver's relative risk.	Strong correlation with use of alcohol.
5 Share of driving in built-up areas and speed limit		
Large share associated with a higher relative risk.	Higher relative risk on roads with speed limit below 60 km/h.	
6 Familiarity of route		
No information in the data.	Drivers familiar with route had a lower relative risk.	
7 Relative speed		
No information in the data.	Exceeding the speed limit increased the relative risk; more so on high speed-limit roads.	
8 Wintertime kilometreage		
Relative risk increased with kilometreage.	No information in the data.	
9 Annual vehicle kilometreage		
Relative risk increased with annual vehicle kilometreage.	Relative risk decreased with increase in kilometreage.	There was a systematic difference between "primary" and "other involved parties".
10 Vehicle age		
Young drivers and old vehicles had a high relative risk. No vehicle effect for 25–50-year old drivers. Old drivers and old vehicles had a low relative risk.	Users of old vehicles had somewhat lower risk.	Not a statistically significant variable.
11 Vehicle weight		
No information in the data.	Users of light vehicles had a higher relative risk.	The impact of the vehicle weight was more important than that of traction type.
12 Tyre type		
Users of non-studded winter tyres had a higher relative risk than users of studded tyres.	Users of non-studded winter tyres had a higher relative risk than users of studded tyres. Users of old vehicles with non-studded tyres had the highest risk.	Not a statistically significant variable in the models.
13 Tyre's tread depth		
No information in the data.	Small tread depth increases relative risk.	Not statistically very significant variable.

7.2 Can driver involvement in wintertime accidents and exposure be examined with survival models?

The first main question was that can driver involvement in wintertime accidents be examined on the basis of exposure over time using survival models and is the amount and nature of wintertime exposure a major risk factor.

In chapter 5 (preliminary analysis) and chapter 6 (Kaplan–Meier estimates) we have seen that there are many factors having a statistically significant effect on the accident probability. Here we first test if we can find any significant predictor variables by using the survival models.

We test at the significance level α=0.05 as follows (chapter 6.1.3):

H_0: $b_1 = b_2 = \ldots = b_k = 0$ e.g. all parameters of the models are 0.

H_1: $b_i \neq 0$ for at least one i, $i = 1, 2\ldots, k$.

In order to get the answer to the question concerning the nature of wintertime exposure as a major risk factor we test at the significance level α=0.05 as follows:

H_0: the coefficient of driver wintertime kilometreage does not differ from zero in the models.

H_1: the coefficient of driver wintertime kilometreage differs from zero in the models.

In the first problem the hypothesis H_0 for all of the main survival models (chapter 6.4.2 table 17 and chapter 6.5.2 table 21) is rejected. The main significant factors found in chapters 5 and 6 are statistically significant here, too, and in the last test H_0 is also rejected., i.e. driver wintertime kilometreage is a statistically significant risk factor.

In the survey data exposure was described by wintertime kilometreage reported by the drivers and in VALT data by annual vehicle kilometreage. According to the models of the survey data accident probability increased with wintertime kilometreage (statistically significant effect, p-value $\leq$ 0.05). According to the VALT data models the probability to be primary party in a fatal accident decreased as the annual vehicle kilometreage increased.

The conflicting results appear to be caused by the differences between the sources of data. In the survey data, wintertime kilometreage indeed represented the length of exposure to risky conditions, and in terms of risk it favoured those who were exposed less. In the VALT data, all drivers were already involved in accidents and, therefore, annual vehicle kilometreage represented driving experience. It distinguished between those with little experience who got involved as “guilty parties”, and the more experienced whose involvement was as “others”.

It is likely that more detailed and accurate information on actual wintertime exposure might improve the models. The nature and quality of exposure is also relevant. This was indicated by the methodological analysis (see section 6.6.3).

About 73% of the accidents in the postal survey and 63% of the accidents in the VALT data had occurred on snowy, slushy or icy road surfaces. Young drivers and female drivers were particularly vulnerable in such conditions.

Relative risks, influencing the baseline accident risk, were higher in built-up areas than outside built-up areas, reflecting the nature of traffic environment as well as the characteristics of drivers and vehicles on different types of roads. The share of driving in built-up areas reflected systematic differences between drivers in terms of amount and nature of exposure (Roine 1996).

Driver involvement in wintertime accidents can be examined with survival models on the based on exposure over time. The amount and nature of wintertime mobility are specific risk factors. With better exposure data more accurate models can be specified.

7.3 Are driver age and sex major risk factors?

The second main question of the study was that driver age and sex are major background risk factors for differences in accident probability. Several other factors are expressed through sex and age.

We test the nature of driver age and sex as major risk factors at the significance level α=0.05 as follows:

H_0: the coefficients of driver age do not differ from zero in the models.

H_1: the coefficients of driver age differ from zero in the models.

The coefficients of driver age differ from zero in the main models (chapter 6.4.2 table 17 and chapter 6.5.2 table 21) for survey and VALT data. H_0 is rejected., i.e. driver age is statistically significant risk factor.

H_0: the coefficients of driver sex do not differ from zero in the models.

H_1: the coefficients of driver sex differ from zero in the models.

In the postal survey data models the coefficients of driver sex differ from zero in the main models. The effect of sex is mainly contributed by middle aged drivers (chapter 6.4.2 table 17). In the VALT data models (chapter 6.5.2 table 21) the coefficients of driver sex differ from zero only partly for the main models (model 1 and 4). We conclude that H_0 can only be rejected for the survey data models.

In the postal survey, the average wintertime accident rate by age was in agreement with previous studies. Young drivers had the highest rate, which decreased for the older age groups. The rate increased somewhat for the very old drivers. Also in the VALT data models, the relative risk decreased with age, but went up again in the oldest driver group to a level higher than that of middle age.

It is usually difficult to separate the effects of driver age and driving experience due to the strong correlation between them. The above results suggest that as the driver age and driving experience increase, the risk of mild accidents decreases, but the risk factors for the most serious accidents accumulate with both young and old drivers (Elvik & Vaa 1990). In survival models of the survey data, dealing with all accidents, driver age and experience were closely linked. In VALT data models, dealing with fatal accidents, experience was better represented by annual vehicle kilometreage and less linked to age.

The relative risk of middle-age females, in the postal survey, was larger than that of male drivers. In the VALT data, females had a higher relative risk across all age groups. A case-control analysis of the survey data confirmed the difference between female and male drivers in relative risk, even when wintertime accident exposure was controlled. A logit model also resulted in the same conclusion, although the results were statistically weaker.

Logit models of the VALT data showed difference in the relative risk between female and male drivers, when driver age and annual vehicle kilometreage were kept constant (Roine 1996).

Previous studies demonstrated that differences in risks between the sexes could be explained by differences in mobility. Some recent results suggest additional explanatory factors (e.g. Massie et al. 1994; Ernvall & Pirtala 1992; Mannering 1993).

In the survey data, female drivers were younger than male drivers, drove fewer kilometres, both during the winter and during the entire year, and drove a greater share of their kilometreage in built-up areas. In the VALT data female drivers drove fewer kilometres, were less often under the influence of alcohol, used safety belt more often than male drivers. Female drivers had fewer accidents and committed fewer traffic offences during the last five years.

The analysis of the VALT data indicated that driving under the influence of alcohol doubled driver's relative risk. Almost all drivers, in the VALT sample, who drove under the influence, were also categorised as the accident's main guilty parties. There may have been an obvious bias in the categorisation, however.

According to other previous studies, driving under the influence of alcohol has a significant effect on driver's accident risk (e.g. Elvik & Vaa 1990, Evans 1991, Häkkinen & Luoma 1990). The use of alcohol is often associated with young drivers' lifestyle and other risk factors related to it (Rolls & Ingham 1992, Lynam & Twisk 1995, Rajalin et al. 1989). Young drivers' high relative risk in the postal survey data may have been related to alcohol use but there was no data on alcohol in that survey.

Older drivers may also be thought of as having a “deviant” life-style in terms of mobility and driving behaviour. They were driving less, using vehicles with appropriate winter tyres, less likely to be drinking and speeding, and more likely to wear a safety belt. All these are risk reducing behaviours. Previous studies have stated that old drivers’ driving style appears to compensate for their general deteriorating skills (Holopainen 1994).

To sum up, the statistical testing and other available information support the conclusion that driver age is a major risk factor. The role of sex as a risk factor is less conclusive, as in many of the models the effect of driver sex was minimised when other factors, correlated with it, where considered in the models. These included amount and nature of mobility, specific driving experience, access to different types of vehicles and certain behaviours. However, to the extent that some strong risk factors, such as drink and drive, entail a choice made more often by males or females, the sex of drivers may remain a practically useful explanatory variable.

7.4 Do vehicle characteristics contribute to accident risk?

The third main question of the study was that specific characteristics of vehicles can contribute to accident risk probabilities. In the postal survey vehicle age was the main characteristic; in the VALT data models the variables were vehicle age and weight.

We test the nature of vehicle age and weight as major risk factors at the significance level $\alpha=0.05$ as follows:

H_0: the coefficients of vehicle age do not differ from zero in the models.

H_1: the coefficients of vehicle age differ from zero in the models.

In the postal survey data models the coefficients of vehicle age differ from zero in the main models. The effect of vehicle age is mainly contributed by young drivers (chapter 6.4.2 table 17). In the VALT data models (chapter 6.5.2 table 21) the coefficients of vehicle age do not differ from zero. We conclude that H_0 can only be rejected for the survey data models.

Then for vehicle weight, only VALT data included this variable:

H_0: the coefficients of vehicle weight do not differ from zero in the models.

H_1: the coefficients of vehicle weight differ from zero in the models.

In the VALT data models (chapter 6.5.2 table 21) the coefficients of vehicle weight do not differ from zero. We conclude that H_0 cannot be rejected for VALT data models.

In the survey models, old vehicles and young drivers was a high risk combination, whereas old vehicles and old drivers was a low risk combination. In the VALT data models, vehicle age was not a clear risk factor, perhaps because of the strong correlations with alcohol and safety belt use. Users of old vehicles included many alcohol users, young drivers, and non-users of belts.

In both data sets, young drivers tended to drive older vehicles. Drivers of newer vehicles were more likely to be involved in speeding prior to the accident. Drivers of lightweight vehicles were more likely to be the guilty party in a fatal accident in the VALT database. Vehicle characteristics were highly correlated with driver age and sex, and with vehicle kilometreage.

The case-control analysis, tabulations, and Kaplan–Meier estimation indicated that vehicle weight was more important to driver's accident risk and accident time than was vehicle's traction (front or rear drive).

The analysis of the third main question of the study pointed out that the features of a vehicle, its accessories, the manner of driving and the nature of exposure are all interrelated. However, some vehicle characteristics can be separated into their own individual risk factors but they seemed no to have strong explanatory power in these survival models.

7.5 Do studded tyres reduce drivers wintertime accident risk?

The fourth main question of the study was that does the use of studded tyres reduces driver wintertime accident risk ?

We use studded tyres as basis in the tests. We test the nature of use of non-studded tyres as major risk factor at the significance level $\alpha=0.05$ as follows:

H_0: the coefficients of use of no-studded tyres do not differ from zero in the models.

H_1: the coefficients of use of non-studded tyres differ from zero in the models.

In the postal survey data and in the VALT data models the coefficients of use of non-studded tyres do not differ from zero in the main models (chapter 6.4.2 table 17 and chapter 6.5.2 table 21). We conclude that H_0 cannot be rejected.

In the postal survey survival models, the relative risk of drivers who used non-studded winter tyres, was 1.4 times (95% confidence interval: 0.75–2.64) that of the drivers who used studded tyres. A case-control and cohort analyses, controlling for driver age and wintertime kilometreage (Roine 1996) gave similar results. So did logit models of the survey data. (see Table 30 for a comparison between the methods).

VALT data survival models, estimated the relative risk of driving with non-studded tyres, controlling for tyre condition, as 1.27 times (95% confidence interval: 0.75–2.01) that of driving with studded tyres (Table 30). (To be precise, the relative risk here refers to the probability of being an accident's primary involved party given that a fatal accident has occurred).

The variable that described tyre condition in the VALT data was the worst tyre's tread depth. The models pointed out that the smaller tread depth was, the higher was the relative risk. This variable was included in the models, but it did not have a large contribution to the explanatory power of the models after vehicle age was included in the same models. Nevertheless, tyre condition appears to be a better explanatory factor than vehicle age, as such. Poor tyre condition is more likely, of course, with old vehicles. Data on tyre condition were not available in the postal survey. Vehicle age served, indirectly, as a proxy for it.

In the survey data models the parameter estimate for use of non-studded tyres is about 0.35 and standard error about 0.32 (chapter 6.4, table 17). If we describe with μ the expected value of the parameter and the estimate as X, we can calculate how great should the value of μ be in order to get statistically significant proof with 95 % probability:

$P(X/0.32 > .645) = 0.95$, and

$P(X > 0.5264) = 0.95$, then

$P(X > 0.5264 |\mu) = P((X - \mu)/0.32 > (0.5264 - \mu)/0.32)) = 0.95$, we get

$(0.5264 - \mu)/0.32 = -1.645$, (where -1.645 = one-sided test, normal distribution, probability 95 %), and

$\mu = 1.0528$ and the relative risk should be $e^{1.0528} = 2.87$, when calculated by the survey data it is only about $e^{.0.32} = 1.38$.

If we adopt as μ = the estimated value by the models = 0.35 we can calculate that the probability is then 0.30 meaning that there is about 30 % chance to get statistically significant proof with 95 % probability.

In summary, according to all the estimated models of both the postal survey and the VALT data, driving with non-studded (and worn out) winter tyres was consistently more risky than driving with studded winter tyres in good condition. However, despite the consistency of the results the differences were not statistically significant. Therefore, the fourth main question of the study cannot be solved because of statistical grounds. However, we can conclude as showed above that with this data including only small amount of non-studded drivers, this statistical result is strongly expected.

Table 30. The effect of tyre type on driver's relative accident risk, according to different models applied to two data sets.

RESEARCH DATA **Research method** Other variables taken into consideration		Effect of tyre type 95% confidence interval (lower limit–***expected value***–upper limit)	Amount of observ-ations	Statistical significance (p)
POSTAL SURVEY DATA				
Case-control				
All drivers		0,76–***1,43***–2,69	4356	0,271
Wintertime kilometreage	≤ 8000	0,75–***1,65***–3,66	2894	0,212
	> 8000	0,39–***1,12***–3,19	1462	0,835
Driver age ≤ 30	≤ 8 000 km/winter	0,54–***2,00***–7,30	446	0,291
	> 8 000 km/winter	0,58–***2,26***–8,81	348	0,229
Driver age >30	≤ 8 000 km/winter	0,48–***1,36***–3,81	2434	0,561
	> 8 000 km/winter	0,06–***0,44***–3,31	1106	0,416
Logit model				
Driver age, wintertime kilometreage		0,73–***1,39***–2,64	4333	p > 0,05
Driver age, wintertime kilometreage and vehicle age			4333	
	vehicle age ≤10	0,90–***1,93***–4,13		p > 0,05
	vehicle age > 10	0,13–***0,53***–2,27		p > 0,05
VALT DATA				
Case-control, Basic tables				
All drivers		0,88–***1,81***–3,73	596	0,102
Tyre's tread depth	≤ 4 mm	0,86–***1,95***–7,33	125	0,318
	> 4 mm	0,71–***1,78***–4,45	454	0,212
Road surface condition				
not winter road surface conditions		0,76–***1,69***–3,80	408	0,196
winter road surface conditions		0,45–***2,30***–11,75	167	0,304
Road surface conditions and tyre's tread depth				
Not winter road surface conditions	≤ 4 mm	0,18–***1,94***–21,12	32	0,581
	> 4 mm	can't be estimated	131	
Winter road surface conditions	≤ 4 mm	0,39–***1,96***–9,78	90	0,403
	> 4 mm	0,55–***1,44***–3,76	306	0,455
Case-control, Model data				
All drivers		0,96–***2,33***–5,62	446	0,054
Tyre's tread depth	≤ 4 mm	0,41–***2,07***–10,43	92	0,368
	> 4 mm	0,74–***2,15***–6,25	349	0,149
Road surface conditions				
Not winter surface conditions		0,42–***3,86***–35,62	117	0,202
Winter road surface conditions		0,76–***1,996***–5,25	314	0,154
Road surface conditions and tyre's tread depth				
Not winter road surface conditions	≤ 4 mm	0,11–***1,33***–17,28	23	0,825
	> 4 mm	can't be estimated	94	
Winter road surface conditions	≤ 4 mm	0,32–***2,85***–25,37	66	0,329
	> 4 mm	0,55–***1,66***–5,03	244	0,362

continued

Table 30. Continued.

RESEARCH DATA **Research method** Other variables taken into consideration	Effect of tyre type 95% confidence interval (lower limit–***expected value***–upper limit)	Amount of observ-ations	Statistical significance (p)
Logit models, Basic tables			
Driver age in two classes, annual vehicle kilometreage, tyre condition	0,75–***1,89***–4,74	465	p >0,05
Driver age in three classes, annual vehicle kilometreage	0,237–***2,04–***4,79	478	p >0,05
Logit models, Model data			
Driver age in two classes, annual vehicle kilometreage, and tyre condition . Observations with leverage eliminated from model	0,64–***1,64***–4,21	446	p >0,05
Driver age in three classes, annual vehicle kilometreage. Observations with leverage eliminated from model	0,81–***2,02***–5,02	446	p >0,05
POSTAL SURVEY DATA			
Survival models			
In models: Driver age, sex, wintertime kilometreage, vehicle age and correlations			
<u>Cox model</u>	0,75–***1,40***–2,64	4099	0,293
VALT DATA			
Survival models			
In models: Driver age, sex, annual vehicle kilometreage, alcohol, safety belt, vehicle weight, tyre's tread depth, vehicle age and correlations			
<u>Cox model</u>	0,80–***1,27***–2,01	446	0,311
In models: Driver age, sex, annual vehicle kilometreage, alcohol, safety belt, vehicle weight, tyre's tread depth, vehicle age, route familiarity, speed limit and correlations			
<u>Cox model</u>	0,72–***1,24***–2,15	407	0,443
In models: Driver age, sex, annual vehicle kilometreage, alcohol, safety belt, tyre's tread depth, vehicle age, tyre type and correlations			
<u>Cox model</u>		463	
summer tyres	0,92–***1,68***–3,07		0,09
non-studded winter tyres	0,82–***1,31***–2,07		0,26
In models: Driver age, sex, annual vehicle kilometreage, alcohol, safety belt, tyre's tread depth, vehicle age			
<u>Cox model:</u>		446	
Studded tyres in other road surface conditions	0,90–***1,21***–1,63		0,21
Non-studded winter tyres in other road surface conditions	0,61–***1,74***–4,95		0,30
Non-studded winter tyres in winter road surface conditions	0,80–***1,40***–2,44		0,24

8. SUMMARY AND CONCLUSIONS

8.1 Objectives and general research approach

The objective of the study was to clarify which variables best explain drivers' involvement in wintertime accidents and to demonstrate how survival modelling can contribute to understanding of the risks of winter driving. The theory of hierarchical systems provided the study's conceptual framework. Wintertime accidents were viewed as an end result of process in which the characteristics of drivers interact with vehicle characteristics, with roadway conditions and with the amount and nature of travel. This process produces safe mobility as well as wintertime accidents. A specific objective of the study was to estimate the impact of driving with non-studded winter tyres on accident risk.

The variation in accidents can be divided into systematic and random variation. Accident models were developed to explain the systematic variation, which was described by regression functions estimated from actual data. The random variation around the estimated expected value, is usually determined by the characteristics of the probability function describing the accident phenomenon. However, in some types of models, e.g. Cox models, a distribution assumption is not necessary.

The multivariate modelling techniques applied here were of the disaggregated type. Observations were related, therefore, to individual drivers: their characteristics, the vehicles they used, how much they drove and what accidents they were involved in.

The research approach can be considered as an epidemiological one concentrating on risk analysis. Drivers' probability of being involved in an accident was modelled in terms of the direct effects of various conditions and factors (risk factors) and their interactions.

8.2 Research method

The specific modelling technique used here is survival modelling. Survival models attempt to explain which factors influence the temporal appearance of an exceptional occurrence such as a defect in a device, the emergence of a disease or the occurrence of an accident. In this study, survival models (accident time models) examined the involvement of drivers in wintertime accidents. Factors that influence the involvement were expressed here in terms accident probability or accident risk. The risk factors are not thought to be the accidents' direct causes, but they describe statistically the effects of these variables on drivers' accident probability.

The hazard functions of the models represent drivers' conditional accident risk as a function of time. In other words, how likely is it for a driver to be involved in an accident at time *t*, given that the driver has survived to that time. Similarly, the survival function represents the proportion of drivers who have not been involved in accidents, as a function of time.

The data for modelling came from two sources. One was a questionnaire survey sent to 10,000 vehicle owners (59% response rate), who were selected at random from the Motor Vehicles Registry and provided information about drivers' basic characteristics and driving experiences in the years 1991–93. Although preliminary estimates suggested that the sample size needed would be 42,000 cases, practical economic consideration limited it to the above size. After a process of data cleaning and removing non-usable cases, 4,352 driver questionnaires were used in the final analysis.

The second source of data were wintertime fatal accidents investigated in-depth and documented by special teams during 1987–91 (VALT data). All passenger-car accidents, involving at least two cars, were selected. The five year period was chosen to insure relative homogeneity in traffic conditions across the years. The sample consisted of 658 accident cases.

Several of the variables in the two data sources were identical but some were unique to each source. Both the survey data and the VALT data contained information about the driver age and sex, annual vehicle kilometreage, vehicle age, vehicle weight and engine size, and tyre type. The survey data also included: driver's employment status and wintertime kilometreage. The VALT data also included: tyre condition, vehicle traction (drive) type, road surface conditions, speed limit at site of the accident, vehicle speed, use of alcohol, use of safety belt, route familiarity, driving licence duration, and recorded accidents and violations.

The probability of drivers being involved in an accident was estimated by comparing drivers who had been involved, along with their characteristics, to drivers who had not been involved in accidents. The survey provided data on drivers who were, as well as on drivers who were not, involved in traffic accidents during the reference period. However, in the fatal accident data-base, VALT data, all the drivers were involved in accidents. The method of "induced exposure" was applied to these data (distinguishing so called "guilty drivers" from other involved drivers) in order to have a reference group for relative risk estimates.

Wintertime accidents were defined as those that occurred from the beginning of November to the end of March. Thus the examination period in the survey consisted of 390 winter days (winter periods 1991, 1992 and the beginning of 1993 (January–March). In the VALT data, winter period consisted 760 days in 1987–1991. Driver's survival time or accident time was counted from the beginning of the examination period.

8.3 Survival models

The most successful survival models were based on Cox type model (see section 6). A base-line survival or hazard function is first formed based on the original data, and then the relative effects of the variables and their interactions are estimated from the base-line function. The relative effects are assumed multiplicative. The Cox model does not require a distribution assumption concerning survival, because the basic hazard is formed with a distribution-free method. There is a cost to that, however. The likelihood of verifying the statistical hypotheses can be reduced, as is typical with using distribution-free test methods.

In Cox proportional survival models the effects of variables have to be proportional, when the variables are directly included in the models. The correctness of these assumptions was tested during the compilation and evaluation of the models and no substantial deviations were found. The variables could then be used together with other variables in a joint model.

The following issues had to be considered and resolved in the process of developing the survival models for the data in this study. Wherever possible, data-based testing and statistical analysis support the methodological choices.

- Has the chosen period (time slice) for the analysis (which was generally shorter than the exposure period of most drivers) caused any significant errors in estimations?
- Did the models consider adequately differences in exposure between drivers?
- Did survival models provide additional information that could not have been obtained with other methods, which do not depend on time?
- Could survival distributions be taken into account more reliably?
- How serious problem is the issue of multi-collinearity?
- How to deal with missing data?
- Can accidents of the same driver be assumed independent observations?
- How reliable and accurate were drivers' reports of accidents and exposure?
- What is the goodness of fit of the models?

These issues were examined by testing random samples of data with appropriate assumptions and comparing the performance of the various models. It was concluded that Cox type distribution-free survival models with a time-slice examination period, selected carefully, could result in a model that is not significantly distorted.

Because of limitations of exposure data, time-slice models did not provide major advantage over time independent models, such as logit models. When exposure was defined in terms of "hours of wintertime spent in traffic" the models improved. This means that better (more detailed, more relevant) exposure data could improve all models. Typical multi-collinearity problems were encountered. Small changes in the set of observations caused changes in the coefficients of variables and their standard errors. The problems were especially pronounced with the following variables: driver age, kilometreage, vehicle age, sex, use of alcohol and safety-belt. These problems were analysed and taken into account in the models and in the interpretation of the results.

In the postal survey, the extent of missing data varied from one variable to another, but much of it was not systematic. Consequently, the number of observation in the models varies, depending on the variables used. In VALT data information was missing about killed drivers, and those cases had to be selected out.

The implications of missing data and case selection were analysed in detail and it was determined that they had no significant effect on the interpretation of the results.

Few drivers had been involved in more than one accident during the reference periods. These accidents were treated as independent events, although it is possible that they were not. Further research is needed on this issue.

Reliability tests indicated that drivers in the survey may have had either difficulties to remember all their accidents during the three year period or they did not want to report all of them. The accuracy of reported kilometreage was in line with average group or annual data in other studies. However, we do not know how accurate were individual reports.

Despite the differences in the two sets of databases and the models' compilation principles, the essential outcomes of the models were quite similar. The effects of driver characteristics, such as the age and sex, were represented credibly and in accordance with earlier studies. Pseudo R-squared values of the survival models varied between 0.24–0.30. For satisfactory modelling it is commonly considered that this index of quality should exceed 0.2. Certainly, there is room for improvement in the models.

8.4 Wintertime risk factors and their effects

Each data source included many variables that were examined during the process of modelling. Many are interrelated. In practice, only a smaller number of variables and their interactions entered the models as significant factors contributing independently to the hazard and survival functions. The first three variables listed below are common to main models based on either the survey data, or the VALT data. The next ten variables are unique to either one or the other group of models, depending on the availability of the variable in the data source and its significance (variables marked with * are common):

- *driver age,
- *driver sex,
- *tyre type,
- *vehicle age,
- wintertime vehicle kilometreage,
- annual vehicle kilometreage,
- being under the influence of alcohol,

- use of safety belt,
- route familiarity,
- vehicle weight,
- speed limit.

A number of other variables were included in the models, despite their not necessarily being statistically significant, because of their consistent and logical relationship with other factors, thus helping interpreting the models.

The model that best described the data of the postal survey was a Cox-type model (Table 17, p. 71). The main risk factors were driver age, wintertime kilometreage, interaction between driver age and sex and interaction between driver age and vehicle age and tyre type. VALT data models (Table 21, p. 84) included in addition to the above the variables driving under the influence of alcohol, use of safety belt, vehicle weight and tyre condition.

Figure 20 shows the relative risks and their 95% confidence intervals for the variables in the two main survival models (model 6 for the postal survey data and model 1 for the VALT data). The relative risk is the ratio of a variable's coefficient to the reference coefficient.

Tables 31 and 32 show how each relative risk factor expressed in the survival function, contributes an additional share (measured in %) of accidents relative to the base level (average variable values of the other factors.) Table 31 refers to the postal survey model (model 6) and Table 32 to the VALT data model (model 1).

Table 31 shows that driver age has a major effect on the share of accident drivers. The added mean share of young drivers was 7% (alternatively stated, if all drivers were young, and other factors held constant, there would have been 14% accident drivers). On the other hand, the added mean share of old drivers is –2%.

Table 32 shows that the share of primary parties (accident drivers) in VALT data was manly influenced by drivers' annual average kilometreage, the use of alcohol, the use of safety belt and tyre's tread depth. Use of alcohol, for example, added 26% to the mean share of primary parties (accident drivers), whereas using tyres in good condition reduced the mean share of primary parties by 9%.

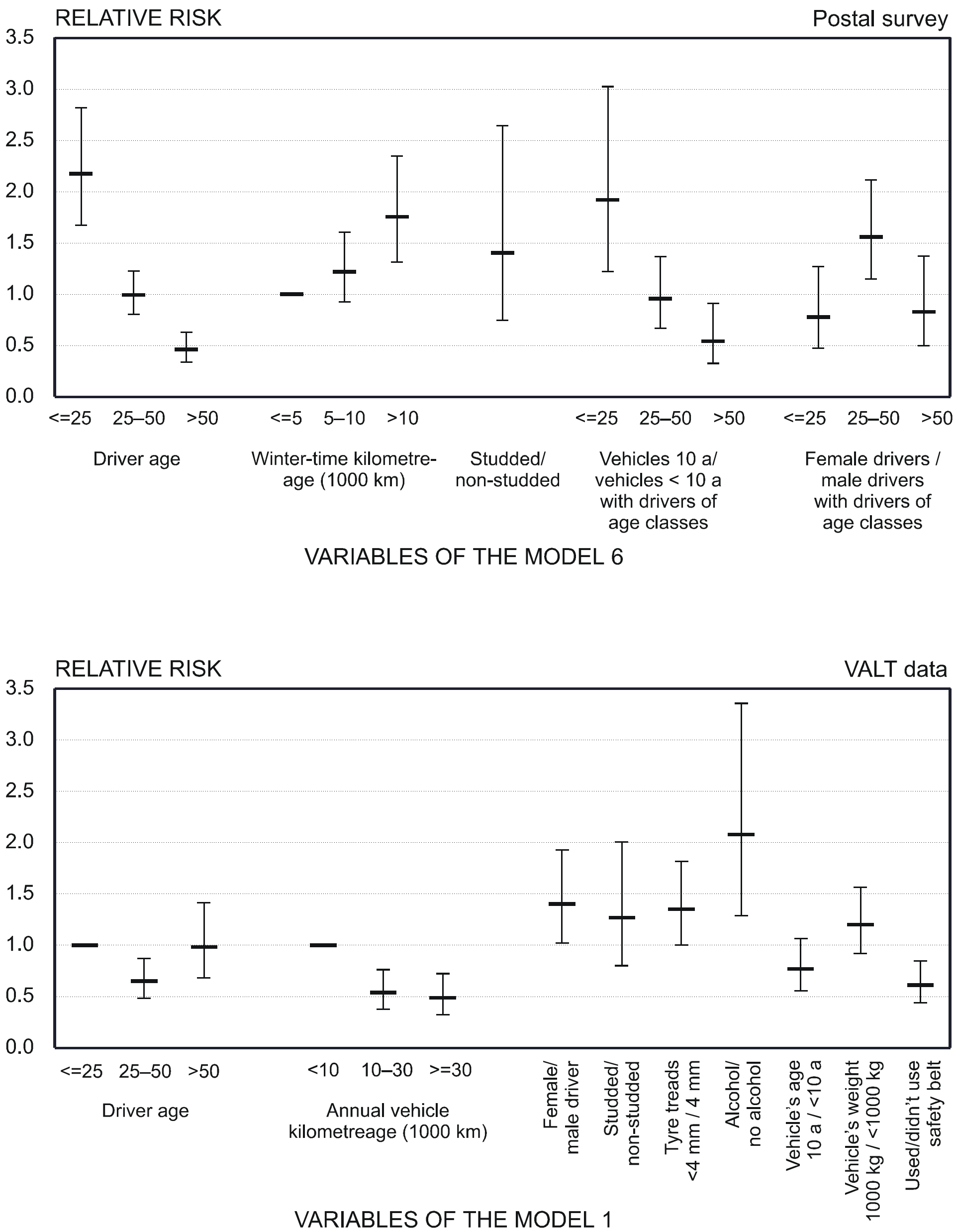

Figure 20. Relative risks and their 95% confidence intervals according to the postal survey and VALT data (model 6 and model 1).

Table 31. Share of accident drivers contributed with the explanatory variables of survival model 6, of survey data.

Variable	Variable category and the share of accident drivers					
	Category	Share	Category	Share	Category	Share
Driver age	≤ 25 years	14%	26–50 years	7%	> 50 years	5%
Wintertime kilometreage	≤ 5 000 km	5%	5 001–10 000 km	7%	> 10 000 km	9%
Sex	Female	7%	Male	6%		
Vehicle age	≤ 10 years	7%	> 10 years	6%		
Tyre type	Studded tyres	6%	Non-studded tyres	9%		

Table 32. Share of accident drivers (primary parties) contributed with the explanatory variables of survival model 1, of VALT data.

Variable	Variable category and the share of accident drivers					
	Category	Share	Category	Share	Category	Share
Driver age	≤ 25 years	65%	26–50 years	50%	> 50 years	65%
Annual vehicle kilometreage	≤ 10 000 km	78%	10 001–29 999 km	56%	> 30 000 km	49%
Sex	Female	66%	Male	53%		
Alcohol	No	55%	Yes	81%		
Vehicle age	≤ 10 years	58%	> 10 years	49%		
Vehicle weight	≤ 1 000 kg	61%	> 1 000 kg	55%		
Use of safety belt	No	71%	Yes	54%		
Tyre type	Studded tyres	57%	Non-studded tyres	65%		
Tyre treads	1 mm	69%	5 mm	60%	10 mm	57%

8.5 Solving the main questions

In addition to demonstrating and developing the application of survival models, four specific main questions about the accident risk of drivers in wintertime traffic were analysed with the help of statistical hypothesis testing.

The first main question was dealing with the amount and nature of wintertime mobility (or exposure) as a major (driver) risk factor in wintertime accidents. The second main question was dealing with driver age and sex as major background risk factors for differences in accident probability. The third main question was concerning with vehicle weight and age and other vehicle characteristics as separate risk factors. The fourth main question of the study was dealing with the effects of using studded tyres.

The stated main questions reflect a system theory view of wintertime accident probability as an interaction of mobility (or exposure), driver characteristics, vehicle characteristics, and the connection between the vehicle and the roadway, mediated by tyres.

The questions were solved with statistical hypothesis testing as a tool with survival models. These results demonstrated again the complex interactions between various risk factors. Exposure to wintertime driving had a clear impact on wintertime accidents. The larger the kilometreage, the higher the probability of being involved in an accident. Models for the VALT data suggested that specific inexperience in winter driving might also increase accident risk.

Young and old drivers are at higher risk during winter time driving. This may reflect lack of experience (and perhaps riskier driving style) of young drivers, and reduced capabilities of old drivers. Winter traffic conditions set greater demands on all drivers; the young and very old may have a harder time meeting the greater demands.

Driver sex did not have a clear impact on wintertime accident probability, once exposure and experience were taken into account. There were significant differences between female and male drivers in age distribution, mobility, driving experience, the type of cars used, and in driving related behaviour. Female drivers generally drove less than male and got involved in fewer accidents. However, many of the females who were involved in wintertime accidents, lacked driving experience in wintertime conditions.

Vehicle weight was related inversely to risk, and so was vehicle age. However, these variables were strongly related to driver and exposure factors, such that the unique effect of these vehicle factors was rather small and it depended on the exact combinations of other factors. For example, although it was generally safer for large and older vehicles, the most risky combination was of young drivers using old, and rear-wheel driven vehicles. Models based on the fatal accidents database also included a strong influence of driver behaviour – use of alcohol, non-use of safety belt, and excessive speed – all of which added to the probability of causing an accident.

The statistical hypothesis about the beneficial impact of studded tyres was consistent with the effects estimated in survival models (as well as with loglinear models and other examinations of the data). However, the difference between the relative risks of users and non-users of studded tyres was statistically insignificant. The difficulty in obtaining a strong confirmation of the hypothesis is because only small number of drivers in Finland used non-studded tyres in winter time, and non-use was related to other driver and exposure factors. In addition, tyre condition proved to be as important a factor as tyre type in determining the overall risk associated with tyres.

In conclusion, the application of survival models to the accident data appears to be a promising approach. The models apply well to the examination of risk factors. This research method should be developed further. Improving the models will require better information on drivers’ exposure times, more detailed information on the exposure and its quality, and further development of the statistical tools themselves.

8.6 Tyre type and condition in wintertime accidents

In the survey data, the number of drivers who reported using non-studded tyres was small. Therefore, although the findings about the risk of driving with non-studded tyres were consistent with expectations and with previous results, they were statistically non-significant. Due to the importance of the issue, more detailed examination was carried out on this variable and its relationships with other driver and vehicle characteristics.

Studies in the Nordic countries have generally reported reduced accident risk when driving with studded winter tyres or with the more recent type of "friction tyres" (Öberg et al. 1993; Carlsson & Öberg 1995 in Sweden; Ingebritsen & Fosser 1991; Fosser & Saetermo 1995, in Norway; Roine 1996, in Finland). According to the latest Swedish studies, the relative risk of using studded tyres, in icy and snowy conditions on rural roads, was 0.60 of that of using summer tyres. In built-up areas, the ratio was 0.65. Using non-studded winter tyres raised the relative risks to 0.75 and 0.80, respectively. In the present study, survival models of the VALT data, estimates the relative risk of studded tyres, on icy or snowy roads, at about 0.60 that of summer tyres,

Norwegian accident studies in the early 90's proved that the condition of tyres, especially their tread depth, is at least as important as having studs. A similar conclusion was derived here from the survival models of the VALT data. The Norwegian study estimated that a 1 mm decrease in the tyre's tread depth increased the accident probability by about 4%, which is totally consistent with the result obtained in the present study (3%–6%).

A Norwegian postal survey study (Fosser & Saetermo 1995) did not find significant differences in reported accidents by drivers using studded or non-studded winter tyres. A logistic regression analysis controlled for various other roadway, vehicle and driver characteristics. However, the direction of the difference was in favour of the studded tyres; they reduced the accident probability in snowy and icy road surface conditions. The same result was obtained in the present study. (However, according to the same Norwegian study, studded tyres increased the accident probability in non-icy or non-snowy road surface conditions.)

The present study, (as well as an earlier Finnish study, Heinijoki 1994) found that users of non-studded tyres tended to drive older vehicles with worn out tyres (sometimes worn out studs). They were more likely to have used alcohol prior to the accident, had more previous accidents in both wintertime and non wintertime conditions.

A number of inherent methodological problems make it difficult to estimate the effect of tyre type on relative accident risk. Studded tyres were introduced many years ago, so there is no clear time mark for a "before /after" type study.

The currently small proportion of users of non-studded tyres in Nordic countries create practical problems in compiling suitably large samples of drivers or of accidents, and a greater theoretical difficulty of interpreting results of a special, self-selected population.

Roads in wintertime are not always covered by snow or ice so that the relative advantage of studded tyres is in effect only part of the time. For example, only 30% of the time during the winter season of 1992–93 the country's main roads were covered by ice or snow (Alppivuori et al. 1995).

Tyre condition – such as tread depth, air pressure, balance, and match between all the tyres is important for their effective functioning. Perhaps some of the advantage of studded tyres (or friction winter tyres) is a function of their being in better condition – either because they are used less or because car owner are more particular about their condition.

There is little knowledge about how drivers' decisions regarding change of tyres to studded winter tyres are influenced by the condition of their summer tyres, the characteristics of their car, their travel plans, or their travel style. Similarly, we don't know enough how having one type of a tyre or another, influences peoples driving plans or driving behaviour.

Behaviour adaptation (e.g. Elvik & Vaa 1990; OECD 1990) can either increase or decrease the difference in the relative risks between using studded and non-studded tyres. Norwegian studies claim that increased speeds by drivers of vehicles with studded tyres most likely compensated for some of the safety benefits of studded tyres. A Finnish study (Mäkinen 1994) found that drivers driving vehicles with studded tyres travelled at lower speed on non-slippery roads than drivers with non-studded vehicles. It was noted that most of the time roads during the winter are not covered by ice, slush, or snow.

In conclusion, there is sufficient evidence that using studded tyres at least on icy or snow covered road surfaces is, on balance, beneficial to safety. The magnitude of the effect is less clear. It probably varies considerably according to regional and local winter, road and traffic conditions.

REFERENCES

Aitkin, M. & Anderson, D. 1989. Statistical modelling in GLIM. Oxford: Oxford Statistical Science Series. Clarendon Press. 374 p. ISBN 0-19-852204-5

Alppivuori, K., Kanner, H., Mäkelä, K. & Kallberg, V.-P. 1995. Nastarenkaiden käytön ja talvikunnossapidon yhteiskunnallinen optimointi. Helsinki: Tielaitos. (Tielaitoksen tutkimuksia 4/1995). 82 p. (In Finnish) ISBN 951-726-081-4

Björnskau, T. 1994. Spillteori, trafikk og ulykker: En teori om interaksjon i trafikken. (Game theory, road traffic and accidents: A theory of road user interaction). Oslo: Transportøkonomisk institutt. (TØI rapport 287/1994). 405 p. (In Norwegian) (ISBN 82-7133-927-3

Björnskau, T., Midtland, K. & Sagberg, F. 1994. Beskrivelse og dröftning av aktuelle modeller for bilföreres atferd. Oslo: Transportøkonomisk institutt. (Arbeidsdokument TST/0472/93). 70 p. (In Norwegian)

Broughton, J. 1988. The variation of car drivers' accident risk with age. Crowthorne: Transport and Road Research Laboratory. 27 p. ISSN 0266-5247

Bunday, B. D. 1991. Statistical methods in reliability theory and practice. London: Ellis Horwood limited. 252 p. ISBN 0-13-85397-6

Carlsson, A. & Öberg, G. 1995. Vinterdäck. Effekter av olika regelförslag. (Winter tyres. Effects of proposals for rules). Linköping: Väg- och transportforskningsinstitutet. 36 p. (VTI Meddelande Nr. 757). (In Swedish) ISSN 0347-6049

Central Bureau of Statistics 1992. Tieliikenneonnettomuudet 1991. Helsinki: Tilastokeskus. 49 p. (Liikenne 1992:16.). (In Finnish) ISSN 0785-6245

Cerrelli, E. 1972. Driver exposure the indirect approach for obtaining relative measures. Accident Analysis and Prevention, Vol. 5, pp. 147–156. ISSN 0001-4575

Chatfield, C. 1994. Problem solving. A statistician's guide. Second edition. London: Chapman & Hall. 317 p. ISBN 0-412-60630-5

Collet, D. 1991. Modelling Binary Data. London: Chapman & Hall. 369 p. ISBN 0-412-38800-6

Crowder, M. J. et al. 1991. Statistical analysis of reliability data. London: Chapman & Hall. 250 p. ISBN 0-412-30560-7

Elvik, R. & Vaa, T. 1990. Human factors, road accident data and information technology. Oslo: Institute of Transport Economics. (Report 0067/1990).155 p. ISBN 82-7133-674-6

Elvik, R. 1991. Ulykkesteori. Historisk utvikling og status i dag. Oslo: Transportøkonomisk institutt. (Metode-teori 0006/1991). 69 p. (In Norwegian) ISBN 82-7133-701-7

Ernvall, T. & Pirtala, P. 1992. Kuljettajan iän ja kokemuksen sekä automallin vaikutus onnettomuusriskiin. (The effect of driver's age and experience and car model on accident risk). Oulu: Oulun yliopisto. 48 p. (Tie- ja liikennetekniikan laboratorion julkaisuja 16). (In Finnish) ISBN 951-42-3245-3

Estlander, K. 1995. Sään ja kelin vaikutukset eri ajoneuvoryhmien nopeuksiin. (Effects of weather on driving speeds). Helsinki: Tielaitos. 91 p. + app. 16 p. (Tielaitoksen selvityksiä 23/1995). (In Finnish) ISBN 951-726-057-1

Evans, L. 1983. Accident involvement rate and car size. Michigan: General Motors, Research Laboratories. 36 p. (IRRD: GM Research Publication).

Evans, L. 1987. Fatal and severe crash involvement versus driver age and sex. New Orleans: 31st Annual Proceedings, American Association for Automotive Medicine. 18 p.

Evans, L. 1991. Traffic safety and the driver. New York: Van Nostrand Reinhold. 405 p. ISBN 0-422-00163-0

Evans, L. 1993. Driver injury and fatality risk in two-car crashes versus mass ratio inferred using newtonian mechanics. Accident Analysis and Prevention, Vol. 26, No. 5, pp. 609–616. ISSN 0001-4575

FinnRA 1992. Liikenneonnettomuudet yleisillä teillä 1991. Helsinki: Tielaitos. 63 p. (Tielaitoksen selvityksiä 6/1992.) (In Finnish) ISBN 951-47-6965-1

FinnRA 1993. Henkilöliikennetutkimus 1992. Helsinki: Tielaitos. 81 p. (Tielaitoksen selvityksiä 58/1993.) (In Finnish) ISBN 951-47-8104-X

Fosser, p. & Saetermo, I. 1995. Vinterdekk med eller uten pigger – betydning for trafikksikkerheten. (Winter tires with or without studs – The effect on traffic safety). Oslo: Transportøkonomisk institutt. 49 p. (In Norwegian) (Rapport 310/1995). ISBN 82-7133-954-0

Freedman, D. et al. 1991. Statistics. New York: W.W. Norton & Company, Inc. 514 p. ISBN 0-393-96043-9

Greene, W. H. 1990. Econometric analysis. New York: Macmillan Publishing Company. 791 p. ISBN 0-02-346391-0

Hatakka, M., Keskinen, E., Katila, A. & Laapotti S. 1995. Kuljettaja, ajoneuvo ja ajotapa: nuorten mieskuljettajien suhtautumistekijöiden yhteys automallin valintaan ja liikennevahinkoihin. (Driver, car and driving habits: the relation between young male drivers' attitudinal factors and choice of car and traffic accidents). Turku: Turun yliopisto. 42 p. (Psykologian tutkimuksia 99). (In Finnish) ISBN 951-29-0446-2

Hauer, E. 1997. Observational before-after studies in road safety. Estimating the effect of highway and traffic engineering measures on road safety. New York: Elsevier Science Ltd. 289 p. ISBN 0-08-0430538

Heinijoki, H. 1994. Talvi ja tieliikenne -projekti, Kelin kokemisen, rengaskunnon ja rengastyypin vaikutus nopeuskäyttäytymiseen. (Influence of the type and condition of tyres and drivers' perceptions of road condition on driving speed). Helsinki: Tielaitos. 99 p. (Tielaitoksen selvityksiä 19/1994). (In Finnish) ISBN 951-47-9098-7

Hemdorff, S., Leden, L., Sakshaug, K., Salusjärvi, M. & Schandersson, R. 1989. Trafiksäkerhet och vägytans egenskaper (TOVE). Slutrapport. (Traffic safety and the properties of road surfaces. Final report). Espoo: Valtion teknillinen tutkimuskeskus. 75 p. (VTT Tiedotteita 1075). (In Swedish) ISBN 951-38-3626-6

Hensher, D. A. & Mannering, F. L. 1993. Hazard-Based duration models and their application to transportation analysis. Working Paper ITS-WP-93-1. Sydney: Institute of Transport Studies. 32 p.

Holopainen, A. 1994. Iäkkäät henkilöauton kuljettajat kuolemaan johtaneissa onnettomuuksissa. Vakuutusyhtiöiden liikenneturvallisuustoimikunnan rahoittama apurahatutkimus. 33 p. (In Finnish)

Huhtala, M. & Kallberg, V.-P. 1978. Nastarenkaiden vaikutus liikenneturvallisuuteen, Onnettomuustutkimus. (The influence of the use of studded tyres on traffic safety. Accident investigation). Espoo: Valtion teknillinen tutkimuskeskus. 83 p. + app. 20 p. (VTT Tie- ja liikennelaboratorio, tiedonanto 33). (In Finnish) ISBN 951-38-0543-3

Huttula, J. & Ernvall, T. 1994. Henkilöauton koon ja vetotavan vaikutus kuolemaan johtaneissa onnettomuuksissa. (The influence of size and drive of the car on fatal accidents). Oulu: Oulun yliopisto. 86 p. (Tie- ja liikennetekniikan laboratorion julkaisuja 24). (In Finnish) ISBN 951-42-3836-2

Häkkinen, S. 1978. Tapaturmateoriat ja niiden kehittäminen. Espoo: Teknillinen Korkeakoulu. 87 p. (Teollisuustalouden ja työpsykologian laboratoriot, Report No. 36/1978). (In Finnish) ISBN 951-751-383-6

Häkkinen, S. & Luoma, J. 1990. Liikennepsykologia. Espoo: Otatieto Oy. 171 p. (534 Otatieto). (In Finnish) ISBN 951-672-110-9

Häkkinen, S., Lehtimäki, R., & Saharinen, L. 1996. Liikennepsykologia. Espoo: Weilin+Göös. 129 p. (In Finnish) ISBN 951-35-3568-1

Ingebrigtsen, S. & Fosser, S. 1991. Dekkestandardens berydning for trafikulykker om vintern. (The effect of tire standard on traffic accidents during wintertime) Oslo: Transportøkonomisk institutt. 50 p. (Rapport 0075/1991). (In Norwegian) ISBN 82-7133-682-7

Joly, P. et al. 1993. Exposure for different licence categories through a phone survey: validity and feasibility study. Accident Analysis & Prevention. Vol 25, No. 5, pp. 529–536. ISSN 0001-4575

Jovanis, P.P. & Chang H.L. 1989. Disaggregate model of highway accident occurrence using survival theory. Accident Analysis & Prevention. Vol 21, No. 5. pp. 445–458. ISSN 0001-4575

Kalbfleisch, J. & Prentice, R. L. 1980. The statistical analysis of failure time data. New York: John Wiley & Sons. 321 p. ISBN 0-471-05519-0

Kallberg, V.-P. 1983. Kestopäällysteen vaikutus liikenneonnettomuuksiin. Espoo: Valtion teknillinen tutkimuskeskus. 61 p. (VTT Tie- ja liikennelaboratorio, Tutkimusselostus 391/1983). (In Finnish)

Kallberg, V.-P., Kulmala, R., Laukkanen, K., Mäkelä, K., Pellinen, T., Rämä, P. & Unhola, T. 1991. Talvi ja tieliikenne, tutkimusohjelma. (Winter and road traffic in Finland). Espoo: Valtion teknillinen tutkimuskeskus. 73 p. (VTT Tiedotteita 1308). (In Finnish) ISBN 951-38-4041-7

Karttunen, R. 1994. Selvitys tutkijalautakuntien arvioimista kuolonkolarien avaintapahtumista, riskitekijöistä ja turvallisuuden parantamisehdotuksista. Työraportti. Helsinki: Tielaitos, Vakuutusyhtiöiden liikenneturvallisuustoimikunnat (VALT). 42 p. (In Finnish)

Keskinen, E., Hatakka, M., Katila, A. & Laapotti, S. 1994. Nuorten mieskuljettajien ajoneuvon valinta ja ajotapa. (Young male drivers' choice of car and driving habits). Turku: Turun yliopisto. 44 p. (Psykologian tutkimuksia 97). (In Finnish) ISBN 951-29-0365-2

Kulmala, R. & Peltola, H. 1985. Traffic safety in the dark on public roads in Finland. Espoo: Technical Research Centre of Finland. 51 p + app. 14 p. (VTT Research Notes 301). ISBN 951-38-1925-6

Kulmala, R. 1995a. Safety at rural three- and four-arm junctions. Development and application of accident prediction models. Espoo: Technical Research Centre of Finland. 104 p. + app. 42 p. (VTT Publications 233). ISBN 951-38-4771-3

Kulmala, R. 1995b. Tuntiliikenteen vaikutus liikenneturvallisuuteen. Helsinki: Tielaitos. 42 p. (Tielaitoksen selvityksiä 37/1995). (In Finnish) ISBN 951-726-083-0

Laapotti, S. 1991. Uusien kuljettajien kuolemaan johtaneet liikenneonnettomuudet Suomessa vuosina 1978–87. (Fatal motor vehicle accidents of new drivers in Finland during the years of 1978–87). Turku: Turun yliopisto. 65 p. (Psykologian tutkimuksia 92). (In Finnish) ISBN 951-880-662-4

Lee, E. T. 1992. Statistical methods for survival data analysis. London: John Wiley & Sons Inc, second edition. 482 p. ISBN 0-471-61592-7

Lundell, M. 1993. Vinterdäckens inverkan på trafiksäkerheten. Diplomarbete. (The effect of winter tires on traffic safety. Master’s thesis) Helsinki: Teknillinen korkeakoulu, konetekniikan osasto, kuljetusvälinetekniikan laitos. 89 p. (In Swedish)

Lynam, D. & Twisk D. 1995. Car driver training and licensing systems in Europe. Crowthorne: Transport Research Laboratory, Forum of European Road Safety Research Institutes. 65 p. (TRL Report 147). ISSN 0968-4107

Mannering, F. L. 1993. Male/Female driver characteristics and accident risk: some new evidence. Accident Analysis & Prevention. Vol 25, No. 1. pp. 77–84. ISSN 0001-4575

Massie, D., Campbell, K. L. & Williams, A. 1994. Traffic accident involvement rates by driver age and gender. Accident Analysis and Prevention, Vol. 27, No. 1, pp. 73–87. ISSN 0001-4575

Maycock, G., Lockwood C. R & Lester, J. 1991. The accident liability of car drivers. Transport and Road Research Laboratory. 34 p. ISSN 0266-5247

Milton, J. S. & Arnold, J. C. 1986. Probability and statistics in the engineering and computing sciences. MacGraw-Hill.643 p. ISBN 0-07-100574-9

McPherson, G. 1990. Statistics in scientific investigation. Its basis, application and interpretation. New York: Springer-Verlag. 666 p. ISBN 0-387-97137-8

Mäkelä, K., Anila, M. & Kuusola J. 1993. Henkilöautojen kylmäkäyttö. Espoo: Valtion teknillinen tutkimuskeskus. 38 p. (VTT Tie- , geo- ja liikennetekniikan laboratorio, tutkimusraportti 192). (In Finnish)

Mäkinen, T. 1985. Riskin käsite ikäryhmien välisessä liikenneturvallisuusvertailussa. Helsinki: Liikenneturva. 53 p. (Tutkimusosaston julkaisuja 72/1985) (In Finnish) ISBN 951-9431-89-6, ISSN 0357-9751

Mäkinen, T. 1994. Talvi ja tieliikenne -projekti. Nastarenkaiden vaikutus matkoihin ja kuljettajan riskinottoon. (Road traffic in winter: Effects of studded winter tyres on trip making and risk taking). Helsinki: Tielaitos. 28 p. (Tielaitoksen sisäisiä julkaisuja 1/1994). (In Finnish)

Niiniluoto, I. 1983. Tieteellinen päättely ja selittäminen. Helsinki: Kustannusosakeyhtiö Otava. 416 p. (In Finnish) ISBN 951-1-07379-6

Norusis, M. J. 1993. SPSS for Windows. Advanced Statistics. Release 6.0. Chicago: SPSS Inc. 578 p. ISBN 0-13-178823-X

Näätänen, R. 1988. Miten muut tienkäyttäjät suhtautuvat nuoriin kuljettajiin. Esitys XIII:ssa liikenneturvallisuusalan tutkijaseminaarissa Vääksyssä 27.4.1988. (In Finnish)

Näätänen, R. & Summala, H. 1976. H. Road-user behaviour and traffic accidents. Amsterdam: North-Holland Publishing Company. 270 p. ISBN 0-7204-0338-3

OECD 1990. Behavioural adaptations to changes in road transport system. 1990. Paris: Organisation for Economic Co-operation and Development (OECD). 123 p. ISBN 92-63-13389-5

Pirtala, P. & Ernvall, T. 1992. Henkilöautojen omistus, ajosuoritteet ja käyttöalueet. Helsinki: Tielaitos. 61 p. (Tielaitoksen selvityksiä 53/1992). (In Finnish) ISBN 951-47-6518-4

Polvinen, P. 1984. Talvikelien onnettomuusriskit. Helsinki: Tie- ja vesirakennushallitus, Ins.tsto Pentti Polvinen Ky. 43 p. (TVH 74822). (In Finnish) ISBN 951-46-7288-7

Rajalin, S., Auranen, T. & Sipinen, L. 1989. Ongelmakuljettajat osa II. Kuljettajien riskikäyttäytyminen. Helsinki: Liikenneturva. 124 p. (Liikenneturvan tutkimuksia 100/1989). (In Finnish) ISBN 951-9151-63-X

Rajalin, S. & Immonen, T. 1993. Vakavasta onnettomuudesta selviytyneiden myöhempi ajokäyttäytyminen. (Driving behaviour subsequent to involvement in a serious accident). Helsinki: Liikenneturva. 56 p. (Liikenneturvan tutkimuksia 110/ 1993). (In Finnish) ISBN 951-560-010-3

Roine, M. 1993a. Talvi ja tieliikenne -projekti. Kuljettajakäyttäytyminen kaarre- ja jonoajossa. (Driver behaviour in sharp curves and queues on main roads). Helsinki: Tielaitos. 34 p. (Tielaitoksen selvityksiä 87/1993). (In Finnish) ISBN 951-47-8771-4

Roine, M. 1993b. Yksilömallit ja kuolemaan johtaneet liikenneonnettomuudet. Espoo: Valtion teknillinen tutkimuskeskus. 39 p. (VTT Tie-, geo- ja liikennetekniikan laboratorio, tutkimusraportti 164). (In Finnish)

Roine, M. 1994. Talvi ja tieliikenne -projekti. Nastattomia talvirenkaita käyttäneiden kuljettajien onnettomuusriskit. (Accident risks in winter traffic among drivers using non-studded winter tyres). Helsinki: Tielaitos. 49 p. (Tielaitoksen selvityksiä 87/1993). (In Finnish) ISBN 951-47-8771-4

Roine, M. 1996. Nastattomia talvirenkaita käyttäneiden kuljettajien ominaisuudet ja talviajan onnettomuusriskit. Espoo: Valtion teknillinen tutkimuskeskus. 101 p. (VTT Yhdyskuntatekniikka, tutkimusraportti 340). (In Finnish)

Rolls, G. & Ingham, R. 1992. ‘Safe’ and ‘unsafe’ – a comparative study of younger male drivers. Southampton: AA Foundation for Road Safety Research. 80 p. (University of Southampton, AA Foundation for Road Safety Research).

Roosmark, P.-O., Andersson, K. & Ahlqvist, G. 1976. Dubbdäcks effekt på trafikolyckor. Linköping: Statens väg- och trafikinstitut (VTI). 107 p. + app. 36 p. (Rapport Nr 72). (In Swedish)

Rämä, P., Kulmala, R. & Heinonen, M. 1996. Muuttuvien kelivaroitusmerkkien vaikutukset liikennekäyttäytymiseen Turun tiepiirissä talvella 1993–1994. (Behavioural effects of slippery road variable message signs in Turku road district in winter 1993–1994).

Helsinki: Tielaitos. 39 p. (Tielaitoksen selvityksiä 36/1996). (In Finnish) ISBN 951-726-082-2

Saastamoinen, K. 1993. Talvi ja tieliikenne -projekti. Kelin vaikutus ajokäyttäytymiseen ja liikennevirran ominaisuuksiin. (Effect of road conditions on driving behaviour and properties of the traffic flow). Helsinki: Tielaitos. 49 p. (Tielaitoksen selvityksiä 80/1993). (In Finnish) ISBN 951-47-8139-2

Saastamoinen, K. 1994. Talvi- ja tieliikenne -projekti. Talvirengastutkimus. Helsinki: Tielaitos. 49 p. (Tielaitoksen selvityksiä 34/1994). (In Finnish) ISBN 951-47-9422-2

Schlesselman, J. 1982. Case-control studies. New York: Oxford University Press. 354 p. ISBN 0-19-502933-X

Sheppard, P. 1992. Experience of an accident and its influence on driving. Transport Research Laboratory, Department of the Environment, Department of Transport, TRRL Supplementary Report 750.

Siegel, S. 1956. Nonparametric statistics for behavioural sciences. Tokyo: McGraw-Hill Kogakusha, Ltd. 312 p.

Snedecor, G. & Cochran, W. 1980. Statistical methods. Iowa: The Iowa State University Press. 507 p. ISBN 0-8138-1560-6

Spolander, K. 1983. Bilförares olycksrisker. En modell testad på män och kvinnor. Linköping: Statens väg- och trafikinstitut. 28 p. (VTI rapport nr 268.) (In Swedish) ISSN 0347-6030

Stamatiadis, N. & Deacon, J. 1995. Trends in highway safety. Effects of an aging population on accident propersity. Accident Analysis and Prevention. Vol. 27, No 4, pp. 443–459. ISSN 0001-4575

Tapio, J., Pirtala, P. & Ernvall, T. 1994. Onnettomuus- ja loukkaantumisriskit automalleittain. (The accident involvement and injure risk rates of car models). Oulu: Oulun yliopisto. 96 p. (Tie- ja liikennetekniikan laboratorion julkaisuja 28). (In Finnish) ISBN 951-42-4080-4

TVH 1982. Liikenneonnettomuustilastojen edustavuustutkimus 1982, pääraportti. Helsinki: Tie- ja vesirakennushallitus, Liikennevakuutusyhdistys, Kehittämistoimisto Oy ERG Ab. 64 p. (TVH 741939). (In Finnish) ISBN 951-46-5529-X

Uusitalo, P. 1974. Suunnittelun tavoitteet ja keinot. Helsinki: Kustannusosakeyhtiö Tammi. 267 p. (In Finnish) ISBN 951-30-2954-9

VALT 1994. Vakuutusyhtiöiden liikennevahinkotilasto 1993. Helsinki: Liikennevakuutuskeskus, Vakuutusyhtiöiden liikenneturvallisuustoimikunta (VALT). 94 p. (In Finnish) ISBN 951-9330-44-5

Valtonen, J. 1986. Tutkimus nastarenkaiden liikenneturvallisuus- ja kustannusvaikutuksista. Helsinki: Liikenneturva. 218 p. (Tutkimusosaston julkaisuja 79/1986). (In Finnish) ISBN 951-9151-06-0, ISSN 0782-2421

Wasielewski, P. & Evans, L. 1983. Do drivers of small cars take less risk in everyday driving. Michigan: General Motors, Research Laboratories. 17 p. (IRRD: GM Research Publication).

Wilde, G. S., & Murdoch, A. 1982. Incentive systems for accident-free and violation-free driving in the general population. Ergonomics, Vol. 25, No. 10, pp. 879–890.

Zador, P. 1991. Alcohol-related relative risk of fatal driver injuries in relation to driver age and sex. New Jersey: (IRRD: Journal of studies on alcohol. Vol. 52, No.4,). pp. 302–310. ISSN 0333-5649

Öberg, G., Junghard, O. & Wiklund, M. 1993. En studie av metoder för att beräkna samband mellan dubbdäcksanvändning och trafiksäkerhet. (A study of methods of calculating the connection between the use of studded tyres and road safety). Linköping: Väg- och transportforskningsinstitutet. 58 p. (VTI Meddelande nr 722). (In Swedish) ISSN 0347-6049

APPENDIX A: Questionnaire layout
(translated from Finnish original)

ACCIDENT SURVEY

1	Your birth year	└──┴──┴──┴──┘
2	The municipality you live in	....
3	Your sex	1 man 2 woman
4	Are you	1 working 2 not working 3 unemployed 4 pensioner
5	Which of the following groups to you consider yourself to be in :	1 worker 2 functionary 3 executive 4 entrepreneur 5 farmer 6 housewife, working at home 7 student or schoolboy/girl 8 other, what ....
6	The nature of your driving license	1 no driving license, 2 motorcycle or tractor (A, T) 3 car (B) 4 truck or bus (C, D, E)
7	Availability of vehicle	0 no vehicle 1 own vehicle (number) 2 others in family have own vehicle (number) 3 company car (number) 4 something else, what and number ....
8	What vehicle do you primarily use?	0 nothing 1 motorcycle 2 car 3 van 4 truck 5 bus 6 other, what ....
9	How much do you drive with the vehicle annually?	0 not at all 1 less than 5,000 km/a 2 5,000–15,000 km/a 3 15,000–25,000 km/a 4 25,000–40,000 km/a 5 over 40,000 km/a

10 Your current vehicle's data (e.g. on the basis of the note of registration)

Make: ______________________

Model: ______________________

date, when vehicle was first taken into use: ______________________

engine volume (l or cm3): ______________________

motor power (kW): ______________________

11 If you have changed vehicles during the last two years, what were your former vehicle's

make: ______________________

model ______________________ year ______________________

12 If you have changed vehicles during the last two years, when did it occur?

month.. |__|__|
year.. |__|__|
haven't changed vehicles .... |__|

13 How many kilometres have been driven with your current vehicle

kilometreage .... |__|__|__|__|__|__| km

14 How much did you yourself drive with your vehicle last winter between November and March (1.11.1992-31.3.1993)?

kilometreage .... |__|__|__|__|__|__| km

15 How big a share of your annual kilometreage did you drive in built-up areas (speed limit 60 km/h at maximum)?

the share is about....................................... |__|__|__| %

16 How big a share of your annual kilometreage is

1 trips to and from work .... |__|__| %
2 business trips .... |__|__| %
3 shopping etc. trips .... |__|__| %
4 vacation or leisure trips............................ |__|__| %

17 Did the vehicle you used during the winter (1.11.-31.3.) have

1 studded tyres
2 non-studded winter tyres
3 others what tyres

18 What do you think of the necessity of studded tyres in winter traffic

1 absolutely necessary
2 quite necessary
3 quite unnecessary
4 totally unnecessary
5 cannot say

19 You need new tyres for your vehicle. You have the opportunity to buy non-studded winter tyres for half the price of studded tyres. Would you choose

1 definitely the studded tyres
2 quite surely the studded tyres
3 quite surely the non-studded winter tyres
4 definitely the non-studded winter tyres
5 cannot say

Fill one column for each traffic accident you've been in as your vehicle's driver in or after 1991. The tables have room for four accidents. If you've been in just one accident, only fill column 1. If you've had several accidents, fill one column for one accident, so that a certain accident number in the different tables means the same accident.

20 Year, month and day of accident

	Accident nr 1	2	3	4
year				
month				
day				

21 What were the consequences of the accident for yourself (driver, check square)

	Accident nr 1	2	3	4
vehicle was damaged				
minor injury				
serious injury				

22 What were the consequences to the other people in your vehicle?

	Accident nr 1	2	3	4
number of minor injuries				
number of serious injuries				
number of fatalities				

Minor injury: bruises, stitches, a visit to the medical centre; Serious injury: treated at medical centre or hospital

23 All of the accident's consequences (altogether for different involved parties)

	Accident nr 1	2	3	4
number of damaged vehicles				
number of injured				
number of fatalities				

24 Scene of the accident (check square)

	Accident nr 1	2	3	4
street or road in built-up area				
rural road				
courtyard, parking area				
private road				
other, what?				

25 Was the scene of the accident (check square):

	Accident nr 1	2	3	4
junction				
other section				

26 What were the accidents' other involved parties (number). If there were no involved parties, check square "no other involved parties"

	Accident nr 1	2	3	4
motorcycle				
moped				
car				
van				
truck				
bus				
tram				
train				
pedestrian				
bicyclist				
tractor				
animal (e.g. moose)				
no other involved parties				
other, what				

27 Did you yourself use a safety belt, when the accident occurred (check square)?

	Accident nr 1	2	3	4
yes				
no				

28 When the accident occurred, did your vehicle have (check square)

	Accident nr 1	2	3	4
studded tyres				
non-studded winter tyres				
summer tyres				
other tyres				

29 When the accident occurred, were your vehicle's tyres (check square)

	Accident nr 1	2	3	4
new or good as new				
somewhat worn				
quite worn				
very worn				

30 What was the purpose of the trip, during which the accident occurred (check square)

	Accident nr 1	2	3	4
trip to or from work				
business trip				
shopping or errand trip				
leisure trip				

31 What were the road surface conditions when the accident occurred (check square)

	Accident nr 1	2	3	4
dry				
wet				
wet, salted				
snowy				
slushy				
icy				
snowy, slushy or icy, but vehicle paths bare				

32 What was the weather like when the accident occurred (check square)

	Accident nr 1	2	3	4
sunny				
partly cloudy				
raining				
snowing				
foggy				

33 What was the lighting like when the accident occurred (check square)

	Accident nr 1	2	3	4
light				
dim				
dark, no road lighting				
dark, road lighted				
foggy				

34 For how long had you been on the trip when the accident accident occurred (e.g. 7 min or 1 h 45 min)

	Accident nr 1	2	3	4
time (hours and minutes)	........h........min	........h........min	........h........min	........h........min

Can you add an account of the causes of the different accidents and a drawing of what happened?

Accident 1

Accident 2

Accident 3

Accident 4

If we need additional information, can you write your

name ________________________________

address ________________________________

zip code and post office ________________________

phone number (also area code) ____________________

Return address: VTT/Road, Traffic and Geotechnical Laboratory
Additional information: Matti Roine and Veli-Pekka Kallberg

APPENDIX B: Survival models compiled from the postal survey data

The following is a short summary of the most important models. Table B1 gives a summary of the models and the other tables provide more detailed descriptions. Table B2 presents the statistical significance of including different variables into the models.

Model 1. A basic model, the variables were selected through backward elimination. The variables of the maximal first model were driver age, wintertime kilometreage, sex, vehicle age and the second degree interactions between the variables (section 6.4).

Model 2. A basic model, the variables were chosen through backward elimination. The variables of the maximal first model were driver age, wintertime kilometreage as a logarithmic direct variable, sex, vehicle age and the second degree interactions between the variables, except for the correlations of wintertime kilometreage (section 6.4).

Model 3. A basic model, the variables were chosen through backward elimination. The variables of the maximal first model were driver, wintertime kilometreage, sex, vehicle age, share of driving in built-up areas and the second degree interactions between the variables.

Model 4. A model consistent with basic model 1, in which driver age was categorised into six categories for more detailed examination of the effects. Tyre type was also added to the model.

Model 5. A model composed for the examination of the impact of correlations. The most important variables, including tyre type, were included in the model as individual direct variables.

Model 6. The model was formed from basic model 1 by adding tyre type into the model as a variable. The model included driver's wintertime kilometreage as a categorised variable with three classes.

Model 7. The model was formed from basic model 2 by adding tyre type into the model. The model included driver's wintertime kilometreage as a continuous variable.

Model 8. The model was formed from basic model 3 by adding tyre type into the model. The model included driver's wintertime kilometreage as a categorised variable with three classes.

Model 11. The model was based on model 1, adding the effect of tyre type into it. Wintertime kilometreage was replaced with driver's annual vehicle kilometreage. Backwards elimination and forwards selection led to the same kinds of models, except for model 2. There, the share of driving in built-up areas and its correlations were added to the model.

In models that were composed through forwards selection, share of driving in built-up areas was included in the models only as a correlation with the sex.

Table B1. Proportional survival models for the postal survey data. The variables of the models, log-likelihood values and number of observations.

Model	**Variables**	**-2Log-likelihood**		
		Initial situation	**Model**	**Number of observations**
1	CKIKA2 + CTAKM + CKIKA2×SPUOLI + CKIKA2×AUTONI	5054.0	4980.3	4116
2	CKIKA2 + LNSUO + CKIKA2×SPUOLI + CKIKA2×AUTONI	5054.0	4978.9	4116
3	CKIKA2 + CTAKM + CTAJA + CKIKA2×SPUOLI + CKIKA2×AUTONI + CKIKA2×CTAJA + CTAKM×CTAJA	5044.7	4956.5	4056
4	CKIKA + CTAKM + CKIKA×AUTONI + CKIKA×SPUOLI	4952.4	4866.8	4099
5	CKIKA2 + CTAKM + SPUOLI + CTAJA + AUTONI + RENGAS	4943.4	4875.1	4040
6	CKIKA2 + CTAKM + RENGAS + CKIKA2×SPUOLI + CKIKA2×AUTONI	4952.4	4875.0	4099
7	CKIKA2 + LNSUO + RENGAS + CKIKA2×SPUOLI + CKIKA2×AUTONI	4952.4	4874.4	4099
8	CKIKA2 + CTAKM + RENGAS + CKIKA2×SPUOLI + CKIKA2×AUTONI + CKIKA2×CTAJA + CTAKM×CTAJA	4943.4	4850.5	4040
11	CKIKA2 + CAJOKM + SPUOLI + RENGAS + CKIKA2×AUTONI + CAJOKM×SPUOLI	4952.4	4877.6	4099

CKIKA2	Driver age
CTAKM	Wintertime kilometreage
SPUOLI	Driver sex
AUTONI	Vehicle age
LNSUO	Ln (wintertime kilometreage/1000)
CTAJA	Share of driving in built-up areas
RENGAS	Tyre type
CAJOKM	Annual vehicle kilometreage

Table B2. Proportional survival models for the survey data (models 1-3) and the statistical significance of each included variable.

The models' variables, the effect of the variables on log-likelihood value, degrees of freedom and the statistical significance based on the Wald test. Log-likelihood value and number of observations of model's initial situation and the final model.

Model	Variable	Contents	Change: Chi-Square (Change of –2LL)	Change: Degrees of freedom	Significance p
MODEL 1					
1	CKIKA2	Driver age	49.281	2	0.000
2	CTAKM	Wintertime kilometreage	10.538	2	0.005
3	CKIKA2×SPUOLI	Driver age & sex	7.124	2	0.028
4	CKIKA2×AUTONI	Driver & vehicle age	6.719	2	0.035
	Initial situation –2LL		5054.0		
	–2LL for model		4980.3		
	Number of observations		4116		
MODEL 2					
1	CKIKA2	Driver age	49.281	2	0.000
2	LNSUO	Ln(wintertime kilometreage /1000)	11.448	1	0.001
3	CKIKA2×SPUOLI	Driver age & sex	8.083	2	0.018
4	CKIKA2×AUTONI	Driver & vehicle age	6.238	2	0.044
	Initial situation –2LL		5054.0		
	–2LL for model		4972.1		
	Number of observations		4116		
MODEL 3					
1	CKIKA2	Driver age	48.122	2	0.000
2	CTAKM	Wintertime kilometreage	10.065	2	0.007
3	CKIKA2×SPUOLI	Driver age & sex	7.494	2	0.024
4	CKIKA2×AUTONI	Driver & vehicle age	6.499	2	0.039
5	CTAJA	Share of driving in built-up areas	3.536	1	0.060
6	CTAKM×CTAJA	Wintertime kilometreage & driving in built-up areas	6.595	2	0.037
7	CKIKA2×CTAJA	Driver age & driving in built-up areas	5.858	2	0.054
	Initial situation –2LL		5044.7		
	–2LL for model		4851.6		
	Number of observations		4056		

Table B3. Proportional survival model 4 estimated for postal survey data. Estimated parameters and their standard errors in parenthesis.

Variables **Name**	**Value**	**Estimated parameters** **Model 4**
CKIKA Driver age	≤ 25	1.0120 (0.1582)
	26–35	0.3466 (0.1560)
	36–45	–0.0816 (0.2104)
	46–55	0.0816 (0.2104)
	56–65	0.8566 (0.3601)
	> 65	0.1984 (0.1600)
CTAKM Wintertime kilometreage	≤ 5 000 km	1.000*
	5 000–10 000 km	0.2028 (0.1414)
	> 10 000 km	0.5740 (0.1509)
CKIKA×SPUOLI Driver age & sex	CKIKA(1)×SPUOLI	–0.1315 (0.3047)
	CKIKA(2)×SPUOLI	0.5846 (0.2423)
	CKIKA(3)×SPUOLI	0.2951 (0.3003)
	CKIKA(4)×SPUOLI)	–0.2305 (0.5192)
	CKIKA(5)×SPUOLI	–0.9075 (0.6819)
	CKIKA(6)×SPUOLI	0.3898 (0.2598)
CKIKA×AUTONI Driver age & vehicle age	CKIKA(1)×AUTONI(2)	0.7600 (0.2597)
	CKIKA(2)×AUTONI(2)	–0.0646 (0.2852)
	CKIKA(3)×AUTONI(2)	0.3862 (0.3915)
	CKIKA(4)×AUTONI(2)	–0.9698 (0.5965)
	CKIKA(5)×AUTONI(2)	0.4015 (0.4875)
	CKIKA(6)×AUTONI(2)	0.2590 (0.2797)
RENGAS Tyre type	studded tyre	0.0000*
	non-studded tyre	0.3206 (0.3227)
Initial situation's -2log-likelihood		4952.4
Model's -2log-likelihood		4866.8
Model's degrees of freedom		18
Number of observations		4099

* Reference level

Table B4. Proportional survival model 5 estimated for the postal survey data, Estimated parameters and their standard errors in parenthesis.

Variables		Estimated parameters
Name	**Value**	**Model 5**
CKIKA2	Driver age ≤ 25	0.000 *
Driver age	Driver age 26–50	–0.8020 (0.1539)
	Driver age > 50	–1.1343 (0.1881)
CTAKM	≤ 5 000 km	0.000 *
Wintertime kilometreage	5 000–10 000 km	0.2355 (0.1430)
	> 10 000 km	0.6221 (0.1554)
SPUOLI	Male	0.000 *
Driver sex	Female	0.2704 (0.1357)
CTAJA	Driving in built-up areas ≤ 50%	0.000 *
Share in built-up areas	Driving in built-up areas > 50%	0.2538 (0.1266)
AUTONI Vehicle age	Vehicle age > 10	0.1184 (0.1426)
RENGAS	Studded tyre	0.000 *
Tyre type	Non-studded tyre	0.2849 (0.3238)
Initial situation's -2log-likelihood		4943.4
Model's -2log-likelihood		4875.1
Model's degrees of freedom		8
Number of observations		4040

* Reference level

APPENDIX C: Survival models compiled from the VALT data

The following is a summary of the most important models. Table C1 provides a summary of the models and the other tables more detailed descriptions. Table C2 presents the statistical significance of including different variables into the models.

Model 1. A model estimated on the basis of the first group of variables using backwards elimination. Tyre's tread depth, and the use of alcohol and safety belt were variables without correlations in the maximal first model (section 6.5).

Model 2. A model estimated on the basis of both the first and the second group of variables using the backwards elimination. Tyre's tread depth, the use of alcohol and safety belt, route familiarity, relative speed and speed limit were included without interactions, in the maximal first model (section 6.5).

Model 3. A model estimated on the basis of the second group of variables. Variables from model 2 were added later, except for the interaction between speed limit and relative speed.

Model 4. A model based on the first group of variables, plus the interaction of vehicle age and tyre type in order to estimate the effect of tyre type.

Model 7. A model defined through backwards elimination in order to examine the influence of the sex. A basic risk function was calculated separately for males and females, with the other variables that best explained the accident time included later in the model.

Model 8. A model consistent with model 1 of the first group of variables, but with driver age categorised into six age groups in order to analyse more precisely the effects of driver age.

Model 9. A model based on backwards elimination, in which the driver age, annual vehicle kilometreage, the use of alcohol and safety belt, vehicle age and weight, tyre's tread depth and tyre type were variables in the initial situation. Tyre type had three (summer tyres, non-studded winter tyres and studded tyres). Interactions with described tyre's tread depth and use of safety belt were not included

Model 10. A model consistent with model 1 with tyre type and road surface condition added to the model.

Table C2. Proportional survival models (models 1 and 2) estimated for VALT data. the statistical significance of each included variable.

The table contains the models' variables, the impact of the variables on the log-likelihood value, the degrees of freedom, the variables' statistical significance based on Wald test score, the log-likelihood values and number of observations of the models' initial and final situations.

Variable's		**Content**	**Change**		
Nr	**Name**		**Chi-Square (Change of -2LL)**	**Degrees of freedom**	**Significance p**
MODEL 1					
1	VAKM	Annual vehicle kilometreage	23.181	2	0.000
2	ALKO	Alcohol	16.017	1	0.001
3	TVYO	Safety belt	8.711	1	0.003
4	CKIKA	Driver age	9.724	2	0.008
5	SPUOLI	Female driver	6.182	1	0.013
6	RENURA	Tyre's tread depth	2.888	1	0.089
7	AUTONI	Vehicle age	2.607	1	0.106
8	KEVYT	Vehicle weight	1.651	1	0.199
9	RENGAS	Tyre type	0.961	1	0.311
	Initial situation -2LL		2881.7		
	-2LL for model		2809.8		
	Number of observations		446		
MODEL 2					
1	VAKM	Annual vehicle kilometreage	21.067	2	0.000
2	ALKO	Alcohol	12.369	1	0.000
3	NRAJA	Speed limit	7.795	1	0.005
4	TVYO	Safety belt	5.197	1	0.023
5	TUTTU	Familiarity of route	5.062	1	0.025
6	CKIKA	Driver age	6.484	2	0.039
7	SPUOLI	Sex	4.034	1	0.045
8	NRAJA×SNOPE	Relative speed	2.566	1	0.109
9	RENURA	Tyre's tread depth	2.452	1	0.117
10	AUTONI	Vehicle age	2.315	1	0.128
11	KEVYT	Vehicle weight	1.904	1	0.168
12	RENGAS	Tyre type	0.686	1	0.401
	Initial situation -2LL		2524.9		
	-2LL for model		2453.0		
	Number of observations		407		

Table C3. Proportional survival model 7 (stratified on sex) for VALT data. Estimated parameters and their standard errors in parenthesis.

Variables		**Estimated parameters**
Name	**Value**	**Model 7**
CONSTANT		
VAKM	< 10 000 km	0.000*
Annual vehicle kilometreage	10 000–29 999 km	-0.5042 (0.1805)
	≥ 30 000 km	-0.6133 (0.2057)
CKIKA2	Driver ≤ 25	0.0811
Driver age	Driver 25–50	-0.3720 (0.1034)
	Driver > 50	0.2909 (0.1194)
ALKO Use of alcohol	Effect of alcohol	0.7819 (0.2460)
TVYO Use of seat belt	Safety belt in use	-0.4837 (0.1670)
KEVYT Vehicle weight	Vehicle's weight ≤ 1 000 kg	0.1799 (0.1366)
RENURA Tyre's tread depth	Tyre's tread depth / mm	-0.0540 (0.0295)
RENGAS Tyre type	Non-studded winter tyres	0.2418 (0.2348)
CKIKA2 ×AUTONI	CKIKA(1)×AUTONI	-0.2614*
Driver age & vehicle age	CKIKA(2)×AUTONI	-0.3014 (0.2024)
	CKIKA(3)×AUTONI	0.5628 (0.2292)
Initial situation's -2log-likelihood		2570.7
Model's -2log-likelihood		2507.5
Number of observations		446

* Reference level

Table C4. Proportional survival model 8 for VALT data. Estimated parameters and their standard errors in parenthesis.

Variables		**Estimated parameters**
Name	**Value**	**Model 8**
VAKM Annual vehicle kilometreage	< 10 000 km	0.000 *
	10 000–29 999 km	-0.5614 (0.1809)
	≥ 30 000 km	-0.6332 (0.2088)
CKIKA Driver age	Driver ≤ 25	0.000 *
	Driver 26–35	-0.3506 (0.1797)
	Driver 36–45	-0.6335 (0.1961)
	Driver 45–55	-0.1683 (0.2039)
	Driver 55–65	-0.4766 (0.3091)
	Driver > 65	0.4003 (0.2400)
SPUOLI Driver sex	Driver sex	0.3778 (0.1625)
ALKO Use of alcohol	Effect of alcohol	0.6873 (0.2452)
TVYO Use of safety belt	Safety belt in use	-0.5159 (0.1682)
KEVYT Vehicle weight	Vehicle weight ≤ 1000 kg	0.1184 (0.1371)
RENURA Tyre's tread depth	Tyre's tread depth / mm	-0.0573 (0.0305)
AUTONI Vehicle age	Vehicle age >10 v	-0.2696 (0.1676)
RENGAS Tyre type	Non-studded winter tyres	0.1490 (0.2370)
Initial situation's -2log-likelihood		2881.7
Model's -2log-likelihood		2800.1
Number of observations		446

* Reference level

Table C5. Proportional survival model 9 for VALT data. Estimated parameters and their standard errors in parenthesis.

Variables		**Estimated parameters**
Name	**Value**	**Model 9**
VAKM	< 10 000 km	0.000 *
Annual vehicle kilometreage	10 000–29 999 km	-0.5556 (0.1825)
	≥ 30 000 km	-0.6592 (0.2145)
CKIKA2	Driver ≤ 25	0.000 *
Driver age	Driver 25–50	-0.5433 (0.1748)
	Driver > 50	0.1561 (0.2060)
SUKUPUOLI Driver sex	Female driver	0.3178 (0.1601)
ALKO Use of alcohol	Effect of alcohol	0.7685 (0.2410)
TVYO Use of seat belt	Safety belt in use	-0.4931 (0.1638)
KEVYT Vehicle weight	Vehicle weight ≤ 1 000 kg	0.3277 (0.1437)
RENURA Tyre's tread depth	Tyre's tread depth / mm	-0.0442 (0.0304)
AUTONI Vehicle age	Vehicle age > 10 v	-0.1797 (0.1636)
RENGASTYYPPI	Summer tyres	1.000 *
Tyre type	Non-studded winter tyres	-0.2509 (0.3740)
	Studded tyres	-0.5178 (0.3077)
VAKM×KEVYT	VAKM(1)×KEVYT	0.000 *
Annual vehicle kilometreage & vehicle weight	VAKM(2)×KEVYT	-0.5928 (0.3324)
	VAKM(3)×KEVYT	-0.1776 (0.3921)
CKIKA2×AUTONI	CKIKA(1)×AUTONI	0.000 *
Driver age & vehicle age	CKIKA(2)×AUTONI	-0.2317 (0.3469)
	CKIKA2(2)×AUTONI	0.5994 (0.4002)
Initial situation's -2log-likelihood		3046.9
Model's -2log-likelihood		2962.1
Number of observations		463

* Reference level

Table C6. Proportional survival model 10 for VALT data. Estimated parameters and their standard errors in parenthesis.

Variable's Name	**Value**	**Estimated parameters Model 10**
VAKM	< 10 000 km	0.000 *
Annual vehicle kilometreage	10 000–29 999 km	-0.6152 (0.1800)
	≥ 30 000 km	-0.7185 (0.2040)
CKIKA2	Driver ≤ 25	0.000 *
Driver age	Driver 25–50	-0.4530 (0.1518)
	Driver > 50	-0.0388 (0.1870)
SPUOLI Driver sex	Driver sex	0.3071 (0.1635)
ALKO Use of alcohol	Effect of alcohol	0.7532 (0.2466)
TVYO Use of seat belt	Safety belt in use	-0.4974 (0.1709)
KEVYT Car weight	Vehicle weight ≤ 1 000 kg	0.1871 (0.1369)
RENURA Tyre's tread depth	Tyre's tread depth / mm	-0.0530 (0.0306)
AUTONI Vehicle age	Vehicle age > 10 years	-0.2482 (0.1646)
NTONTKE	Non-studded winter tyres and winter road surface conditions	0.3349 (0.2834)
NASTKE	Studded tyres and winter road surface conditions	0.1920 (0.1519)
NTONPKE	Non-studded winter tyres and other road surface conditions	0.5557 (0.5326)
Initial situation's -2log-likelihood		2881.71
Model's -2log-likelihood		2808.05
Number of observations		446

* Reference level

APPENDIX D: Calculation example of a survival model

By using model 6 on the basis of the postal survey data, we get (see section 6.4):

$S_0(t)$ = The baseline of the survival function = 0.941

When using studded tyres:

$\exp(XB) = 1.091$

When using non-studded winter tyres:

$\exp(XB) = 1.520$

The share of accident drivers is calculated as follows (see sections 6.1–6.4):

Users of studded tyres:

$S(t, X) = S_0(t)^{\exp(XB)} = 0.941^{1.091} = 0.936$

The share of accident drivers = 1– 0.936 = 0.064 (6.4%)

Users of non-studded winter tyres:

$S(t, X) = S_0(t)^{\exp(XB)} = 0.941^{1.520} = 0.912$

The share of accident drivers = 1– 0.912 = 0.088 (8.8%)

Published by

Vuorimiehentie 5, P.O.Box 2000, FIN-02044 VTT, Finland
Phone internat. +358 9 4561
Fax +358 9 456 4374

Series title, number and report code of publication

VTT Publications 401
VTT–PUBS–401

Author(s)
Roine, Matti

Title

Accident risks of car drivers in wintertime traffic

Abstract

The wintertime accident risks of drivers and the factors affecting the risks were analysed using statistical accident models. The evaluation method was based on reliability theory and on survival modelling. The data consisted of two parts: responses to a postal questionnaire, addressed to 10,000 vehicle owners, about driving and accidents during wintertime in the years 1991–1993; detailed records of fatal accidents, in the years 1987–1991, generated by in-depth accident investigation teams and compiled by the Motor Insurers' Road Safety Committee (VALT). Replies to the postal questionnaires were received from 5,881 vehicle owners, giving a response rate of 59%. The replies included 296 self-reported accident involved drivers. The VALT data on fatal wintertime accidents contained 658 drivers involved in fatal accidents.

The analyses of the two data sources confirmed that driving conditions and kilometreage driven during the wintertime contribute to amount of accidents. The best explaining variables in both survival and risk models of wintertime driving were driver characteristics (age, driver behaviour, kilometreage driven, speed, and use of safety-belt), drivers' state (driving under the influence of alcohol) and vehicle characteristics (vehicle weight and condition of tyres).

Young and inexperienced on the one hand, or old (and experienced) drivers, on the other hand, had the highest fatal accident risks. There were no direct sex-related differences in accident risks between female and male drivers. Drivers using non-studded winter tyres had a somewhat greater accident risk than drivers using studded tyres, but the difference was not statistically significant.

Survival modelling with two distinct data sources, a postal survey data and an in-depth accident data, were used in the analysis. Modelling methods were also analysed by simulation and it was concluded that survival modelling promises to be a useful tool for safety analysis but the method needs further development.

Keywords
winter traffic, vehicular traffic, cold environments, automobiles, drivers, traffic accidents, risk factors, research, models

Activity unit
VTT Communities and Infrastructure, Transport and Urban Planning,
Metallimiehenkuja 10, P.O.Box 1902, FIN–02044 VTT, Finland

ISBN			Project number
951–38–5546–5 (soft back ed.) 951–38–5547–3 (URL: http://www.inf.vtt.fi/pdf/)			
Date November 1999	Language English	Pages 137 p. + app. 20 p.	Price D
Name of project Winter and Road Traffic Project		Commissioned by Finnish National Road Administration (FinnRa)	
Series title and ISSN VTT Publications 1235–0621 (soft back ed.) 1455–0849 (URL: http://www.inf.vtt.fi/pdf/)		Sold by VTT Information Service P.O.Box 2000, FIN–02044 VTT, Finland Phone internat. +358 9 456 4404 Fax +358 9 456 4374	