12TH INTERNATIONAL CONFERENCE: TRAFFIC SAFETY ON THREE CONTINENTS
MOSCOW, 19-21 SEPTEMBER 2001
NIIAT, CSIR, FERSI, VTI, TRB, TRANSDEKRA, MTRF

DEALING WITH LACK OF EXPOSURE DATA IN ROAD ACCIDENT ANALYSIS

Dr. John Golias
Associate Professor

Dr. George Yannis
Lecturer

Department of Transportation Planning and Engineering
National Technical University of Athens
phone: +30.1.7721276, e-mail: igolias@central.ntua.gr

Abstract

Road accident analysis at national and international level is limited today by a number of problems inherent to the availability, the reliability, the comparability and the disaggregation level of exposure data. When adequate data are not available, then the use of alternative types of road accident analysis may produce reliable and useful results. This work identifies the basic insufficiencies inherent usually to the traffic data available at national and international level and the implications of this fact on accident analysis results. The use of absolute numbers and trends of values as well as of severity indices is generally free of the basic insufficiencies of exposure data but without useful information on accident rates. The use of induced exposure method and of accident type related percentages can provide useful information eliminating partially the need for exposure data. However, these methods which may answer a number of questions concerning road safety at international level should always be used with great care as interpretation of results may sometimes be a difficult exercise.

Key-words: road safety, road accident, accident risk exposure, accident analysis, traffic data

1. Introduction

Recent developments in Europe (integration of the European Community internal market, opening of the Eastern European markets) gave a new dimension in traffic and related road accidents in Europe. The increase of international traffic made the international dimension of road accidents a very important parameter of the problem. But road accident analysis at European level can not show today, results comparable to those of accident analysis at national level[1]. A number of difficulties, such as the unavailability and the incomparability of exposure data, limits significantly accident analysis results at European level[2].

The objective of this paper is to propose a typology of alternative accident analysis methods in order to deal with existing insufficiencies of exposure data. This theoretical approach is based on experience from road accident analysis carried out at national and European level using existing aggregate and disaggregate data on exposure (traffic) and related accidents. Particular emphasis is given to the international dimension of the problem as well as to the analysis of disaggregate data.

2. About Insufficiencies of Exposure Data

Road accident analysis at international level is very often limited not only by the incomparability of the national accident data but also by a number of insufficiencies of the respective exposure data. These insufficiencies refer to poor availability and reliability, to comparability problems and to insufficient or inappropriate disaggregation.

2.1. Poor Availability and Reliability of Traffic Data

Road accident rates can better describe the road accident phenomenon than absolute numbers because they take into consideration the actual traffic patterns (exposure). Their use implies the combination of accident data with respective traffic data which are not always available; even if they are available very often they are not reliable. Traffic data are usually estimates based on surveys and on a number of assumptions. Furthermore, they are not always available for all types of traffic; and even if they are, their precision is not the same for all types of traffic. In most cases, the use of new methods demonstrate the insufficiency of the previously used and traffic data concerning previous periods are rectified (backward extrapolation) in the light of the new methods[3].

For example, traffic on motorways is well defined in most of the countries as there exist well-established count systems (tolls, permanent counters on the road, etc.). The situation becomes less bright as the type of network is less important due to the fact that adequate counts are lacking at regional and local level for obvious reasons.

2.2. Incomparability of Traffic Data

The use of traffic data in road accident analysis presents serious difficulties at international level due to the existing incomparability of traffic data in the various countries[4]. Several different traffic estimation methodologies (sample counts, surveys, use of fuel sales, etc.) are used in the various countries, with important differences in the statistical methodology used (calculation of the sample size, etc.) and the frequency of updates. This incomparability of traffic data leads to analysis results, which are followed by large confidence intervals not allowing for reliable and really useful comparisons.

For example, due to the above incomparability two countries with similar safety rate in the national road network may present a safety rate difference of at least 1:10 in their regional road network, which is too large to be attributed to different safety behaviour given that this difference is negligible in all other road network types.

Even though the problem of incomparability of traffic data concerns mainly international road accident analysis, it is also found sometimes at national level when different methodologies are used for the estimation of traffic in the various types of road network, and for the various vehicle types and road user characteristics (age, sex, etc.).

2.3. Inappropriate Disaggregation of Traffic Data

The level of disaggregation of traffic data defines also the level of detail of possible road accident analysis at both national and international level. For example, it is impossible to produce accident rates for the several vehicle types for which accident data exist if respective traffic data exist only for very few vehicle types. Consequently, the rather general level of disaggregation of traffic data observed in most of the countries limits significantly the level of detail of accident rates used in road accident analysis. Additionally, for certain accident characteristics such as the use of seat belt and helmet, drinking and driving and the respect of speed limits, most often there is no respective traffic data available allowing for the extraction of a number of useful accident rates, although sufficient information exist for these parameters in relation to the observed accidents. It is noted, however, that for certain other characteristics such as daylight-night and weather conditions, traffic data can be extracted from other sources and be used for the formation of accident rates.

For example, it is very interesting to analyse accident rates of young persons driving cars or motorcycles during the night inside urban areas, but this is impossible because there are disaggregate traffic data available to be combined with the existing respective disaggregate accident data. The rates, which can be produced in the best case, are aggregate ones dealing separately with young drivers, with vehicle type, with the time of the day, and the type of area.

All the above problems of insufficient or inappropriate disaggregation of traffic data are more acute when it comes to road accident analysis at international level, where detailed comparable traffic data are scarcely available. As a consequence, today at European level, only very few and general accident rates are used due to the fact that only limited, general and hardly comparable traffic data exist for several countries.

3. Accident Analysis Alternatives

On the basis of the above presented insufficiencies of exposure data, a number of road accident analysis alternatives is considered. The shortcomings and advantages of each alternative as well as its capability of overcoming some of the above problems without significant loss of the value of the analysis results is investigated. The presentation of the alternatives is accompanied by examples of road accident analysis demonstrating their use. These examples use data from various national and international data sources.

3.1. Absolute Numbers

Analysis of aggregate or disaggregate absolute numbers of road accidents is the most basic analysis concerning road accident data. The results can be very detailed (multi-dimension Tables) and may refer to one or more countries. If common definition values exist, multi-country comparisons are possible; if not, only country-specific results can be derived[5]. These absolute number statistics can be used for the general description of the road safety level without taking into consideration the related traffic. In fact, they rather reflect the existing traffic situation than the actual accident rates and their use in road accident analysis should be considered with care. Even though, this kind of results are rather easy to produce at international level (a lot of easily comparable data exist: age, sex, etc.), their use should be limited.

For example, analysis showing that there is a higher number of road accidents on regional network during summer Saturday nights, where young drivers are involved, leads to no valid conclusion about corresponding accident rates, as this information reflects mainly the fact that the driver population consists basically of young drivers, i.e. it is a product of the young driver behaviour. Possibly, police can use this result for the intensification of law enforcement (speed limit, drinking and driving) for the reduction of the number of accidents, but no valid accident analysis conclusion can be derived for the relation among the road type, the seasonal effect and the day of the week.

It is noted that for the improvement of comparability of accident absolute numbers at international level, special transformation rules are sometimes used, by the application of factors to the values of data with different definitions in order to produce common definition data values. The transformation rules can be either simple or advanced[6]. Simple transformation rules can be the union (value 1 OR value 2), the intersection (value 1 AND value 2) or the exclusion (value 1 NOT value 2). Advanced transformation rules can be the use of a coefficient (value 1 x coefficient) or of a specific algorithms [e.g. day of the week = function(date of accident)]. Simple transformation rules can easily be applied even by the end-user, whereas the coefficients and algorithms of advanced transformation rules require an important work effort through specialised studies. The use of transformation rules presupposes the detailed knowledge of the definition of the data to be transformed.

In the following Table, an example of absolute numbers of accident data converted to a common definition through the use of transformation rules is presented. The definition of a person killed in a road accident in the EU countries is not uniform. For the conversion of the existing data to data obeying to the common 30-days definition for a killed person, transformation rules in the form of coefficients are applied for each country. These coefficients are the result of specific research comparing police and hospital data and are subject to periodic changes[7, 8]. The use of these transformation rules allows the comparison of the number of fatalities in the various EU countries.

Table 1. Number of persons killed in the EU countries (1991-2000)[9]

	B	DK	D	GR	E	F	IRL	I	L	NL	A	P	FIN	S	UK	EU 15
1991	1.873	606	11.300	2.112	8.836	10.483	445	8.083	80	1.281	1.551	3.218	632	745	4.753	55.998
1992	1.671	577	10.631	2.158	7.818	9.900	415	8.014	73	1.253	1.403	3.084	601	759	4.379	52.736
1993	1.660	559	9.949	2.159	6.378	9.867	431	7.163	76	1.235	1.283	2.700	484	632	3.957	48.533
1994	1.692	546	9.814	2.253	5.615	9.019	404	7.091	74	1.298	1.338	2.504	480	589	3.807	46.524
1995	1.449	582	9.454	2.411	5.751	8.891	437	7.020	68	1.334	1.210	2.711	441	572	3.765	46.096
1996	1.356	514	8.758	2.058	5.483	8.541	453	6.676	72	1.180	1.027	2.730	404	537	3.740	43.529
1997	1.364	489	8.549	2.199	5.604	8.444	472	6.712	60	1.163	1.105	2.521	438	541	3.743	43.404
1998	1.500	499	7.792	2.226	5.957	8.918	458	6.837	57	1.066	963	2.425	400	531	3.581	43.210
1999	1.397	514	7.772	2.131	5.738	8.487	417	7.150	58	1.090	1.079	2.231	431	580	3.564	42.639
2000	1.475	527	7.487	2.072	5.510	8.036	415	6.923	67	1.135	1.016	2.201	385	573	3.451	41.274

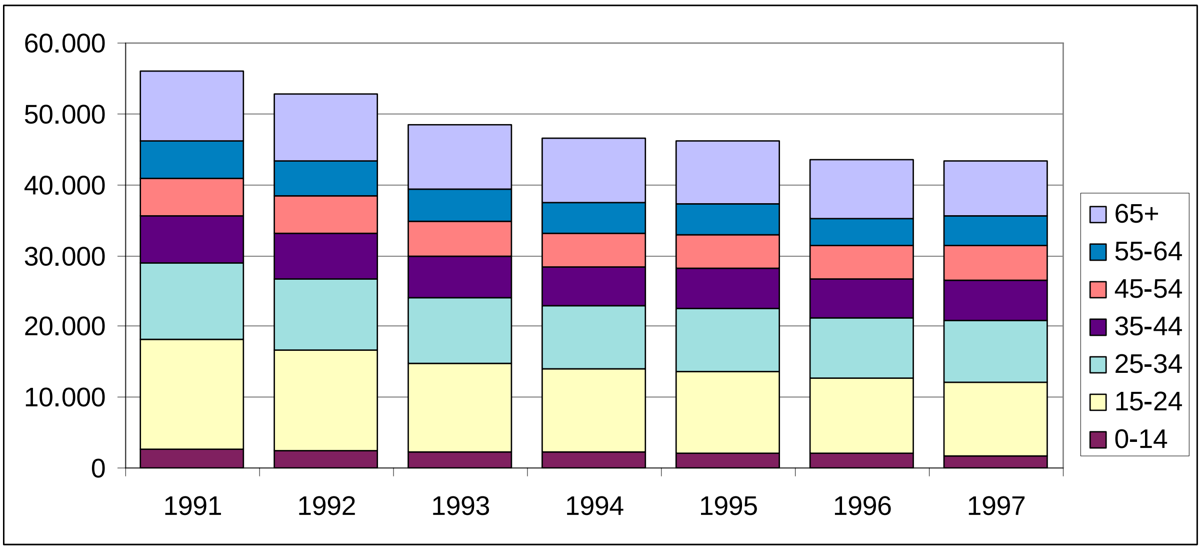
Some figures for 1999 and 2000 are estimations based on the EC Road Safety Quick Indicator
Killed: 30-day period except: GR (1 day up to 1995) +18%, E (24 hours) +30%, F (6 days) +9% up to 1993 and +5,7% 1994 onwards, I (7 days) +7,8%, A (24 hours) +12% up to 1991, P (24 hours) +30% up to 1998

It is obvious, that analysis of road accident absolute numbers can only give a general description of the road accident phenomenon.

3.2. Trends

Trends of road accident data in any form (absolute numbers, percentages etc.) and at any disaggregation level can be used in order to show the variation over time of the various accident characteristics. Obviously, trends do not provide sufficient information about the accident risk exposure but they provide very interesting information about the development of the road safety level and its parameters. This information is very interesting in the process of road safety policy planning and evaluation.

Figure 1. Number of persons killed in road accidents in the 15 EU countries by age group (1991-1997)



In Figure 1, trends in the number of persons killed in road accidents in the 15 EU countries by age group (in absolute numbers) are presented. By use of these data it is impossible to derive which age group is more dangerous as there aren't any exposure data for each age group. Only conclusions of general character can be extracted from the chart. For example, there is a trend of reduction in the number and the percentage of persons killed of age group 15-24 (28% in 1991 to 24% in 1997) whereas there is a limited increase of the number and the percentage of persons killed of age group 25-64. This information could be useful for the identification of target groups of road safety campaigns (44% of the total persons killed belong to age group 15-34).

3.3. Severity Indices

The use of severity indices can provide interesting results on both aggregate and disaggregate level without any need for traffic data. These indices provide information about the accident severity by the use of ratios in which the traffic data are not necessary anymore as they are contained both in the nominator and the denominator of the ratio (number of killed persons per injury accidents or per fatal accidents, number of injured per injury accidents, etc.).

Incomparability among the national definitions for persons injured (seriously, slightly) and related accidents involving injury limits significantly the possibilities for international comparisons. The only European-wide comparable severity index which can be used today, is the number of persons killed (30-days definition) per fatal accidents. In the future, possible use of an harmonised definition like e.g. "24-hour hospitalised injured person", could lead to the use of more comparable severity indices. Of course, indices using the number of injured persons or injury accidents can be used, without any particular problem, in disaggregate road accident analysis at national level.

Table 2 presents accident severity indices expressed as the ratio of number of persons killed per 100 persons injured. Such analyses do not require exposure data as the related exposure is the same in both the nominator (persons killed) and the denominator (persons injured) of the ratio[10]. It is interesting to observe in Table 2 that accidents with pedestrian involvement are very serious in the national and the departmental road network (26 and 16 persons killed respectively per 100 persons injured) whereas the most severe accidents in the municipal/communal network correspond to cases that the vehicle comes off the road (11 persons killed per 100 persons injured). Accidents involving collision of vehicles at angle are the less severe in all types of networks.

Table 2. Ratio of persons killed per 100 persons injured in road accidents in Greece (1985-99)

Accident type	National road	Dept road	Municipal road	Total
Head-on collision	15	6	3	**8**
Lateral colission	8	4	1	**3**
Collission at angle	5	4	1	**2**
Rear end collission	5	5	2	**3**
Collission with parked car/fixed object	14	10	7	**9**
Pedestrian involvement	26	16	6	**9**
Came off the road	9	9	11	**9**
Total	10	7	3	**6**

It is obvious that extraction of interesting results for accident severity by the use of appropriate ratios do not need exposure data. This kind of analysis can be produced for any type of road accident data.

3.4. Induced Exposure

The induced exposure method is based on the assumption that in every road accident in which two vehicles are involved there is one driver responsible for the accident and one innocent driver involved randomly from the total population of drivers. Consequently, the innocent driver can be considered as a sample of the total population of the drivers and reflects the exposure of any specific driver population defined on the basis of certain characteristics[11].

The basic requirement for the use of this method is the identification of the driver who provoked the accident. Accidents in which more than one drivers are responsible should not be taken into consideration. Accident indices are the ratio of the “guilty” drivers percentage with a certain characteristic (age, sex, vehicle or network type, etc.) divided by the percentage of “innocent” drivers of the same characteristic group. The relative involvement ratio (RIR), which is the ratio of the two relative accident indices, is representative of the tendency of the driver groups to provoke an accident. Ratios higher than 1 show that the relative driver group with the accident index as the nominator provokes more accidents than the other group. This method has been tested in several occasions and its statistical validity has been verified[12].

It is obvious that the use of the induced exposure method overcomes the need for traffic data. But the most interesting feature of the induced exposure method is the fact that it allows for disaggregate analysis to the level of disaggregation of the existing accident data. Thus, it overcomes insufficiencies and inadequacies due to the disaggregation of traffic data, widening substantially the possibilities for detailed road accident analysis at both national and international level.

However, the use of the induced exposure method is limited by the fact that it concerns only drivers and not all road users (passengers and pedestrians) and that it requires the knowledge of the “guilty” and “innocent” drivers. Additionally, this method concerns mainly accidents in which at least two vehicles were involved whereas its use in single-vehicle accidents should be considered carefully.

Table 3 shows an example of how the induced exposure method is applied. If the distribution of alcohol level of "guilty" drivers (driver A) is considered then it appears that 42% of the drivers provoking accidents are under the influence of alcohol (> 0,5 g/lt). However, from the distribution of alcohol level of "innocent" drivers (driver B) it appears that only 11% of the drivers on the roads are under the influence of alcohol (> 0,5 g/lt). The relative accident indices can be calculated, as the ratio of driver A percentage on the driver B percentage. For the drivers under the influence of alcohol this ratio is 42%/11%=3,974 and for those not under the influence of alcohol is 58%/89%=0,645. Consequently, the relative ratio of involvement in an accident of drivers under or not under the influence of alcohol in comparison with that of the sober drivers is 6.16 (=3,974/0,645).

Table 3. Distribution of alcohol level of drivers involved in road accidents in Greece (1995)[13]

	< 0,5 g/lt	> 0,5 g/lt	Total
Driver A	916	675	1.591
	58%	42%	
Driver B	1421	170	1.591
	89%	11%	
Relative Accident Index	0,645	3,971	

It is very interesting to observe that it is possible to extract the very useful relative accident involvement ratios without using any exposure data. The use of this method in road accident analysis at international level can provide an appropriate solution for overcoming to a certain degree the lack of exposure data.

3.5. Percentages Related to Accident Type

A wide number of meaningful and useful results can also be extracted by the use of specific accident related percentages eliminating the need for exposure data. This elimination is based on the fact that for a percentage referring to a certain factor in total (e.g. collision type) and for a percentage referring to a certain sub-category of this factor (e.g. head-on collisions) corresponding exposures are equal.

For example it would be possible - and very useful - to know whether rainy weather conditions affect seriously the percentage distribution of accidents among accident collision types, without the use of any corresponding traffic data. Such useful results can obviously be obtained only by analysis on disaggregate level of specific accident data (collision type, accident type, vehicle manoeuvre, person manoeuvre). The use of this method overcomes satisfactorily in certain cases the need for traffic data at national and international level, eliminating thus, problems related to the poor availability, reliability and comparability of traffic data. However, this method does not provide information concerning actual accident rates.

Table 4 presents percentages of fatal accidents in three European countries (NL, IRL, I) by vehicle type and collision type. Lack of exposure data is not a problem for the extraction of meaningful results by analysing this Table. It can be observed that the percentages are significantly different for all vehicle types when examining a certain collision type (e.g. single-vehicle or head-on) instead of the total number of fatal accidents (all collision types). For example in Ireland cars participate in the 56% of the total number of accidents, but this percentage increases to 65% in single - vehicle collisions and decreases to 50% in head - on collisions. One could also observe that in the Netherlands the percentage of cars in head-on collisions (62%) is greater that the average in all collision types (56%), whereas in Ireland and in Italy the respective percentage of cars in head-on collisions (50% and 46% respectively) is lower than the average in all collision types (56% and 54% respectively).

It should however be noted that no conclusion concerning the safety level can of course be drawn from the direct comparison among the three countries for any collision type, as the percentage of accidents per vehicle category is obviously influenced by the relative exposure.

Table 4. Percentages of fatal accidents in three European countries by vehicle type and collision type (1991-93)[14].

	All Collision Types			Single-Vehicle Accidents			Head-On Collisions		
Vehicle Type	Netherlands	Ireland	Italy	Netherlands	Ireland	Italy	Netherlands	Ireland	Italy
Car	56%	56%	54%	73%	65%	66%	62%	50%	46%
Lorry	10%	21%	15%	5%	11%	8%	14%	30%	23%
Bus	2%	2%	1%	2%	0%	0%	2%	2%	3%
Two - Wheeled	12%	11%	19%	12%	19%	21%	12%	11%	21%
Bicycle	14%	6%	7%	3%	2%	2%	7%	4%	4%
Other	6%	4%	4%	5%	3%	3%	3%	3%	3%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%

For the analysis on the basis of Table 4 data, traffic data are not necessary if comparison of the vehicle type distribution in the various collision types is only required. Additionally, the use of these percentages allows for certain comparisons between countries independently of their different exposure figures. It is noted that this possibility applies only for data related to collision type, accident type, vehicle manoeuvre and person manoeuvre.

4. Conclusion

Accident rates are very useful parameters in road accident analysis. The production of such rates depends directly on the availability, the reliability, and the disaggregation level of exposure data. When adequate exposure data are not available, then the use of alternative types of road accident analysis is the only way to produce reliable and useful analysis results. This work identified the basic insufficiencies concerning exposure data at national and international level and investigated a number of alternative ways to face some of the corresponding difficulties. The significance of these alternatives is greater for analysis at international level and analysis of disaggregate data, where exposure data insufficiencies are commonly met.

The use of absolute numbers and trends of values may lead to conclusions on traffic safety, which are in general of limited significance due to lack of exposure information. The use of severity indices overcomes the need for exposure data but corresponding results are obviously limited only to accident severity characteristics. The application of the induced exposure method is certainly more useful as it allows the identification of relative risk exposure without the use of data other than those concerning accidents. Finally, the use of percentages related to certain accident parameters (e.g. accident type) gives useful information without using any traffic data.

These methods can be used separately or in combination in order to overcome efficiently the difficulties, which are inherent to the exposure data available at national and international level. The current situation in road accident analysis at international level can thus be improved. However, these methods should be used with great attention if all conditions for their appropriate functioning are not fulfilled and the interpretation of their results should always be considered carefully in an attempt to get the most from existing data.

Finally, it is worth mentioning that some of the above-described insufficiencies of international exposure data could eliminated or limited - at least at European level - if some actions of data harmonisation took place at this level[15, 16]. A basic action, and not necessarily very difficult to implement, could be the adoption by all European countries of a common road accident data collection form containing uniform basic information on the accident, allowing for direct international comparisons. Furthermore, the adoption of common methodologies for traffic estimations by all European countries could also be very useful for the effective solution of comparability problems concerning exposure data. Possibly, the execution of frequent Europe-wide traffic surveys using a unique methodology could also be a positive approach for the availability of reliable and comparable traffic data at European level. All these harmonisation actions could be implemented progressively; first the common approach should be defined in detail, then each country could optionally implement it so that common data collection methods are used in all European countries after some years.

References

1. ETSC, (2001), Transport accident and incident investigation in the European Union. European Transport Safety Council, Brussels.
2. OECD, (1995), International road traffic and accident databases. OECD Seminar, Helsinki.
3. IRTAD, (1992), Definitions and data availability. OECD/IRTAD special report, BASt, Bergisch Gladbach, Germany.
4. TRB, (1993), Accident data quality: a synthesis of highway practice. Transportation Research Board, NCHRP Synthesis 192, Washington DC.
5. YANNIS, G., GOLIAS, J., FRANTZESKAKIS, J., (1996), Report on national road accident analyses in the EU countries. Journal of IATTS, Vol.20, No.2.
6. CETE SO, (1997), CAREPLUS: A proposal for improving the comparability of road accidents within EU member states using the CARE database. CAREPLUS consortium, Bordeaux.
7. DIRECCION GENERAL DE TRAFICO, (1997), Follow-up of traffic victims during the 30 days period after the accident. IRTAD- BASt, Bergish-Gladbach, Germany.
8. TRRL, (1997), Time interval between road accident and death 1985. Leaflet LF 1051, Transport Research Laboratory, Crowthorne, United Kingdom.
9. EUROPEAN COMMISSION, (2000), CARE - Community Road Accident Data Base - Summary Statistics. EC DG TREN E-3, Brussels.
10. GOLIAS, J., MATSOUKIS, E., YANNIS, G., (1997), An analysis of factors affecting road safety: the Greek experience. ITE Journal.
11. HAIGHT, F., (1973), Induced exposure. Accident Analysis and Prevention, Vol.5.
12. HODGE, G. A., RICHARDSON, A. J., (1985), The role of accident exposure in transport system safety evaluations II: Group exposure and induced exposure. Journal of Advance Transportation, Vol.19:2.
13. GEORGIOPOULOS, S., (1997), Investigation of driver's risk in relation their blood alcohol concentrations. Diploma thesis at the Department of Transportation Planning and Engineering of the National Technical University of Athens, Athens.
14. NATIONAL TECHNICAL UNIVERSITY OF ATHENS / DTPE, (1996), Current and future potential of a European road accident data base with dissaggregate data. NTUA/DTPE, Athens.
15. YANNIS, G., GOLIAS, J., KANELLAIDIS, G., (1998), A comparative analysis of the potential of international road accident data files. IAATS Research, Vol. 22 No. 2.
16. ETSC, (2001), EU transport accident, incident and casualty database: -current status and future needs. European Transport Safety Council, Brussels.