CADaS - A common road accident data framework in Europe

George Yannis	Petros Evgenikos	Antonis Chaziris
Assistant Professor	Research Associate	Researcher
National Technical	National Technical	National Technical
University of Athens	University of Athens	University of Athens
Athens, Greece	Athens, Greece	Athens, Greece
geyannis@central.ntua.gr	pevgenik@central.ntua.gr	achaz@central.ntua.gr

Abstract

The objective of this research is the formulation of a common framework for road accident data collection across European countries. In this way, harmonisation of accident data at both national and EU level could be achieved and progressively, additional comparable road accident data from all EU countries will be available, allowing for more interesting and complete analyses. Information on existing national road accident data collection systems was gathered, and the needs of the various stakeholders in different EU countries were identified, in order to define the necessary data for road accident analysis. The national collection systems from all EC member states were analysed and through an iterative process, taking into account both data availability and usefulness, and the National Experts views, the recommendation for a Common Accident Data Set was formulated, consisting of a minimum set of standardised data elements, which will allow for comparable road accident data to be available in Europe. The adoption of this Common Accident Data Set is flexible according to any national needs and/or particularities and it could be a solid basis for the development of the respective common data set to be used at global level.

Key-words : Road accident data; road accident data collection; road safety; variables; values.

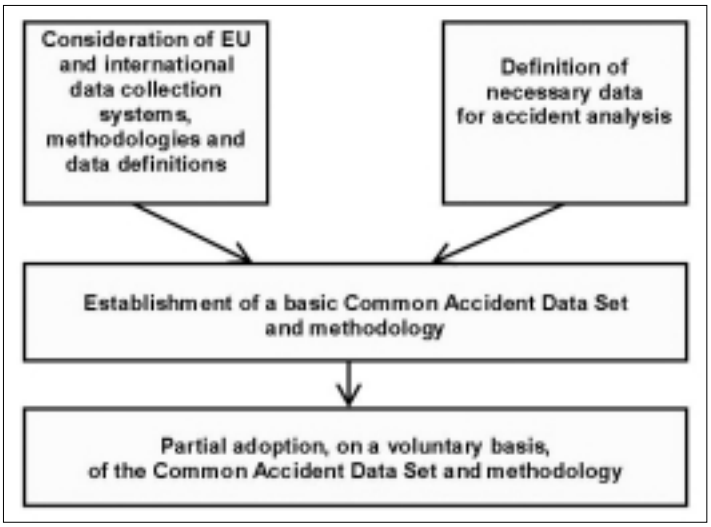
Introduction

Existing European road accident data are not always comparable among the various countries, mainly due to the different national accident data collection systems (Golias et al., 2001). Data variables and values are currently collected under different definitions in the EU countries, the various accident data collection forms have different structures and the relevant data fill-in systems cannot be compared (Yannis et al., 1996, Frantzeskakis et al., 1995). Both accident data quality and availability are affected and consequently, data analyses and comparisons among the various EU countries are not always reliable, even for some of the common variables and values of CARE, the European road accidents' database with disaggregate data (Yannis, 2000).

Harmonisation of accident data at national level (apart from the EC level) could be very beneficial for road accident analysis, using more common variables and values across the European countries, and in view of improving the accident data compatibility throughout Europe, a Common Accident Data Set (CADaS) and methodology were established, to be used by any EU country that wishes to update their national road accident collection system (NTUA, 2008).

A two-stage approach, which takes into account both data availability (bottom-up approach) and usefulness (top-down approach), was adopted as it can be seen at the following Figure 1 (NTUA, 2008). On one hand, the data required for road accident analysis in several EU countries was identified and on the other hand, the current potential of the national data collection systems was recorded (ETSC, 2006). The basic common accident data collection set and methodology were derived through an iterative process that took into account both data availability and usefulness, with the participation of experts and Governmental representatives.

Figure 1: CADaS methodology



In order to establish a basic accident data collection set and methodology, information concerning the existing national collection systems, as well as the identification of the needs for road accident data are required. A questionnaire to collect information about the national accident collection forms, methodologies and data definitions in all EU countries was initially prepared and at a next phase, the recording and examination of national road accident data took place. Data elements, as well as the respective definitions used in each national system, were gathered and analysed in order to identify good practices in general, but also detailed variables and values for accident analysis.

Moreover, the identification of the needs for road accident data was considered important for the establishment of a concrete proposal and on that purpose, the needs of the main stakeholders from several EU countries were recorded through an appropriate Grid. By filling-in this Grid for several stakeholders, the maximum needs were defined for each country and these were further compared, in order to identify the minimum/common needs for all countries examined.

National data collection systems

The questionnaire that was prepared to collect information about the national accident collection forms, methodologies and data definitions in all EU countries concentrated mainly on the national road accident collection forms and collection methodologies and information from 25 European countries were provided. It was divided into four different sections: National road accident collection system, National road accident data validation, Underreporting and Road accident data analysis. Several conclusions have been drawn about the current situation on accident data collection in European countries, based on the answers received.

Regarding the length of time over which the collection systems have been developed, the content of national databases, the national procedures established for reviewing the system and the institutional arrangements which facilitate its successful operation several differences were identified among the countries. Original introduction of national road accident databases and collection forms varied widely among the participating European countries and a number of changes such as definition changes, incorporation of additional information, i.e. on alcohol consumption, more detailed information, usage of electronic collection form and linkage of databases were recorded in many European countries, affecting both the collection form and the respective database. Revisions are rarely undertaken in most European countries, only when it is considered necessary and in some of them no review procedures were recorded.

In all countries, the authority responsible for entering national road accident data is the Police force, while other authorities are ordinarily responsible for maintaining road accident databases. However, more than one authority is responsible for developing the respective database, which may interfere with the potential harmonisation of the collection procedures among the European countries. Furthermore, more than one road accident database is maintained in many countries, mainly by hospitals, insurance companies, or Ministries and these are rarely linked.

Even though recording of road accident data is an important task of the Police, no such specific training takes place. Road safety data collected by the Police at National level are transferred to the responsible authority and entered in the respective database. Some countries apply a time limit to this data transfer, while in other countries there is no such limit. Usually, data entered in the databases relate to casualties with fatal, serious or slight injuries, while additional data are collected in some countries and incorporated into the respective databases.

Electronic filling-in of accident forms would not only accelerate the entire collection procedure, and thus improve it, but also minimise the potential for data inconsistency. Electronic methods to collect data are already used in some European countries, and electronic data entry across Europe could additionally facilitate harmonisation of the form and the related collection procedures at European level.

As far as validation of the road accident data's accuracy and reliability is concerned, among the most important data limitations identified in many European countries is the issue of underreporting, which mainly affects accidents in which vulnerable users are involved (pedestrians, two-wheelers, etc.) and accidents that occurred outside urban areas. Furthermore, in most countries certain items of road accident data are missing or are considered to be of low quality. These limitations interfere with reliable and constructive comparisons at both National and European level.

In most European countries the data collected at national level are considered reliable, as far as road accident fatalities are concerned, while the reliability of data for the generality of road accidents is mostly considered of medium or in some cases of low quality. The data that are most frequently missing relate to material damage only accidents, information on use of drugs and BAC level, or the existence of safety equipment (airbags, seatbelts, etc.), and data on speed. Most countries apply a validation methodology that mainly consists of either internal validation checks in the database, checks for illogical or contradictory statements, cross-checking with other databases, or a combination of all these.

Currently, the only figures that could be considered comparable at international level are those referring to fatal accidents, as only these are considered reliable and have common definition among the European countries. Figures relating to accidents and injuries cannot be considered either reliable or comparable at international level, due to different definitions used as well as underreporting. Within the questionnaire information on definitions of road accident injuries, levels of underreporting, as well as relevant recent national studies were collected.

Even though underreporting is considered as a major limitation in most countries, it has not been thoroughly investigated over the last decades. Recent underreporting studies are available in very few participating countries, mostly regarding the number of fatalities (Petridou et al., 2009). However, plans for underreporting studies are under way in several countries, in some cases consisting of linkage of or comparison between road accident and medical databases.

The main national road accident statistics are usually published annually in both electronic form and hardcopy. Furthermore, in almost all of the participating countries national road accident data are also available at the internet, while in many cases an English version of these data is available. In most European countries, disaggregate data analysis is usually carried out by the competent authority, as well as other organisations that have access to the disaggregate data files.

Regarding the main users of road accident data, a variety of users were identified. National public administration, research and scientific institutions, as well as accident involved bodies are ordinarily the main users of road accident data. At the same time, industries and professional associations are the least common accident data users. Furthermore, in all countries road accident data are used by local / regional authorities, in order to improve local / regional road safety, by carrying out high risk sites analyses, cost benefit analyses, or design educational campaigns.

As far as the improvement of road safety analysis is concerned, it is necessary to improve road accident data quality in order to improve analyses. Thus, common definitions should be used throughout Europe and exposure data should be collected at national level in each country. At the same time, the use of electronic collection forms could contribute to improving data quality. Furthermore, easier access to disaggregate data files should be provided, while accident databases should be linked with medical and other relevant ones. This will ensure better exploitation of road accident data. Other important actions for improving road safety analysis include better cooperation among all involved authorities and better funding for studies of road safety issues.

Moreover, recording and examination of the national road accident data took place during the second phase of the project. Variables and values as well as the respective definitions used in each national system were collected and examined in order to identify good practices in data recording (how data are entered, whether several values may be attributed to the same variable etc) as well as identify important variables and values for accident analysis. The results from this study were exploited for the formulation of a recommendation for a common accident data set.

National road accident data collection systems differ significantly, as some countries collect more information than others, but even the common accident data are not always compatible among the countries, due to different definitions of the collected variables and values. Moreover, the type and the level of detail of the collected data are different in each national system. Consequently, at a next phase a thorough examination of national road accident data collection systems was carried out. Specific information on the collected data (national road accident data collection forms), the way these data are recorded (instructions for the completion of the collection forms), detailed description of the data elements (list of variables and values) and detailed data definitions were gathered from 26 countries in both native and English language for completeness reasons. In this way the most commonly used variables and values for accident data collection and analyses throughout Europe were identified, as well as useful features in national collection systems, including data elements, methodologies, fill-in systems etc. and were exploited for the formulation of the proposal for a Common Accident Data Set. Finally, all variables and values that are currently included in the CARE database were also used for the identification of necessary data, but other international data files were also considered (US- MMUCC, WHO) (NHTSA, 2008).

Identification of necessary data

Identification of the needs for road accident data is important for the establishment of a concrete recommendation for a uniform road accident data set at EU level. On that purpose, the needs of the main stakeholders from seven EU countries were recorded. According to the specific circumstances in each country and the specific needs of each stakeholder, different needs were expected to be identified, thus this activity took place at both national and local level. The main interest groups were Public Services (Police, Hospitals etc), Central Governmental Authorities (Transport, Health), Local Governmental Authorities, Research Institutes and Industry (including transport associations). An appropriate Grid was developed to establish a list of various stakeholders by country and then identify their needs for accident data.

Based on the answers received, a ranking of the identified variables was performed, by calculating the number of stakeholders (percentage of the stakeholders) in all European countries who considerably use these variables. At the same time, all identified stakeholders were also ranked according to the average number of road accident variables that they seem to use frequently (High use).

Analysis showed that road accident variables related to the category of accident type seem to be more important for stakeholders in general, as on average they are highly used more frequently, followed by the respective of the road environment and road user category. Furthermore, variables regarding the vehicle category are not considered highly important, with the exception of the vehicle type variable which is widely used by most of the stakeholders, followed by information on the existence of security equipment in the vehicle.

Among the variables related to the road user category, age, gender, injury severity, and person class are frequently used by the stakeholders, unlike nationality and hit and run accidents. As far as the road environment variables are concerned, speed limits and road and area type are among the most highly used road accident data, while lighting conditions and weather conditions seem to be considered less important by most of the stakeholders.

Variables regarding accident type (accident type, collision type, pedestrian movement, vehicle manoeuvre) seem to be considered as important, as they are all highly used by more than 50% of stakeholders.

In the following Table 1, road accident variables are absolutely ranked, regardless their category, according to their average ratio allowing identifying which variables are more frequently used by the various stakeholders, independently of the country. 22 out of 29 variables are highly used by half or more of the stakeholders in the six countries, while it seems that the variables mostly used by the stakeholders (over 70% of the stakeholders use them) are those related to the road user's age, gender, injury severity and accident type, as well as to the collision type. Other important road accident variables for the various stakeholders (over 50% of the stakeholders use them) concern: person class, vehicle type, speed limits, road and area type, alcohol/drug test, road surface conditions, region, junction control, security equipment, road markings, junction type, number of lanes, vehicle manoeuvre, carriageway type, lighting conditions and pedestrian movement. At the same time, the least used variables when conducting road accident data analyses, are related to driving licence age, nationality and hit and run accidents.

As far as the several stakeholders are concerned, grid analysis revealed that the same breakdown of road accident data stakeholders is not applicable in all countries. For example 25 stakeholders were identified in Greece, whereas in the United Kingdom there were only 3 important ones (Ministry of Transport, Police, and Local Authorities). Additionally, the total number of accident variables used by each stakeholder in the same country varies significantly, according to their domain of interest and also the extent of the data analyses performed.

The importance of the stakeholders, according to the frequency and the extent of accident data use, was further considered. The analysis of the Grid for several stakeholders, allowed for defining the maximum needs for each of the examined countries and these were further compared, in order to identify the minimum/common needs for road accident data.

Table 1. Number of road accident data stakeholders per road accident variable and country - ranked according to average ratio

	Austria					Greece					Netherlands					Hungary					Czech Republic					United Kingdom					
Variables	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	High	Low	-	Total*	Ratio	Av. Ratio
Age	13	1	9	23	57%	16	6	3	25	64%	12	0	0	12	100%	10	1	0	11	91%	23	1	0	24	96%	3	0	0	3	100%	79%
Sex	13	1	9	23	57%	14	6	5	25	56%	12	0	0	12	100%	7	4	0	11	64%	23	1	0	24	96%	3	0	0	3	100%	73%
Injury severity	13	1	9	23	57%	15	7	3	25	60%	6	0	6	12	50%	10	0	1	11	91%	23	1	0	24	96%	3	0	0	3	100%	71%
Accident type	12	1	10	23	52%	15	7	3	25	60%	11	0	1	12	92%	10	0	1	11	91%	19	4	1	24	79%	3	0	0	3	100%	71%
Collision type	12	1	10	23	52%	15	7	3	25	60%	11	0	1	12	92%	9	1	1	11	82%	19	4	1	24	79%	3	0	0	3	100%	70%
Person class	13	1	9	23	57%	16	6	3	25	64%	10	0	2	12	83%	3	1	7	11	27%	23	1	0	24	96%	3	0	0	3	100%	69%
Vehicle type	13	1	9	23	57%	13	7	5	25	52%	11	0	1	12	92%	11	0	0	11	100%	16	8	0	24	67%	3	0	0	3	100%	68%
Speed limits	12	2	9	23	52%	12	4	9	25	48%	10	0	2	12	83%	9	0	2	11	82%	20	0	4	24	83%	3	0	0	3	100%	67%
Road type	11	2	10	23	48%	12	6	7	25	48%	10	0	2	12	83%	9	0	2	11	82%	20	4	0	24	83%	3	0	0	3	100%	66%
Area type	11	2	10	23	48%	16	5	4	25	64%	10	0	2	12	83%	7	2	2	11	64%	20	4	0	24	83%	1	0	2	3	33%	66%
Alcohol/drug test	13	2	8	23	57%	14	8	3	25	56%	9	0	3	12	75%	10	0	1	11	91%	14	9	1	24	58%	3	0	0	3	100%	64%
Road surface conditions	12	1	10	23	52%	9	8	8	25	36%	10	0	2	12	83%	8	1	2	11	73%	17	3	4	24	71%	2	0	1	3	67%	59%
Region	11	2	10	23	48%	15	4	6	25	60%	11	0	1	12	92%	1	6	4	11	9%	16	6	2	24	67%	2	1	0	3	67%	57%
Junction control	11	2	10	23	48%	10	4	11	25	40%	9	1	2	12	75%	8	1	2	11	73%	15	5	4	24	63%	2	0	1	3	67%	56%
Security equipment	13	1	9	23	57%	15	6	4	25	60%	8	1	3	12	67%	9	0	2	11	82%	9	9	6	24	38%	0	0	3	3	0%	55%
Road markings	11	2	10	23	48%	10	3	12	25	40%	9	1	2	12	75%	6	3	2	11	55%	18	2	4	24	75%	0	0	3	3	0%	55%
Junction type	11	2	10	23	48%	10	6	9	25	40%	9	1	2	12	75%	6	2	3	11	55%	15	5	4	24	63%	2	0	1	3	67%	54%
Number of lanes	11	2	9	23	48%	9	4	12	25	36%	9	1	2	12	75%	6	3	2	11	55%	16	4	4	24	67%	2	0	1	3	67%	54%
Vehicle manoeuvre	12	1	10	23	52%	14	8	3	25	56%	9	1	2	12	75%	6	3	2	11	55%	9	10	5	24	38%	2	0	1	3	67%	53%
Carriageway type	11	2	10	23	48%	9	4	12	25	36%	8	2	2	12	67%	6	3	2	11	55%	15	5	4	24	63%	2	0	1	3	67%	52%
Lighting conditions	12	1	10	23	52%	12	6	7	25	48%	10	0	2	12	83%	8	1	2	11	73%	6	13	5	24	25%	2	0	1	3	67%	51%
Pedestrian movement	11	2	10	23	48%	14	8	3	25	56%	10	0	2	12	83%	8	1	2	11	73%	5	14	5	24	21%	2	0	1	3	67%	51%
Psychophysical circumstances	13	1	9	23	57%	13	9	3	25	52%	2	0	10	12	17%	6	3	2	11	55%	12	11	1	24	50%	0	0	3	3	0%	47%
Weather conditions	12	1	10	23	52%	9	7	9	25	36%	10	0	2	12	83%	7	2	2	11	64%	6	13	5	24	25%	2	0	1	3	67%	47%
Vehicle age	13	1	9	23	57%	13	6	6	25	52%	2	0	10	12	17%	7	1	3	11	64%	9	12	3	24	38%	0	2	1	3	0%	45%
Car passenger type	12	1	10	23	52%	11	8	6	25	44%	2	1	9	12	17%	3	6	2	11	27%	12	12	0	24	50%	3	0	0	3	100%	44%
Driving license age	13	1	9	23	57%	11	9	5	25	44%	5	1	6	12	42%	5	4	2	11	45%	8	7	9	24	33%	0	0	3	3	0%	43%
Nationality	12	2	9	23	52%	11	10	4	25	44%	6	1	5	12	50%	4	5	2	11	36%	7	14	3	24	29%	0	0	3	3	0%	41%
Hit and run	12	1	10	23	52%	11	9	5	25	44%	2	0	10	12	17%	3	5	2	11	27%	5	6	13	24	21%	1	1	1	3	33%	35%

Total number of national road accident data users

Scale of significance: H high importance L low importance - Not used

A recommendation for a Common Accident Data Set

After thorough co-examination of all information collected and analysis results of the national collection systems and the Grids, the formulation of a complete recommendation for a Common Accident Data Set (CADaS) was carried out though an iterative process, considering both data availability and usefulness.

The recommendation for a Common Accident Data Set consists of a minimum set of standardised data elements, which will allow for comparable road accident data to be available in Europe. In this way, more variables and values with a common definition will be added to those currently included in the CARE database, maximising thus the potential of CARE and allowing for more detailed and reliable analyses at European level. CADaS is structured in a simple way, without levels of hierarchy, constituting in fact the record layout of the data set to be transferred to the EU (NCHRP, 2007). It also refers to the set of data to be voluntarily transmitted by each country to the EU, which should be derived from the national road accident data collection system. Moreover, the variables and values of CADaS may be considered as recommendations for national police road accident data collection reports.

CADaS consists of 73 variables and 471 values. The selection of these variables and values resulted from the balanced co-consideration of some basic criteria, taking into account that variables and values must be comprehensive, concise and useful for road accident analysis at EU level, the level of detail of the variables and values should correspond to all data useful for macroscopic data analysis and that each country should have the possibility to choose alternative level of detail of the various variables and values. Data which are impossible or very difficult to be collected are not retained in the CADaS, however, the future perspective of using certain variables and values was also taken into account, even though those data are not currently collected by most of the countries. Existing CARE variables and values are of first priority within CADaS and additionally, CADaS variables and values refer to casualty road accidents.

The CADaS variables are divided into four basic categories: Accident related variables, Road related variables, Traffic unit related variables and Person related variables. Several variables include two distinct types of values, referring to different level of detail: Detailed values, concerning information at the highest level of detail and alternative values, concerning information at a more aggregate level of detail, when more detailed values are not available.

The number of variables and values contained in the CADaS are presented at the following Table 2:

Table 2: CADaS variables and values in numbers

Variable	Code	Number of Variables			Number of Values		
category		High (H) importance	Lower (L) importance	Total	Detailed values	Alternative values (A)	Total
Accident	A	7	5	12	86	13	98
Road	R	11	15	26	106	13	119
Traffic Unit	U	7	10	17	137	15	152
Person	P	11	7	18	91	10	102
Total		36	37	73	420	51	471

In the following Table 3 the variables included in CADaS are presented, grouped into the four basic categories:

Table 3: CADaS variables

PERSON VARIABLES	ROAD VARIABLES	TRAFFIC UNIT VARIABLES	PERSON VARIABLES
A-1 ACCIDENT ID	R-1 LATITUDE	U-1 TRAFFIC UNIT ID	P-1 PERSON ID
A-2 ACCIDENT DATE	R-2 LONGITUDE	U-2 TRAFFIC UNIT TYPE	P-2 AGE
A-3 ACCIDENT TIME	R-3 E-ROAD	U-3 VEHICLE SPECIAL FUNCTION	P-3 GENDER
A-4 NUTS	R-4 E-ROAD KILOMETRE	U-4 TRAILER	P-4 NATIONALITY
A-5 LAU	R-5 FUNC. CLASS - 1st ROAD	U-5 ENGINE POWER	P-5 ROAD USER TYPE
A-6 WEATHER CONDITIONS	R-6 FUNC. CLASS - 2nd ROAD	U-6 ACTIVE SAFETY EQUIPMENT	P-6 INJURY SEVERITY
A-7 LIGHT CONDITIONS	R-7 AADT - 1st ROAD	U-7 VEHICLE DRIVE	P-7 ALCOHOL TEST
A-8 ACCIDENTS WITH PEDESTRIANS	R-8 AADT - 2nd ROAD	U-8 MAKE	P-8 ALC. TEST SAMPLE TYPE
A-9 ACCIDENTS WITH PARKED VEHICLES	R-9 SPEED LIMIT - 1st ROAD	U-9 MODEL	P-9 ALCOHOL TEST RESULT
A-10 SINGLE VEHICLE ACCIDENTS	R-10 SPEED LIMIT - 2nd ROAD	U-10 REGISTRATION YEAR	P-10 ALCOHOL LEVEL
A-11 AT LEAST TWO VEHICLES - NO TURNING	R-11 MOTORWAY	U-11 TRAFFIC UNIT MANOEUVRE	P-11 DRUG TEST
A-12 AT LEAST TWO VEHICLES - TURNING OR CROSSING	R-12 URBAN AREA	U-12 FIRST POINT OF IMPACT	P-12 DRIV. LICENSE ISSUE DATE
	R-13 JUNCTION	U-13 FIRST OBJECT HIT IN	P-13 DRIVING LICENSE VALIDITY
	R-14 REL TO JUNCTION / INTERCH.	U-14 FIRST OBJECT HIT OFF	P-14 SAFETY EQUIPMENT
	R-15 JUNCTION CONTROL	U-15 INSURANCE	P-15 POSITION IN/ON VEHICLE
	R-16 SURFACE CONDITIONS	U-16 HIT & RUN	P-16 DISTRACTED BY DEVICE
	R-17 OBSTACLES	U-17 REGISTRATION COUNTRY	P-17 PSYCOPHYS / PHYS. IMP.
	R-18 CARRIAGEWAY TYPE		P-18 TRIP/JOURNEY PURPOSE
	R-19 NUMBER OF LANES		
	R-20 EMERGENCY LANE		
	R-21 MARKINGS		
	R-22 TUNNEL		
	R-23 BRIDGE		
	R-24 WORK ZONE RELATED		
	R-25 ROAD CURVE		
	R-26 ROAD SEGMENT GRADE		

For each of the variables included in the CADaS, the following information is presented:

Variable Label : The label of the proposed variable, consisting from the category identifier (A, R, U or P), the numbering and the name of the variable. The importance of the variable for road safety analysis is also added: (H) for variables of high importance and (L) for variables of lower importance.

Variable definition and scope : A brief description of the variable is provided, followed by the importance and usefulness of the variable, explaining the rational lying behind its selection.

List of values : The attribute values to each variable are listed.

Value labels : Each value is identified by the code of the variable, followed by a number which corresponds to each value and its name. The (A) code is added next to the variable category code for the alternative value, when is the case.

Value definitions: The definition of each value of the variable is provided, indicating also any particularities of the value and any relevant assumptions regarding its collection.

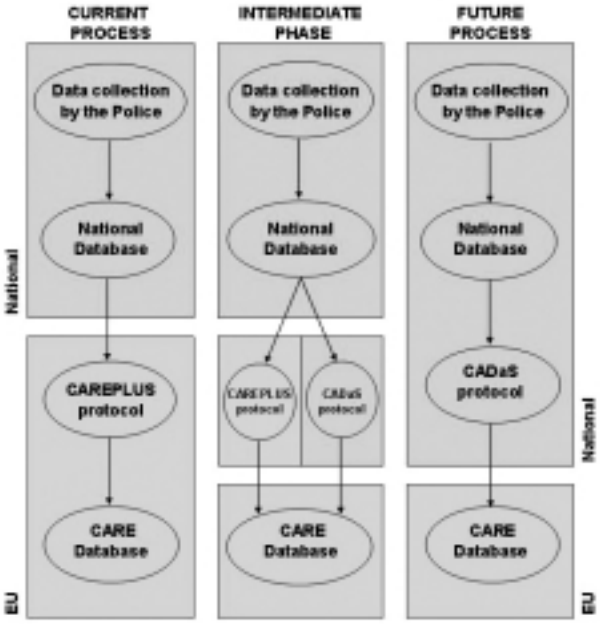
Data Format : The way in which each variable has to be provided. Data formats concern:
- -the possibility to attribute one or more values to a variable,
- -the format of the value (code, number, text).

Adoption of the common data collection set and methodology

The adoption of the CADaS recommendation by the European countries is a very important step towards the success of this Task. One of the CADaS advantages is that it can be adopted gradually by EU countries, without presupposing any changes in a country's national data collection system; however, any part of it (variables, values, definitions and data formats) can be implemented within an existing national collection system, increasing thus the compatibility of the national road accident data with the respective CARE data. If one country decides to start using the CADaS protocol, it can transform its national data into the CADaS data by using appropriate transformation rules and eventually transmit the transformed data to the EC. Consequently, the level of adoption of the CADaS can vary according to any national needs and/or particularities and can be performed during any time in the future.

In the following Figure 2, the current, intermediate and future (based on the CADaS adoption) processes of the national road accident data files are presented. Using both (current and future) approaches ensures compatibility of the accident data among EU countries and the main difference of these two approaches is related to the degree of involvement of the country in the process.

Figure 2 : Accident data transformation processes



Taking into account that many Member States may partially adopt CADaS, an intermediate phase is also necessary, during which, countries may use a part of the CADaS in order to transform specific variables and values at national level and transmit the rest of the data in the current format in order to be transformed using the CAREPLUS protocol. According to the proposed future process, transformation of the national accident data (based on the CADaS protocol) will be performed at the national level and the derived CADaS variables and values will be transmitted to the EC, where they will be included in a more automatic way into the CARE database. This process will allow for more common variables and values but also for higher quality, given that the national authorities better perceive any particularities related to national data collection, and subsequently can better identify the interrelation between the collected and the CADaS variables.

Conclusion

A holistic approach was used for the formulation of a common framework for road accident data collection across European countries, based on thorough examination of both data and systems availability, and usefulness to meet the needs of the different stakeholders for accident data analyses (OECD, 2000).

Initially, information on existing national road accident data collection systems was gathered, allowing for the exploitation of available experience, but also the identification of any specific requirements in some countries. At the same time, the needs of the main stakeholders in different EU countries were identified, in order to define the necessary data for road accident analysis. The national collection systems from all EC member states were analysed and through an iterative process, and the feedback received by several national road accident data Experts, the recommendation for a Common Accident Data Set was formulated, a complete proposal consisting of 73 variables and 471 values, grouped into four basic categories (accident, road, traffic unit, person), for which the scope and definition are provided, but also the format of the data, as well as the structure of the Data Set are defined. Having a flexible format, the Common Accident Data Set can be adopted gradually by the countries, but any part of it (variables, values, definitions and data formats) can be implemented within an existing national system according to any national needs and/or particularities.

Introducing this minimum set of standardised data elements, harmonisation of accident data at both national and EU level could be achieved and progressively, additional comparable road accident data from all EU countries will be available, allowing for more interesting and complete analyses (Yannis et al., 1996). Moreover, for countries that have recently introduced a road accident data collection system or consider revisions and improvements to their existing national system, CADaS can provide an invaluable tool for benchmarking road accident data collection and subsequently improve the efficiency of a national data collection system.

As a further step, a pilot phase for the implementation of the CADaS could be considered. Such a pilot can enable the identification of possible weaknesses in the recommendation that could only be tracked through an actual application of the system; as a result, it would significantly strengthen the content of the Common Accident Data Set. This pilot phase could be implemented in countries wishing to revise their national systems or countries with less experience in road safety wishing to exploit the experience of other countries through the CADaS recommendation.

CADaS could prove to be a solid basis for the development of the respective common data set to be used at global level (World CADaS). International organisations such as WHO, UN, etc, are planning such initiatives for developing countries, which could be further extended worldwide. Its adoption should be supported by many different parties in order to maximise its acceptance by the countries and national stakeholders at various levels could contribute by promoting the recommendation in their own countries.

Acknowledgment

This paper is based on work carried out within the scope of the SafetyNet project (The European Road Safety Observatory) of the 6th Framework Program for Research, Technological Development and Demonstration of the European Commission (Thomas et al., 2005).

References

1. Cete-Sud Ouest (2000). “A proposal for improving the comparability of road accidents within the EU Member States using the CARE Database”, CAREPLUS 2 Consortium, Bordeaux.
2. ETSC (2006). “Road accident data in the enlarged European Union - Learning from each other”. Report prepared by the ETSC Data Working Party, European Transport Safety Council, Brussels.
3. Frantzeskakis J., Golias J., Yannis G. (1995). “National and international official publications for road accident statistics”, Proceedings of the international conference “International road traffic and accident databases”, OECD, Helsinki, pp. 121-133.
4. Golias J., Yannis G., “Dealing with lack of exposure data in road accident analysis”, Proceedings of the 12th international conference “Traffic Safety on three Continents”, Federation of European Road Safety Institutes, Transportation Research Board, Moscow, September 2001.
5. National Cooperative Highway Research Programme (NCHRP) - (2007). “Technologies for Improving Safety Data - A Synthesis of Highway Practice”, Transportation Research Board (TRB) 2007 Executive Committee, Washington D.C.
6. National Highway Traffic Safety Administration (NHTSA) - (2008). “Model Minimum Uniform Crash Criteria Guidelines - Third Edition”, U.S. Department of Transportation, By the Governors Highway Safety Association, Washington D.C.
7. NTUA - National Technical University of Athens (2008). “CADaS - The Common Accident Data Set”, Deliverable 1.14 of SafetyNet WP1 - Task 1.4, SafetyNet Consortium, Athens.
8. OECD (2000). “International seminar on Road Traffic and Accident Data needs for the new century”, OECD seminar, Vienna.
9. Petridou E., Yannis G., Terzidis A., Dessypris N., Germeni E., Evgenikos P., Tselenti N., Chaziris A., Skalkidis I. (2009). “Linking Emergency Medical Department and Road Traffic Police casualty data: a tool in assessing the burden of injuries in less resourced countries “, Traffic Injury Prevention, Vol. 10.
10. Thomas P., Morris A., Yannis G., Lejeune P., Wesemann P., Vallet G., Vanlaar W. (2005). “Designing the European Road Safety Observatory”, International Journal of Injury Control and Safety Promotion, Vol. 12, No. 4, pp. 251-253.
11. Yannis G., Golias J., Frantzeskakis J., (1996). “Report on national road accident analyses in the EU countries”, Journal of International Association for Traffic and Safety Sciences (IATSS), special feature on traffic accident analysis and traffic safety policies, Vol.20, No.2, pp. 60-68.
12. Yannis G. (2000). “Traffic and accident analysis priorities and data availability in European countries”, Proceedings of the international seminar “Road traffic and accident data needs for the new century”, OECD, Vienna.