EUROPEAN COMMISSION

DIRECTORATE GENERAL
HEALTH & CONSUMER PROTECTION
(SANCO)

FINAL REPORT

Transformation of
EHLASS Data for 1997-1999
and
for Latecomers for 1986-1996

Institute of Public Health
North Rhine–Westphalia
(lögd)

Agreement N^{o} SI2.134737
(99CVF3-314)
IPP 1040 (1999)

April 2001

EUROPEAN COMMISSION
DIRECTORATE-GENERAL
HEALTH & CONSUMER PROTECTION
(SANCO)

Injury Prevention Programme (IPP)

Final Report

Transformation of EHLASS Data for 1997-1999 and for Latecomers for 1986-1996

Institute of Public Health
North Rhine - Westphalia

Agreement N°.SI2.134737
(99CVF3-314)
IPP 1040 (1999)

Originator:	ISSUE DATE:	Version:
lögd		1.0

SUBJECT:

Transformation of EHLASS Data for 1997-1999 and

for Latecomers for 1986-1996

FINAL REPORT SI2.134737

SIGN OFF

Name	Position	Date	Signature
Dr.Helmut Friza	IPP Project Officer		

0. Summary and conclusions

0.1 Results of the project

In the EUPHIN HIEMS Project (Sub 50 10 26), the loegd collected EHLASS case data from the participating countries, assessed the quality of coding, transformed the data into the final data structure for EHLASS data, documented the transformation and uploaded the aggregated data into the database. This work was done with EHLASS data for the years 1986-1996. In the project "Transformation for EHLASS data for 1997-1999 and for latecomers 1986-1996" the EUPHIN HIEMS database was completed for the years 1997 to 1999 with the exception of data from Finland and the United Kingdom for the year 1999. The data collection and transformation process has up to now yielded more than 5 million accidents aggregated into 130 data files.
The main tasks to be solved were:

- Updating the aggregated data structure and adding the items: Date of injury, time of injury and sports
- Implementation of the conversion from IREQ 1991 to EHLASS version 1996
- Construction of conversion routines for NOMESCO and EHLASS v2000 versions
- Transformation of survey data for two countries
- Carrying out 30 Factory Acceptance Tests for Quality Assurance so that in all 45 FATs have been completed
- Documentation of the transformation process in reports and of the agreements reached with the participating countries
- Completion of the EUPHIN HIEMS Data Dictionary
- Storing all individual and aggregated EHLASS data and transformation routines on CD-ROM, sorted by countries

0.2 Implications on the programme (lessons learned)

- It should be ensured that the countries use as few country-specific codes as possible in order to unify the data collection and to minimise frictions by inconsistencies in coding.
- Meta-information about the collection methodology, hospitals involved and also about data holder and national responsibility was very rare. This information is essential for the further work with the data, e.g. statements on the representativity. For the future it is recommended to increase efforts for collecting meta-data. Special questionnaires should be used to support the data holders in their collection of meta information.

0.3 Added value for European Injury Monitoring

- First, it was possible to pool aggregated EHLASS data of the participating countries up to the year 1999. On this basis an analysis of aggregated and non-aggregated HLA data for 1986-1998 was possible (project MiaMa).
- A unique database on home and leisure accidents in Europe was built up and, in spite of limitations, can be used as a basis for analyses and development of strategies.
- The quality assurance tests (FATs) demonstrate that the quality of transformation is quite good. The countries agreed to the transformation routines. This is a good prerequisite for the further completion and use of the database.

Table 1: Overview of the data stored in EUPHIN HIEMS by countries, April 2001

Country	Short form of country name	Number of databases	Sum of accident case data per country
Austria	AT	4	39.823
Belgium	BE	9	168.067
Denmark	DK	11	738.724
France	FR	14	521.601
Finland	FI	4	36.649
Germany	DE	1	9.639
Greece	GR	10	207.700
Iceland	IS	2	49.464
Ireland	IE	9	80.866
Italy	IT	12	110.333
Luxembourg	LU	4	3.031
The Netherlands	NL	14	798.247
Norway	NO	4	81.554
Portugal	PT	12	496.747
Spain	ES	4	15.571
Sweden	SE	4	84.950
United Kingdom	UK	12	1.982.126
Total		**130**	**5.425.092**

For more details see annex XIII: Available EHLASS case data and aggregated data uploaded into the EUPHIN HIEMS database.

1. Introduction

Unintentional injuries consume a considerable amount of health care resources world-wide. The European Union is concerned with the public health problem, too. Beside the development of injury prevention measures, e.g. by a high standard of consumer protection, it was necessary to monitor the epidemiological situation for further studies. With the European Home and Leisure Surveillance System (EHLASS) which the European Union started in 1986 data in the field of home and leisure accidents (HLA) were collected. In addition to the countries of the European Union, Norway and Iceland participated in the EHLASS project. EHLASS was included into the Health Information Exchange and Monitoring Information Network (HIEMS), which belongs to the European Public Health Information Network (EUPHIN). The overall objective of EUPHIN HIEMS is the development of a community-wide system for the exchange of health data and indicators by telematic means. With the help of this network the EU Ministries of Health, the European Commission and relevant international institutions can exchange information in support of the Community´s Health Monitoring Programme (HMP) (see Final Documents on Data Transformation for EUPHIN HIEMS, subvention N°. 501026).
Within the HMP of the European Union, the Institute of Public Health in North Rhine-Westphalia (loegd) assumed the task of data transformation for the four different databases (mortality, demography, hospital data and EHLASS) of EUPHIN HIEMS (contract No. 501026) which was carried out from 1998 - 2000. As a part of the EUPHIN HIEMS project, the Institute of Public Health transformed EHLASS data for the years 1986 to 1996, stored them into the EUPHIN HIEMS database (the latest online version as of April 2001 is HIEMS version 1.1) and compiled a Data Dictionary for each country. In the follow-up project "Transformation of EHLASS Data for 1997-1999 and for latecomers for 1986-1996" (IPP-1040) contracted with DG SANCO (contract VS/1999/5294-DE) the data transformation work on EHLASS accident data was continued by loegd. EUPHIN HIEMS and IPP-1040 were part of the injury prevention projects of the European Commission.

1.1 Objectives of IPP-1040

It was the aim of the IPP-1040 project to complete the accident part of the HIEMS database with latecomers of the years 1986 to 1996 and with data available for the years under report 1997 to 1999 so that finally the EHLASS data from 1986 to 1999 of all participating countries (PC) were available in the transfer structure for uploading into the EUPHIN HIEMS database. Another objective was the completion of the Data Dictionary with meta-data from participating countries like collection methodology, participating hospitals, national responsibilities, national data item definitions. Also meta-data related to the data transformation had to be added to the Data Dictionary. The third objective was the data transformation of individual data for the future Injury Home and Leisure Accidents (IHLA) case database.

1.2 Tasks of the loegd

Due to the IPP-1040 project loegd was in charge of the following tasks:

- to complete the EHLASS database for latecomers for the years 1986 to 1996
- to add available data from the participating countries for the years 1997 to 1998 to the HIEMS database
- to collect and transform accident data for 1999 from all participating countries
- to ensure quality assurance standards by Factory Acceptance Tests (FAT)
- documentation of the data transformation
- to complete the Data Dictionary
- upload of aggregated data into the HIEMS database
- to support the participating countries in the transformation of national data for 1999 into the HIEMS structure, quality assurance (FAT) and documentation, if desired by the participating countries
- to prepare individual data for input into the IHLA database for accident case data and upload of these data
- revision of the HIEMS aggregation programme by extension to include three additional variables
- processing of IREQ version 1991 coded data

Work on the IPP-1040 project started after the contract had come into effect on March 15, 2000 and ended on April 30, 2001.

2. Changes to the data structures involved

Due to the increasing number of databases which contained all variables of EHLASS v96, loegd was confronted with the additional task of integrating these variables. Before all variables of EHLASS v96 could be processed, an update of the final EHLASS data structure became necessary. Inconsistencies between single categories of the final data structure and the EHLASS v96 data structure, the lack of single categories and reasons of practicability of the defined categories required these changes (for more details see annex I: Update of the aggregated data structure).

2.1 Update of the aggregated data structure

Home and leisure accident data were available in the EHLASS version 1986 case data structure and for later years under report also in the EHLASS version 1996 case data structure. The 1986 EHLASS case data version was transformed into the EHLASS 1996 case data version to achieve a uniform basis. In order to respect the privacy of the persons injured, the data collecting countries were requested to make their data anonymous. Therefore a new proposal for an anonymous data structure with so-called accident classes was introduced by loegd and the Community Services. At the meeting of the EHLASS working party on July 16, 1998 in Brussels, the proposal for the final data structure for EHLASS data, version 1.0, was accepted. This structure does not operate anymore with case data. Each record in this accident data model gives the overall number of a category of one accident class for one sex by age-groups. Every variable of the EHLASS case data finds its expression in an accident class of the aggregated data structure. The code lists (categories) of the aggregated data structure have been taken from the EHLASS version 1996 case data structure, as far as possible. Only the following variables have been aggregated:

- Mechanism of injury
- Place of accident
- Activity at the time of injury
- Part of body injured
- Product involved in the accident
- Product causing the injury
- Other product
- Sports

The aggregation of product codes and of the other variables was carried out differently. Product codes in the case data structure have a 5-digit code. They were aggregated in the main categories of product codes by reducing the codes to the first position. The same was done to 3-digit sports codes. For the other variables the aggregation was realised by bringing the variables' code to the main categories (from 2-digit to 1-digit).

For this data structure some changes in the code lists and to the characteristics of some accident classes became necessary for reasons of inconsistencies and practicability. The changes were:

- Re-definition of accident class 08 - "accident according to the length of time between injury and contact to emergency department"

- Re-naming of categories of accident class 13 - "accident according to sport activities and type of injury"
- Adding "unknown" to accident classes 8 and 9

These changes were settled at the meeting with the Advisory Group in Bielefeld on July 11, 2000. The topical version of the final data structure for EHLASS data is version 1.2 (see annex II: Final data structure of EHLASS data, version 1.2).

2.2 Extension of the working version

Case data of home and leisure accidents were aggregated into the "Final data structure for EHLASS data " for all participating countries by loegd.
The aggregation of EHLASS case data into the "Final data structure of EHLASS data" is based on the coding manual of the EHLASS 1996 version. The main part of available historical accident case data was coded in the EHLASS version 1986 structure. For this reason the aggregation in the EUPHIN HIEMS project was reduced to the available variables of the EHLASS 1986 version. The EHLASS v1986 data structure (see 3.1) was used as input structure for the aggregation programme. The variable "sports" does not belong to EHLASS version 1986. Over the last years under report, the number of databases which included the variable "sports" increased under the Latecomer project so that at least 13 databases contain the variable "sports". Following the update of the final data structure for EHLASS data, version 1.2, it became possible to process EHLASS data including the variable "sports". Therefore an update of the HIEMS aggregation programme was necessary. For this work loegd contracted Symmedia in Bielefeld (see also annex III: Extension of the aggregation programme). The extended aggregation programme needs as input EHLASS accident case data as a text-file coded with the EHLASS version 1996 code manual in the structure of the extended input structure (see table 2). The aggregation programme delivers output data which were in the transfer structure for upload into the HIEMS database respectively the final data structure for EHLASS data.

Table 2: Description of the extended input structure and coding of the aggregation programme for EHLASS case data

No.	Name of variables	Field	Structure	Type of data	Coding manual
1	Country code	2	1-2	Alphanumeric	1996
2	Hospital record number	10	3-12	Alphanumeric	1996
3	Date of treatment YYMMDD	6	13-18	Alphanumeric	1996
4	Time of attendance	2	19-20	Numeric	1996
5	Number of hospitalised days	2	21-22	Numeric	**1986**
6	Treatment and follow-up	1	23	Numeric	1996
7	Age of patient	3	24-26	Numeric	**1986**
8	Sex of patient	1	27	Numeric	1996
9	Accident mechanism	2	28-29	Numeric	1996
10	Site of accident	2	30-31	Numeric	1996
11	Activity at the time of accident	2	32-33	Numeric	1996
12	Injury diagnosis 1	2	34-35	Numeric	1996
13	Injury diagnosis 2	2	36-37	Numeric	1996
14	Part of the body affected 1	2	38-39	Numeric	1996
15	Part of the body affected 2	2	40-41	Numeric	1996
16	Product involved in an accident	5	42-46	Alphanumeric	1996
17	Product causing injury	5	47-51	Alphanumeric	1996
18	Other product	5	52-56	Alphanumeric	1996

19	Accident description	90	57-146	Text	1996
20	Date of injury YYMMDD	6	147-152	Alphanumeric	1996
21	Time of injury	2	153-154	Numeric	1996
22	Sports	3	155-157	Alphanumeric	1996

The following variables were added to the EHLASS version 1986 structure:

- Date of injury (position 20)
- Time of injury (position 21)
- Sports (position 22)

The variables “hospitalised days” and “age of patient” are EHLASS version 1986 variables. From the EHLASS version 1996 variables are computed as these values by:

- Age of patient = date of birth – date of injury
- Hospitalised days = date of attendance – date of discharge

All data processed by loegd were standardised to the above extended input structure or to the EHLASS version 1986 structure which can be distinguished from the extended structure by the missing of the positions 20 to 22.

The additional variables of the extended input structure for EHLASS case data allowed to generate additionally four classes of the aggregated structure. These classes had been part of the final data structure for EHLASS data, but they were not generated before the extension of the aggregation programme. The following classes are additionally available:

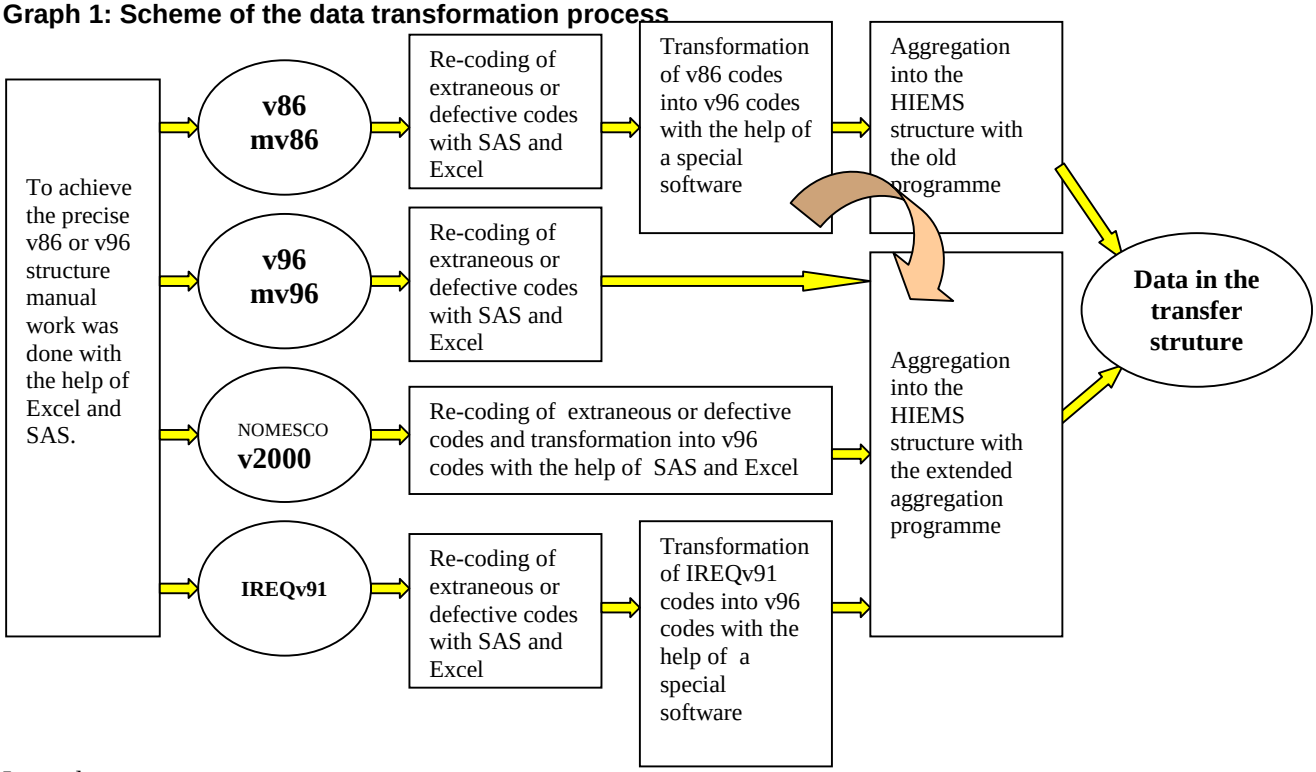
- Accident class 8 – accident according to the length of time between injury and contact to emergency department
- Accident class 12 – accident according to the time of injury
- Accident class 13 – accident according to sport activities and type of injury
- Accident class 18 – accident according to type of injury and length of time between injury and contact to emergency unit

Following the update of the aggregation programme it was possible to generate all 21 classes of the final data structure for EHLASS data.

3. Data transformation process

For the period from 1986 to 1999, loegd received 130 EHLASS case databases with more than 5 million cases gathered by the participating countries (see 3.1 and 4.1). This enormous amount of data had to be analysed with regard to their structure and coding. For standardising data to the input structure (see 2.2) to prepare the input of the aggregation programme it was necessary to re-code all defective codes into valid codes and to transform the EHLASS v1996 data structure into the input structure of the aggregation programme. In cases in which the original data were coded with the EHLASS v86 manual, the variables had to be transformed from EHLASS v86 code into the EHLASS v96 code as described in the final report on Transformation for EUPHIN HIEMS (graph 1). All re-coding done was documented. The documents were submitted for agreement to the data holder at least within the final reports (see annex VI: List of final reports).

Graph 1: Scheme of the data transformation process



Legend:

- v86 - EHLASS version 1986 manual
- mv86 – modified EHLASS version 1986
- v96 - EHLASS version 1996
- mv96 - modified EHLASS version 1996
- NOMESCO - NOMESCO classification for external causes of injuries
- V2000 - EHLASS version 2000
- IREQv91 – IREQ version 1991
- HIEMS data – data in the final data structure for EHLASS data, version 1.2 (transfer structure for the upload of aggregated accident case data into the EUPHIN HIEMS database)

It was possible to aggregate data in the EHLASS version 1986 structure with the extended aggregation programme by adding dummies (blanks) at positions 20 to 22 of the extended input structure.

3.1 Data structures involved

Home and leisure accident data are available in the EHLASS version 1986 case data structure and for later years also in the EHLASS version 1996 case data structure for the years under report 1986 to 1999. For the provision of data the EHLASS version 1986 structure was proposed by the Institute of Public Health, and this structure was defined as obligatory by the project steering committee (Final Documents on Data Transformation for EUPHIN HIEMS, subvention N°.501026). Data in the EHLASS version 1986 structure with coding of EHLASS version 1996 were also accepted. Below you find two brief descriptions of these relevant data structures (tables 3 and 4):

Table 3: Description of the structure and coding of the original EHLASS version 1986

No.	Name of variables	Fields	Structure	Coding definition
1	Country codes	2	1-2	A two-digit code for each country
2	Hospital record number	10	3-12	Letters or numbers
3	Date of treatment (YYMMDD)	6	13-18	Date when the patient came to the hospital
4	Time of attendance	2	19-20	From 00:00-24:00; min. not taken into account
5	Number of hospitalised days	2	21-22	Includes last day and first day if admission <20:00 hr
6	Treatment and follow-up	1	23-23	Kind of the first treatment-follow-up; 7 categories
7	Age of patient	3	24-26	In years. Under 2 years: a number two + # months
8	Sex of patient	1	27-27	Male, female and unknown
9	Accident mechanism	2	28-29	Includes 24 different two-digit code categories
10	Site of accident	2	30-31	Includes 39 different two-digit code categories
11	Activity at time of the accident	2	32-33	Includes 13 different two-digit code categories
12	Injury diagnosis 1 (Type)	2	34-35	Includes 18 different two-digit code categories
13	Injury diagnosis 2 (Type)	2	36-37	
14	Part of the body affected 1	2	38-39	Includes 47 different two-digit code categories
15	Part of the body affected 2	2	40-41	
16	Product involved in an accident	5	42-46	Includes a list of >1500 five-digit code categories
17	Product causing injury	5	47-51	
18	Other product	5	52-56	
19	Accident description	90	57-156	Brief description of how the accident occurred

Table 4: Description of the structure and coding of the original EHLASS version 1996 version

No.	Name of variables	Fields	Structure	Coding definition
1	Country codes	2	1-2	A two-digit code for each country
2	Hospital number	6	3-8	Only numbers
3	Case number	10	8-18	Letters or numbers. Free spaces filled with zeros
4	Sex of patient	1	19-19	Male, female and unknown
5	Date of birth	6	20-25	The reported birthday date
6	Date of injury (YYMMDD)	6	26-31	The reported date when the person was injured
7	Time of injury	2	32-33	From 00:00-24:00; min. not taken into account
8	Date of attendance (YYMMDD)	6	34-39	Date when the patient came to the hospital
9	Time of attendance	2	40-41	From 00:00-24:00; min. not taken into account
10	Treatment and follow-up	1	42-42	Kind of the first treatment-follow-up; 8 categories
11	Date of discharge (YYMMDD)	6	43-48	Date when the person left the hospital
12	Place of occurrence	2	49-50	1st digit – 10 major; 2nd digit – 80 minor categories
13	Mechanism of injury	2	51-52	1st digit – 10 major; 2nd digit – 71 minor categories
14	Activity	2	53-54	1st digit – 9 major; 2nd digit – 27 minor categories
15	Sports	3	55-57	1st letter-16 sports, 1st digit-73, 2nd-180 subcategories
16	Type of injury 1	2	58-59	Includes 22 different two-digit code categories
17	Type of injury 2	2	60-61	

18	Part of the body injured 1	2	62-63	1st digit – 9 major; 2nd digit – 52 minor categories
19	Part of the body injured 2	2	64-65	
20	Product involved in an accident	5	66-70	1st letter – 21 major groups
21	Product causing injury	5	71-75	1st letter + 1st digit – 77 major categories
22	Other product	5	76-80	1st letter + 1st digit + 3-5th digits – >1500 categories
23	Accident description	120	81-200	Brief description of how the accident occurred

The EHLASS version 1996 case data structure contains some additional variables which did not appear in EHLASS v86:

- Hospital number
- Date of injury
- Time of injury
- Date of discharge
- Sports

The following variables from EHLASS version 1986 are no longer part of the EHLASS version 1996 data structure, but they are still computable (see also 2.2).

- Age of patient
- Number of hospitalised days

The EHLASS version 1986 data structure was used as input structure for the aggregation programme. The data at this common basis were changed in the coding of the variables from EHLASS version 1986 to version 1996 before they were put into the aggregation programme. This data transformation was described in the Final report on Transformation for EUPHIN HIEMS. The extension of the input structure by three additional variables (see 2.2) allowed the procession of the variable “sports”. All accident data were standardised to this input structure by the loegd. Neglecting coding and the question of comparability, it would be possible to pool all available case data in the input structure for the HIEMS aggregation programme.

Despite the data structure and coding defined for the provision of accident data, not all participating countries have the means to precisely stick to these limitations with regard to codes and structure. Therefore the data from all participating countries had to be analysed with regard to their structure and coding. Some data delivering countries support this work by supplying structural descriptions and proposals for the re-coding of extraneous codes. In other cases the structure had to be reconstructed by identifying the variables and re-organising the structure or it was necessary to extract the EHLASS variables from another data structure which was used by the data holder for national purposes. For most of the databases, a re-arrangement of variables was necessary to achieve an EHLASS v86 structure or EHLASS v96 structure. In the next step of the data procession an agreement about extraneous and defective codes had to be found with the data-holder, as far as possible. For all participating countries the re-coding carried out by loegd was agreed with the corresponding national EHLASS authority. These transformations were documented and sent to the participating countries.
If we distinguish three groups by the efforts caused through the re-construction of the structure and coding of the 40 databases of IPP-1040 with the following definitions :
The first group defined less efforts, .i.e. no changes were necessary, type of data was not appropriate, single codes were defective or extraneous, only single re-arrangements of variables.
The second group defined medium efforts, i.e. re-arrangement of variables, some codes of different variables were not valid. The third group defined great efforts, extensive structural changes, extensive code changes, i.e. completely re-configuration from another data structure,

a lot of non valid codes in most of the variables, conversion from another code system to EHLASS v96, we get the following distribution:

Less efforts: 10 databases
Medium efforts: 18 databases
Great efforts: 12 databases

Especially the extraneous codes which were used for national needs from national authorities led to systematic inconsistencies in the data. Here it was an advantage that the categories of most of the EHLASS version 1996 variables were hierarchically structured so that it was possible to add sub-categories to such variables without any statistical impact on calculations only in the main categories. Nevertheless by this coding practise the data transformation became more costly. As a result, we have beside the official data structure and codes of EHLASS version 1986 and version 1996 the modified versions 1986 (mv86) and modified versions 1996(mv96).
Some countries were not able to deliver data in version 1986 or version 1996 codes because they used a different code manual (table 5). Beside the above-mentioned code systems, loegd received data in EHLASS version 2000 codes (Denmark, 1999), NOMESCO classification (Iceland 1998 and 1999)(see annex V: Transfer NOMESCO code to EHLASS v96 code) and IREQ 1991 code (see below 3.2 and annex VI: Transfer IREQ code to EHLASS v96 code). For EHLASS version 2000 data only a few re-codings were necessary with the exception of the product codes for which only the main groups had been re-coded. Data in NOMESCO classification contains ICD-10 codes for the injury diagnoses and body part affected. For loegd it was not possible to transfer the ICD-10 code to the EHLASS system. The other available variables, including sports, were re-coded. In the annex no. V you can find a description of the re-coding from data in the NOMESCO classification of external causes for injuries.

Table 5: Availability of data and the code system found

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	
AT											86	86	86	86	AT
BE						86	86	86	86	86	86	86	86	86	BE
DE											86				DE
DK				86	86	86	86	86	86	86	86	86	V2000	v2000	DK
ES											IREQ	IREQ	IREQ	96	ES
FI										86	86	86	86		FI
FR	86	86	86	86	86	86	86	86	86	86	86	86	86	86	FR
GR					86	86	86	86	86	86	86	86	86	96	GR
IE					86	86	86	86	86	86	86	86	86		IE
IS													NOMES	NOMES	IS
IT						86	86	86	86	86	86	86	86		IT
LU											IREQ	IREQ	IREQ	IREQ	LU
NL	86	86	86	86	86	86	86	86	86	86	86	86	86	86	NL
NO											86 /96	86 /96	86 /96	86 /96	NO
PT		86	86	86	86	86	86	86	86	86	86	86	86		PT
SE											86	86	96	96	SE
UK	96	96	96	96	96	96	96	96		96	96	96	96		UK

Available accident case No data available

Legend:

- 86 - EHLASS version 1986 or modified v86
- 96 - EHLASS version 1996 or modified v96
- 86/96 – the Norwegian data were coded with the v86 manual with the exception of product codes which were coded with 2-digit v96 codes
- NOMES - NOMESCO classification for external causes of injuries
- V2000 - EHLASS version 2000
- IREQ – IREQ version 1991

3.2 Conversion from IREQ 1991 to EHLASS version 1996

Beside the mentioned case data structures, loegd also received case data in IREQ 1991 structure and coding from Luxembourg and Spain. The IREQ version 1991 manual was only available in French so that a translation of the categories of the variables was necessary before an analysis could be started. The IREQ 1991 code system is similar but not identical to both systems of the EHLASS version 1986 and EHLASS version 1996 code systems (see annex VI: Transfer IREQ code to EHLASS v96 code). Due to the fact that these data were not directly compatible with EHLASS v96, the loegd constructed a transformation routine from IREQ 1991 case data structure into EHLASS v96 case data structure. In annex V you can find more details in the description of the conversion from IREQ 1991 case data to EHLASS v96 case data. In contrast to version 1986 case data, the IREQ coded data also included the variable “sports” so that an additional benefit is achieved by transferring these data into the HIEMS database. Of 13 databases which contained the variable “sports” 7 databases were coded with the IREQ 1991 manual. The rest of 6 databases was coded with EHLASS v96 and all these databases belong to the years under report 1998 or 1999. The following databases coded as IREQ 1991 case data were available

- Spain 1996 and 1998
- Luxembourg 1996 to 1999

Since 1999 survey data in Spain have been collected with the EHLASS v96 manual, the only country which collected accident data with the IREQ manual in 1999 and 2000 was Luxembourg.

4. Data transformation

Accident case data were provided during the EUPHIN HIEMS project and IPP-1040. From the 130 EHLASS case databases with historical data up to the year 1999 only 40 databases were processed under the IPP-1040 project (table 6). The transformation of EHLASS data became more difficult for the last years under report because further participating countries had to be integrated and because of the development of the code manual for national purposes. Historical data from the beginning of the EHLASS project were coded exclusively with EHLASS version 1986. In later years, i.e. after the publication of the IREQ 1991 manual, some countries used noticeably more extraneous codes. In the IPP-1040 project 11 databases were added to the HIEMS database, which was based on completely different code manuals.

Table 6: Provision of data for IPP-1040 and EUPHIN HIEMS

	1986	1987	1988	1989	1990	1991	1992	1993	1994	1995	1996	1997	1998	1999	
AT															AT
BE															BE
DE											survey				DE
DK															DK
ES											survey	survey	survey	survey	ES
FI														○	FI
FR															FR
GR															GR
IE														X	IE
IS															IS
IT														X	IT
LU											survey	survey	survey	survey	LU
NL															NL
NO															NO
PT															PT
SE															SE
UK														○	UK

Data processed in EUPHIN HIEMS project
Data processed in IPP-1040
No data available
○ No data delivered in 1999
X No contract in 1999

Starting with the year under report 1999 Finland changed the collection method from hospital-based accident data collection to interview-survey data collection.

4.1 EHLASS latecomers' data for 1986 - 1996

Data from the United Kingdom for 1986 to 1993 and 1995/1996 were transformed into the HIEMS structure. Tests on completeness and integrity of the data were carried out, the data transformation was documented and the data were uploaded into the HIEMS database. Factory Acceptance Tests (FATs, see 5.1) had been done for 2 years under report. The Consumer Safety Unit, Department of Trade and Industry in the United Kingdom, was contacted to check the availability of data for the year 1994 and 1999. The Department of Trade and Industry informed us on April 4, 2001 that at the moment it was not possible to

deliver British accident data for 1999 in the case structure. Should the data be made available, the transformation work on this database will be a task for the planned future follow-up project "Transinjury".

Data from Italy for the year 1996 were received, analysed, transformed, documented, aggregated and uploaded. The data transformation was documented. Tests on completeness and integrity of the transformation were done by FATs. The data were uploaded. The responsible national authorities have been contacted to provide data from the year 1990. Data for the years 1986 - 1989 and for 1991 – 1995 were transformed and uploaded into the EUPHIN HIEMS database. In 2000, the national responsibility for Italian EHLASS data changed from the Italian Ministry of Industry to the Italian Ministry of Health. Meanwhile, we have got the information from the Ministero della Sanità that Italy attempts to recover EHLASS data from the year 1990. Transformation work on this database would be a task for the planned future follow-up project "Transinjury".

The Spanish survey data from the year under report 1996 were re-coded from IREQ 1991 codes to EHLASS v96 codes. For this purpose, translation work from Spanish into English for the analysis of the description of the variables was done. Problems occurred due to the coding of the Spanish data which differed by extraneous codes from the IREQ1991 code system. Numerous contacts with the Spanish responsible person were necessary to get some information about the availability of data, the Spanish coding practice and to clarify how to handle the data.
The data were analysed, transformed, aggregated into the final data structure and uploaded into the EUPHIN HIEMS database. The transformation was documented and the FAT was done.

4.2 EHLASS data for the years 1997 – 1998

The following databases have been processed by loegd.

- For the year under report 1997: Italy, Luxembourg, Norway, the Netherlands, Sweden and additionally Spain and the United Kingdom.
- For the year under report 1998: Belgium, Denmark, Ireland, Italy, Luxembourg, the Netherlands, the United Kingdom and additionally Iceland, Norway, and Spain.

All these databases were transformed into the final data structure of EHLASS data, version 1.2. The transformation was documented and tested on completeness and integrity by FATs. Two databases had to be sent again because the data were wrongly generated by the data delivering countries.

First contacts with the participating countries for requesting data for the years 1997 and 1998 were made by DG SANCO. Further requests to the national EHLASS authorities were initiated by loegd.

4.3 Transformation of EHLASS data for 1999

Accident data for 1999 from Austria, Belgium, Denmark, France, Greece, Iceland, Luxembourg, The Netherlands, Norway, Spain and Sweden were processed by loegd. All these databases were transformed into the final data structure of EHLASS data, version 1.2. The transformation was documented and tests of completeness and integrity were carried out with the help of FATs.
The delivery of data by the participating countries was co-ordinated by DG SANCO. In cases of delayed data delivery, loegd contacted the participating countries and called upon them to deliver the data.

Data upload

The available accident case data were aggregated into the transfer structure of the HIEMS database respectively into the final data structure for EHLASS data. In this structure, the upload into the central database was made possible by a client installed in the loegd and access to the central database with the help of an administrator password. After completion of the Factory Acceptance Tests and confirmation of the transformation by the country, at least by the final report sent to the data delivering countries, the data were uploaded into the HIEMS database.

5. Documentation of the transformation and transfer reports

5.1 Factory Acceptance Tests (FATs)

The FATs had to back up the data transformation process by tests on completeness, integrity and correctness end to end for all transformed data. The FATs had been worked out for the EUPHIN-HIEMS project (see annex VII: Factory Acceptance Test on Test Data Collection, Version 1.3) and were used again to support the documentation of the data transformation process and for quality assurance. Information on country-specific codes, variables etc and the changes carried out during the transformation process were documented.

In the EUPHIN HIEMS project at least one FAT for every country was done. For 30 databases processed in IPP-1040 FATs were done so that for all processed databases FATs were done, with the exception of the data from the United Kingdom. The EHLASS data from the United Kingdom of the years under report 1986 to 1998 were extracted from the Home Accident Surveillance System (HASS) with the help of a specific routine. This routine converted the codes of the HASS database into EHLASS version 96 codes. This transformation was done for every year under report delivered to loegd so that the extraction of EHLASS data from the HASS database was done in the same way for all available years under report comprehensively in 1999. For economical reasons, loegd performed the transformation process on the input structure and the aggregation without carrying out FATS for all databases of the United Kingdom because loegd assumed that their performance in this case was not unconditionally necessary (table 7):

Table 7: Availability of FATs

For the years under report 1995 to 1999

45 FATs

are available.

	1995	1996	1997	1998	1999
AT					
BE					
DE					
DK					
ES					
FI					
FR					
GR					
IE					
IS					
IT					
LU					
NL					
NO					
PT					
SE					
UK					

Available FAT
No Fat available
No data delivered

5.2 Final reports to the data delivering countries

After completion of the data transformation process for a country, this country received a final report from the loegd. In the report all data sent to the loegd were listed. Furthermore, a documentation of all changes to the data which were transformed into IPP-1040 was included. The data delivering countries could confirm the final report or they could make their objections to the re-coding done by loegd. At least in the final report every data delivering country was informed about the data transformation process. Each country received a final report and up to the end of April 2001 seven participating countries confirmed the report. In annex IV you find the list of final reports on the transfer of HLA data between Member States and loegd and upload of the aggregated data into the HIEMS database.

5.3 Data Dictionary

During the EUPHIN-HIEMS project, informations on the different data sets (mortality, demography, hospital and EHLASS), on the national register structure, the institutions involved, national item definitions and specific classification information were collected to ensure that a potential user of the HIEMS system is able to interpret the data in the right way and that he becomes aware of the limits of comparability (see annex VIII: Structure of the Data Dictionary). For the IPP 1040 project, loegd filled the existing database with available information. While information on the collection methods or comparability of data was still marginal, the meta-database on the aggregated HLA data was completed by loegd with available information for all participating countries. The information was filled into the HIEMS system under the national register information (see annex VIII) by loegd. Still the structure of the Data Dictionary is not user-friendly. All information on HLA data was loaded into the national register information. Unfortunately the layout did not allow to use separate register cards in the national register for information on the four databases of the HIEMS system. The national data items (see annex VIII) have no context relation so that for a potential user it is very inconvenient to look into every position by selecting each data item. In annex IX: “Data Dictionary”, you find the complete list of input into the Data Dictionary uploaded into the HIEMS database by loegd.

5.4 Data archive

Beside the original case data, which were delivered by the participating countries to loegd and the standardised case data in the input structure for the aggregation programme, the aggregated data were stored on CD ROM by loegd. Additionally, the available FATs, SAS scripts for calculations of the FATs and documentation of the data transformation were stored on CD ROM. All data of each participating country were stored on a single CD ROM so that 17 different CD ROMs were produced. In annex X: CD ROM, you can find a description of the contents of the CD ROMs with EHLASS data.

6. Data transformation for the Injury Home and Leisure Accident Database (IHLA)

6.1 Upload of case data into the IHLA database

Individual data were available for upload. They were generated from the standardised case data pool of the data in the input structure of the aggregation programme with the help of a SAS-script on the original EHLASS version 1996 data structure which will be the input structure for the future IHLA database for case data. The upload work could not be done by loegd in IPP 1040 because an interface for upload into the IHLA database does not yet exist.

6.2 Transformation of historical case data into the EHLASS v1996 data structure

Transformation from standardised case data in the input structure into EHLASS v96 for the variables “age of patient” (3-digit field) of the input structure to “date of birth” of EHLASS v96 was done by reducing the “date of attendance” with the “age of patient” to have a virtual birth-date which still delivers the same age. The same procedure was done for “number of hospitalised days” (2-digit field) of the input structure and “date of discharge” of EHLASS v96. The “number of hospitalised days” was added to the “date of attendance” to get the virtual “date of discharge” for the EHLASS v96 structure. Nevertheless a problem remains with the coding of hospitalised days. For those data which were in the original EHLASS version 1986 data the maximum for “hospitalised days” was 39 days. All hospitalised days of more than 38 days were coded as 39 in EHLASS v86. For IREQ coded data the maximum was 69 days. As far as available, in single databases loegd coded a higher number of “hospitalised days” up to 99, with 99 standing for all “hospitalised days” of more than 98 days. These differences had no impact on the aggregated data because in accident class 9 “hospitalised days” of more than 30 days were assigned to category “9” “more than 30 days”. This different coding practise of “hospitalised days” becomes relevant for the analysis of case data.

7. Management and co-ordination activities

7.1 Meetings

The state of progress of the IPP 1040 project was continuously discussed at the meetings of the project co-ordinators in Luxembourg on June 27-28, 2000 and on December 14, 2000 in Luxembourg as well as at the IPP Network Meetings on September 6, 2000 in Paris and on March 14, 2001 in Vienna.
At the first meeting on June 27-28, 2000 a problem of data transformation was discussed. The number of items between the code manuals v86 and v 96 differs by three positions. So there are the following additional items in the extended input structure:

- Date of injury
- Time of injury
- Sports

In the report of the meeting it was stated that: „The three new data elements in the 1996-manual (in particular the sports codes) should not be lost during the transformation process. This is part of the transformation routine and therefore loegd should ensure within the current project that on the basis of the code manual EHLASS 1996 the transformed data contains these three items."
Before involving a software enterprise, some details had to be clarified at the meeting of the advisory group, which was established for the other MiaMa – IPP 1041 project.
The problems to be solved concerned:

- Time categories for the accident class 08
- The identifier of the categories „sports practised at the time of the injury" between class 13 and EHLASS v96
- The category „sports" did not have the item „unknown" at the first digit level.

The advice of the advisory group was the following:

1. The length of time between injury and contact to emergency department was agreed on the basis of the coding used in the Netherlands.
   These are the following codes for the number of hours between injury and contact with the emergency department:

   00 = < 1 hour
   01 = 1 to <2 hours
   02 = 2 to <3 hours
   03 = 3 to <6 hours
   04 = 6 to <24 hours
   05 = 24 hours and more
   99 = unknown

2. It was agreed to use the coding of the code manual v96 without changing the sequence of letters by the variable sports.
3. Additional codes for „Missing value“ and „unknown value“ will be included.
4. Arrangements will be made with Cap Gemini to inform them about the modifications.
5. The Member States should be informed that data on the three additional items can be supplied immediately.

(Minutes of the advisory meeting held on July 11, 2000 in Bielefeld; see also annex I).

7.2 Project co-ordinator activities

The main task was the co-ordination of data delivery between DG SANCO, the participating countries and loegd. Concerning data transformation for latecomers, the main problems arose with Spanish and Luxembourg EHLASS data. Both countries used the IREQ code and besides some translation and language problems the responsible persons and institutions changed. During the running of the project in five countries (Luxembourg, Italy, Spain, Finland, United Kingdom) the responsible persons and / or institutions, or addresses incl. telephone numbers or e-mail addresses changed. A special problem was caused by the Icelandic data because there was no agreement with the EC for data delivery. That means the data structure deviated from the code manuals. A special problem was the use of the NOMESCO classification for external causes of injuries, especially related to the variables “injury diagnoses” and “body part affected”, based in NOMESCO classification on ICD-10. Contacts concerned data security, too. Some data deliveries had to be repeated. Repeated data transformation processes were also necessary when the Member State revised the data transfer reports.

7.3 Reports

The state of progress was given at four meetings of the project co-ordinators and at the IPP Network Meetings. One presentation was prepared for Newsletter 1, 2000, EU-Injury Prevention Programme, May 2000. A first summary of the project IPP 1040 was handed over for the IPP Network Meeting on March 14, 2001 in Vienna.

8. Additional tasks

8.1 Remote database

For reasons of data confidentiality it was decided to use remote databases for storing national data from those countries which by national law were not able to hand over their data to private institutions. Due to this requirement a remote database on a Unix-AIX operating system with an Oracle database was installed in the loegd during the EUPHIN HIEMS project (for more details see Final Documents on Data Transformation for EUPHIN HIEMS, subvention N°. 501026). In 2001, the loegd supported the migration from TESTA I to TESTA II, which means that the central database located at Cap Gemini in Diegem, Belgium and connected to the remote database in Bielefeld by a virtual private network changed to community services. Beside the maintenance of the system, a manual was worked out which describes how to handle the Oracle database on the Unix machine and how to check the network functionality especially related to the LAN installed in Bielefeld (see annex XII: Manual for handling the remote database in Bielefeld).

8.2 Requests of interested persons or institutions

Numerous requests from institutions or individual researchers or other single private persons arrived at the loegd. The requests were mainly related to the availability of data, to the data structure and its available categories, to national responsibilities or other contacts related to accident data. These requests were answered to the best of our abilities.

8.3 Data delivery to other institutions

Also accident case data in the standardised version and in the original version (delivered data) were requested by our sub-contractor in the second project performed by loegd “Home and Leisure Accidents - Micro and Macro Analysis of Data (HLA MiaMa, Project 1999/IPP1041) Psytel in Paris. Loegd sent these data with the agreement of the European Union and the involved data delivering countries. Also the Austrian Institut Sicherleben in Vienna requested case data in the standardised and original version only for the purpose of working with them for the “Development, testing and diffusion of a common Software for quality control of home and leisure (HLA) data” within the working programme 2000 of the European Union. Loegd delivered data to the Institut Sicherleben with the permission of DG SANCO.

Commission of the European Communities: The European Home and Leisure Accident Surveillance System (EHLASS). Coding Manual 1986.

Commission of the European Communities: The European Home and Leisure Accident Surveillance System (EHLASS). Coding Manual 1996. Draft July 1996.

Commission of the European Communities: The European Home and Leisure Accident Surveillance System (EHLASS). Coding Manual 2000. Draft 2000.

Manuel de Codage EHLASS (V2.0/04/91) from L'Institut de Recherche et D'Etudes Quantitatives à Paris (IREQ 1991).

Nordic Medico Statistical Committee (1997): NOMESCO Classification of External Causes of Injuries . Third edition. Copenhagen.

Deutsches Institut für medizinische Dokumentation und Information, DIMDI (2000): ICD-10-SGB V - Internationale statistische Klassifikation der Krankheiten und verwandter Gesundheitsprobleme. Band 1 - Systematisches Verzeichnis. Verlag Hans Huber, Bern.

Institute of Public Health North Rhine-Westphalia (2000): Final Documents on Data Transformation for EUPHIN HIEMS. Bielefeld.

Bardehle, D.; Mönch, M.; Fuhr, A.; Philippsen, D.; Heyer, C.-M.; Brand, H. (1999): Creation of a Public Health Database for Europe. The Data Transformation Process for EUPHIN HIEMS within the Health Monitoring Programme of the European Union. (4. Implementation of Public Health Policy, Use of new Technology) in: Annual Meeting of EUPHA. Prague December 9-11, 1999. Prague, pp.24 - 25.

Bardehle, D.; Rössler, G.; , Fuhr, A.; Heyer, C.-M. (2000): Final report on data transformation for EUPHIN HIEMS. Institute of Public Health North-Rhine Westphalia. Bielefeld.

Nectoux, M.; Barthelemy, L.-H.; Bay-Nielsen; H., Duval, Ch.; Frimodt-Moller, B.: Evaluation du fonctionnement du sytème Européen pour la surveillance des accidents domestiques et de loisirs. Rapport final 1997
Bauer, R. (1999): EHLASS Austria. Jahresbericht 1998. Institut "Sicher Leben". Wien.

Consumer Safety Institute (1996): EHLASS Annual Report 1996, The Netherlands. Amsterdam.

Federal Institute for Occupational Safety and Health (1998): European Home and Leisure Accident Surveillance System. Jahresbericht EHLASS - Deutschland 1996. Dortmund.

Ministerio de Sanidad y Consumo – Instituto Nacional de Consumo (2000): Injury Prevention Programme (HLA). Results Report 1999. Spain.

Teutsch, S. M.; Churchill, R.E. (ed.) (1998): Principles and Practice of Public Health Surveillance. 2nd Edition. Merck & Co.

Selvin, S. (1991): Statistical Analysis of Epidemiologic Data. Monographs in Epidemiology and Biostatistics. Oxford University Press, New York.

lögd

Annexes

I Update of the aggregated data structure

Discussed on the Meeting with the Advisory Group 11th of July,2000

Problems with the Final Data Structure of EHLASS Data by adding three Variables

On the last working meeting in Luxembourg on June 26-27th , 2000, the Institute of Public Health in North Rhine-Westphalia (lögd) made a commitment to include the three missing variables for the aggregation of the HIEMS Database.

- Date of injury
- Time of injury
- Sports

The data structure of the input format will be as follows

	Field	Number of character		type of data
1	Country code	2	0-2	Alphanumeric
2	Hospital record number	10	3-12	Alphanumeric
3	date of treatment YYMMDD	6	13-18	Numeric
4	time of attendance	2	19-20	Numeric
5	Number of hospitalised days	2	21-22	Numeric
6	Treatment and follow-up	1	23	Numeric
7	age of patient	3	24-26	Numeric
8	sex of patient	1	27	Numeric
9	Accident mechanism	2	28-29	Numeric
10	site of accident	2	30-31	Numeric
11	Activity at the time of accident	2	32-33	Numeric
12	injury diagnosis 1	2	34-35	Numeric
13	injury diagnosis 2	2	36-37	Numeric
14	part of the body affected 1	2	38-39	Numeric
15	part of the body affected 2	2	40-41	Alphanumeric
16	Product involved in an accident	5	42-46	Alphanumeric
17	Product causing injury	5	47-51	Alphanumeric
18	other product	5	52-56	Alphanumeric
19	Accident description	90	57-146	String
20	Date of injury YYMMDD	6	147-152	Numeric
21	Time of injury	2	153-154	Numeric
22	Sports	3	155-157	Alphanumeric

With the inclusion of these variables it will be possible to calculate all the required accident classes (i.e. the four missing accident classes).

The additional available accident classes are:

- **class 8 – accident according to the length of time between 'injury and contact to emergency department'**
- **class 12 – accident according to the time of injury**
- **class 13 – accident according to sport activities and type of injury**
- **class 18 – accident according to type of injury and length of time between 'injury and contact to emergency unit'**

Problems related to the additional classes

accident class		Variable X	Variable Y
08	accident according to the length of time between injury and contact to emergency department	**Code for the number days between injury and contact to emergency department:** 00=none hospitalisation 01=one day 02=two days 03=three days 04=four days 05=five days 06=6-10 days 07=11-20 days 08=21-30 days 09=more than 30 days	Blank

For accident class 8 one problem appears that have to be solved urgently, because its calculation must be programmed for the extension of the aggregation programme.

- The categories within this class are listed on days. It would be more useful to take hours instead. If one uses days, it remains only a very small variation, as most patient attend the hospital within the first day after the accident occurred. Therefore the categorisation of this variables must be defined precisely.

Accident class		Variable X	Variable Y
13	Accident according to sport activities and type of injury	**code of sport activities** A Athletics B Gymnastics C Sports with racket, bat or stick D Team sports with ball E Combat sports F Non-motorised wheel sports G Motor sports H Animal sports I Winter sports J Water sports K Air sports L Weapon sports M Sports with solid balls N Climbing sports	**Type of injury** See coding accident class 02

		O Dance sports P Combined sports	

The item 'Sports practised at the Time of the Injury' from EHLASS v96 will be used to generate the class 13. This item contains a 3 digit code, from which the first position will be used for the aggregation to the HIEMS Structure. Unfortunately the identifier of the categories differ, while the meaning is the same.

	Code of sport activities in HIEMS Structure class 13		**First digit of code of sport activities in EHLASS v96**
A	Athletics	A	Athletics
B	Gymnastics	B	Gymnastics
C	Sports with racket, bat or stick	C	Sports with racket, bat or stick
D	Team sports with ball	D	Team sports with ball
E	Combat sports	E	Combat sports
F	Non-motorised wheel sports	F	Non-motorised wheel sports
G	Motor sports	G	Motor sports
H	Animal sports	H	Animal sports
I	Winter sports	J	Winter sports
J	Water sports	K	Water sports
K	Air sports	L	Air sports
L	Weapon sports	M	Weapon sports
M	Sports with solid balls	N	Sports with solid balls
N	Climbing sports	P	Climbing sports
O	Dance sports	Q	Dance sports
P	Combined sports	X	Combined sports

Finally for Sports in EHLASS v96 the category 'unknown' is missing (First digit of Sports). 'Unknown' can only be found on a lower level (second or third digit).
Similarly, for classes 8 and 9 the category 'unknown' is missing, so that in these classes the missing values were interpreted as 'none hospitalisation'.

Results of the meeting:

The advice of the advisory group was the following:

6. The length of time between injury and contact to emergency department was agreed on the basis of the coding used in The Netherlands.
   These are the following codes for the number of hours between injuy and contacts to emergency department:
   00 = < 1 hour
   01 = 1 to <2 hours
   02 = 2 to <3 hours
   03 = 3 to <6 hours
   04 = 6 to <24 hours
   05 = 24 hours and more
   99 = unknown

7. It was agreed to use the coding of the code manual v96 without changing the sequence of letters by the variable sports.
8. Additional codes for „Missing value" and „unknown value" will be included.
9. Arrangements will be made with Cap Gemini to inform them about the modifications.
10. The Member States should be informed, that data on the three additional items can be supplied immediately.

(Minutes of the advisory meeting held on July 11, 2000 in Bielefeld).

FINAL DATA

STRUCTURE

FOR

EHLASS DATA

Version 1.2
19th July 2000

II FINAL DATA STRUCTURE FOR EHLASS DATA

Data item	Position	Format	Definition and Classification
Country	1-2	AN2	Appendix 1, 1.1
Reporting year	3-6	N4	Appendix 1, 1.2
Sex	7-7	N1	Appendix 1, 1.3
Accident class	8-9	AN2	Appendix 1, 1.4
Variable X	10-11	AN2	Appendix 1, 1.4
Variable Y	12-13	AN2	Appendix 1, 1.4
Number of accidents per age-group			
0-<1 years	14-20	N7	
1- 4 years	21-27	N7	
5- 9 years	28-34	N7	
10-14 years	35-41	N7	
15-19 years	42-48	N7	
20-24 years	49-55	N7	
25-29 years	56-62	N7	
30-34 years	63-69	N7	
35-39 years	70-76	N7	
40-44 years	77-83	N7	
45-49 years	84-90	N7	
50-54 years	91-97	N7	
55-59 years	98-104	N7	
60-64 years	105-111	N7	
65-69 years	112-118	N7	
70-74 years	119-125	N7	
75-79 years	126-132	N7	
80-84 years	133-139	N7	
85-89 years	140-146	N7	
90-94 years	147-153	N7	
95 and over	154-160	N7	
unknown age	161-167	N7	

Appendix 1

1.1 Country and area code

Country (2 characters)	Country and area code. The first two characters are devoted to a country abbreviation. The remaining 2 characters are reserved for the sub-national level code extension and should be left blank if data corresponds to the whole country. The country codes are: AT = Austria BE = Belgium DE = Germany DK = Denmark ES = Spain FI = Finland FR = France GR = Greece IE = Ireland IS = Iceland IT = Italy LI = Liechtenstein LU = Luxembourg NL = Netherland NO = Norway PT = Portugal SE = Sweden UK = United Kingdom Source: ISO 3166 The last two codes are the sub-national levels (NUTSII)

1.2. Reporting year

Year (4 character)	Year of reporting.

1.3. Sex

Sex (1 character)	1= Male 2= Female 9= Unknown

1.4. Accident class

accident class		**Variable X**	**Variable Y**
01	Accident according to the product causing injury	**Code of main product:** A Raw materials, structural elements and particles B Stationary equipment outside, processed surface and natural surface C Part of building and stationary furniture D Installations for water, sanitation and electricity E Electric equipment for use in household F Furniture and textile in building G Domestic appliances and equipment H Machinery for industry , handicraft and hobby I Office and shop furniture J Medico-technical equipment K Means of transport L Toys M Musical instrument, optical equipment N Sports equipment P Clothing, baby caring articles Q Food, beverages, tobacco R Chemical products S Packaging, containers T Human beings, animals, animal articles U Natural element V Other and unspecified product	Blank

Accident class		**Variable X**	**Variable Y**
02	Accident according to the type of injury	**Code of main types of injury:** 01=Concusion 02=Contusion, bruise 03=Abrasion 04=Open wound 05=Fracture 06=Luxation, dislocation 07=Distortion, sprain 08=Lesion of nerve(s) 09=Lesion of the blood vessel(s) 10=Lesion of tendon(s) and/or muscle(s) 11=Crushing 12=Amputation 13=Poisoning 14=Burns, scalds (thermal) 15=Corrosion 16=Electrocution 17=Radiation (sunlight, x-rays) 18=Frostbite 19=Asphyxia 97=No injury diagnosed 98=Other specified injury 99=Unspecified injury	Blank

Accident class		**Variable X**	**Variable Y**
03	Accident according to the mechanism	**Code of mechanism** 0 Struck, hit by, fall 1 Struck, hit by contact with other object, person or animal 2 Crushing, cutting, piercing 3 Foreign body in natural orifice 4 Suffocation 5 Chemical effect 6 Thermal effect 7 Electric/radiation 8 Acute overexertion of body 9 Other and unspecified mechanism of injury	Blank

Accident class		Variable X	Variable Y
04	Accident according to the place of occurrence	**Code of place** 0 Transport area 1 Residential area 2 Production and workshop area 3 Retail, commercial and service area 4 School, institutional area and public premises 5 Sports area 6 Amusement, entertainment and park area 7 Open nature 8 Sea, lake and river 9 Place, other and unspecified	Blank

Accident class		Variable X	Variable Y
05	Accident according to activity at time of injury	**Code of activity** 1 Domestic work 2 Do-it-yourself work 3 Educational activity 4 Play and leisure activity 5 Sports, athletics, exercise 6 Vital activity 8 Other specified activity 9 Unspecified activity	Blank

Accident class		Variable X	Variable Y
06	Accident according to the part of body injured	**Code of part of body injured** 1 Head 2 Neck, throat 3 Thorax 4 Abdomen, lower back, lumbar spine and pelvis 5 Upper extremities 6 Lower extremities 7 Multiple body parts/whole body affected 9 Other and unknown body part	Blank

Accident class		**Variable X**	**Variable Y**
07	Accident according to treatment and follow up	**Code of treatment and follow up:** 1=Examined and sent home without treatment 2=Sent home after initial treatment 3=Treated and referred for further treatment by general practitioner 4=Treated and referred for further treatment as an outpatient 5=Treated and admitted to this hospital 6=Transfered to an other hospital 7=Deceased 9=Unknown	Blank

Accident class		**Variable X**	**Variable Y**
08	Accident according to the length of time between injury and contact to emergency department	**Code for the number hours between injury and contact to emergency department:** 00= < 1 hours 01=1 to < 2 hours 02=2 to < 3 hours 03=3 to < 6 hours 04=6 to 24 hours 05=More than 24 hours 99=Unknown	Blank

Accident class		**Variable X**	**Variable Y**
09	Accident according to the length of hospitalisation (time between hospital admission and discharge)	**Code for the number of hospitalised days:** 00=None hospitalisation 01=One day 02=Two days 03=Three days 04= Four days 05=Five days 06=6-10 days 07=11-20 days 08=21-30 days 09=More than 30 days 99=Unknown	Blank

Accident class		**Variable X**	**Variable Y**
10	Accident according to month of injury	**Code for the month of injury** 01=January 02=February 03=March 04=April 05=May 06=June 07=July 08=August 09=September 10=October 11=November 12=December 99=Unknown	Blank

Accident class		**Variable X**	**Variable Y**
11	Accident according to the day of injury	**Code for the day of injury** 01=Day 1 02=Day 2 . . 30=Day 30 31=Day 31 99=unknown	Blank

Accident class		**Variable X**	**Variable Y**
12	Accident according to the time of injury	**Code for the time of injury** 00=From midnight to 0.59 01=From 1.00 to 1.59 . . 23=From 23.00-23.59 99= Unknown	Blank

Accident class		**Variable X**	**Variable Y**
13	Accident according to sport activities and type of injury	**Code of sport activities** A Athletics B Gymnastics C Sports with racket, bat or stick D Team sports with ball E Combat sports F Non-motorised wheel sports G Motor sports H Animal sports J Winter sports K Water sports L Air sports M Weapon sports N Sports with solid balls P Climbing sports Q Dance sports X Combined sports Z No sports specified V Unknown, unknown specified sports	**Type of injury** See coding accident class 02

Accident class		Variable X	Variable Y
14	Accident according to product causing injury and type of injury	**Product causing injury** Code see accident class 01	**Type of injury** Code see accident class 02

Accident class		Variable X	Variable Y
15	Accident according to product causing injury and mechanism of accident	**Product causing injury** Code see accident class 01	**Mechanism of accident** Code see accident class 03

Accident class		Variable X	Variable Y
16	Accident according to type of injury and mechanism of accident	**Type of injury** Code see accident class 02	**Mechanism of accident** Code see accident class 03

Accident class		Variable X	Variable Y
17	Accident according to type of injury and treatment and follow up	**Type of injury** Code see accident class 02	**Treatment and follow up** Code see accident class 07

Accident class		Variable X	Variable Y
18	Accident according to type of injury and length of time between ‘injury and contact to emergency unit	**Type of injury** Code see accident class 02	**Length of time between ‘injury and contact to emergency unit** Code see accident class 08

Accident class		Variable X	Variable Y
19	Accident according to type of injury and length of hospitalisation	**Type of injury** Code see accident class 02	**Length of hospitalisation** Code see accident class 09

Accident class		Variable X	Variable Y
20	Accident according to mechanism and place of occurrence	**Mechanism of accident** Code see accident class 03	**Place of occurrence** Code see accident class 04

Accident class		Variable X	Variable Y
21	Accident according to place of occurrence and activity at the time of accident	**Place of occurrence** Code see accident class 04	**Activity at the time of accident** Code see accident class 05

III Extension of the aggregation programme

Contract between loegd and Symmedia Projekte in order to develop an extended aggregation Programme

Symmedia Projekte
Ravensberger Straße 10F
33602 Bielefeld

Beschreibung der Aufgabe:
Erweiterung des Programms HIEMS Konverter v1.1 zur Aggregation von EHLASS Falldaten in die HIEMS Datenstruktur durch Symmedia

Die Firma Symmedia aus Bielefeld hat im Sommer 1998 für das LöGD ein Programm in Delphi 1.0 erstellt. Dieses Programm aggregierte Unfall-Falldaten in EHLASS Struktur in die HIEMS Transferstruktur.
Das vorhandene Aggregationsprogramm soll auf der Inputseite um drei weitere Variablen ergänzt werden, aus denen dann vier zusätzliche Unfallklassen generiert werden können. Hierzu sollen dem unten aufgeführten Inputformat (der Falldatenstruktur) drei Variablen angefügt werden.

Verwendete Dateiformate und Betriebssystem für die Software

Die Inputdateien und Outputdateien sind Text bzw. einfache ASCII-Dateien. Die vorgesehenen Extension für den Input ist '.out' und für den Output lautet sie '.hie'. Hier besteht kein Änderungsbedarf.

Die Software soll lauffähig sein ab Windows 95 und Windows NT 4.0 aufwärts.

Verwendete Daten, Datenformate und Datenstrukturen

Das bestehende Programm nutzt die unten dargestellte Inputdatenstruktur.

EHLASS data structure - hier zugleich die Inputstruktur

	Field	Number of characters		Type of data
1	Country code	2	1-2	Alphanumeric
2	Hospital record number	10	3-12	Alphanumeric
3	Date of treatment YYMMDD	6	13-18	Alphanumeric
4	Time of attendance	2	19-20	Numeric
5	Number of hospitalised days	2	21-22	Numeric
6	Treatment and follow-up	1	23	Numeric
7	Age of patient	3	24-26	Numeric

8	**Sex of patient**	**1**	**27**	**Numeric**
9	**Accident mechanism**	**2**	**28-29**	**Numeric**
10	**Site of accident**	**2**	**30-31**	**Numeric**
11	**Activity at the time of accident**	**2**	**32-33**	**Numeric**
12	**Injury diagnosis 1**	**2**	**34-35**	**Numeric**
13	**Injury diagnosis 2**	**2**	**36-37**	**Numeric**
14	**Part of the body affected 1**	**2**	**38-39**	**Numeric**
15	**Part of the body affected 2**	**2**	**40-41**	**Numeric**
16	**Product involved in an accident**	**5**	**42-46**	**Alphanumeric**
17	**Product causing injury**	**5**	**47-51**	**Alphanumeric**
18	**Other product**	**5**	**52-56**	**Alphanumeric**
19	**Accident description**	**90**	**57-146**	**String**

Die Kategorien der Unfallklassen der HIEMS Datenstruktur (Datenstruktur der aggregierten Daten) sind eine strenge Untermenge der möglichen Werte der oben aufgeführten Variablen der Falldatenstruktur (Inputstruktur). Deshalb können die Kategorien der Inputvariablen und der Unfallklassen beide der “Final Data Structure for EHLASS Data” entnommen werden. Die weitergehenden bzw. detaillierteren Kodes der Inputdaten gehen durch die Aggregation verloren und sind deshalb hier nicht von Interesse.

Die Verarbeitung der oben aufgeführten Inputvariablen durch den HIEMS Konverter v1.1 bleibt unberührt von den hinzukommenden Variablen, die weiter unten vorgestellt werden. Trotzdem sind kleine Änderungen im bisherigen Programm erforderlich, diese sind am Ende zusammengefaßt.

Hinzu kommen folgende Variablen

- Date of injury
- Time of injury
- Sports

20	**Date of injury YYMMDD**	**6**	**147-152**	**Alphanumeric**
21	**Time of injury**	**2**	**153-154**	**Numeric**
22	**Sports**	**3**	**155-157**	**Alphanumeric**

Beschreibung der Inputdatenstruktur

Date of Injury	YYMMDD; gleiches Format wie ‚Date of Attendance'. Ein unbekannter Monat oder unbekannter Tag wird jeweils mit "99" kodiert.Datumsformate mit Leerstelle im Datum (99 1 1) oder mit vorstehender Null (990101) sind möglich.
Time of Injury	NN; diese Variable hat das gleiche Format wie „Time of Attendance". Zulässige Werte sind "00 – 23, 99 und 2 Leerzeichen". "99" und Leerzeichen stehen für unbekannt.
Sports	CCC; der Kode ist 3-stellig, alphanumerisch und linksbündig. An der ersten Stelle steht ein Buchstabe, danach können bis zu zwei Ziffern folgen. Nur der Buchstabe wird in der HIEMS Datenstruktur ausgewertet. Zulässige Kategorien sind: A, B, C, D, E, F, G, H, J, K, L, M, N, P, Q, V, X, Z

Die Aggregation der Variablen

Die hinzukommenden Klassen in der HIEMS Struktur sind:

- Accident Class 8 – accident according to the length of time between ‚injury and contact to emergency department'
- Accident Class 12 – accident according to the time of injury
- Accident Class 13 – accident according to sport activities and type of injury
- Accident Class 18 – accident according to type of injury and length of time between 'injury and contact to emergency unit'

Die genaue Beschreibung der Unfallklassen können der „Final Data Structure for EHLASS Data" entnommen werden.

Beschreibung der hinzukommenden Unfallklassen

Accident class 08	Die Kategorien dieser Unfallklasse sind noch nicht endgültig festgelegt. In der nächsten Woche erhalten Sie hierzu eine klärende Mitteilung. Zur Zeit sind folgende Aussagen möglich: Die Zahl der Stunden ergibt sich aus „Date of Injury“ (Pos. 20 der Inputstruktur) und „Time of Injury“ (Pos. 21 der Inputstruktur) einerseits und „Date of Treatment“ (Pos. 3 der Inputstruktur) und „Time of Attendance“ (Pos.4 der Inputstruktur). Falls der Tag im Datum von „Date of Injury“ oder „Date of Treatment“ unbekannt ist oder ein Zeitpunkt von „Time of Injury“ oder „Time of Attendance“ unbekannt ist, soll die Kategorie „Unknown“ ausgewählt werden (d.h. also 99). Die genaue Berechnung entnehmen Sie bitte dem „Anhang zur Berechnung der Klassen 8 und 13“.
Accident class 12	Die Kategorien dieser Unfallklasse sind in der „Final Data Structure for EHLASS Data“ beschrieben. Die Werte für diese Unfallklasse sind in den Inputdaten unter Pos. 21 zu finden.
Accident class 13	Die Kategorien wurden oben bereits aufgeführt und stehen auch in der „Final Data Structure for EHLASS Data“. Die Werte für diese Unfallklasse stehen in der Inputstruktur unter Pos. 22. Die genaue Berechnung entnehmen Sie bitte dem „Anhang zur Berechnung der Klassen 8 und 13“.
Accident class 18	Da "Accident Class 8" nun verfügbar ist, kann auch diese Unfallklasse berechnet werden. Die Beschreibung dieser Unfallklasse kann wieder der „Final Data Structure for EHLASS Data“ entnommen werden.

Folgende Änderungen am bestehenden Programm (Inputstruktur Pos.1-19) sind erforderlich:

1. Die Variable „Age of Patient“ (Pos.7 der Inputstruktur) kann den Wert „999“ für unbekanntes Alter annehmen. Der Konverter berücksichtigt bisher diesen Wert nicht in der Kategorie für „unbekanntes Alter“, sondern fügt diese Datensätze der letzten Altersgruppe zu (95 and over).
2. Der alte Konverter erzeugt in der „Accident Class 6“ die Kategorie „9“ für „unbekannt“ nicht. Kategorie „9“ gehört jedoch zur Kodeliste dieser Unfallklasse.
3. In der „Accident Class 9“ wurde die Kategorie „unbekannt“ mit der Kodierung „99“ hinzugefügt, ferner sollen 2 Leerzeichen als "unbekannt" ausgewertet werden.
4. Das Berichtsjahr in Pos. 3 soll für jeden Datensatz geprüft werden. Fehlerhafte Berichtsjahre sollen nicht bearbeitet, sondern übersprungen werden und zu einer Fehlermeldung führen. Als Vergleichswert kann das Berichtsjahr des ersten eingelesenen Datensatzes verwandt werden (Widersprüche sollten dann jedenfalls erkannt werden).

- Alle deutschsprachigen Strings sollen ins Englische übersetzt werden. Dies gilt für die Programmoberfläche

1. *"Projekt",*
2. *"Statuszeile",*
3. *"Konvertiere File",*
4. *"Ende" und*
5. *das Dialogfeld „Inputfile öffnen“*

und Rückmeldungen des Programms:

1. *„Eingelesene Zeilen: (..) !!! Es sind Fehler beim Einlesen der Inputdatei aufgetreten. Siehe Logfile !!! 33% geschrieben. 66% geschrieben. 99% geschrieben. Outputdatei geschrieben.“*
2. *„!Geschlecht nicht konvertierbar: "(..)"“*
3. *„Konvertierungsfehler in Klasse: (..) Datenfeld wird übersprungen!“*
4. *„- Zeile: (..) Offset: (..) Länge: (..) Inhalt:"(..)"“*
5. *„- Folgefehler in Klasse: (..) Feld Var(..) s.o.“,*
6. *"Laenge der Zeile (..) ist nicht gleich 157, ueberspringe !!! "*
7. *„Eingelesene Zeilen: (..) Keine Fehler beim Einlesen der Inputdatei.“* und alle weiteren Meldungen, soweit sie hier nicht aufgeführt sind.

Folgende Vereinbarungen wurden getroffen:

- Die Aufwendungen für die Übersetzung ins Englische sollen extra ausgewiesen werden, um eine getrennte Abrechnung zu ermöglichen.
- Es werden Symmedia Testdaten in der alten Inputstruktur zur Vorbereitung des Angebotes zur Verfügung gestellt (Textdatei, <30KB).
- Wenn Symmedia die Erweiterung des HIEMS Konverter programmiert, werden Testdaten in der erweiterten Inputstruktur konstruiert und Symmedia zur Verfügung gestellt. Diese Daten ermöglichen es Symmedia alle neu programmierten Features zu prüfen (Textdatei, <20KB), so daß der Testaufwand für Symmedia minimiert wird.
- Symmedia unterbreitet auf der Grundlage der Aufgabenbeschreibung und des am 13. Juli 2000 geführten Gespräches mit Herrn Kinner ein Angebot.

Anhang zur Aufgabenbeschreibung

Die Berechnung der Unfallklassen 8 und 13 der HIEMS Datenstruktur im erweiterten HIEMS Konverter von Symmedia

Für die Erweiterung des HIEMS Konverters durch Symmedia sind zur Berechnung der Klassen 8 und 13 mehrere Variablen der Inputstruktur zu berücksichtigen. Bei der Auswertung der Variablen können verschiedene Situationen eintreten, die einer genaueren Erklärung bedürfen.

Berechnung der Unfallklasse 8 der HIEMS Datenstruktur

Zur Ermittlung der Unfallklasse 8 „Length of Time between injury and contact to emergency department“ müssen insgesamt 4 Variablen herangezogen werden. Diese sind

1. „Date of Injury“ (Position 20 der Inputstruktur; unten DI genannt)
2. „Date of Attendance“ (Position 4 der Inputstruktur; unten DA genannt)
3. „Time of Injury“ (Position 21 der Inputstruktur; unten TI genannt)
4. „Time of Attendance“ (Position 4 der Inputstruktur; unten TA genannt)

Die gesuchte Zeitdauer ergibt sich aus folgender Rechnung:

gesuchte Zeitdauer = (DA+TA)-(DI+TI)

Da alle Variablen der Inputstruktur unbekannte Werte enthalten können, ist eine Fallunterscheidung erforderlich. Für die folgende Fallunterscheidung ist es ausreichend, wenn nur ein Datum oder ein Zeitpunkt unbekannt ist.

Fall	Date of Attendance	Date of Injury	Time of Attendance	Time of Injury	Berechnung möglich ? Sonst ist „unknown“ zu setzen.
1	**MMDD**	**MMDD**	**hh**	**hh**	**Ja (Normalfall)**
2	**MM99**	**MM99**	**hh**	**hh**	**Nein**
3	**9999**	**9999**	**hh**	**hh**	**Nein**
4	**99DD**	**99DD**	**hh**	**hh**	**Ja**
5	**MMDD**	**MMDD**	**99**	**99**	**Ja, falls** **DA – DI >0**

Beispiel: Sei DA = 960301, TA = 09, DI = 960228, TI = 22, so ist nach (1) eine Berechnung möglich.

Sei DA = 999901, TA = 09, DI = 999928, TI = 22, so ist eine Berechnung nach (4) möglich. Die Kategorien der Unfallklasse 8 finden Sie in der „Final Data Structure for EHLASS Data Version 1.2."

Berechnung der Unfallklasse 13 der HIEMS Datenstruktur

Zur Ermittlung der Unfallklasse 13 „Sport Activities and Type of Injury" müssen zugleich die Variable „Activity at the Time of Injury" (Position 11 der Inputstruktur) und die Variable „Sports" (Position 21 der Inputstruktur) ausgewertet werden. Die Werte 50, 51, 58 und 59 der Variablen „Activity at the Time of Injury" (unten als ATI bezeichnet) stehen für Sportaktivitäten.

Folgende Fälle müssen unterschieden werden:

1. Falls ATI keine Sportaktivität anzeigt, also kein Wert von 50, 51, 58 oder 59, sondern eine andere Aktivität, ausgenommen „unknown", und die Variable „Sports" keinen Wert enthält, wird der Fall zu Kategorie Z der Unfallklasse 13 gezählt.
2. Falls ATI keine Sportaktivität oder „unknown" anzeigt, trotzdem aber unter „Sports" ein Wert gefunden wird, so kann dieser Wert für die Unfallklasse 13 ausgewertet werden.
3. Falls ATI eine Sportaktivität oder „unknown" anzeigt, aber die Variable „Sports" keine Sportaktivität enthält, d.h. „Missing Value", wird der Fall zu der Kategorie V der Unfallklasse 13 gezählt.

Testdaten zur Prüfung der Modifikationen zur Erweiterung des HIEMS Konverters

In den Dateien „**Testdaten_1.out** " und „**Testdaten_2.out** " sind die in den Tabellen aufgeführten Werte enthalten. Mit den Datensätzen der zuerst genannten Datei werden problematische Kategorien geprüft. Die Testdaten der zweiten Datei zielen auf einen Check der Berechnungen für die Unfallklassen 8 und 13. In der linken Spalte der Tabellen sind die Zeilen durchnumeriert und in der mittleren Spalte sind die zu prüfenden Werte aufgeführt. Die rechte Spalte enthält Kommentare zum erwartetem Ergebnis. Diese Form der Darstellung wurde in Anlehnung an das Ergebnisprotokoll des Konverter gewählt, da auch im Ergebnisprotokoll die Zeile und die Position der Variablen, die eine Fehlermeldung verursachte, ausgegeben wird und so direkt eine Gegenüberstellung gemacht werden kann.

Tabelle zu prüfender Kategorien

Zeile	Variablenbezeichnung in der Inputstruktur/ Kategorien	
	Date of Treatment	
	990101	
1	99 1 1	
2	000101	
3	Time of Attendance	→Fehlermeldung, da Berichtsjahr nicht 1999
	25	
4	99	→Fehlermeldung, da nur 0 bis 23 und 99 zulässig sind
5		→als “unknown“ erkennen
6	Age of Patient	
	225	→Fehlermeldung, da von den Werten über 200 nur 201 bis 223 und 999 zulässig sind
7	999	
8		→999 und als “unknown“ erkennen
	Date of Injury	
	990101	
9	99 1 1	
10	990199	
11	9999 1	
12	99 199	
13	Time of Injury	Wie „Time of Attendance“
	00	
14	23	→Fehlermeldung
15	25	
16	99	
17		
18	Sports	
	A00	
19	B00	
20	C0	
21	D0	
22	E00	
23		

24	F0	
25	G00	
26	H0	
27	J00	
28	K0	
29	L00	
30	M0	
31	N0	
32	P00	
33	P0	
34	Q00	
35	Q0	
36	X0	
37	Z0	
38	Y0	
39		→Fehlermeldung; W ist keine zulässige Kategorie
40	W0	

Es werden nur 35 Datensätze der insgesamt 40 Datensätze ausgewertet. Die übrigen 5 Datensätze werden übersprungen, weil sie einen Fehler in der Kodierung haben.

In Unfallklasse 6 gibt es genau einen Fall bei Männern und Frauen (sex=1 und sex=2) mit Kode „9“. In der höchsten Altersgruppe (age group= 95 and over) darf es keinen Fall geben.

Auswertung der Unfallklassen 8 und 13 für Testdaten2.out

Nach der Aggregation von Testdaten2.out sollten folgende Kategorien der Unfallklassen 8 und 13 in den unten angegebenen Häufigkeiten (Fälle) erscheinen.

Unfallklasse 8		Unfallklasse 13	
Kategorie	Anzahl	Kategorie	Anzahl
02	1	A	5
05	1	B	1
99	8	C	1
		V	2

Final remarks

- **The software was accepted by loegd and used for the aggregation of more than 20 databases**
- **The source code was not bought by loegd, but Symmedia offered it for 1800 DM plus VAT**

IV List of Final Protocols
on the transfer of HLA Data between
the Participating Countries (PC) and loegd

IPP 1040 Transformation of EHLASS Data 1997-1999 and for Latecomers

Country	Date of the first sending of the final protocol to the PC	Date of confirmation of the final report by PC
Austria	27 June 2000	23 August 2000[*1]
Belgium	16 November 2000	21 November 2000
Denmark	22 December 2000	
Finland	17 November 2000	
France	26 June 2000	
Greece	20 December 2000	7 February 2001
Iceland	23 March 2001	
Ireland	4 September 2000	
Italy	17 November 2000	26 April 2001
Luxembourg	23 March 2001	12 April 2001
The Netherlands	20 December 2000	11 January 2001
Norway	17 November 2000	
Portugal	1 September 2000	
Spain	23 March 2001	
Sweden	22 February 2001	10 April 2001
United Kingdom	19 April 2001	

[*1]The confirmation was done by mail

V Transfer NOMESCO code into EHLASSv96 code

Transfer of Icelandic accident case data in the NOMESCO classification into EHLASS Version 1996 case data codes

The Icelandic accident case data which were delivered to loegd are coded in the NOMESCO Classification of External Causes of Injuries (Nordic Medico-Statistical Committee (1997): NOMESCO Classification of External Causes of Injuries, third edition, Copenhagen; The first edition of NOMESCO´s classification for registration of external causes of unintentional injuries was published in 1984). Dr. Mogensen from the University of Iceland informed us that more than 98 percent of the Icelandic accidents are represented by these data. Before the aggregation of the Icelandic accident data it was necessary to re-code the data into the EHLASS 1996 version, as far as possible. The accident data from Iceland contain the following variables, which can be used as EHLASS version 1996 codes:

- Date and time of attendance and injury (4 variables)
- Age of patient
- Sex of patient
- Accident mechanism
- Place of accident
- Activity at the time of injury
- Sports

In the NOMESCO classification also product codes are available, but unfortunately none were delivered. Not available are the following variables:

- Number of hospitalised days
- Treatment and follow up
- Injury diagnoses (2 variables)
- Body part affected (2 variables)
- Product codes (3 variables)
- Description

The variables „injury diagnosis“ and „body part affected“ are explicitly not available, the Icelandic data have ICD-10 codes. The ICD-10 code is a complex historically developed code system which integrates the content of the EHLASS variables „injury diagnosis“, „body part affected“ and sometimes “accident mechanism” into one code. Loegd was not able to transform the ICD-10 codes into appropriate EHLASS version 1996 codes.

Only a few variables in the NOMESCO classification have to be re-coded. With the exception of “activity at the time of injury” the re-codings of the following variables are limited to a few categories:

- Sex of patient
- Place of occurrence
- Activity at the time of injury
- Mechanism of accident
- Sports
- Product codes (not available)

In the NOMESCO classification for sex other codes are used than in the EHLASS 1996 version. The variable "place of occurrence" is identical with those of the EHLASS 1996 version. The same applies to "accident mechanism", with the exception of category "01" of the NOMESCO nomenclature, which merges categories "01" and "02" of the EHLASS version 1996. This should be refer to the data dictionary only for case data, because after the aggregation of case data to the HIEMS structure the above mentioned merge of two categories to one have no impact to the aggregated data.

The variable 'activity at the time of injury' in the NOMESCO classification differs from the EHLASS 1996 version. Only the main groups can be assigned to the appropriate main groups of the EHLASS 1996 version. In the tables below you can find this re-coding. Data sets in the NOMESCO classification with the category 'working for income' for the variable 'activity at the moment of injury' were deleted by loegd from the database because these accidents do not belong to Home and Leisure Accidents.

Category in NOMESCO classification	code	Appropriate category in EHLASS v96 (main groups)	Code Of the main group	valid 2-digit EHLASS v1996
Working for income	0	These cases have to be deleted because they are no home and leisure accidents		
Sports, athletics, exercises	1	Sports, athletics, exercises	5	59
Play, hobby and other leisure	2	Play and leisure activity	4	49
Educational activity	3	Educational activity	3	39
Unpaid work	4	Domestic work	1	19
Vital activity	5	Vital (basic) work	6	69
Harvesting natural resources (For instance: Fishing, hunting, setting traps, search of gold)	6	Other specified activity	8	88
Activity, other specified	8	Other specified activity	8	88
Activity, unspecified	9	Unspecified activity	9	99

The variable "sports" is almost identical with the EHLASS 1996 version. It already contains a part of the categories of the present EHLASS Version 2000 code manual. Therefore the following codes had to be changed:

Nomesco Classification	EHLASS version 1996	Nomesco Classification	EHLASS version 1996
C25	C2	D12	D1
C26	C2	N11	N1
C3	C0	N12	N1
C30	C01	N83	N8

Product codes are not available in the Icelandic accident data, but it can be stated that similarly to the variable “sports” the product codes in the NOMESCO classification merge EHLASS version 1996 codes with EHLASS version 2000 codes.

Finally it can be summarised that Icelandic accident case data are representative from the point of view of comprehensiveness and that the necessity of re-coding does not require much effort. Especially the lack of the important variables “injury diagnoses” and “body part affected” is a disadvantage which will be overcome if it becomes possible to integrate the ICD-10 code into the EHLASS system.

VI Transfer IREQ code into EHLASS v 96 code

Table of contents

0. Introduction and Objective

The following documents have been used

- Report of EHLASS data transformation
- European Home and Leisure Accident Surveillance System, coding manual 1986 (EHLASS version 1986)
- Manuel de Codage EHLASS (V2.0/04/91) from L'Institut de Recherche et D'Etudes Quantitatives à Paris (IREQ 1991)
- European Home and Leisure Accident Surveillance System, coding manual 1996, draft (EHLASS version 1996)
- Transformation tables from Psytel for all variables from the EHLASS version 1986 code to the EHLASS version 1996 code
- Final data structure for EHLASS data Version 1.2

The EUPHIN HIEMS database for EHLASS data is part of the Community-wide information system for the transfer and sharing of health data and indicators using telematic interchange of data as the principal means. The EHLASS project was started in 1986. Case data on home and leisure accidents were collected and aggregated into the "Final data structure for EHLASS data " for all participating countries by loegd. On 28 February 2001, EHLASS data from 17 countries (EU, Iceland and Norway) were available.

The basis for the aggregation of EHLASS case data into the "Final data structure of EHLASS data" is the coding manual of the EHLASS 1996 version. Most of the available historical case data are coded with the manual on the EHLASS 1986 version. Therefore, in a first step loegd had to transform all EHLASS version 1986 case data into EHLASS version 1996 case data. The next step consisted in the aggregation of case data into the "Final data structure of EHLASS data". In addition to case data in the EHLASS version 1986 coding case data in the IREQ 1991 structure were supplied by Luxembourg and Spain. The IREQ 1991 code system is similar to but not identical with the EHLASS version 1986 and EHLASS version 1996 code system, as you can see below. The aim was to develop a transformation routine from the IREQ 1991 case data structure to the EHLASS version 1996 case data structure

1. Data availability and comparison of the code systems IREQ 1991 and EHLASS version 1996

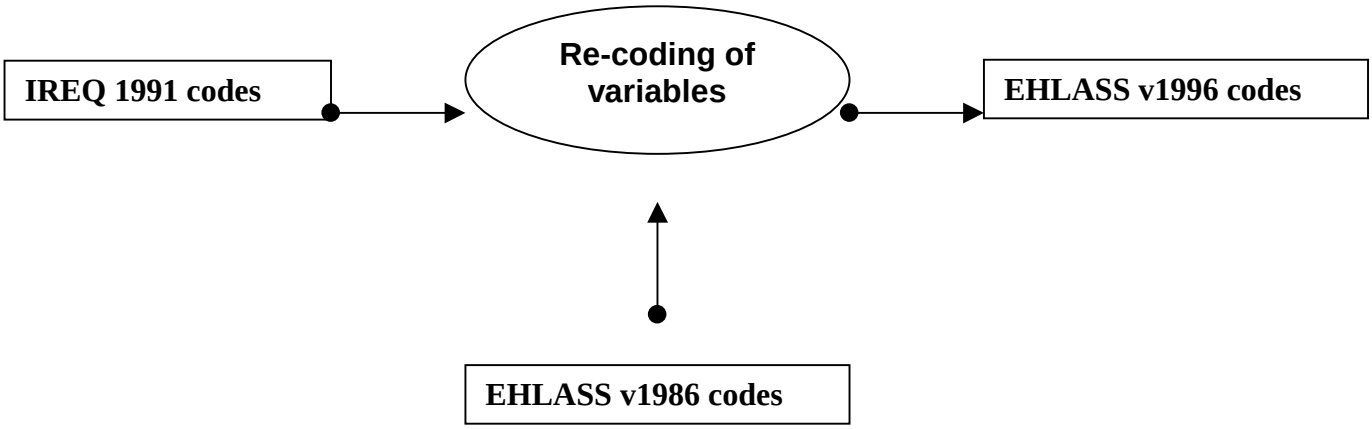
Loegd holds both home and leisure accident case data coded in accordance with the IREQ 1991 code system. In contrast to the EHLASS version 1986 case data these data also included the variable "Sports" so that additional benefit can be achieved by transferring these data into the HIEMS database. The following databases coded as IREQ 1991 case data are available

- Spain 1996 and 1998

- Luxembourg 1996 to 1999

These data were not directly compatible with the EHLASS 1996 version. For a transformation from IREQ 1991 to the EHLASS 1996 version it is useful to find the common elements between the different platforms and then analyse the differences between the EHLASS 1986 version, IREQ 1991 and the EHLASS 1996 version according to the variables and the code lists. It was useful to also concentrate on the coding system of the EHLASS 1986 version because the product codes of the EHLASS 1986 version and IREQ 1991 are identical. Therefore it is possible to use the transformation tables for product codes from Psytel. With the transformation tables for product codes we can change the EHLASS version 1986 product code and also IREQ 1991 product codes into EHLASS version 1996 product codes. The following diagramme illustrates the transformation.

Related code manuals for re-coding IREQ 1991 codes into EHLASS version 1996 codes



Not for all variables in the EHLASS 1996 version does an equivalent variable exist in IREQ 1991. The variables belonging to the EHLASS 1986 version, IREQ 1991 and the EHLASS 1996 version are shown in the next table.

Data structure of EHLASS version 1986, IREQ 1991 and EHLASS version 1996

<table>
<tr><th>EHLASS v1986</th><th>IREQ 1991</th><th>EHLASS v1996</th></tr>
<tr><td colspan="3">Country</td></tr>
<tr><td colspan="3">Date of treatment or attendance</td></tr>
<tr><td colspan="3">Time of treatment or attendance</td></tr>
<tr><td colspan="3">Sex of patient</td></tr>
<tr><td colspan="3">Accident description</td></tr>
<tr><td colspan="3">Treatment and follow-up</td></tr>
<tr><td colspan="3">Activity at the time of accident</td></tr>
<tr><td colspan="3">Place of occurrence</td></tr>
<tr><td colspan="3">Mechanism of injury</td></tr>
<tr><td colspan="3">Type of injury</td></tr>
<tr><td colspan="3">Part of body affected</td></tr>
<tr><td colspan="3">Product involved in the accident</td></tr>
<tr><td colspan="3">Product causing the injury</td></tr>
<tr><td colspan="3">Other product</td></tr>
<tr><td>Not available</td><td colspan="2">Sport</td></tr>
<tr><td colspan="2">Number of hospitalised days</td><td>Date of discharge +day of attendance</td></tr>
<tr><td colspan="2">Hospital record number</td><td>Case number</td></tr>
<tr><td colspan="2" rowspan="2">Age of patient</td><td>Date of birth</td></tr>
<tr><td>Date of injury</td></tr>
<tr><td colspan="2" rowspan="3">Not available</td><td>Hospital number</td></tr>
<tr><td>Date of injury</td></tr>
<tr><td>Time of injury</td></tr>
</table>

Only a few of the above-listed variables have a different coding in IREQ 1991 and the EHLASS 1996 version. We should pay our attention to these variables. They are shown below. They are identical, but have different codes in IREQ 1991 and EHLASS version 1996.

- Treatment and follow-up
- Activity at the time of accident
- Place of accident occurrence
- Type of injury (diagnosis)
- Mechanism of injury
- Part of the body affected
- Sports
- Product involved in the accident
- Product causing the injury
- Other products

The variables are listed in the tables below. They are distinguished by the following criteria:

- Variables not available in each system
- Identical variables
- Identical variables with a different coding
- Identical variables with a partially different coding and equivalent variables

The variables marked grey have to be re-coded . For these variables tables with the IREQ 1991 code and the appropriate EHLASS version 1996 code organised in pairs are given below.

1.1 Variables not available in every system

EHLASS v86	IREQ	EHLASS v96
Not available		Hospital number
		Date of injury
		Time of injury
Not available	Sports	

1.2 Identical variables

EHLASS v86	IREQ	EHLASS v96
Country codes		
Date of treatment or attendance		
Sex of patient		
Time of attendance		
Accident description (length of variable not identical)		

1.3 Identical variables with different coding

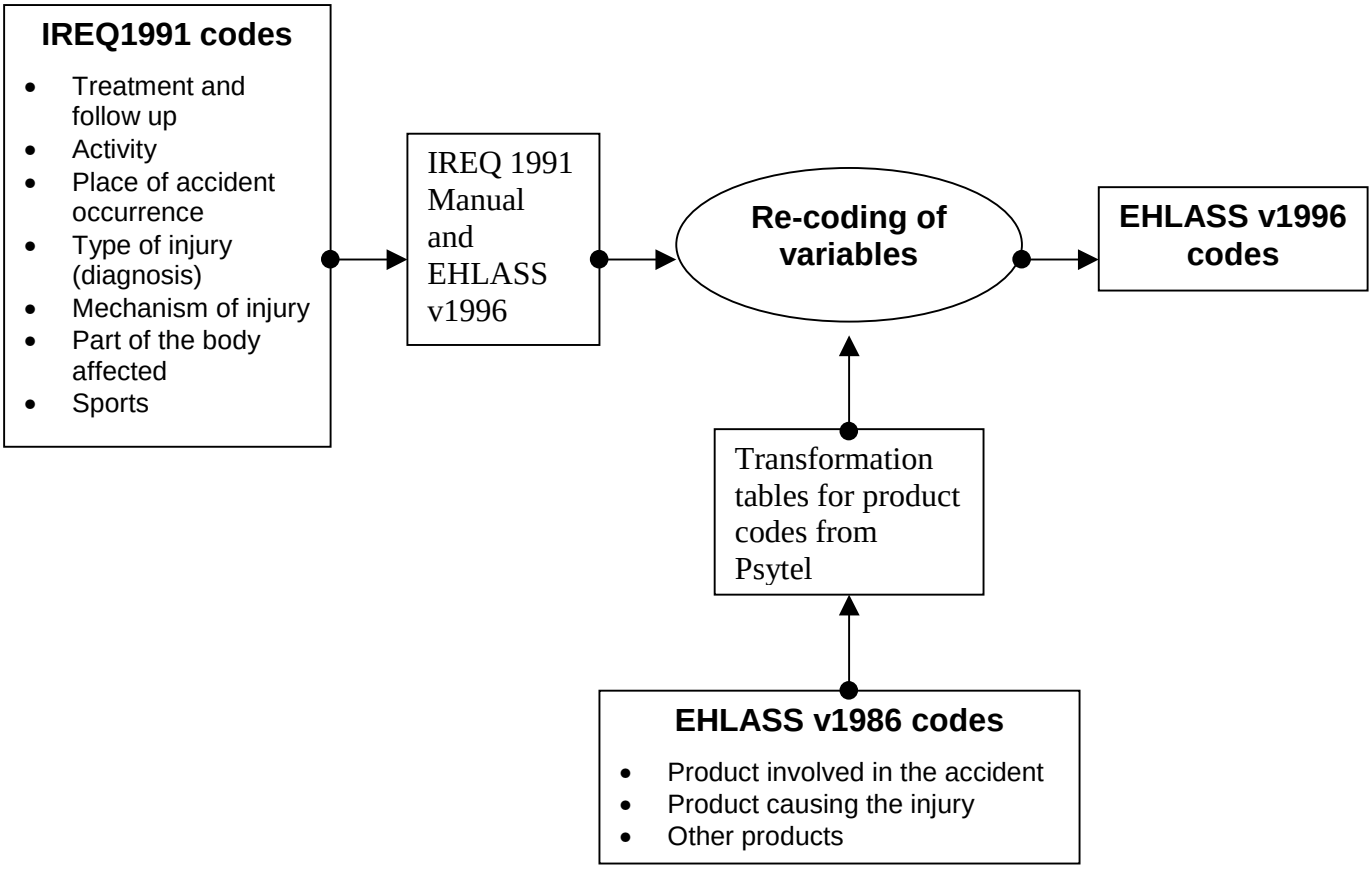
EHLASS v86	IREQ	EHLASS v96
Treatment and follow-up		
Activity at the time of accident		
Place of accident occurrence		
Type of injury (diagnosis)		
Mechanism of injury		
Part of the body affected		

1.4 Identical variables with a partially different coding and equivalent variables

EHLASS v86	IREQ	EHLASS v96
Hospital record number		Case number
Age of patient		Date of birth + day of attendance
Number of hospitalised days		Date of attendance + day of discharge
Product involved in an accident*)		
Product causing the injury*)		
Other Product*)		
Not available	Sports	

*) The product codes of EHLASS version 1986 and IREQ are identical

Re-coding of IREQ 1991 codes to EHLASS v1996 codes

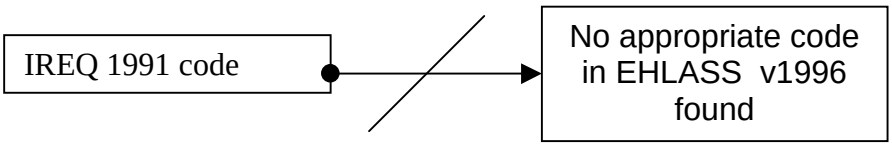


2. Practise of re-coding and loss of information on the case data level

The re-coding of an IREQ 1991 variable into an EHLASS version 1996 variable is realised by re-coding every category of one variable from IREQ 1991 to EHLASS version 1996.

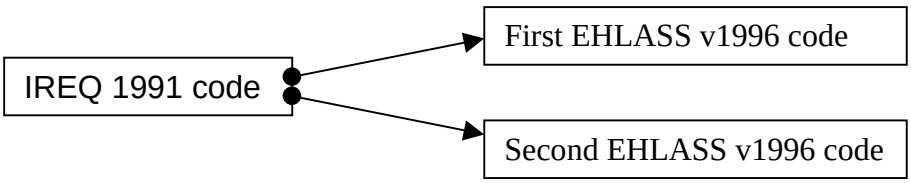
You should be aware that it is not possible to assign every category appearing in IREQ 1991 directly to the appropriate code in EHLASS version 1996. So a category of IREQ 1991 had to be set to “Unknown” in the EHLASS 1996 version because the precise category does not appear anymore in the EHLASS 1996 version. It is also possible that two or more categories of IREQ 1991 appear in the EHLASS 1996 version, joined into one and vice versa. If one category of IREQ is split in two in the EHLASS 1996 version, we add the value to one of them in the EHLASS 1996 version with a corresponding reference given in the Data Dictionary.

Loss of information by the disappearing of a category



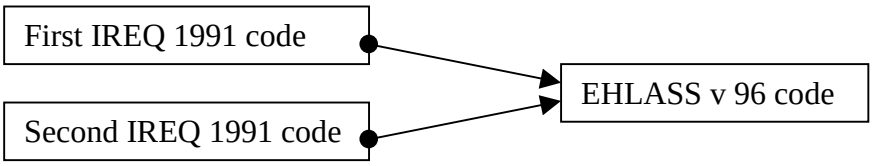
Coded as "Unknown" in EHLASS v1996

Loss of information by dividing one category into two



Coded into one category in EHLASS v96 with reference given in the DD

Loss of information by melting two categories into one



Loss of information by decreasing the "categories in use"

In the transformation tables below, you can see the actual amount of categories re-coded from IREQ 1991 into the EHLASS 1996 version as "Unknown or other specified" in the main category of the EHLASS 1996 version:

- Place of accident 3 categories as "other specified" in a sub-category
- Place of accident 4 categories as "unknown" in a sub-category
- Accident mechanisms 4 categories as "other specified" in a sub-category
- Accident mechanisms 7 categories as "unknown" in a sub-category
- Activity at the time of injury 5 categories as "other specified" in a sub-category
- Activity at the time of injury 4 categories as "unknown" in a sub-category
- Type of injury (diagnosis) 2 categories as "Type of injury, other specified" in a main category
- Body part affected 2 categories as "other specified" in a sub-category

- Body part affected 4 categories as "Body part, other specified" in a main category
- Sports 2 categories set to missing value

Nevertheless, we recommend not to use case data after the transformation of the IREQ 1991 case data code into the EHLASS version 1996 case data code, except for the purpose of aggregation into the HIEMS data structure because after the aggregation most of the re-coding does not have any impact on the aggregated data.

The following table illustrates the "evolution" or refinement of codes from EHLASS version 1986 and IREQ 1991 to EHLASS version 1996.

Frequencies of codes from the affected variables

Name of variable	EHLASS v1986 # of categories	IREQ 1991 # of categories	EHLASS v1996 # of categories
Treatment and follow-up	7	8	8
Activity at the time of accident	13	19	27
Place of accident occurrence	39	48	80
Type of injury (diagnosis)	18	22	22
Mechanism of Injury	24	52	71
Part of the body affected	47	55	52
Sports	--	161	253

The variable "Sports" is not available in the EHLASS 1986 version. 73 categories of the "Sports" variable in the EHLASS 1996 version result from the main categories as two-digit codes. Only the other 180 categories have a three-digit code.

Loss of information by the aggregation of EHLASS case data into the HIEMS data structure

In addition to the above-mentioned loss of information by re-coding IREQ 1991 case data into EHLASS version 1996 case data, information is also lost by the aggregation of case data into the HIEMS data structure. It is possible that the loss of information on the case data level has no impact on the information on the aggregated level. To analyse this, it is useful to distinguish between loss of information by re-coding as "Unknown" and by "decreasing the categories in use". The following variables are aggregated :

- Mechanism of injury
- Place of accident
- Activity at the time of injury
- Part of body injured
- Product involved in the accident
- Product causing the injury
- Other product
- Sports

The aggregation of product codes and sports is done in the same way as the other variables. The product codes in the EHLASS version 1996 case data structure are aggregated into the HIEMS database structure by reducing the categories to the main categories of product codes. The 5-digit product codes of the case data structure are reduced to the first position.
The codes of sports in the EHLASS version 1996 case data structure are two digit or three digit codes. They are reduced to the first position by the aggregation. For the other variables the aggregation is realised by bringing the variables' code into the main categories (from 2-digit to 1-digit).

For these variables, loss of information by transformation from IREQ 1991 to EHLASS version 1996 only affects the aggregated data when a subcategory of IREQ 1991 has to be coded as "unknown" or "other specified" in the main category of the EHLASS version 1996 variable. Loss of information by "decreasing the categories in use" only affects the case data, but the aggregated data do not lose any information by this re-coding of case data.
The following variables are not subjected to an aggregation

- Treatment and follow-up
- Type of injury (diagnosis)

For only two categories of the variable "type of injury" is it necessary to re-code them to "type of injury, other specified". For "treatment and follow-up" the re-coding from IREQ to EHLASS version 1996 is unambiguous.

3. Transformation tables from IREQ 1991 to EHLASS version 1996

3.1 Treatment and follow-up

IREQ 1991	EHLASS v96	3	3	6	7
1	1	4	4	7	6
2	2	5	5	9	9

3.2 Activity at the time of accident

IREQ 1991	EHLASS v96	41	39	58	51	72	61
10	18	48	39	59	59	78	68
20	14	49	39	60	48	79	69
30	13	50	58	70	60	98	88
40	39	51	50	71	80	99	99

3.3 Site of accident

IREQ 1991	EHLASS v96	19	19	50	45	80	76
00	13	20	00	51	46	81	70
01	10	28	06	58	45	82	88
02	12	29	09	59	45	88	78
03	11	30	30	60	50	89	79
04	11	31	31	61	51	90	20
08	14	32	32	68	58	91	25
09	19	38	38	69	59	92	29
10	17	39	39	70	63	98	98
11	18	40	48	71	68	99	99
12	18	41	42	72	65		
13	19	48	48	78	68		
18	19	49	49	79	69		

3.4 Type of Injury (diagnosis)

IREQ 1991	EHLASS v96	06	07	12	14	18	98
01	02	07	08	13	16	19	98
02	03	08	09	14	17	97	97
03	04	09	10	15	18	98	98
04	05	10	11	16	19	99	99
05	06	11	13	17	12		

3.5 Mechanism of Injury

IREQ 1991	EHLASS v96	32	28	63	53	79	79[*)]
01	01	33	25	64	54	80	89
02	02	39	29	65	55	81	89
09	09	40	30	66	56	82	89
10	03	41	39	67	57	83	89
11	04	49	39	69	59	84	89
12	05	50	40	70	68	85	89
19	09	51	41	71	62	89	89
20	10	52	42	72	63	90	09
21	11	53	43	73	65	98	98
22	18	59	49	74	70	99	99
29	19	60	50	75	72		
30	21	61	51	76	73		
31	28	62	52	77	74		

[1)] The IREQ manual does not distinguish between thermal effect and electric or radiation effect. Therefore this category contains the accumulated values of the categories “69” and “79”. For case data and aggregated data is it necessary refer to this in the Data Dictionary.

3.6 Part of body affected

IREQ 1991	EHLASS v96	28	28	44	45	61	62
10	10	29	29	45	42	62	63
11	11	30	31	48	42	63	64
12	12	31	30	49	43	64	65
13	13	32	50	50	51	65	66
14	14	33	33	51	52	68	68
15	15	34	34	52	53	69	69
16	18	35	32	53	54	70	98
17	18	38	38	54	55	71	98
18	18	39	39	55	56	72	98
19	19	40	40	56	57	79	98
20	20	41	41	58	58	80	70
21	21	42	43	59	59	98	98
22	22	43	44	60	60	99	99

3.7 Sport

IREQ 1991	EHLASS v96	218	C2	422	E22	732	J42
000	A00	219	C2	423	E23	737	J43
001	A01	220	C20	424	E24	738	J4
002	A02	221	C21	425	E25	739	J4
003	A03	222	C23	428	E2	740	J60
004	A04	223	C22	429	E2	780	J98
005	A05	224	C24	430	E3	799	J99
008	A0	228	C2	480	E98	800	K00
009	A0	229	C2	499	E99	801	K01
010	A10	280	C98	500	F00	802	K02
011	A11	299	C99	501	F01	806	K03
012	A12	300	D00	507	F03	807	K10
013	A13	301	D01	508	F0	808	K0
018	A1	302	D02	509	F0	809	K0
019	A1	308	D0	510	G10	810	K50
020	A20	309	D0	511	G11	820	K20
021	A21	310	D10	512	G12	821	K21
022	A22	320	D20	518	G1	822	K2
023	A23	330	D30	519	G1	828	K2
028	A2	380	D98	520	G70	829	K2
029	A2	399	D99	530	G0	830	K31
030	A30	400	E00	570	F30	831	K30
031	A31	410	E10	571	F31	835	K40
038	A3	411	E11	572	F32	836	K42
039	A3	418	E1	580	F98[1)]	838	K4
040	A4	419	E1	599	F99[2)]	839	K4
080	A98	420	E20	600	H00	840	K98
099	A99	421	E21	601	H10	849	K9
100	B09	422	E22	602	H11	850	L10
110	B10	423	E23	603	H01	851	L20
111	B11	424	E24	604	H12	852	L00
112	B12	425	E25	606	H20	853	L01
113	B13	428	E2	607	H13	854	L02
114	B14	429	E2	608	H2	858	L98
115	B15	430	E3	609	H3	859	L99
116	B16	480	E98	610	H51	868	L6
117	B17	499	E99	611	H50	869	L6
118	B1	500	F00	680	H98	899	L9
119	B1	501	F01	699	H99	900	M10
120	B20	507	F03	700	J0	901	M12
121	B21	508	F0	701	J01	902	M00
122	B22	509	F0	702	J02	903	M01
123	B23	510	G10	703	J00	904	M03
128	B2	511	G11	704	J03	905	M02
129	B2	512	G12	705	J04	908	M98
180	B98	518	G1	706	J05	909	M99
199	B99	519	G1	707	X00	910	N00
200	C00	520	G70	708	J0	920	Q0

201	**C01**	**530**	**G0**	**709**	**J0**	**930**	**B30**
202	**C02**	**570**	**F30**	**710**	**J30**	**940**	**N80**
203	**C03**	**571**	**F31**	**711**	**J31**	**950**	**P20**
208	**C0**	**572**	**F32**	**718**	**J3**	**960**	**Blank**
209	**C0**	**580**	**F98**[1)]	**719**	**J3**	**980**	**Blank**
210	**C10**	**599**	**F99**[2)]	**720**	**J5**	**999**	**A99**
211	**C11**	**600**	**H00**	**730**	**J40**		
217	**C12**	**601**	**H10**	**731**	**J41**		

[1)] The IREQ manual does not distinguish between non-motorised wheel sports and motor sports. Therefore this category contains the accumulated values of the categories “F98” and “G98”. For case data and aggregated data is it necessary refer to this in the Data Dictionary.

[2)] In this category the values of the categories “F99” and “G99” are accumulated in accordance with the above-given explanation.

3.8 Product codes

As mentioned above, the product codes of the EHLASS 1986 version and IREQ 1991 were coded in the same way. Therefore we could use the transformation tables for product codes from Psytel. They allow to re-code IREQ 1991 product codes into EHLASS version 1996 product codes. This re-coding is again performed by category and in pairs as you can see in the tables above for the other variables.

You can find the transformation tables for product codes from the EHLASS 1986 version respectively IREQ 1991 into the EHLASS 1996 version in the “Report on EHLASS-Data Transformation – Final Report”. It is available in the IRC. The address (URL) in the WWW is http://forum.europa.eu.int/Members/irc/sanco/Home/main

After login in, select the “HIEMS” Group. From there the following links are: Library > Various Project Documents > EHLASS Report Enclosure.

4. Annotations on the variables “Hospitalised days” and “Product codes”

In the IREQ 1991 system, “Hospitalised days” of more than 68 days were coded as “69”. The EHLASS version 1986 system codes all “Hospitalised days” of more than 38 days as “39”. This different coding has no impact on the aggregated data because in “class 09” of the Final data structure of EHLASS data hospitalisation times of more than 30 days are coded as “09”, “more than 30 days”.

Remarks on Product Codes in EHLASS v1986, IREQ 1991 and the transformation tables from Psytel

The comparison of product codes from IREQ 1991 and the EHLASS 1986 version revealed three single inconsistencies. Apart from this, the product codes of the EHLASS 1986 version and IREQ 1991 are identical. For this reason, a transformation of IREQ 1991 product codes into the EHLASS 1996 version is possible based on the transformation tables from Psytel.

IREQ code not belonging to EHLASS v1986

- IREQ 1991 contained the category “Vapeur” with code 26440. This code did not belong to the EHLASS 1986 version. Therefore 26440 in IREQ 1991 had to be changed to 26488 “Kitchen utensil, other specified”
- IREQ 1991 contained the category “Huile de bain” with code 37510. This code did not belong to the EHLASS 1986 version. The same category “Huile de bain” appeared under IREQ 1991 with code 35710. This code was also a valid EHLASS version 1986 code for “Bath oils”, so that 37510 had to be changed to 35710.
- IREQ 1991 contained the category “Poubelle à pédales” with code 38000. This code did not belong to the EHLASS 1986 version. Therefore code 38000 had to be changed to 9 “unknown”.

EUROPEAN COMMISSION
DIRECTORATE-GENERAL III - DIRECTORATE-GENERAL V

Ida Euphin Project

Health Information Exchange and Monitoring System

VII FACTORY ACCEPTANCE TEST ON TEST DATA COLLECTION

Institute of Public Health NRW, Bielefeld, Germany

Version 1.3
24th June 1999

Table of Contents

1. Basis for the document

This paper is based on the following documents:

(Ref1) Contract between LOEGD and EU DG III/ REF. SUB/501026

(Ref2) IDA EUPHIN PROJECT, Functional Specification Version 1.0 of November 18th, 1998, Cap Gemini

(Ref3) EUPHIN HIEMS Global Implementation Plan, IHO20, Issue 3, Draft 1

(Ref4) A Test Strategy for HIEMS, IHOGK036, Version 1.1, 12th Feb, 1999

(Ref5) Minutes of the Health Monitoring Programme,Working Group B,Luxembourg, March 3rd, 1999

(Ref6) Factory Acceptance Test on Test Data Collection, Version 1.2, 28th May, 1999

(Ref7) System Test Plan, Version 0.21, 3rd May, 1999, Cap Gemini

(Ref8) W2Working Paper, IHOGK044, comments on "Factory Acceptance Test on Test Data Collection", 6th June, 1999

2. Test Principles

The deliverable product worked out by the lögd is the test data collection (Ref3, 3.4). This test data is to be collected, transformed and transferred for each of the six data sets on EHLASS, overnight patients, health care facilities, mortality and demography data and on the data dictionary. The task of the lögd as part of the Project Quality Plan is to ensure that data has been collected, transformed and transferred in a proper way. The methods used to ensure the quality of test data are carried out for each data set with emphasis on

- completeness
- integrity
- correctness end - to - end.

2.1 Completeness

The completeness test belongs to the quantity of the delivered data, e.g. it is checked whether the inquired data has been delivered by all responsible national authorities (RNA) or international data holders. The completeness test ensures that all mandatory data items such as different years, nuts-level and age-groups and other RAD-specific categories are taken into consideration. This test will be carried out on the original data delivered to the lögd. The test of completeness comprises different checks, according to the

- availability of the requested and mandatory data items
- period of reporting
- aggregation level
- data structure
- coding of the delivered data
- # records delivered

This is not actually a test, but a description of the delivered data. All results are written down in a corresponding test sheet and additionally in a transfer report which are handed over to the RNAs/international data holders.

2.2 Integrity

The task of the integrity test is to ensure that the

- data transformed into the HIEMS structure corresponds to the upload structure defined in the Functional Specification, Version 1.0, by CAP GEMINI (Ref2)
- agreed coding for the HIEMS data base is applied
- content of the data is plausible from a statistical point of view
- totals, sub-totals and single values coincide with the defined aggregation rules

The statistical calculations will be done on the data delivered by the data holders before transformation into the HIEMS structure. The calculations must comprise every mandatory dimension of a RAD set, e.g.

- different age groups
- different genders
- different categories

The output of these calculations will be used for the evaluation of the data quality delivered by the RNAs/international data holders and for the supervision of the transformation process.

For the test of integrity the rules for transformation of the original data into the HIEMS structure must be adhered too. All codes of the original data have to be checked and, for the case of miscodings, be recoded according to these rules.

After transformation, frequencies of the HIEMS codes in use will be calculated so that these results could be compared after the upload.

The results of all these tests will be recorded in seperate test sheets on data integrity.

2.3 Correctness end - to - end

Testing the correctness end - to - end gives a statement upon the

- quality of the transformation process of the original data
- result of the upload into the HIEMS database
- # records transformed
- # records uploaded

The aim of the correctness end-to-end test is to supervise the transformation and transfer processes. This test must ensure that the contents of the data in the HIEMS structure does not differ from the original data, that calculations on both levels will lead to the same results and that the transformation of the original data into HIEMS data structure was done without a loss of information. Furthermore, the tests must ensure that the quantity of transformed/transferred data equals the quantity at the input of the process or that the differences are reasonable.

For these tests the same statistical indicators will be calculated as for the original data.

If the tests on the HIEMS data lead to the same results as before the transformation, the data is correct end-to-end. If this is not the case, the transformation process must be checked.

If the quantity on input and output of a transformation/transfer process is equal or differences are reasonable (for example if observations are excluded from a file due to defined threshold values), the process is qualified as correct.

The results of these calculations could be handed over together with the test data delivery to CAP GEMINI so that the tests can be repeated after the upload of the data.

3. Data Dictionary

The input of the data dictionary is designed for using the functionality of the HIEMS data base. Neither an upload structure nor a transformation process exist. Therefore, the tests for the data dictionary have to evaluate the

- completeness and
- integrity

of the information given by the RNAs/international data holders in the data dictionary. The correctness end – to – end is guaranteed by the HIEMS system. Therefore, the test on the data dictionary consists of two steps:

- The availability of all requested descriptive information on the national register structure, on the national item definition and on the code information will be checked
- The information given on the national item definition in the data dictionary will be proved as far as possible by comparing it with the corresponding data delivered.

The availability of the requested information will be checked using a list of items. The assessment of the contents is mainly based on the knowlegde of the RAD in question.

4. Test Reporting

4.1 Deliverables

The delivarables of the lögd are the test data collection divided into the following 6 parts: mortality, demography, EHLASS, overnight patients, health care facilities and data dictionary.

The tests will be carried out seperately for all of these data sets. The test requirements described above will be adopted to the specific data structures of the 6 areas and will serve the quality assurance of the deliverable products.

4.2 Reporting

All results will be written down in seperate test reports. Consequently, six reports will be produced summarising and evaluating all tests for the different RAD sets and the data dictionary:

- Report on the completeness, integrity and correctness end – to – end for the Mortality RAD.
- Report on the completeness, integrity and correctness end – to – end for the Demography RAD.
- Report on the completeness, integrity and correctness end – to – end for the EHLASS RAD.
- Report on the completeness, integrity and correctness end – to – end forthe Overnight patients RAD.
- Report on the completeness, integrity and correctness end – to – end for the Health care facilities RAD.
- Report on the completeness and integrity of the Data Dictionary

The structure of all six reports is similar. The reports summarise and assess the test on the RAD under review and include all test sheets as an appendix, taking into consideration every country / data holder. With these reports the data holders are able to check - on a certain level - the whole process of data transformation and data transfer.
The following structure will be applied for the reports:

- Overview on the delivered, transformed and uploaded data
- Evaluation of the data completeness
- Evaluation of the data integrity
- Evaluation of the data correctness end to end
- Appendix: signed test sheets

4.3 Description of the test sheets

The basis for the test reports mentioned are standardised test sheets on completeness, integrity and correctness end – to – end. The contents of the sheets are as follows, whereby the sign (*) indicates that the contents of this field depend on the RAD set under review.

4.3.1 Test sheet on completeness

- Completeness of the original data containing the following information (Test Sheet No. 1):
  - RAD set
  - Country
  - Year of reporting
  - Data structure and coding(*)
  - Availability of mandatory items (*)
  - Comment on completeness
  - # of records
  - Filename

4.3.2 Test sheets on integrity

- Statistics calculated on the original data containing the following information (Test Sheet No. 2):
  - RAD set
  - Country
  - Year of reporting
  - Statistics that depend on the RAD set under review(*)
  - Comment on the calculated statistics
  - # of records
  - Filename

- Transformation process of the original data containing the following information (Test Sheet No. 3):
  - RAD set
  - Country
  - Year of reporting
  - Transformation rules(*)
  - Records transformed
  - Records rejected during transformation
  - Comment on the transformation process
  - Filename

- Frequencies of the HIEMS codes used in the RAD set containing the following information (Test Sheet No. 4):
  - RAD set
  - Country
  - Year of reporting
  - Table of code frequencies of the corresponding RAD set(*)
  - Comment on the coding
  - # of records
  - Filename

4.3.3 Test sheets on correctness end – to – end

- Descriptive statistics calculated on the data in HIEMS structure containing the following information (Test Sheet No.5):
  - RAD set
  - Country
  - Year of reporting
  - Table of descriptive statistics on the corresponding RAD set(*)
  - Evaluation on the statistics calculated on the original and on the HIEMS data
  - # of records

Filename

- Data uploaded into HIEMS database containing information on (Test Sheet No. 6):
  RAD set
  Country
  Year of reporting
  # records prepared for upload
  # records uploaded
  Comment on upload
  Filename

4.3.4 Test sheets on data dictionary

- Completeness of the data dictionary containing (Test Sheet No. 7):
  RAD set
  Country
  National register information
  National item description (*)
  Code Informations (*)
  Qualitativ assessment of the information given in the data dictionary
  Filename

- Integrity of the data dictionary containing (Test Sheet No. 8):
  RAD set
  Country
  Comment on integrity
  Filename

4.4 Overview on the test sheets

DECCXX00	**DE**mography data
EHCCXX00	**EH**LASS data
MECCXX00	**M**ortality, **E**urostat
MWCCXX00	**M**ortality **W**HO
OPCCXX00	**O**vernight **P**atients
HFCCXX00	**H**ealth care **F**acilities
DDDE	**D**ata **D**ictionary **D**emography
DDEHCC	**D**ata **D**ictionary **EH**LASS
DDME	**D**ata **D**ictionary **M**ortality, **E**urostat
DDMW	**D**ata **D**ictionary **M**ortality, **W**HO
DDOPCC	**D**ata **D**ictionary **O**vernight **P**atients
DDHFCC	**D**ata **D**ictionary **H**ealth care **F**acilities

where: CC is the country code,
XX is the year under review,
00 is the test number.

4.5 Test Range

The tests will be carried out on each RAD set and selected data.

Demography Data:	EUROSTAT, selected countries
EHLASS Data:	Data holder, selected years
Mortality Data EUROSTAT:	EUROSTAT, selected countries
Mortality Data WHO:	WHO, selected countries
Overnight Patients Data:	Data holder, selected years
Health Care Facilities Data:	Data holder, selected years.

TEST CODE	EHCCXX00
TEST TITLE	**Completeness:** Completeness of the original data
TEST GOAL DESCRIPTION	The goal of the test is to verify that the requested data is delivered and that all mandatory data items are available
RAD	EHLASS
COUNTRY	
YEAR	
Data holder	
Data structure	☐ EHLASS 86 ☐ EHLASS 96 ☐ NATIONAL STRUCTURE
Data coding	☐ EHLASS 86 ☐ EHLASS 96 ☐ WITH ADDITIONAL NATIONAL CODING
Availability of the data items (A/NA)	
date of treatment	☐
time of attendance	☐
number of hospitalised days	☐
treatment and follow up	☐
age	☐
sex	☐
accident mechanism	☐
site of accident	☐
acitivity at the time of accident	☐
product involved in accident	☐
product causing injury	☐
Comment on availability	
Number of records	
Filename	
Tester / Date	

TEST CODE	EHCCXX00
TEST TITLE	**Integrity:** Statistics calculated on the original data
TEST GOAL DESCRIPTION	The goal is to calculate statistics on the original data in order to verify the transformation process and to identify data problems on different aggregation levels
RAD	EHLASS
COUNTRY	
YEAR	
# accidents by all ages, male # accidents by all ages, female # accidents by all ages, unknown sex	
# accidents by age group 0-19, male # accidents by age group 20-64, male # accidents by age group 65+, male # accidents by age group, unknown age, male	
# accidents with treatment=5, male, all ages # accidents with treatment=5, female, all ages # accidents with treatment=5, unknown sex, all ages	
Comment	
Number of records	
Filename	
Tester / Date	

TEST CODE	EHCCXX00
TEST TITLE	**Integrity:** Transformation process of the original data
TEST GOAL DESCRIPTION	The goal is to describe the rules for the transformation process of the national data in order to fulfil the HIEMS requirements and to quantify the # records transformed
RAD	EHLASS
COUNTRY	
YEAR	
Transformation rules for the data structure and data coding	
Final data structure HIEMS	
# records transformed to HIEMS coding	
# records rejected during transformation to HIEMS coding	
# records aggregated into HIEMS structure	
# records rejected during transformation into HIEMS structure	
Comment on transformation process	
Filename	
Tester/Date	

TEST CODE	EHCCXX00	
TEST TITLE	**Integrity:** Frequencies of HIEMS codes used in the EHLASS RAD	
TEST GOAL DESCRIPTION	The goal is to review all codes and to calculate frequencies of these codes	
RAD	EHLASS	
COUNTRY		
YEAR		
class 01	variable x A Raw materials, structural elements and particles B Stationary equipment outside, processed surface and natural surface C Part of building and stationary furniture D Installations for water, sanitation and electricity E Electric equipment for use in household F Furniture and textile in building G Domestic appliances and equipment H Machinery for industry , handicraft and hobby I Office and shop furniture J Medico-technical equipment K Means of transport L Toys M Musical instrument, optical equipment N Sports equipment P Clothing, baby caring articles Q Food, beverages, tobacco R Chemical products S Packaging, containers T Human beings, animals, animal articles U Natural element V Other and unspecified product	#
class 02	01 Concussion 02 Contusion, bruise 03 Abrasion 04 Open wound 05 Fracture 06 Luxation, dislocation 07 Distortion, sprain 08 Lesion of nerve(s) 09 Lesion of the blood vessel(s) 10 Lesion of tendon(s) and/or muscle(s) 11 Crushing 12 Amputation 13 Poisoning 14 Burns, scalds (thermal) 15 Corrosion 16 Electrocution 17 Radiation (sunlight, x-rays) 18 Frostbite 19 Asphyxia 97 No injury diagnosed 98 Other specified injury 99 Unspecified injury	

class 03	0 Struck, hit by, fall 1 Struck, hit by contact with other object, person or animal 2 Crushing, cutting, piercing 3 Foreign body in natural orifice 4 Suffocation 5 Chemical effect 6 Thermal effect 7 Electric/radiation 8 Acute overexertion of body 9 Other and unspecified mechanism of injury	
class 04	0 Transport area 1 Residential area 2 Production and workshop area 3 Retail, commercial and service area 4 School, institutional area and public premises 5 Sports area 6 Amusement, entertainment and park area 7 Open nature 8 Sea, lake and river 9 Place, other and unspecified	
class 05	1 Domestic work 2 Do-it-yourself work 3 Educational activity 4 Play and leisure activity 5 Sports, athletics, exercise 6 Vital activity 8 Other specified activity 9 Unspecified activity	
class 06	1 Head 2 Neck, throat 3 Thorax 4 Abdomen, lower back, lumbar spine and pelvis 5 Upper extremities 6 Lower extremities 7 Multiple body parts/whole body affected 9 Other and unknown body part	
class 07	1 Examined and sent home without treatment 2 Sent home after initial treatment 3 Treated and referred for further treatment by general practitioner 4 Treated and referred for further treatment as an outpatient 5 Treated and admitted to this hospital 6 Transfered to an other hospital 7 Deceased 9 unknown	
sex	1 male 2 female 3 unknown	
Comments on coding		
Number of records		
Filename		
Tester / Date		

TEST CODE	EHCCXX00
TEST TITLE	**Correctness:** Descriptive statistics calculated on the data in HIEMS structure
TEST GOAL DESCRIPTION	The goal of the test is to calculate statistics on the data transformed for input into HIEMS and to compare the results with the same calculation done on the original data
RAD	EHLASS
COUNTRY	
YEAR	
# accidents by all ages, male # accidents by all ages, female # accidents by all ages, unknown	
# accidents by age group 0-19, male # accidents by age group 20 – 64, male # accidents by age group 65+, male #accidents by age group, unknown age, male	
# accidents with treatment=5, male, all ages # accidents with treatment=5, female, all ages # accidents with treatment=5, unknown sex, all ages	
Evaluation of the statistics calculated on the original and on HIEMS data	
Number of records	
Filename	
Tester / Date	

TEST CODE	EHCCXX00
TEST TITLE	**Correctness:** Data uploaded into the HIEMS data base
TEST GOAL DESCRIPTION	The goal of the test is to verify that the records are uploaded completely and that no records have been rejected due to incorrect data structure or item format
RAD	EHLASS
COUNTRY	
YEAR	
# records prepared for upload	
# records uploaded	
Comment on upload	
Filename	
Tester/Date	

TEST CODE	DDEHCC00
TEST TITLE	**Completeness:** Completeness of the Data Dictionary on the EHLASS RAD
TEST GOAL DESCRIPTION	the goal of the test is to verify that the requested data are delivered, that all mandatory data items are available
DD-RAD	DD-EHLASS
COUNTRY	
Availability of the data items (A/NA)	
Register name	☐
Purpose	☐
Collection method	☐
Collection method description	☐
Volume observation	☐
Volume variables	☐
Update	☐
Reliability (not mandatory)	☐
Timeliness (not mandatory)	☐
Comprehensiveness (not mandatory)	☐
Comparability (not mandatory)	☐
Comment on availability	
Filename	
Tester/Date	

TEST CODE	DDEHCC00
TEST TITLE	**Integrity:** Integrity of the Data Dictionary on the EHLASS RAD
TEST GOAL DESCRIPTION	the goal of the test is to verify that the information given in the Data Dictionary corresponds to the data delivered
DD-RAD	DD-EHLASS
COUNTRY	
Comment on integrity	
Filename	
Tester/Date	

VIII STRUCTURE OF THE DATA DICTIONARY

1. NATIONAL REGISTER STRUCTURE

2. DATA ITEM DESCRIPTION

3. INFORMATION ON THE CLASSIFICATION

version 1.1
22th February 1999

Argumentation for changes in the Data Dictionary

The structure of the Data Dictionary is based on the concept proposed by the Danish Ministry of Health (Ref.: 22-03-V1665).
Some changes -accepted at the Kick-Off-Meeting in Bielefeld (Sept. 16/17, 1998)- concerned the items:
- reliability
- timeliness
- comprehensiveness
- comparability.

They were moved from the “National Register Structure“ to the “Data Item Descripion“ part because these items are mainly related to idividual items.

Additionally, it was proposed and accepted to add the item: “Last Alteration of Item Definition“ and to fix the period of validity of a certain definition.

1: NATIONAL REGISTER STRUCTURE

Data Item	Position	Format
Country and area	1-4	AN4
Register name	5-24	AN20
Purpose	25-1048	AN1024
Collection method	1049-1049	N1
Collection method description	1050-2073	AN1024
Volume observation	2074-2083	N10
Volume variables	2084-2086	N3
Update	2087-2094	N8

1.1 Country and area

Country and area (4 characters)	The first two characters are devoted to a country abbreviation. The remaining 2 characters are reserved for the sub-national level code extension and should be left blank if data corresponds to the whole country. The delivery of these data, related to the sub-national regions is optional. The country codes are: AT = Austria BE = Belgium DE = Germany DK = Denmark ES = Spain FI = Finland FR = France GR = Greece IE = Ireland IS = Iceland IT = Italy Li = Liechtenstein LU = Luxembourg NL = Netherlands NO = Norway PT = Portugal SE = Sweden UK = United Kingdom Source: ISO 3166 EU = Used for the HIEMS Definition

1.2. Register name

Register name (20 characters)	Name of the national register

1.3 Purpose

Purpose (1024 characters)	Textual description of the purpose and contents (in broad terms) of the original/national register:

1.4 Collection method

collection method (1 character)	Indication of data collection method used to populate the register (1= survey, 2= register, 9= other)

1.5. Collection method description

collection method description (1024 characters)	Textual description data collection method used to populate the register

1.6. Volume observations

Volume observation (10 characters)	Number of observations in the register

1.7. Volume variables

Volume variables (3 characters)	Number of variables in the register

1.8. Update

Update (8 characters)	Date of latest update of the register Date: yyyymmdd e. g. 19981019 = 19.10.1998 = oct. 19, 1998

2. DATA ITEM DESCRIPTION

Data Item	Position	Format
Country and area	1-4	AN4
Item name	5-24	AN20
Item format	25-34	AN10
Item definition	35-1058	AN1024
Register name	1059-1078	AN20
Item classification ID	1079-1098	AN20
Conversion information	1099-2122	AN1024
Reliability	2095-3118	AN1024
Timeliness	3119-4142	AN1024
Comprehensiveness	4143-5166	AN1024
Comparability	5167-6190	AN1024
Last alteration of item definition	6191-6198	N8
Previous item definition	6199-7223	AN1024

2.1 Country and area

Country and area (4 characters)	The first two characters are devoted to a country abbreviation. The remaining 2 characters are reserved for the sub-national level code extension and should be left blank if data corresponds to the whole country. The delivery of these data, related to the sub-national regions is optional. The country codes are: AT = Austria BE = Belgium DE = Germany DK = Denmark ES = Spain FI = Finland FR = France GR = Greece IE = Ireland IT = Italy LU = Luxembourg NL = Netherlands PT = Portugal SE = Sweden UK = United Kingdom Source: ISO 3166 EU = Used for the HIEMS Definition

2.2 Item name

Item name (20 characters)	Identification of the data item in question

2.3 Item format

Item format (10 characters)	Format of the data item in question, e. g. AN10

2.4. Item definition

Item definition (1024 characters)	Textual definition of the data item in question

2.5. Register name

Register name (20 characters)	Name of the register as a link to the „National Register Structure“

2.6. Item classification ID

Item classification ID (20 characters)	Foreign key for the table defining the classification of the data item in Question.

2.7. Conversion information

Conversion information (1024 characters)	Textual description of the conversion method, problems or shortcomings, if any, related to the conversion to the HIEMS format of the data item in question

2.8. Reliability

Reliability (1024 characters)	Textual description of reliability

2.9. Timeliness

Timeliness (1024 characters)	Textual description of timeliness

2.10. Comprehensiveness

Comprehensiveness (1024 characters)	Textual description of comprehensiveness

2.11. Comparability

Comparability (1024 characters)	Textual description of comparability

2.12. Last alteration of item definition

Last alteration of item definition (8 characters)	Date of latest alteration of item definition. These item definitions are effective from the date shown. Date: yyyymmdd e. g. 19981019 = 19.10.1998 = oct. 19, 1998

2.13. Previous item definition

Previous item definition (1024 characters)	Textual description of content and time period a previous used item definition was in force. Only fill in, if a change of definition occured within the space of time the supplied data covered.

3. INFORMATION ON THE CLASSIFICATION

Data Item	Position	Format
Item classification ID	1-20	AN20
Item value	21-30	AN10
Item value description	31-158	AN128
HIEMS Item value	159-168	AN10

3.1. Item Classification ID

Item classification ID (20 characters)	Link to „Data Item Description“

3.2. Item value

Item value (10 characters)	Identification of a certain value in the classification of the data item in question

3.3. Item value description

Item value description (128 characters)	Textual description related to the item value

3.4. Hiems item value

Hiems item value (10 characters)	Identification of the corresponding value in the HIEMS classification of thr data item in question.

IX Data Dictionary
Information on EHLASS within the EUPHIN-HIEMS Database
(April 2001)

AT	Data are collected at the a&e units by specially trained interviewers with the injury patients. Oral interviews by specially trained interviewers with injury patients at the a&e dept., data are collected at general a&e units; 4 hospitals participated throughout the study period (3 general, 1 accident), one hospital was replaced in 1997.
BE	The EHLASS register is a hospital register in which each person exposed to an accident at home or during leisure time as well as at school and applying to the emergency service of a co-operation hospital enters the register.4 hospitals participate.4,7 % of all accidents treated in hospitals are covered by EHLASS. 4 hospitals delivered data from 1991-1997, one hospital left the study in 1998. All hospitals were general hospitals.
DE	Consumer protection by estimating the total accident figure; collection of survey data by interview:face to face, telephone and mail. 1996, a survey on accidents with over 100,000 households has been carried out. For home and leisure accidents, data from a survey from 1988/1989 for West-, and from 1991/1992 for East-Germany was extrapolated on the basis of the 1996 survey.
DK	EHLASS.DK collects data at 5 emergency departments. The emergency departments' uptake area cover 14,5% of the population in a representive way. Data for Ehlass are collected at 5 emergency department, covering approx. 14.5% of the population in a representative way. All participating hospitals are general public hospitals.
ES	For the years 1996 - 1999 surveys on the basis of face-to-face interviews were carried out.
FI	**:**two non random selected hospitals collect the data, EHLASS is a limited and non-representative sample, EHLASS covers less than 5 % of all accidents treated in hospitals For Ehlass, data is collected in two non-random selected hospitals, one of them providing paediatric care. The sample is not representative and covers less than 5% of all accidents treated in hospitals.
FR	information is collected within the normal hospital registration procedure, the France EHLASS data is not representative, the sample of hospitals is not constant during 1987-1997 Information is collected within the normal hospital registration procedure. 13 general public hospitals participated over the time, 5 of them over the whole study period. For one hospital, data from a padeatric emergency unit was collected. Data is not representative.
GR	Hellenic Ministery of Health and Welfare/Centre for Research and Prevention of Injuries among the Young. Abbr: CEREPRI Data for Ehlass are collected since 1986, participating hospitals changed over time. Since 1994, data are delivered from 2 general public in Athens and one trauma and one paediatric hospital. Interviewers are specially trained on coding and data entry.
IE	Collection of data from hospitals relating to accidents in the home or in the leisure activities to fulfil Ireland's EHLASS requirements. Each qualifying case presented at the accident and emergency unit of two regional hospitals is reported on a pro forma questionaire. In 1998, 2 general public hospitals collect the data at a&e departments
IS	Data is collected by the only a&e department at the University Hospital in Reykjavik. For the Coding of 1998 and 1999 data, the NOMESCO Classification was used. Nearly all accidents were collected.
IT	Five general public hospitals collected data in 1998.
LI	no information
LU	Data supply by local administrations. Surveys have been carried out by different national and international institustions. Surveys on the basis of telephone interviews have been

	carried out from 1996 to 1999. The sample is representative, covering about 800 accidents for 1996.
NL	LIS is a registration of injury-patients treated at Emergency Departments of hospitals in the Netherlands. It is developed on the one hand to provide information for the developments and evaluation on injury-prevention policies, and on the other hand to supply medical information for the improvement and evaluation of the relief and treatment of patients. The majority of participating hospitals also use LIS to register information about non-injury-patients (e.g. patients who come in for a checkup or with a disease). In those hospitals LIS gives insight into the total population visiting the Emerg. Department. ADD: The national surveillance system LIS collects data from 14 a&e units, only seven of them are part of EHLASS. The data in the EUPHIN database is not representative.
NO	4 hospitals participated in 1998. All hospitals were public, and provided general care. The NIR classification system is based on the "Classification for Accident Monitoring" from Nordic Medico-Statistical Committee (NOMESCO, 2nd revised edition, 1990), with some minor modifications of codes and definitions.
PT	6 hospitals participate, 8% of all cases are covered by the 6 hospitals Since 1986, data for EHLASS is collected in several hospitals which changed throughout the study period. In 1998, 5 hospitals participated.
SE	4 hospitals collect and code the data. The data are stored together in a national database. 2,5% of all accidents treated in hospitals are covered EHLASS. The number of hospitals changed over time, in 1998 three hospitals participated. All hospitals were public, and provided general care.
UK	Random sample and population based. Data delivery by hospitals (samples) and office of national statistics. Hospitals participating in EHLASS were not the same over the study period. In 1998, twelve general hospitals delivered data.

X Description of the CD ROMs

The CD ROMs which the loegd produced for each of the 17 participating countries contain the follow descriptions:

Front of the Cover:

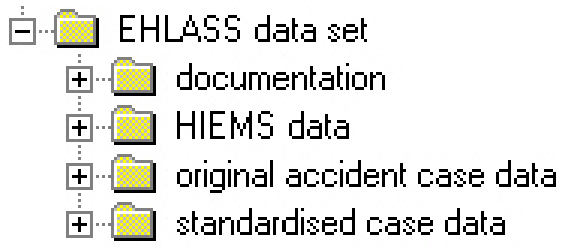
EUPHIN HIEMS / IPP
Project IPP-1040

For instance:

Complete EHLASS data set
Original and transfer data

Austria

Years under report 1996-1999



Backside of the Cover:

Content of the folders:

Documentation: This folder contains the FATs (Factory acceptance tests), the documentation to the data transformation, SAS scripts which carry out elementary statistical analysis for the FATs and the output of the scripts.

HIEMS data: This folder contains aggregated data in the transfer data structure of the Final data structure for EHLASS data. Files with data which are generated by the extended conversion programme have in their name the letter e in front of the dot.

Original accident case data: Here are the original accident case data which were sent from the participating country to the loegd.

Standardised case data: You can find here the accident case data in the input structure and coding of the aggregation programme. Files with accident data which are in the extended input structure of the extended conversion programme have in their name the letter e in front of the dot.

The folder “HIEMS data” contains ASCII-text files with the extension “hie”. They were generated as output of the aggregation programme for EHLASS data.
The folder “standardised case data” also contains ASCII-text files with the extension “out”.

The list of the participating countries

For each of the following countries were produced 3 CD ROMs**:**

Austria
Belgium
Denmark
Finland
France
Germany
Greece
Iceland
Ireland
Italy
Luxembourg
The Netherlands
Norway
Portugal
Spain
Sweden
United Kingdom

XI EHLASS partners in the participating member states

April 2001

Austria
Robert Bauer
Institut Sicher Leben
Zaunergasse 4
1030 Wien
Austria
Phone: +43 17 15 66 44 317
Fax: +43 17 15 66 44 30
robert.bauer@sicherleben.at

Belgium
Peter Hooft
Administratie Gezondhiedszorg
Ministerie Van De Vlaamse Gemeenschap
Markiesstraat, 1
1000 Brussels
Belgium
Phone: +32 2 553 35 00
Fax: +32 2 553 35 84
peter.hooft@wvc.vlaanderen.be
(Dr Christel Tritsmans
christel.tritsmans@health.fgov.be)

Denmark
Birthe Frimodt-Moller
National Institute of Public Health
Svanemollevej 25
2100 Copenhagen
Denmark
Phone:+45 31 20 77 77
Fax:+45 31 20 80 10
bfm@dike.dk

Finland
Dr Mikko Nenonen
Stakes, National Research
and Development Centre for Welfare and Health
Siltasaarenkatu 18,
P.O. Box 220
00531 Helsinki
Finland
Phone:+358 9 3967 2346
Fax:+358 9 3967 2459
mikko.nenonen@stakes.fi

France
Dr Christine Duval

Ministère des Affaires Sociales, de la Santé et de la Ville
DGS-VS 1
1, Place Fontenoy
F-75350 Paris SP O7
France
Phone:+33-1 40 56 40 27
Fax:+33-1 40 56 50 56

Germany
Frau Henter
Bundesanstalt für Arbeitschutz und Arbeitsmedizin
Friedrich-Henkel-Weg 1-25
44149 Dortmund
Germany
Phone: +49 231 9 07 12 41
Fax: +49 231 9 07 14 65
a.henter@baua.de

Greece
Dr Eleni Petridou
University of Athens
Medical School
Depart. of Hygiene and Epidemiology
75, M. Asias Street
11527 Athens
Greece
Phone:+30 77 73 840 or +30 77 71 165
Fax:+30 77 73 840 or +30 77 04 225
epetrid@cc.uoa.gr

Iceland
Dr Brynjólfur Mogensen
Reykjavík Hospital
University of Iceland
Reykjavík
Iceland
Phone: +354 52 51 707
Fax:+354 52 51 702
brynjolf@shr.is

Ireland
Hugh Magee
Information Management Unit
Department of Health and Children
Hawkins House
Dublin 2
Ireland
Phone:+35 31 6 71 47 11
Fax:+35 31 6 71 19 47
hugh_magee@health.irlgov.ie

Italy
Dott. Mariano Alessi
Ministerio della Sanità
Dipartimento Prevenzione Uff.X
Wia della Sierra Nevada, 60
00187 Roma
Italy
Phone: +39 06 / 59 944 260
Fax:+39 06 / 59 944 249
m.alessi@sanita.it

Luxembourg
Dr Isabelle Portal-Rolland
Centre de Recherche Pour La Sante – CRP – Sante
Route d'Arlon, 57
1140 Luxembourg
Phone: +352 45 32 13. 1
Fax: +352 45 32 19
isabelle.portal@crp-santé.lu

The Netherlands
Saakje Mulder
Consumer Safety Institute
Rijswijkstraat 2
1059 GK Amsterdam
The Netherlands
s.mulder@consafe.nl
Phone:+31 20 511 45 11
Fax:+31 20 669 28 31
s.mulder@consafe.nl

Norway
Arve Sjolingstad
National Institute of Public Health
PO Box 4404 Torshov
0403 Oslo
Norway
Phone:+47 22 04 25 65
Fax:+47 22 35 37 05
arve.sjolingstad@folkehelsa.no

Portugal
Dr. Carla Barata
Instituto do Consumidor
Departemento de Estudo de Mercado
Praca Duque de Saldanha, 31 - 10, 20, 30, 50
1050 Lisboa
Portugal
Phone:+351 356 46 00

Fax:+351 356 47 20
carla.barata@ic.pt

Spain
Nelson Castro
Instituto Nacional del Consumo
Principe de Vergara, 54
28071 Madrid
Espagne
Phone:+34 14 31 18 36 – ext 278
Fax:+34 15 76 39 27
nelson.castro@consumo-inc.es

Sweden
Anders Aberg
Epidemiologiskt Centrum
Söcialstyrelsen
Linnégatan 87
10630 Stockholm
Sweden
Phone: +468-783-31 17
anders.aberg@sos.se

United Kingdom
Dr. Phil Raine
Department of Trade and Industry
1, Victoria Street
SW1H OET London
United Kingdom
Phone:+44 712 15 03 81
Fax:+44 712 15 03 57
phil.raine@cacp.dti.gov.uk

XII Manual for handling the remote database

Betrieb der Remote Database in Bielefeld

1. Start des Betriebssystems UNIX-AIX auf der RS-6000

Nach dem Einschalten des Monitors und der externen Festplatte, wahlweise des externen CD-ROM Laufwerkes und des externen Bandlaufwerkes, kann der Rechner eingeschaltet werden.
In der Regel sollten das CD-ROM Laufwerk und das externe Bandlaufwerk nur bei Bedarf betrieben werden. Zum Start eines zusätzlichen externen Gerätes muß das System erneut gebootet werden, nachdem zuvor das CD-ROM Laufwerk oder das externe Bandlaufwerk eingeschaltet worden ist.

Einige Minuten nach dem Einschalten des Rechners ist das UNIX Betriebssystem betriebsbereit. Es erscheint ein Dialogfenster, in welchem der Benutzername und das Kennwort für den Rechnerzugang eingegeben werden muß.

Benutzername (für UNIX-AIX)**: root**
Kennwort: ****

2. Start der ORACLE Datenbank

Vor dem Abschalten des Rechners ist unbedingt eine ordnungsgemäße Beendigung der ORACLE Datenbank erforderlich, da anderenfalls Schäden an der Softwareinstallation zu erwarten sind. Für die Eingaben zum Start der ORACLE Datenbank muß ein Terminalfenster auf der UNIX-Oberfläche geöffnet werden. Der Wechsel in die ORACLE-Umgebung erfolgt mit dem Befehl

su - oracle

Die ORACLE Datenbank wird an dieser Stelle mit

dbstart

aufgerufen. Die Datenbank ist nun betriebsbereit. Die ORACLE-Umgebung kann wieder mit „**exit**“ verlassen werden.

3. Beenden der ORACLE Datenbank

Zum Beenden der ORACLE Datenbank muß über ein Terminalfenster die ORACLE-Umgebung aufgerufen werden (s.o.). Wird in der ORACLE-Umgebung der Befehl

dbshut

aufgerufen, endet die Ausführung der Datenbank ordnungsgemäß.

4. Herunterfahren des UNIX-AIX Betriebssystems

Bevor das Betriebssystem heruntergefahren werden darf, muß die ORACLE-Datenbank unbedingt geschlossen werden. Die ORACLE-Umgebung wird verlassen mit „**exit**“. Durch den Befehl

cd $home

gelangt man in das Startverzeichnis, in dem das Betriebssystem sich mit der Ausführung des Befehls

shutdown

beenden läßt. Nach dem Aufruf von Shutdown verbleiben etwa 2 Minuten um die aktuelle Sitzung, mit dem EXIT-Icon auf dem Desktop, zu beenden. Wenn die Meldung ‚file system unmounted‘ auf dem Bildschirm erscheint, ist das Betriebssystem heruntergefahren und alle Geräte können ausgeschaltet werden.

5. Einige Befehle zur Überprüfung der Netzwerkfunktionalität

Für den Betrieb der Remote Database im lögd existiert ein LAN (Local Area Network), welches mittels eines Routers an das VPN (Virtual Private Network) angeschlossen ist. Das LAN besteht aus drei Geräten mit folgenden IPs :

- **Dem Router und Gateway mit der IP 194.94.1.10**
- **Die RS6000 mit der IP 194.94.1.1**
- **Der Client mit der IP 194.94.1.2**

Zur Eingabe von Befehlen kann entweder ein Terminalfenster auf der RS6000 oder ein MS-DOS-Fenster auf dem Client geöffnet werden. Der Kontakt zwischen den Rechnern läßt sich prüfen durch den Befehl **ping** IP, wobei IP stellvertretend für die oben angeführten IPs der Geräte des LANs steht. Ist eine Verbindung möglich, erhält man eine entsprechende Rückmeldung. Mit dem Befehl

tracert 10.5.1.70 läßt sich das Routing bis zum Zentralen Rechner jenseits („draussen“) des Routers im VPN verfolgen.
Der Befehl **netstat** oder auch **netstat –r** gibt ebenfalls Auskunft über erreichbare Verbindungen.

6. Starten des Clients

Nach dem Einschalten des Clients erscheint das Dialogfenster zum Anmelden beim Betriebssystem des Clients.

Benutzername: Administrator
Kennwort: ****

7. Zugang zur EUPHIN HIEMS Datenbank

Die EUPHIN HIEMS Datenbank kann mittels des installierten Netscape-Browsers erreicht werden. Die URL im VPN für die EUPHIN HIEMS Datenbank lautet:

http://10.5.1.70:10007/hiems/plsql/hiems_package.index_htm

Diese Adresse ist als Startseite im Browser voreingestellt. Mit dem Aufruf von Netscape erscheint sofort ein Dialogfenster für die Anmeldung beim Betriebssystem der EUPHIN HIEMS Datenbank. Zur Zeit ist eine Anmeldung als einfacher Nutzer oder als Administrator möglich. Letzterer verfügt über mehr Zugangsrechte.

Anmeldung als Nutzer:
Benutzername: Deloegd2
Kennwort: ****

Anmeldung als Administrator:
Benutzername: Loegd
Kennwort: ****

8. Direkter Zugang zu der Remote Database in Bielefeld

Die Remote Database in Bielefeld enthält z.Z. alle verfügbaren deutschen Daten in der EUPHIN HIEMS Datenbank. Auf diese Daten kann bei gestarteter ORACLE Datenbank direkt zugegriffen werden. Dies kann sowohl über die RS6000 als auch über den Client geschehen. Um Abfragen (free queries) direkt in der Remote Database zu richten, muß auf der RS6000 innerhalb der ORACLE Umgebung SQLplus aufgerufen werden mit dem Befehl:

sqlplus

Es erscheinen nun zwei Eingabeaufforderungen zur Anmeldung bei der ORACLE Datenbank:

Enter user-name: hiems

Enter password: ****

Nun können SQLplus Befehle direkt über die Kommandozeile eingegeben werden.

Die zweite Möglichkeit, Abfragen direkt an die Remote Database zu richten, besteht durch das SQL Worksheet auf dem Clientrechner. Nach dem Start des SQL Worksheets werden folgende Informationen zur Anmeldung an die ORACLE Datenbank abgefragt:

Login: hiems
Password: ****
Services: hiems

Nach erfolgter Verbindung mit der ORACLE Datenbank sind SQL plus Befehle im SQL Worksheet ausführbar.

XIII Available EHLASS case data and aggregated data uploaded into the EUPHIN HIEMS database

Country by full name and short form (ISO 3166)	Year under report	Accident case data from participating countries	Sum of accident case data per country	# of records in the final data structure for EHLASS data
Austria	**1996**	12.053		1.484
AT	**1997**	10.613		1.347
	1998	8.079		1.404
	1999	9.078		1.298
			39.823	
Belgium	**1991**	1.7978		1.846
BE	**1992**	18.776		1.867
	1993	17.806		1.859
	1994	20.054		1.860
	1995	21.415		1.878
	1996	20.601		1.834
	1997	21.220		1.845
	1998	16.349		1.754
	1999	13.868		1.798
			168.067	
Denmark	**1989**	58.897		1.773
DK	**1990**	62.764		1.793
	1991	63.888		1.809
	1992	67.177		1.786
	1993	67.531		1.728
	1994	69.435		1.727
	1995	70.651		1.751
	1996	70.714		1.755
	1997	71.565		1.767
	1998	65.671		2.191
	1999	70.431		2.177
			738.724	
France	**1986**	2.885		1.276
FR	**1987**	19.758		2.139
	1988	37.978		2.414
	1989	38.075		1.982
	1990	34.640		2.103
	1991	34.007		2.191
	1992	33.616		2.073
	1993	28.585		1.915
	1994	49.130		2.304
	1995	55.657		2.177
	1996	53.618		2.081
	1997	50.522		1.998
	1998	41.823		1.919
	1999	41.307		2.041
			521.601	
Finland	**1995**	2.389		1.119

FI	**1996**	14.312		1.862
	1997	12.872		1.789
	1998	7.076		1.433
			36.649	
Germany	**1996**	9639		1368
DE			**9.639**	
Greece	**1990**	9.461		1.258
GR	**1991**	6.735		1.108
	1992	8.538		1.091
	1993	7.950		1.019
	1994	16.973		1.643
	1995	26.631		1.648
	1996	32.722		1.514
	1997	32.782		1.578
	1998	33.306		1.586
	1999	32.602		1.882
			207.700	
Iceland	**1998**	22.267		519
IS	**1999**	27.197		508
			49.464	
Ireland	**1990**	10.687		1.465
IE	**1991**	10.009		1.299
	1992	9.627		1.239
	1993	9.208		1.282
	1994	9.556		1.290
	1995	6.874		1.174
	1996	9.271		1.190
	1997	9.421		1.333
	1998	6.213		1.271
			80.866	
Italy	**1986**	918		840
IT	**1987**	9.660		1.393
	1988	8.838		1.342
	1989	706		707
	1991	7.651		1.445
	1992	9.831		1.446
	1993	8.324		1.371
	1994	11.553		1.461
	1995	9.810		1.368
	1996	14.366		1.492
	1997	15.153		1.385
	1998	13.523		1.338
			110.333	
Luxembourg	**1996**	794		735
LU	**1997**	679		964
	1998	777		981
	1999	781		1.046
			3.031	

The Netherlands	**1986**	52.257		2.167
NL	**1987**	54.243		2.128
	1988	52.099		2.058
	1989	56.257		1.916
	1990	57.535		1.998
	1991	56.672		1.867
	1992	56.283		1.803
	1993	56.248		1.771
	1994	58.342		1.821
	1995	59.386		1.873
	1996	58.177		1.814
	1997	60.534		1.928
	1998	60.290		1.726
	1999	59.924		1.621
			798.247	
Norway	**1996**	21.466		1.162
NO	**1997**	20.879		1.188
	1998	20.033		1.153
	1999	19.176		1.151
			81.554	
Portugal	**1987**	55.768		1.769
PT	**1988**	53.953		1.772
	1989	58.586		1.871
	1990	39.467		1.717
	1991	36.583		1.604
	1992	21.636		1.426
	1993	19.073		1.565
	1994	30.308		1.679
	1995	35.480		1.678
	1996	52.677		1.901
	1997	55.275		1.808
	1998	37.941		1.792
			496.747	
Spain	**1996**	3.240		1.537
ES	**1997**	3.724		1.473
	1998	4.190		1.694
	1999	4.417		1.070
			15.571	
Sweden	**1996**	17.436		1.611
SE	**1997**	12.035		1.609
	1998	28.200		2.268
	1999	27.279		2.890
			84.950	
United Kingdom	**1986**	7.369		1.097
UK	**1987**	78.565		2.382
	1988	175.676		2.451
	1989	193.223		2.478
	1990	196.339		2.393

	1991	196.527		2.400
	1992	182.000		2.468
	1993	140.123		2.614
	1995	154.270		2.624
	1996	225.386		2.658
	1997	224.009		2.733
	1998	208.639		2.745
			1.982.126	
Totals of all countries and for all available years under report			**5.425.092**	

The availability of accident case data aggregated and uploaded into the EUPHIN HIEMS database by loegd by country ranges from 3031 cases from Luxembourg and 9639 cases from Germany at the lower end to 1 982 125 cases from the United Kingdom at the upper end. The number of records in the right column describes the size of data in the final data structure of EHLASS data, respectively in the transfer structure for input into the EUPHIN HIEMS database. These aggregated data are anymore case data. They were calculated from the underlying accident case data. The aggregated data can give amazing pieces of advice about the coding of the underlying accident case data. The two factors which influence the number of aggregated records are

- The availability of variables and the availability of the categories of the variables in the case data structure. If a category of a variable in EHLASS v1996 does not appear, the according record in the aggregated structure is not calculated. You can easily find out the relation of available categories to the size of the aggregated data from Icelandic data where nearly half of the variables is not available.
- The number of underlying accident cases. You can easily see this relation from the data of the United Kingdom. Finally it is not certain that the size of aggregated data calculated from British case data is more influenced by the number of underlying case data than by the availability of categories.

Looking at the aggregated data, you can get an impression of the changes made to the coding or coding problems which were actually found, i.e. Denmark 1997/1998, Greece 1998/1999, Iceland 1998 and 1999, Spain 1998/1999, Sweden 1997/1998. Nevertheless, it does not seem to be useful to calculate a ratio as a simple score for decision making, i.e. the number of records in the aggregated structure divided by the number of case data (the higher the better), because an inter-country comparison by such a ratio for instance to compare the coding of underlying case data would not work. Additional calculations for standardising become necessary. The reason is the course of the relation between the number of case data and the number of records of the aggregated data. This course is surely not linear. It is much nearer to a logarithmic one.