

**Comparison of EHIS source questions with national survey  
questions**

**Final report,  
Revised version, June 2011**

Antti Tuomi-Nikula, Sanna Ahonen, Päivikki Koponen  
National Institute for Health and Welfare  
Finland

**Contents:**

1. Introduction .....	3
2. Comparison practice .....	4
3. Classification and rating .....	4
4. Overall coverage and comparability of EHIS questions in national surveys .....	5
Picture 1. Comparability of EHIS questions included in this study in national surveys per country, in % .....	6
4.1. Reasons for missing or different questions .....	7
4.2. Full comparability .....	7
4.3. Partial comparability .....	9
4.4. Non-comparability / Missing questions.....	11
4.5. Complete overview.....	12
Table 1. All EHIS questions included in this study and their comparability rating by country, divided in modules .....	13
Table 2. Distribution of fully, partly and non-comparable questions in the three EHIS modules.....	15
5. Conclusions .....	15

# 1. Introduction

The European Health Interview Survey (abbreviated: EHIS) is a standard created as a joint venture in Eurostat. The main goal of EHIS is to obtain survey-based indicator data that are internationally comparable in Europe. That would require all countries to use the same questions and also same methods in conducting their national surveys.

The EHIS consists of four modules of which all consist of a large set of questions:

- 1) Demographic & socio-economic background variables (“Personal characteristics”)
- 2) Health status & functional capacity (“European Health Status Module”)
- 3) Determinants of health (“European Health Determinants Module”)
- 4) Utilisation of health services (“European Health Care Module”)

To date the first wave of EHIS has been completed in 20 European countries to a varying degree between 2006 and 2010 (detailed list in Appendix, worksheet 1). Most countries have included all EHIS modules in their national surveys, while some have only included selected parts. This paper presents a detailed overview of the scope of EHIS by question in all the 20 countries.

The main goal of this study is to identify to which degree the EHIS standard questionnaire has been used in the European countries in the first wave of EHIS. This study aims to answer these three questions:

- 1) Which EHIS questions are included in national surveys?
- 2) Are the questions identical to the original, or do they differ?
- 3) What kind of differences there are, and what is their degree?

In this comparison study, only those EHIS questions are included that provide data for the European Community Health Indicators (ECHI; [www.echim.org](http://www.echim.org)) and Open Method of Coordination (OMC; [ec.europa.eu/social](http://ec.europa.eu/social)) indicators, based on the document “Priority shortlist of indicators to be computed with the EHIS”. Those questions constitute roughly half of all EHIS questions.

This study is based on a contract between the National Institute for Health and Welfare (THL), Finland and Statistics Austria. It concentrates on the questions used in the national surveys. This study cannot give an answer whether the resulting indicator data are comparable between countries or not, as survey methods have not been compared. This document constitutes one of four different aspects in assessing the EHIS Wave 1, the others being the “Synthesis Report on Use of EHIS Quality Assessment Criteria (Dorothy Gauci), Report on National Problems faced During the EHIS Wave I (Małgorzata Piekarczyńska) and Comparability of EHIS Results and Indicators (Niels Kr. Rasmussen).

A pilot study ([www.euhsid.org/docs](http://www.euhsid.org/docs)) including one module of EHIS in national surveys of four countries was carried out during Spring 2010, and its purpose was to find out whether a

larger study would be feasible, and also to estimate the resources needed for such a study. The pilot demonstrated that the study would be feasible and its results can be expected to provide valuable information, especially in implementing the ECHI shortlist indicators in Europe and also in revising the EHIS questions for the second wave.

This paper constitutes a final report of the comparison work. All questions in the 20 countries are examined in detail, and rated according to their comparability. The results are presented and illustrated with examples in chapter 4. A detailed overview and total figures based on the results are presented in the compilation tables on pages 12–14. All details are listed in Appendix, worksheets 2–4.

## **2. Comparison practice**

The work was started by first entering the sets of questions into an Excel sheet (Appendix, worksheet 2) in which they could be examined in parallel. The original EHIS questions were entered in the left column top down, and the national questions were entered in following columns.

The next phase covered identifying all questions in the national question sets, and considering if they deviate from the original EHIS. Deviations were considered to be differences in phrasing or structure of questions that changed the meaning of the question. Examples include different specifications and different number of options in answer alternatives.

This comparison was based on national versions of questions that had been translated back to English. Institutes responsible for conducting EHIS in each country had their national questionnaire translated to English by professional translators. Due to that, it is presumable that every translation has been somewhat influenced by cultural and personal differences. The back translated questionnaires were also received in multiple different file types and layouts. Therefore, slight differences in wording or in the structure of the question were not considered being deviant from the original, as long as the meaning could be considered similar. These differences may be only caused by the translation. Examples of this will be demonstrated in the following chapters.

All differences in wording were analysed in detail and rated according to a classification protocol which is introduced next.

## **3. Classification and rating**

The comparison was made in two separate parts to bring out all aspects contributing to the comparability. That is, the actual questions and their answer alternatives were examined separately. This was done because these two elements carry a considerably different weight when it comes to the data and indicators the question is meant to yield. The question itself is

naturally the most important element, and the weight of the answer alternatives depends on the number of indicators a single question is expected to produce. In many cases there is only one indicator; prevalence of “Yes” answers. In these cases, all other answers are less important. Still, in this analysis, all possibilities have been taken into account.

In the final rating of the questions, the separate ratings of the actual questions and their answer alternatives were put together so that every question entity was given one rating.

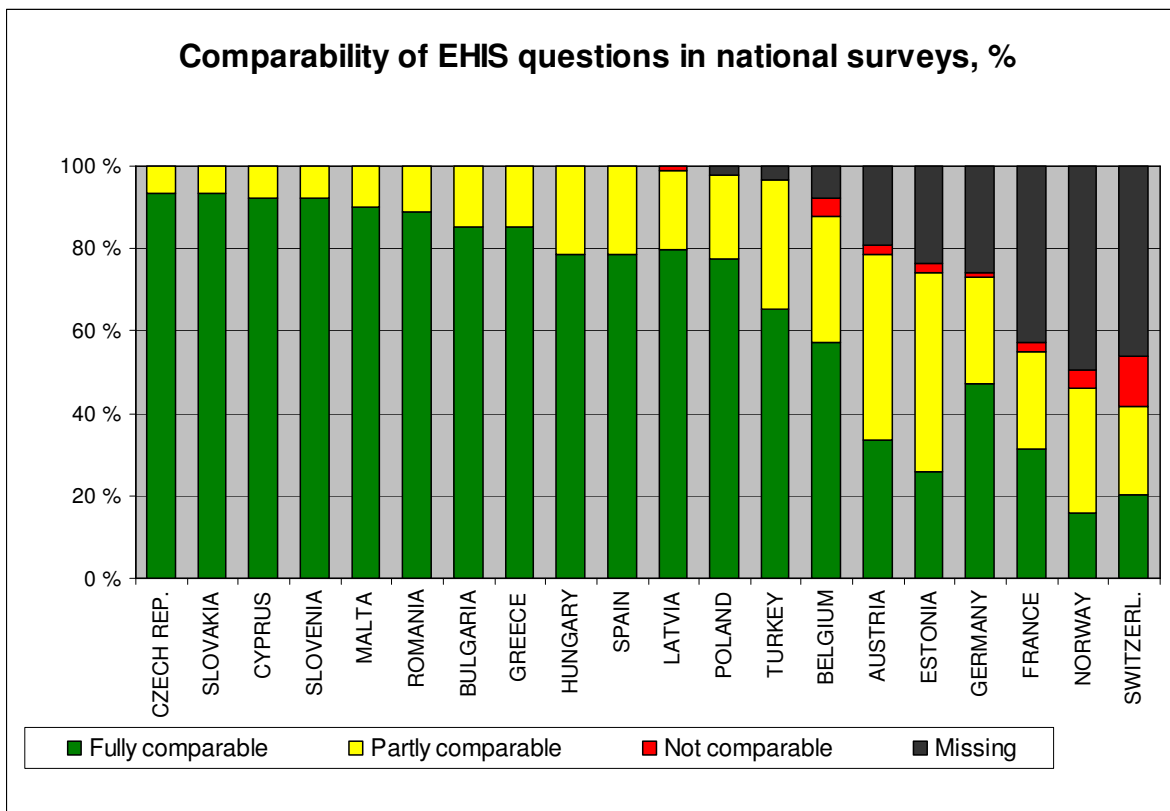
For the evaluation of comparability of the questions, a classification system was created. Its purpose is to rate all the national questions into different levels of comparability. Questions may be fully, partially or non-comparable.

- Fully comparable: National question identical or close to identical with the original EHIS question. Most likely to produce comparable indicators.
- Partially comparable: Different specifications or missing or extra elements or specifications, different number of answer alternatives. Likely to lead to limited comparability of indicators.
- Non-comparable / missing: Different recall period or cut-point of major importance, which leads to total non-comparability of indicators. In most cases, the whole question was missing.

The total coverage and comparability figures of EHIS questions by country are presented next. After that the classification system is introduced in detail with some of the most typical examples found in the national EHIS questionnaires.

## **4. Overall coverage and comparability of EHIS questions in national surveys**

The first wave of EHIS has been included in national surveys of 20 countries. These include 17 EU Member States; 6 “old” Member States (EU-15) and 11 “new” (EU-25 or EU-27). There was also one Candidate Country (Turkey) and two non-EU countries (Norway and Switzerland).



**Picture 1. Comparability of EHIS questions included in this study in national surveys per country, in %, sorted after full and partial comparability combined (Exact figures in Appendix; worksheet 4)**

Picture 1 illustrates that all EHIS questions included in this study were rated as fully or partly comparable in the national surveys of half of the countries (10/20). In three more countries, the figure was 95% or above, and in four, 70–95%. In only three countries the figure was significantly lower, reaching the low point of 41,6% in Switzerland. Based on the picture, we can conclude that EHIS was fully implemented in half of the countries.

In national survey questionnaires of the 10 countries with coverage below 100%, questions were missing or rated as non-comparable quite evenly in all modules. However, certain regularities could be observed, depending on question sets. Simple questions on large entities were only missing very rarely. On the other hand, detailed and strictly focused questions tended to be missing more often. Basic questions on self-perceived health, long-standing illnesses and activity limitations (HS.1–3) were included in all countries. Instead, question HA.2 (Difficulties in doing household activities) was missing in 6 countries. Details are presented in Table 1.

## **4.1. Reasons for missing or different questions**

Even though EHIS was created as an all-European standard, conducting national surveys accordingly has not been mandatory for European countries, nor have they received any funding for data collection. The EHIS is only based on a “gentleman’s agreement”. There are many reasons for countries not strictly following the EHIS, which are presented in more detail in the Synthesis Report.

Seven countries reported using an abridged version of the EHIS, which is naturally the most important reason for questions missing. Two countries, Austria and Estonia, were the first countries to conduct EHIS. They used the older version of the EHIS questionnaire, which was shorter than the newer one, and therefore many questions are missing in those countries.

The questions in the Personal Characteristics Module are different from the other questions in nature; they are background variables for indicator calculation. Those questions differed in detail between countries, and some even were left out. Still, since data for them can be derived from the sampling frame, i.e. population register or census, correct data for them was received from all 20 countries. Therefore the Personal Characteristics Module was not included in this comparison work.

## **4.2. Full comparability**

Most of the national questions fall into this category. The wording does not have to be exactly the same as in the original, as long as the general meaning is the same, the same specifications are included, and the same recall period is concerned. The differences may affect comparability of data, but a detailed analysis of this would require systematic conceptual analysis, which was not feasible in this study. A good example is the question HS.1. on self-perceived health.

EHIS: How is your health in general?  
AT: How is your health generally?  
BE: How is your general state of health?  
CZ: How are you generally feeling as regards health?  
CY: How do you evaluate your health status in general?  
GR: How would you describe your health?  
RO: In general, how would you appreciate your health status?

As regards to the answer alternatives, the same criteria apply. Also the number of answer options included must be similar. Here is an example of the answer alternatives of HS.1.

EHIS: Very good / Good / Fair / Bad / Very bad / Don't know / Refusal  
CZ: Very well / Well / Satisfactorily / Bad / Very bad / Does not know / Refused  
DE: Very good / Good / Moderate / Poor / Very poor / Don't know / N/A  
GR: Very good / Good / Mediocre / Bad / Very bad / I do not know, I am not sure / Refuse to answer  
LV: Very good / Good / Intermediate / Bad / Very bad / I do not know / Refuses to respond  
MT: Very good / Good / Not bad / Bad / Very bad / I don't know / I'd rather not tell you  
PL: Very good / Good / Neither good, nor bad / Bad / Very bad / I do not know / Refusal to answer

As shown above, all answer alternatives range from “Very good” to “Very bad” in five steps, with two additional alternatives outside that range (“Don’t know” and “Refusal”). Although many different words are used for “Fair” (“Moderate”, “Mediocre”, “Intermediate”, “Satisfactorily” etc.), they all carry generally the same meaning and have thus been rated as fully comparable.

The last of the answer alternatives “Refusal” is considered similar to “No answer”. This is because the outcome is “no answer”. In addition, in the EU data analysis, both the “Refusal” and “Don’t know” options are coded in the same category. That is why these cases get the rating “fully comparable”.

There were many cases in which the “Refusal” and “Don’t know” options were missing in the national questionnaires. Still, these may have been recorded in the data entry phase, even if the alternatives are not included in the questionnaire. In these cases, if that is the only deviation from the original, a question is rated as fully comparable, but an additional mark “A” is added in table 1. In the Austrian questionnaire, the Refusal” and “Don’t know” options were missing, although in the interview the respondents had these options. This information was received as feedback to the draft version of this report. It is a good example of why the questionnaire documents alone may not be enough for this kind of an evaluation since some information may be lacking.

Differences in question structure have also been considered, in order to determine if they affect the comparability. Questions which include clearly separate elements are usually rated as fully comparable, despite structural differences, as long as the original parts are included separately. A typical example is the question HS.4 on chronic conditions. Most countries have expanded the list of conditions. Still, as long as all conditions listed in the original EHIS are included, the question is rated as fully comparable, because each condition forms a separate sub-question of its own.

On the other hand, expanded or split questions on a single entity cause a question to be rated as only partly comparable, as can be seen in the following chapter.

Some national questions were more detailed than the original. For example, the question HA.1 on problems doing household activities includes a list of most typical activities with no additional specifications. Some countries have added detailed specifications about each activity. These cases are rated as fully comparable, since the core of the question is the same. Below is an example of the HA.1 sub-question on light housework.

EHIS: Light housework  
GR: Light domestic chores (washing dishes, ironing, making the bed etc.)  
HU: Light housework (e.g. washing up, ironing, sweeping)  
PL: Light household chores (e.g. cooking, washing up, ironing)  
ES: Do light housework such as the laundry, make the bed, clean the house...

Some countries used more detailed answer categories than the original. These cases are rated as fully comparable, as long as ranges congruent with the original EHIS can be compiled

from the national answer categories. The question EN.2 on social support in Turkey is a typical example.

EHIS: How many people are so close to you that you can count on them if you have serious personal problem?

–None / 1 or 2 / 3 to 5 / More than 5 / Don't know / Refusal

TR: If you have a serious personal problem, how many people do you have who are very close to you and whom you can trust?

–None / 1 person / 2 people / 3 people / 4 people / 5 people / More than 5 people / I do not know / I do not wish to respond

There was only one question, HS.1 (health in general) that was rated as fully comparable in all 20 countries. Question sets of particularly high rating included also HS.2–3 (health problems in general), SF-2–10 (mental health) and BMI.1–2 (height and weight). Of the 89 questions in the national questionnaires in 20 countries, 66,5% were rated as fully comparable.

### **4.3. Partial comparability**

The following differences make questions partially comparable:

- Differences in specifications
- Differences in recall period or cut-point
- Missing parts
- Extra parts
- Remarkably different question structure

Different specification such as giving an example is a common reason to rate a question as partly comparable. Question PL.11. is put like this:

EHIS: Can you bite and chew on hard foods such as a firm apple without any aid (for example, denture)?

BE: Can you bite and chew hard food such as an apple?

EE: Are you able to bite and chew hard food such as a fresh apple without any problems?

In the Belgian and Estonian questions the absence of “without any aid (for example, denture)” is causing them to be rated as partially comparable.

Differences in recall period can be found e.g. in question HS.2.

EHIS: Do you have any longstanding illness or longstanding health problem? [By longstanding I mean illnesses or health problems which have lasted, or are expected to last, for 6 months or more].

BE: Do you have any chronic, i.e. permanent illness or a chronic, i.e. permanent health problem?

BG: Do you suffer from a chronic disease or health problem?

The recall period of six months is missing in Belgian and Bulgarian questions. Even though those questions include the specification “chronic” or “permanent”, which in some cases could be considered to be very close to “6 months or more”, the lack of the exact recall period is here considered as a reason to rate these questions as partially comparable. Still, in this question the recall period can be understood to be only an extra specification,

“longstanding” being the priority specification. In questions in which the only specification is the recall period, a difference may result in total non-comparability.

In answer alternatives, a usual deviation from the original EHIS is the different number of answer options. As described before, in most cases it is about the absence of “Don’t know” or “Refusal” options. Those cases are, however, rated as fully comparable with a qualification (see previous chapter).

Cases in which answer alternatives are clearly different from the original, cause a question to be rated as partly comparable. In most cases that is due to missing or extra alternatives in the answer scales, like in the question PC.1, “Do you usually have difficulty doing any of these activities by yourself?”

EHIS: No difficulty / Yes, some difficulty / Yes, a lot of difficulty / I can't achieve it by myself / Don't know / Refusal  
AT: Yes / No / Unsure

In the Austrian question, alternatives for the degree of difficulty are missing. Still, it is possible to derive information on whether there is difficulty or not, regardless of its degree. Therefore the Austrian question is rated as partly comparable.

In some cases, the original answer alternatives include some detailed specification, which was left out from national questions. The question PC.4, “Would you need help”, is a typical example.

EHIS: Yes, at least for one activity / No, I do all these activities by myself / Don't know / Refusal  
AT: Yes / No / No answer  
BE: Yes, for at least one activity / No  
CZ: Yes, at least during one activity / No / Does not know / Refused

In the original, both the “Yes” and “No” options include detailed specifications. Leaving them out from national questionnaires led to rating them as partly comparable.

In the Health Care Module, several differences between the national and EHIS source questions were due to missing specifications in the national question introduction, such as HC.1.

EHIS: The next set of questions is about time spent in hospital. All types of hospitals are included. Visits to emergency departments or as outpatient only should not be included. Interviewer: for women up to age 50 years, add: Also, the time spent in hospital for giving birth should not be included. During the past 12 months, that is since (date one year ago), have you been in hospital as an inpatient, that is overnight or longer?

Clear specifications of not including the visits to emergency departments or stays in hospital due to childbirth were missing in 11 countries. Those national questions were scored as partially comparable. It is important to point out that if the introduction is missing, it usually affects to all questions in the same set. The countries that were missing an important specification in HC.1 were scored as partially comparable also in HC.2 and HC.3, since the introduction is a part of the question.

Answer alternatives were also rated partially comparable, if national answer options covered other aspects in addition to the original EHIS question. Question PL.4 (Do you wear a hearing aid?) is a typical example:

EHIS: Yes / No / I am profoundly deaf / Don't know / Refusal

FR: Yes / No, but I need one / No, I don't need one / Refusal to answer / Doesn't know

In France an extra aspect of “Need for a hearing aid” is included in the answer options. Even without the additional aspect, this question would fall into the category of partially comparable since it is also missing the answer alternative of “I am profoundly deaf”.

Differences in question structure that cause a question to be rated as only partly comparable were quite common especially in the Health Care Module. The Question HC.8 “When was the last time you visited a dentist or orthodontist on your own behalf (that is, not while only accompanying a child, spouse, etc.)?” is split in two parts in Greece; there is a separate question for visits to a dentist and visits to an orthodontist. That case is rated as partly comparable, because the Greek question allows the respondent to report a visit to both a dentist and an orthodontist, while the original only allows reporting a visit to either a dentist or an orthodontist.

7 questions out of 89 were rated as fully or partially comparable in all 20 countries. 21% of all national questions were rated as partly comparable. Thus, 87,5% of the questions were rated as fully or partly comparable.

#### ***4.4. Non-comparability / Missing questions***

When there is no question at all this is rated as missing. When there actually is a question on the topic, but it is so different from the original it cannot produce a similar indicator, it is rated as non-comparable. Major reasons to rate questions as completely non-comparable include a different recall period, when it is essential for the outcome. Here is an example of the question SF.1.

EHIS: Overall during the past four weeks, how much physical pain or physical discomfort did you have?

– None / Mild / Moderate / Severe / Extreme / Don't know / Refusal

AT: Please indicate the average severity of your pain in the past seven days on a scale between 1 and 10. 1 indicates little pain, 10 indicates the severest pain imaginable.

–... (1-10)

In the original EHIS the recall period is four weeks, while in the Austrian question it is one week. In addition, the answer scale is very different.

Another example is the question AL.2 on drinking patterns. In the long and complex original EHIS question, the respondent is required to specify the amounts he/she drank during a week, per day and per type and amount of alcoholic drink. In some countries, the specification per day is missing, which in the results makes it impossible to review the

possible differences in drinking habits within a week. Therefore in those cases the question is rated as non-comparable.

In the case of question PL.6 on walking ability, the Swiss question uses a cut-point of 200 metres instead of 500. The answer alternatives are also clearly different. Thus the Swiss question is considered non-comparable with EHIS.

EHIS: Can you walk 500 metres on a flat terrain without a stick or other walking aid or assistance?  
–Yes, with no difficulty / With some difficulty / With a lot of difficulty / Not at all / Don't know / Refusal  
CH: How far can you walk alone, that is without assistance, without having to stop and without experiencing considerable discomfort?  
–200 metres or more / More than a few steps, but less than 200 metres / Just a few steps / Cannot walk at all / No answer

A common feature in these examples is a major difference in a basic figure: a recall period or a distance. Those differences naturally result in total non-comparability.

A technical detail in the question PA.16 is an example of a different specification of major importance. Different cancer screening practices are a reason to rate the Austrian question as non-comparable, since a screening test and a clinical examination cannot be compared:

EHIS: Have you ever had a faecal occult blood test?  
AT: Have you ever had a colonoscopy for the early detection of cancer of the colon?

However, cases in which an existing question was rated as non-comparable were rather rare. In most cases, the question was simply missing. This is clarified in Table 1: A missing question is marked in grey colour, whereas an existing but non-comparable question is marked in red.

There were only two questions out of 89 that were rated as non-comparable in as many as three countries (PE.4 on moderate physical activities and HC.11 on visits to GP or family doctor). Questions that were missing most often were HA.2 on reasons for difficulties in household activities and SA.1 on satisfaction in health services, both of which were missing in 6 countries. In total, 11% of the national questions was missing, and 1,5% rated as non-comparable.

#### **4.5. Complete overview**

The following table 1 presents the complete overview of all examined EHIS questions in the 20 countries. A green cell indicates a fully comparable question and a yellow cell a partly comparable question. A red cell indicates an existing question that has been rated as totally non-comparable with the original EHIS question, and a grey cell indicates a missing question. The letter “A” in a green cell means that the options “Don’t know” and “Refusal” are missing in the question’s answer alternatives. Still, they are rated as fully comparable.

**Table 1. All EHIS questions included in this study and their comparability rating by country, divided in modules**

EHIS	AT	BE	BG	CY	CZ	EE	FR	DE	GR	HU	LV	MT	NO	PL	RO	SK	SI	ES	CH	TR
<b>European Health Status Module</b>																				
HS.1		A				A							A					A	A	
HS.2													A							
HS.3		A				A												A	A	
HS.4						A														
HS.4.1																				
HS.4.2						A														
HS.4.3						A														
HS.4.5						A														
HS.4.6						A														
HS.4.11						A														
HS.4.19																				
HS.5						A							A							
HS.6						A							A							
HS.7																				
HS.8																				
PL.1		A																		
PL.2		A											A							
PL.3		A											A							
PL.4		A																		
PL.5		A																		
PL.6		A																		
PL.7																				
PL.8		A																		
PL.9		A																		
PL.10		A																		
PL.11																				
PC.1																				
PC.2																				
PC.3																				
PC.4																				
HA.1																				
HA.2																				
HA.3																				
HA.4																				
HA.5																				
SF.1																				
SF.2		A	A																	
SF.3			A			A														
SF.4			A			A														
SF.5			A																	
SF.6		A	A																	
SF.7			A			A														
SF.8		A	A																	
SF.9			A																	
SF.10		A	A			A														
<b>European Health Determinants Module</b>																				
BMI.1						A							A	A						
BMI.2						A							A	A						
PE.1																				
PE.2																				
PE.3																				
PE.4																				
PE.5		A																		
PE.6																				
FV.1																				
FV.2																				

EHIS	AT	BE	BG	CY	CZ	EE	FR	DE	GR	HU	LV	MT	NO	PL	RO	SK	SI	ES	CH	TR
EN.4		A											A							
SK.1																				
SK.2																				
SK.3																				
SK.4																				
AL.1																				
AL.2																				
AL.3																				
CN.2																				
CN.4																				
<b>European Health Care Module</b>																				
HC.1																				
HC.2																				
HC.3																				
HC.8																				
HC.9																				
HC.10																				
HC.11																				
HC.12																				
HC.13																				
HC.14																				
HC.15																				
HC.16																				
MD.1																				
MD.2																				
PA.1																				
PA.2																				
PA.3																				
PA.10																				
PA.11																				
PA.13																				
PA.14																				
PA.16																				
PA.17																				
SA.1																				

As stated before, there were significant differences in comparability between different question sets. Regularly, simple and short questions (e.g. HS.1–3; 5–6, SF.2–10, BMI.1–2) were more often comparable than detailed and complex ones (e.g. PC.1–4, HA.1–5, all of HC, SA.1).

The differences even out when reviewed by EHIS modules (Table 2). The percentages of missing questions are remarkably close in all three modules (Table 2), ranging from 10% to 11,5%. Proportions of non-comparable questions vary more, but on the other hand, their total number is very small. Differences between full and partial comparability are greater. Full comparability is the highest (71%) in the Health Status Module and lowest (57,5%) in the Health Care Module. Partial comparability, on the contrary, ranges from 30% in the Health Care Module to 17% in the Health Status Module. The figures of the Health Determinants Module are very close to the totals of all three modules combined.

What might explain the low comparability figures of the Health Care Module are the special characteristics and different practices of national health care systems. Some countries have obviously wanted to obtain more detailed information by splitting some questions between public and private care, while there is only one question in the original EHIS (e.g. SA.1 on

satisfaction in health care system in Cyprus, Greece, Malta and Turkey). The Health Status and Health Determinants Module questions are less dependent on national characteristics, but still there are some. A good example is the question AL.2, in which different countries have different measures of both doses and types of alcoholic drinks. Also questions on diet (FV.1–2) may have been somewhat influenced by national characteristics. Of course, all other national differences can be said to stem from national characteristics and special needs too, but there are not that clear regularities among them.

**Table 2. Distribution of fully, partly and non-comparable questions in the three EHIS modules, %**

<b>EHIS MODULES</b>	<b>Fully comparable</b>	<b>Partly comparable</b>	<b>Not comparable</b>	<b>Missing</b>
Health Status	71	17	0,5	11,5
Health Determinants	67	18,5	3	11,5
Health Care	57,5	30	2,5	10
<b>TOTAL</b>	<b>66,5</b>	<b>21</b>	<b>1,5</b>	<b>11</b>

## 5. Conclusions

This report presents an overview of to which degree the EHIS standard questions have been used in European countries. Moreover, the practice of comparing the questions and the classification protocol are described. These are demonstrated with examples that clarify the most common deviations from the original EHIS question set.

As the national questions were translated to English before they were evaluated, some nuances in the wordings may have changed. This is why slight differences in wording were agreed not to affect comparability. Actually, there were some obvious errors and typos in the back-translated questionnaires which had to be checked either from country respondents or from the HIS/HES Database before rating the questions.

According to the Synthesis Report, in 13 of 18 countries all the recommended EHIS modules were implemented in their entirety. Close examination of the national questionnaires, however, revealed that only 10 of 20 countries had included all questions that were examined in this study. That may be partly explained by the fact that some countries had modified some questions to a degree they could not be considered as EHIS questions any more.

An interesting observation is that out of the 10 countries in which the questions were fully implemented, 8 were “new” EU Member States. On the other hand, the bottom four countries were all either “old” EU Member States or non-EU countries. Among countries where the coverage was below 95% there was only one “new” EU Member State. An apparent reason for this is the long tradition of surveys in Western and Central European countries, in which compromises had to be made between European comparability and a possibility to follow national trends. In the HIS/HES Database ([www.euhsid.org](http://www.euhsid.org)), the average number of previous national health interview surveys conducted in EU-15 Countries

was 7,5 while it was 4 in the EU-25/27 Countries. But then, it was also 4 in the non-EU countries, in which comparability figures were the lowest.

Two countries, Austria and Estonia, used the older EHIS questionnaire, which is the main reason for their comparability rating to be significantly low. Comparability of questions was rated slightly higher in Austria than in Estonia, but the differences were small (see Picture 1).

Introductions of different question sets were sometimes unclear, making it difficult to know to which very questions they refer to. The Problems Report also pointed out flaws of the same kind. A questionnaire as a guideline may not be enough to assure comparable results. There is an obvious need for instructions and manuals, training, auditing and other quality assurance and co-operation in preparing and conducting the second wave of EHIS.

Results of this study could be pooled with the information in the Synthesis Report, the Results Report and the Problems Report, for deeper analyses of the EHIS data. Also an even more specified study could be done, focusing only on those questions that yield ECHI Indicators; the detailed comparability evaluation (Appendix, worksheet 2) would allow a more indicator oriented rating with little effort. A detailed analysis of the results could be helpful in planning the second wave of EHIS, in order to assure more comparable results and complete time series in the future.

We received comments on this report from some countries, and most of them were about errors found in the back-translated national questionnaires. As the translators were not experts in public health issues, some concepts had been apparently unclear to them. In addition to that, some questions' introductory texts had been left out of the translations. Unfortunately we had no resources to check all national questionnaires for errors, and therefore we decided not to re-rate the questions after correcting the errors found in any of them. If the errors could have been checked and corrected for all surveys the results of this comparison could be slightly better. Additional information and corrections would in most cases have improved the comparability.

The following examples show how the questionnaire alone may not be enough to receive all necessary information for evaluation, and how easily errors can occur in the back-translation process:

- The question SK.4 “Have you ever smoked (cigarettes, cigars, pipes) daily, or almost daily, for at least one year?” is missing the specification “(Cigarettes, cigars, pipes); daily or almost daily for at least one year” in the Austrian questionnaire, but it was included in the interviewer instructions which are read aloud by the interviewer.
- In the Austrian questions HC.8–9 on dental care, the specification “orthodontist or other dental care specialist” is missing and only visits to dentists are asked. However, the Austrian term “Zahnarzt” comprises those specifications. Thus, the question will eventually produce comparable indicator data.

In addition, the worksheet 5 of the Appendix includes several back-translated national questions of Latvia thoroughly checked. In many of them, some additional information is also missing.

Should a same kind of an evaluation study be made in the future, a few prerequisites should be improved. First, back-translations of the national questionnaires should be made by one translation agency with at least some expertise in public health issues, or by giving more detailed information on the questionnaire to the agencies (e.g. asking them to check the translation against the original source questionnaire and/or conceptual cards). Furthermore, the back-translated questionnaires should be thoroughly checked by professionals before evaluation. Detailed information on survey implementation and data collection methods should also be evaluated together with the questionnaires, since some information on e.g. question introductions or answer alternatives may be missing in the questionnaire documents and additional specifications may be introduced in instructions for the interviewer.