APPENDICES

EHLEIS Technical report 2013_5.1

Cross-sectional analysis of health expectancies in 2008

Evaluation of the 2008 implementation of the greater harmonisation of the Mini European Health Module

Carol Jagger
Tony Fouweather
for the WP5 team
## Contents

APPENDIX 1 - Translation issues ....................................................................................................................................... 3

APPENDIX 2 - Prevalence of health indicators in European countries ............................................................................. 5
  1. Prevalence of activity limitation .................................................................................................................................. 5
  2. Prevalence of severe activity limitation ....................................................................................................................... 12
  3. Prevalence of chronic morbidity ..................................................................................................................................... 19
  4. Prevalence of fair or bad self perceived health ........................................................................................................... 25
  5. Prevalence of bad self perceived health ....................................................................................................................... 32

Appendix 3 – The logistic procedure in SAS 9.2 .............................................................................................................. 39
  LINK function (GLOGIT) ........................................................................................................................................... 39
  Model ...................................................................................................................................................................... 39
  The surveylogistic procedure in SAS 9.2 ................................................................................................................... 40
APPENDIX 1 - Translation issues

Translation issues were reported relating to the EU-SILC questions on health in the national surveys. The effect on comparability, in particular for the Healthy Life Years indicator was identified as a problem. A meeting of the Working Group on Living Conditions Statistics in June 2007 recommended harmonisation and corrections for the 2008 EU-SILC questionnaire using revised guidelines and contact with other Health Interview Survey (HIS) representatives.

In January 2008 Eurostat issued a note outlining the problems and comments on the harmonisation between the EU-SILC and European Health Interview Survey (EHIS) questions on health. There were limitations of the comparability between the two surveys for the questions used to assess the three health measures of interest. Revised guidelines for the EU-SILC questions were issued in October 2007 to address the harmonisation issues and give a more comparable data from 2008 onwards. The purpose of the analysis carried out, using 2008 as the reference year in the logistic procedures, was to assess if there is evidence that this has been the case.

The aim of the harmonisation was not to change either EU-SILC or EHIS as these were developed as the standard and adopted by the Working Group on Public Health Statistics. The aim was to provide the EU-SILC questions in both English and the national languages to be comparable with EHIS questions translated from the English source version. This should make both fully harmonised across the member states. The three questions of interest are PH010-PH030 from the Minimum European Health Module (MEHM) and the harmonisation of these should lead to better comparability of the indicators derived. This included the Healthy Life Years (HLY) indicator based on the Global Activity Limitation Indicator (GALI) question, PH030.

It is anticipated that data collection using the EU-SILC survey will be fully harmonised by April 2012. A report produced for the UK by the Office for National Statistics (ONS) concludes that the general health question (PH010) is largely harmonised across the member states. The other two questions still need to be improved to achieve harmonisation. A report for Eurostat by the Working Group on Living Conditions Statistics gives an overview of the progress made with dates so far towards harmonisation in the member states. The status from these reports relating to current progress concludes that the

- General health question (PH010) is more or less harmonised across the member states. The main problems were due to the ‘fair’ category as translation is difficult/ambiguous in some member states. Eurostat recommended that the problems caused by the ‘fair’ category could be solved in most places with use of a more plain English term such as ‘neither good nor bad’.
- Chronic/long standing condition question (PH020). Using ‘long-standing’ rather than ‘chronic’ would be easier to understand and interpret in non-English languages. Repeating the question with translation into national languages should help to harmonise the responses. Directing those suffering with physical or mental health conditions to the activity limitation question (PH030) should also help in harmonising the responses. The report for Eurostat noted that this question mixes concepts of illness and disability and does not have a fixed timeframe. The UK field tested question addresses these issues. Eurostat recommended wording very similar to UK field tested question. Both remove disability or impairment from the question making health conditions or illnesses the focus of the question. This meets the EU-SILC guidelines for harmonisation.
- The activity limitation duration in question (PH030) should be consistent at 6 months as specified in the EU-SILC guidelines. The UK report recommended a separate question/part to capture the duration of activity limitation to improve conformance to the EU-SILC six month standard.
In May 2009 it was noted that the 2008 national EU-SILC questionnaire has not changed where the requirement for changes were identified. Aim is to have EU-SILC questions in national languages as far as possible translated from the English source version.
APPENDIX 2 - Prevalence of health indicators in European countries

Separate line charts for men and women showing the prevalence of activity limitation, chronic morbidity and self perceived health for each of the EU25 countries over time between 2005 and 2009 were produced. The prevalence data was extracted from the Eurohex website.

1. Prevalence of activity limitation

In this section line charts for the prevalence of either mild or severe activity limitation for each EU25 country are presented separately for men and women. The charts show the prevalence for years 2005-2009.


There is consistency in the shape of the plots between gender in most countries. Exceptions can be seen in Austrian men where there is a plateau between ages 55-69 that is much shorter in women. Czech women have a steeper gradient in the line towards the upper age groups than men. Estonian and Slovenian men show more variation in the prevalence of activity limitation at the higher ages than women in some years. Although the general shape of the lines are quite similar. Rates of increase across the age groups are similar in most countries with the exception of Denmark where the rate is smaller in both men and women. This may be explained by the questionnaire as it wasn’t until 2008 that the relevant question was modified to allow respondents to distinguish between mild and severe activity limitation. Note similarities with Sweden where gradient of lines also not as steep as in most countries so may be a cultural or translation/language effect? Oldest age groups in Sweden have higher prevalence of activity limitation than Denmark where men report less than 0.4 prevalence in years 2005-2007, against women where this is not the case.

Most of the plots reveal that there have not been great changes in the prevalence rates for respondents reporting mild or severe activity limitation in the years 2005 to 2009. Variation over the years can be seen in some countries however, most notably in Sweden and Denmark, but also observed in Cyprus, Finland, Greece, Italy, Lithuania, Luxembourg, the Netherlands and Slovakia.

2. Prevalence of severe activity limitation

In this section line charts for the prevalence of severe activity limitation for each EU25 country are presented separately for men and women. The charts show the prevalence for years 2005-2009. In general, the prevalence increases with age group, towards the higher ages. At lower ages, up to [40-44] the prevalence of severe activity limitation is close to zero in most cases. In some countries there is again a trend of a flattening of the lines at higher ages. The trend is observed for Belgian men, Cyprus, both men and women, Czech men in most years, Finnish women, German men, Irish men (2006-2008) and Latvian men (2006, 2008, 2009).

The consistency between men and women in the countries is also apparent for severe activity limitation. Exceptions include Germany and Luxembourg, where the lines show a much steeper rise in the older age groups for women. Consistency between years can be seen in most countries for prevalence of severe activity limitation. Exceptions are Belgium, Cyprus, Estonia, Finland, Greece, Hungary, Latvia, Lithuania, Luxembourg and Sweden where there is some variation in the older age groups across the years.

3. Prevalence of chronic morbidity

In this section line charts are presented for the prevalence of chronic morbidity in both men and women in each of the EU25 countries. In general, the prevalence increases with age group, but there is a flattening of the line at higher age groups in most cases.

There is consistency in gender in most countries in the rate of increase across the age groups.

Most of the plots reveal that there have not been great changes in the prevalence rates for respondents reporting of chronic morbidity in the years 2005 to 2009. Variation over the years can be seen in some countries however, most notably in Austria, Denmark, Greece, Ireland, Luxembourg, Malta, Slovenia and in particular Sweden.

In some cases the prevalence of chronic

in highest age group of 85+ shows a reduction from that seen in age group [80-84]. This is possibly due to smaller sample sizes or the lack of institutionalised respondents taking part in the survey. This can be seen in a number of countries, most notably in Luxembourgian, Slovakian and Slovenian men.
4. Prevalence of fair or bad self perceived health

In this section line charts for the prevalence of fair or bad self perceived health are presented for men and women for each EU25 country.

In general, the prevalence increases with age group, towards the higher ages. For some counties this has increased to more than 90% by age group 85+. Countries where this is the case for all years are; Hungary (women), Latvia (women), Lithuania (men and women) and Poland (women).

The consistency between men and women in all of the countries is also apparent for fair or bad self perceived health.

For Austria, Belgium, Denmark, Finland, Luxembourg, Slovenia and there is a plateau between the age groups [55-59] to [65-69] in all years.

In some cases the prevalence of fair or bad self perceived health in the highest age group of 85+ shows a reduction from that seen in the age group [80-84]. This is possibly due to smaller sample sizes or the lack of institutionalised respondents taking part in the survey. This can be seen in a number of countries, with the most notable occurrences Danish women all years except 2006, and Estonian women also all years except 2006.
5. Prevalence of bad self perceived health

In this section line charts for the prevalence of bad self perceived health are presented for men and women for each EU25 country.

In general, the prevalence increases with age group, towards the higher ages, but to lower levels by 85+ than for those with either fair or bad SPH. At lower ages, up to [40-44] age group, the prevalence of bad self perceived health is close to zero in most cases, with the odd exception where the prevalence has started to increase just before this age. The countries where this is observed are Hungary, Latvia, Lithuania, Portugal and Slovenia.

The consistency between men and women in all of the countries is also apparent for those reporting bad self perceived health. There are some exceptions in some countries such as Luxemburg and Slovakia at the higher ages.

In some cases the prevalence of activity limitation in the highest age group of 85+ shows a reduction from that seen in the age group [80-84]. This is possibly due to smaller sample sizes or the lack of institutionalised respondents taking part in the survey. This can be seen in a number of countries, with the most notable occurrences in Cypriot women (2006-2008), Czech men (2005-2007) and Danish women (all years except 2005).
Appendix 3 – The logistic procedure in SAS 9.2

The first analysis conducted used the LOGISTIC procedure in SAS. The syntax used is shown below;

```
proc logistic data=AL;
class year (ref='2008') agegp (ref='[45-49]') gender;
  Weight weights;
model M_S_AL/TOT = year agegp gender year*agegp year*gender agegp*gender year*agegp*gender / link=glogit expb;
  By country;
run;
```

Categorical variables are specified with the CLASS statement. The reference (ref=) command allows particular values of the categorical variable to be taken as reference values that all others will be compared to in the analysis.

The WEIGHT statement identifies the column in the data set containing the weighting allocated to each row of data.

**LINK function (GLOGIT)**

There are various link functions available for the procedure. GLOGIT is used as this link function specifies the generalised logit function. The LOGISTIC procedure fits the generalised logit model where each of the non-reference categories is contrasted with the reference category.

The procedure allows a reference class to be specified. For this analysis the year 2008 and age group of [45-49] years are specified as reference categories. Gender does not need a reference class as it is a binary response, so one is automatically compared with the other in the procedure. 2008 was fixed as the reference year as this is when the questions were harmonised.

**Model**

Two forms of the MODEL statement can be specified. The first form, referred to as single-trial syntax, is applicable to binary, ordinal, and nominal response data. The second form, referred to as events/trials syntax, is restricted to the case of binary response data. The single-trial syntax is used when each observation in the DATA= data set contains information about only a single trial, such as a single respondent in the EU-SILC instrument.

When each observation contains information about multiple binary-response trials, such as the counts of the number of subjects observed and the number responding, then events/trials syntax can be used. In the events/trials syntax, two variables are specified that contain count data for a binomial experiment. These two variables are separated by a slash. The value of the first variable, events, is the number of positive responses. The value of the second variable, trials, is the number of trials. In the single-trial syntax, one variable is specified on the left side of the equal sign as the response variable. This variable can be character or numeric.

For both forms of the MODEL statement, explanatory effects follow the equal sign. Variables can be either continuous or classification variables. Classification variables can be character or numeric, and they must be declared in the CLASS statement. When an effect is a classification variable, the procedure enters a set of coded columns.

This analysis has to use the 2nd option for the model statement as the data available publically via Eurohex, can be transformed to this format, i.e, the response used is of the form events/trials.

The response with syntax events/trials representing the number of respondents with mild or severe activity limitation (M_S_AL) divided by the total number of respondents in the sample for each category (TOT). The terms in the model are gender, age group and year along with all interactions between these terms.

The EXPB statement displays exponentiated values of the estimates.
The BY country statement allows for subpopulation analysis, with the subpopulations defined as the EU25 countries. Analysis of subpopulations is a common analytic practice. A subpopulation is a subset of a larger population where the inclusion in the subpopulation is dependent on a particular characteristic. For example, a subpopulation from the population of European citizens could be defined by gender, age group, or country of residence.

Although the use of a BY or WHERE statement in the LOGISTIC procedure seems an intuitive approach for analysing the subgroups of interest, it is incorrect. This approach simply considers the subpopulations of the original sample but ignores the variability of the domain sample sizes across the strata of the sample design (Cochran, 1977). This often results in under-estimated variances.

The surveylogistic procedure in SAS 9.2
The SURVEYLOGISTIC procedure fits linear logistic regression models for discrete response survey data by the method of pseudo-maximum likelihood, incorporating the sample design into the analysis. The procedure enables the use of categorical classification variables as explanatory variables, using the syntax for main effects and interactions employed in the GLM and LOGISTIC procedures.

Before the survey procedures were added to SAS, the analysis of sample data was carried out under the assumption that the sample is drawn from an infinite population by simple random sampling. This does not take account of complex survey designs that are more generally used when conducting surveys from large and spread out populations.

The next analysis conducted used the SURVEYLOGISTIC procedure in SAS. The syntax used is shown below:

```
proc surveylogistic data=AL nomcar;
  class year (ref='2008') agegp (ref='[45-49]') gender;
  Weight weights;
  model M_S_AL/TOT = year agegp gender year*agegp year*gender agegp*gender
                   year*agegp*gender / link=glogit expb;
  Domain country;
  run;
```

NOMCAR requests that the procedure treat missing values in the variance computation as not missing completely at random (NOMCAR) for Taylor series variance estimation. The NOMCAR statement treats those missing data on the outcome of interest as a separate domain and hence, performs an unconditional subpopulation analysis of non-missing cases. This ensures that despite missing data, the full structure of the complex sample is reflected in the estimates of statistics computed based on the non-missing data (after case-wise deletion).

The DOMAIN statement is used for the subpopulation analysis. As stated in the description for the LOGISTIC procedure using a BY or WHERE statement before analysing the subgroup of interest seems sensible yet it is statistically incorrect. This approach simply considers the subpopulations of the original sample but ignores the variability of the domain sample sizes across the strata of the sample design (Cochran, 1977). This often results in under-estimated variances.

Using the DOMAIN statement along with the survey procedures for subpopulation analysis first produces an overall model for the whole population, i.e., the EU25 countries pooled together. Then a subpopulation analysis is carried out for each of the 25 countries independently.

A subpopulation analysis should use the entire sample in the analysis and also take the sample size of the created domain into account. This is important because created domains are often not related to the original sample design, as in this case. Using the DOMAIN statement in the SAS Survey procedures ensures that the correct use of the entire data set occurs and separate analyses per domain are performed while accounting for the random variability introduced by domain sample sizes unrelated to sample design.
Co-funded by 10 Member States, the European Commission, DG SANCO and two French institutions: DREES and CNSA.

Contact EHLEIS:
Jean Marie ROBINE, INSERM
Centre Val d’Aurelle
Parc Euromédecine
34298 Montpellier cedex 5, France.
Tel: +33 (0) 467 61 30 29
Fax: +33 (0) 467 61 37 87
Email: jean-marie.robine@inserm.fr
www.eurohex.eu