

## Extrapolation guide ECHI-29

Guide to extrapolation of sample hospital data to national estimates for a given year using the EXCEL spreadsheet "JAMIE extrapolation" using the HDR method (**The other methods do not work properly yet**)  
*Version: July 3, 2012, by Bjarne Laursen*

This guide is relevant for countries recording data from a sample of hospitals, either with the Minimum dataset (MDS) or the full dataset (FDS). The principles are the same.

**NOTE: The calculations are only valid if the hospital sample is representative for the whole country, and the number of hospitals and cases is sufficiently large (see Manual chapter 4). The representativeness of the hospital sample should be documented, see Manual chapter 8.**

The calculation may be done using an excel spreadsheet, and the present guide shows how this is used.

The guide is based on the HDR method, but the spreadsheet also allows for using the EDR and the catchment area methods.

At present, the spreadsheet contains Danish data – these must be substituted with your data. Most of the needed data can be generated automatically by using the MDS quality control software (MDS\_QC.EXE). See the manual to this software.

### What data you need:

- 1) Number of admitted patients related to an injury diagnosis in the whole country, and in each of the hospitals in the sample. This number should be available in 5-year age groups up to 85 years and above.
- 2) The national population by 5-year age groups, up to 85 years and above.
- 3) Injury data from the hospital sample, according to the MDS or FDS description.
- 4) If not all cases in the sample hospitals are recorded, you need the sampling rate (e.g. 0.125 means 1 of 8 cases)

If you want to calculate confidence rates, you also need:

- 5) Number of admitted patients from each sample hospital.
- 6) MDS/FDS data for a list of injury types, for each sample hospital separately.

## How to use the spreadsheet (HDR method)

***In general, all the red figured need to be filled in, while the black should be left unchanged***

- 1) Paste the national numbers of admitted patients by 5-year age group into sheet “HDR methods”, cells D6:D23.
- 2) Paste the numbers of admitted patients in the hospital sample into cells E6:E23. For the calculation of confidence intervals, you must paste the total number of injury-related discharges for each of the sample hospitals into the cells D56,E56 etc. in the sheet “CI estimation”.
- 3) Paste the national population into cells G6:G23.
- 4) Paste the hospital sampling rate (e.g 0.2) into cells K6:K23 (the same number in all cells). If there is no sampling (all cases are recorded), all cells should contain the number 1.
- 5) Write the number of sample hospitals in cell L26.
- 6) For HLI incidence rate calculation, paste the number of the HLI injuries (without and with non-residents) in the hospital sample in the cells L6:L23 and M6:M23 respectively, by 5 year age groups.

Then the following is available:

- National estimates of the number of HLI injuries by age group (cells C30-C47) and the total number (C49) for residents, and I30-I47 and I49 including non-residents.
- Estimated national incidence rate of HLI for each age group (D30-D47), and mean incidence rate (D49). Correspondingly J30-J47 and J49 including non-residents.
- Incidence rates standardized to the European population, in order to compare with other countries (G49 and K49)

How to calculate confidence intervals:

- 1) In the sheet “CI estimation” paste the number of injuries for each sample hospitals in columns D,E,F, etc. (you may have up to about 250 hospitals). For each hospital, the lines 5-55 contains the numbers of injuries related to the shown codes. “Other” and “unspecified” is not used. If only unintentional injuries are collected, the lines 7-8 should be deleted, and if only home and leisure injuries are collected, also the lines 9 and 14 should be deleted. These data are most easily produced by using the MDS\_QC software.
- 2) The number of columns in lines 59-119 must correspond to the number of sample hospitals. Add or clear cells as needed
- 3) The model parameters A and B (D113, D114) are NOT calculated automatically, but should be calculated manually according to the procedure shown in the cells H118-H127.
- 4) When this is done, check that the estimated variances in cells D116-D166 seems reasonable (e.g. all positive), or check the plot (about cell N120) – does the model results reasonably fit the data ?

The A and B parameters are used for calculating the confidence intervals shown in sheet “HDR method”. Please note that the confidence intervals are not statistically “correct”, but are model estimates. The confidence intervals are not available if data are collected from only one hospital, and they may be quite wide with only few hospitals.