UNHCR
The UN Refugee Agency

Things to watch out for during analysis of SENS data & Tips to follow

- **Code « 8 » or « 98 » for « don't know » response:**

  - Make sure to _exclude_ codes « 8 », or « 98 » for « don't know » from the analysis. This should not be part of the denominator in the calculation of indicators (e.g. diarrhoea, IYCF, ANC enrolment). If you follow the standard SENS PGM (analysis) codes for analysis, this will not be an issue.

- **Code « 6 » or « 96 » for « other » response:**

  - Make sure to _include_ codes « 6 » or « 96 » for « other » in the analysis. This should be part of the denominator in the calculation of indicators as it represents one response option (e.g. reason for not having a ration card, source of water). If you follow the standard SENS PGM (analysis) codes for analysis, this will not be an issue.

- **Recoding with the « IF » command in Epi Info software:**

  - Do not forget to take into account missing values when using the **« IF »** command in Epi Info software. This applies to many different variables. Whenever feasible, it is much better to use the « RECODE » command instead of the **« IF »** command for this specific reason. If you follow the standard SENS PGM (analysis) codes, this will not be an issue.

- **Missing data or consent not provided:**

  - Missing data should be excluded from all analysis and should not be accounted for in the denominator. A missing data can occur because of an omission during the interview process or simply because the individual or household did not provide consent for the interview/measurements. If you follow the standard SENS PGM (analysis) codes, this will not be an issue.

- **Confidence intervals:**

  - Different software (including Epi Info) often calculate confidence intervals as being negative and hence below zero or above 100. Negative confidence intervals or CI above 100 are meaningless. In the report, always round negative confidence intervals to « 0 » and round those above 100 to « 100 ».

  - For example, a result of -0.069 for the lower confidence limit (LCL) given by Epi Info should be reported as « 0 ». A result of 100.565 for the upper confidence limit (UCL) given by Epi Info should be reported as « 100 ».

- **Rounding decimal points:**

  - Make sure to round properly decimal points according to basic rules:
    o When decimal is between 1-4, round down.
    o When decimal is between 5-9, round up.

  - For example, a result of 20.569% (16.684-24.423 95% CI) for moderate anaemia given by Epi Info software should appear as: 20.6% (16.7-24.4 95% CI) in the SENS reports.

- **Decimal points in the results:**

  - When the results is a whole number e.g. 30%, make sure to always write 30.0% with the « .0 » in the decimal place in the report. This ensures that the decimal point was not forgotten and is actually equal to zero.

- **Age variable:**

  - When selecting age or creating an age variable category in Epi info software from the 'months' variable generated by ENA, don't forget the '.99' otherwise some children with an exact birth date may be excluded from the analysis. E.g. 6-23.99 (and not 6-23 or 6-23.9). If you follow the standard SENS PGM (analysis) codes, this will not be an issue.

- **When using the « SELECT » command in Epi Info, always make sure you keep track of your selected dataset.** If you follow the standard SENS PGM (analysis) codes, this will not be an issue.

- **Analysis of anthropometric data with ENA for SMART and « Options » screen:**
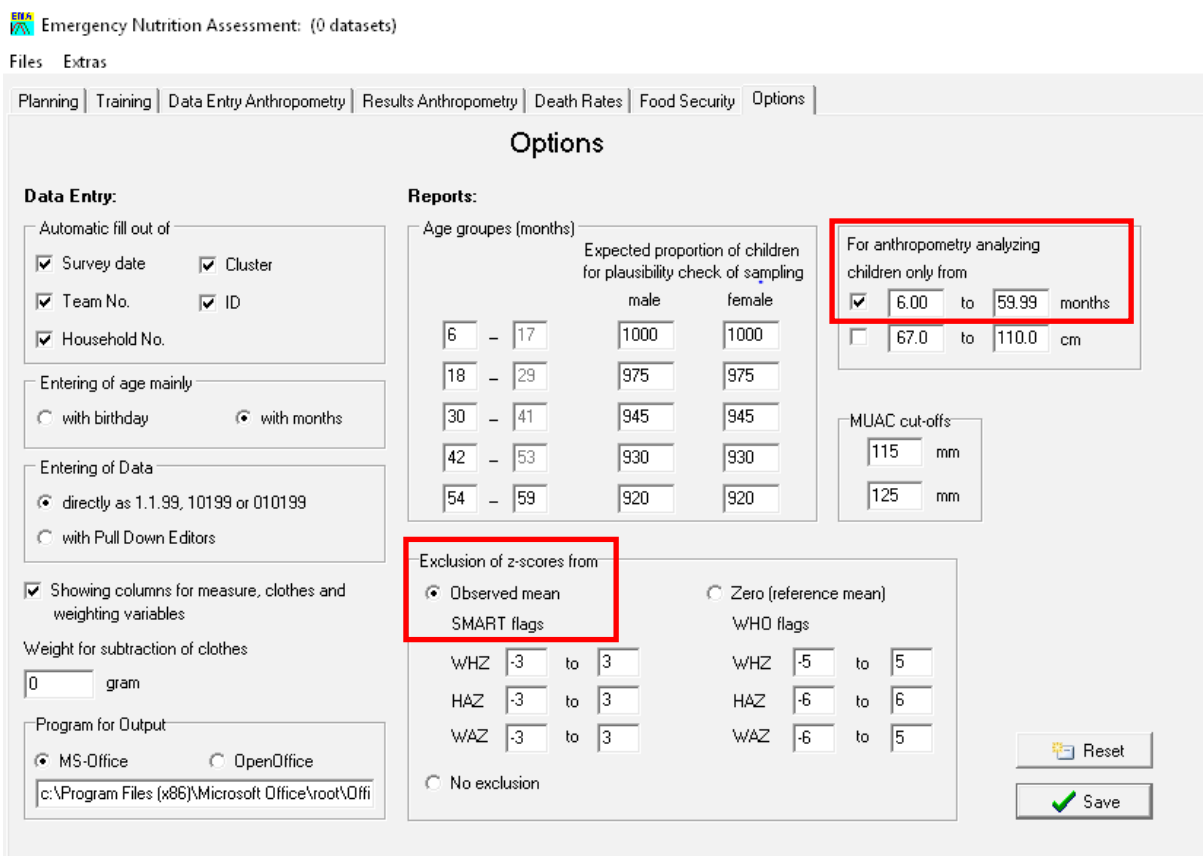
  - Always check the ENA for SMART « Options » screen to make sure that:
    o The « Observed mean » button is selected in the « Exclusion of z-scores from » window.
    o The analysis was done in children aged 6-59.99 months only (i.e. select « 6.00 to 59.99 months » button in «For anthropometry analyzing children only from » window) unless you recruited a significant number of children based on height. Refer to decision tree for process of age determination (Figure 1-Flow chart for collecting age data during a survey, Module 2).

- **Always review the data first before going into analysis:**

  - Frequencies and means should be run on categorical and continuous variables, respectively. This is very crucial to do especially in paper-based surveys because recording errors (i.e. writing a value out of the acceptable range or writing an inexistent code) can occur. In mobile data collection (MDC) surveys, each variable has pre-determined ranges and acceptable codes, and hence there can be no data recorded out of the acceptable ranges or codes.

  - Missing data should be looked at and a record of them should be kept. This should included in the final SENS report. Missing data can occur in both MDC and paper-based surveys.

- **Categorical variables having several code options:**

  - During the data review process, make sure that the codes are the same as those on the standard questionnaire. E.g. vitamin A should be coded « 1 », « 2 », or « 3 » in the questionnaire but code « 4 » may appear in the database. This code « 4 » needs to be excluded from analysis and not included in the denominator as it was a probably mistake: (i) if an MDC survey was conducted, the survey manager probably made an incorrect modification to the data during the data review process (Refer to SENS Pre-module tool 15 for SOP on how to conduct the data review process for MDC surveys); and (ii) if a paper-based survey was conducted, it was probably a recording error on the field.

- **Always keep and save the original, raw data sets from each SENS modules completed. The final dataset reviewed and used for final analysis needs to be sent to UNHCR HQ for storage in the Global SENS Database. Make sure to save newly generated variables into a new data file named following a naming convention to be respected by all involved in the survey data analysis.**