

WHO Reference 2007 STATA macro package

The package **who2007_stata** contains the following items:

1. One macro (**who2007.ado**).
2. Three permanent (read-only) Stata data sets containing the WHO 2007 reference for children aged 5-19 years: *wfawho2007.dta*, *hfawho2007.dta* and *bfawho2007.dta*.
3. The file *Readme.pdf*
4. An example set, *survey_who2007.dta*.
5. An example do-file, *survey_who2007.do*.
6. The example output files: *survey_who2007_z.xls*, *survey_who2007_z.dta* and *survey_who2007_prev.xls*.

Pre-requisites

STATA Version 7.0 Stata/SE (Special Edition of Stata) or higher is required to run two macros (*who2007_standard.ado* and *who2007_restricted*). **Intercooled** Stata has a limit of 2,047 variables and with that the macros will only produce the z-scores output files (see below the macros' section **Exported files**, point 1). The macro requires a STATA data set containing age, sex and the anthropometric measurements. The input variables are specified in the **Parameters** section.

Precautions:

1. Avoid any variable names starting with underscore "_" in the input STATA data set; otherwise they may be replaced by the derived ones created by the macro.
2. Avoid any temporary format names starting with underscore "_"; otherwise they may be replaced by the temporary ones created by the macro.
3. Avoid any STATA global macro variable names starting with underscore "_", except those defined by the system.

Contact for reporting bugs/ comments:

Should you encounter any problems with this package, please send an e-mail with a clear description of the identified problem to "**anthro2005@who.int**", specifying in the subject line that it concerns the **who2007_Stata package**, the name of the macro (*who2007.ado*) and kindly indicate which version of STATA you are using. Thank you.

Recommended setup and run

- Step 1.** Create a sub-directory, for example "D:\WHO 2007 Stata", where you wish to save the package (**who2007_stata.zip**). This directory should be reserved only for the references tables (*who2007.dta) and the macros (*who2007.ado*) that are contained in the zip file.
- Step 2.** Create a sub-directory, for example "D:\WHO 2007 workdata", where the example data (survey_who2007.dta and pertinent output files) and your STATA input data can be stored and where all the macro output files will be written to.
- Step 3.** It is recommended that you start by loading and running the example code below (also found in *survey_who2007.do*) in the STATA do-file editor to see how the data should be prepared and to fill in the macros' parameters according to their requirements.

Note: The macros run on Stata/SE: The Special Edition version of Stata (type "help SpecialEdition"). For users of Intercooled Stata, the macros will only produce the z scores output files and the user gets the message

*Please wait, programme is calculating prevalences.....
.....
no room to add more variables*

Reference

de Onis, M, Onyango, A, Borghi, E, Siyam, A, Nishida, C, Siekmann, J. Development of a WHO growth reference for school-aged children and adolescents. Bulletin of the World Health Organization 2007; 85: 661-668.

The macro (**who2007.ado**)

Description

1. The macro (**who2007.ado**) calculates z-scores for the **three** anthropometric indicators, weight-for-age, height-for-age and body mass index (BMI)-for-age. In this macro, all available (non-missing and non-flagged) z-score values are used for each indicator-specific prevalence estimation (standard analysis).
2. The macro produces sex- and age-specific estimates for the prevalence of under/over nutrition and summary statistics (mean and SD) of the z-scores for each indicator.
3. The macro runs with **three** permanent (read-only) Stata data sets containing the WHO Reference 2007: wfawho2007.dta, hfawho2007.dta and bfawho2007.dta.
4. Extreme (i.e. biologically implausible) z-scores for each indicator are flagged according to the following system:

Weight-for-age z-score (zwfa)	$zwfa < -6$ or $zwfa > 5$
Height-for-age z-score (zhfa)	$zhfa < -6$ or $zhfa > 6$
BMI-for-age z-score (zbfa)	$zbfa < -5$ or $zbfa > 5$

Parameters

The macro requires 10 parameters that **must** be specified without any quotation marks.

- **reflib**: to specify the package directory where the three STATA data sets containing the WHO 2007 reference (age 5-19 years) are stored.
- **datalib**: to specify the working directory where the input STATA data set containing anthropometric measurements is stored.
- **datalab**: to specify the name that will prefix the output files (datalab_z, datalab_prev).
- **sex**: to specify the name of a variable containing sex information. If it is a numeric variable, its values must be, 1 for males and 2 for females. And if it is a character variable, it must be, "m" or "M" for males and "f" or "F" for females. Users must code any missing values as "." (for a numeric variable) or " " (for a character variable), in which case no z-scores will be calculated.
- **age**: to specify the name of a numeric variable containing age information. Age can be in days, months or years. An accurate age without any rounding or truncating is

strongly recommended; ideally it should be derived from date of birth (DOB) and date of measurement (DOM), i.e. DOM-DOB. Users must code any missing values as ".", in which case none of the three age-related z-scores can be calculated.

- **ageunit:** to specify the age unit of the age variable. It must be specified as either "days", "months" or "years" (they are *case sensitive*). The macro converts age in days or years to months, dividing it by 30.4375 days or multiplying it by 12 months, respectively.
- **weight:** to specify the name of a numeric variable containing body weight information, which must be in kilograms. Users must code any missing values as ".", in which case weight-related z-scores are not calculated.
- **height:** to specify the name of a numeric variable containing length (recumbent) or height (standing) information, which must be in centimeters. Users must code any missing values as ".", in which case length- or height-related z-scores are not calculated.
- **oedema:** to specify the name of the character variable containing oedema information. The values of this variable must be "n" or "N" for non-oedema, and "y" or "Y" for oedema. Users must code any missing values as " ", and the macro assumes that they are non-oedema. For oedema cases, weight-related z-scores (`_zwfa` and `_zbfa`) are not calculated, but they are treated as being < -3 SD in their prevalence estimations.
- **sw:** to specify the name of a numeric variable containing the sampling weights. If "**sw=1**" for all records, the un-weighted analysis is performed. If otherwise specified, negative values in the sampling weights are not allowed and in this case no prevalence tables will be produced.

Exported files

1. The macro creates, in the working directory, a STATA data set. The name of the data set is ***datalab_z.dta*** (see the preceding **Parameters** section). This data set retains all the records and variables from the input STATA data set and adds on the following 8 variables derived by the macro:

Variable name	Variable label
<code>_agedays</code>	calculated age in days for deriving z score
<code>_cbmi</code>	calculated bmi=weight / squared(_clenhei)
<code>_zwfa</code>	Weight-for-age z-score
<code>_zhfa</code>	Length/height-for-age z-score
<code>_zbfa</code>	BMI-for-age z-score
<code>_fwfa</code>	Flag for <code>_zwfa</code> < -6 or <code>_zwfa</code> > 5
<code>_fhfa</code>	Flag for <code>_zhfa</code> < -6 or <code>_zhfa</code> > 6

_fbfa Flag for _zbfa < -5 or _zbfa > 5

2. The macro creates, in the working directory, a data set in XLS format that has the same data structure as the one in STATA format. The name of the data set is ***datalab_z.xls***.
3. The macro creates, in the working directory, an MS Excel file that contains the point estimates and 95% confidence intervals for the indicator prevalences and z-scores summary statistics. In this summary analysis, only children aged between 61 to 228 completed months are included. The age groups in yearly and monthly intervals are shown below:

Age		Indicator		
Years	Months	Weight	Height	BMI
5	61-71	✓	✓	✓
6	72-83	✓	✓	✓
7	84-95	✓	✓	✓
8	96-107	✓	✓	✓
9	108-119	✓	✓	✓
10	120-131	✓*	✓	✓
11	132-143		✓	✓
12	144-155		✓	✓
13	156-167		✓	✓
14	168-179		✓	✓
15	180-191		✓	✓
16	192-203		✓	✓
17	204-215		✓	✓
18	216-227		✓	✓
19	228		✓	✓

* For Weight-for-age, age group 10 yrs covers up to age 120 completed months.

An example

An example survey, ***survey_who2007.dta***, is included in the package zip file (***who2007_stata.zip***) that should be used with the stata do-file, ***survey_who.do***.

```

/* Example: survey_who2007.do using survey_who2007.dta */
clear
set more 1
/* Higher memory might be necessary for larger datasets */
set memory 50m
set maxvar 10000

/* Indicate to the Stata compiler where the who2007.ado file is stored*/
adopath + "D:\WHO 2007 Stata/"

/* Load the data file */
use "D:\WHO 2007 workdata\survey_who2007.dta", clear

/* generate the first three parameters reflib, datalib & datalab */
gen str60 reflib="D:\WHO 2007 Stata"
lab var reflib "Directory of reference tables"
gen str60 datalib="D:\WHO 2007 workdata"
lab var datalib "Directory for datafiles"
gen str30 datalab="survey_2007"
lab var datalab "Working file"

/* check the variable for "sex" 1 = male, 2=female */
desc sex
tab sex

/* check the variable for "age" */
desc agemons
summ agemons

/* define your ageunit */
gen str6 ageunit="months" /* or gen ageunit="days", gen
ageunit="years" */
lab var ageunit "=days or =months or =years"

/* check the variable for body "weight" which must be in kilograms*/
/* NOTE: if not available, please create as [gen weight=.]*/
desc weight
summ weight

/* check the variable for "height" which must be in centimeters*/
/* NOTE: if not available, please create as [gen height=.]*/
desc height
summ height

/* check the variable for "oedema"*/
/* NOTE: if not available, please create as [gen str1 oedema="n"]*/
desc oedema
tab oedema

/* check the variable for "sw" for the sampling weight*/
/* NOTE: if not available, please create as [gen sw=1]*/
desc sw
summ sw

/* Fill in the macro parameters to run the command */
who2007 reflib datalib datalab sex agemons ageunit weight height oedema sw

```