WHO 2007
R macro package

The package, *WHO2007_R.zip,* contains:

1. The source codes of the function who2007.r,

2. Three tab-limited txt reference tables containing the WHO 2007 references: wfawho2007.txt, hfawho2007.txt and bfawho2007.txt,

3. Readme.pdf, and

4. An example survey data, survey_who2007.csv, and its related output files:

   survey_who2007_z.csv,
   survey_who2007_prev.csv,

Supported versions: The functions in this package were developed and tested using R version 3.0.1 (2013-10-07) only. Some adaptations might be necessary to run them in older versions.

Contact for reporting bugs/comments:
Should you encounter any problem with this macro, 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* package and also indicate which version of R you are using. Thank you.

Setup and run

1. Download the file **"who2007_R.zip"** to a basic directory [dir] where you want to include the package.

2. Restore reference data sets using the following commands:
   wfawho2007<-read.table("[dir]\\ wfawho2007.txt",header=T,sep="",skip=0)
   hfawho2007<-read.table("[dir]\\ hfawho2007.txt",header=T,sep="",skip=0)
   bfawho2007<-read.table("[dir]\\ bfawho2007.txt",header=T,sep="",skip=0)

3. Source the who2007.r file in the package directory [dir], using the command
   > source("[dir]\\who2007.r")

References

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.

Function who2007

DESCRIPTION

1. The function who2007, calculates z-scores for the ***three*** anthropometric indicators, weight-for-age, height-for-age, and BMI-for-age for children aged 5 to 19 years old.

2. It produces indicator-, sex- and age-specific prevalence estimates including 95% confidence intervals of under/over nutrition and summary statistics (mean and SD) of the z-scores (non-missing and non-flagged); if required, the analysis may be further stratified by a grouping variable such as urban/rural or regions. For oedema cases with non-missing age, weight-related z-scores (weight-for-age and BMI-for-age) are not calculated, but they are treated as being < -3 SD in their prevalence estimations

3. Extreme (i.e. biologically implausible) z-scores are flagged for each indicator 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

4. The function requires three tab-limited txt reference tables containing the WHO 2007 references. They must be stored in the a basic directory [dir] (see step 2 in "Setup and run" above):

Standards	R names
Weight-for-age	wfawho2007
Height-for-age	hfawho2007
BMI-for-age	bfawho2007

USAGE
who2007(FileLab="Temp",FilePath="C:\\Documents and Settings\\", mydf=, sex=, age=, weight=, height, =, oedema=rep("n",dim(mydf)[1]), sw=rep(1,dim(mydf)[1]))

REQUIRED ARGUMENTS:

- mydf: the data frame containing the variables to be used.

- sex: the name of the variable containing gender information; if it is numeric, its values must be: 1 for males and 2 for females. If it is character, it must be "m" or "M" for males and "f" or "F" for females. No z-scores will be calculated if sex is missing.

OPTIONAL ARGUMENTS

The following parameters can be left unspecified, and a default analysis will be performed. Note: If none of the measurement variables (weight, height) is provided, though, the analysis output will be empty. If variable age is not available, also no resulting analysis will be provided.

- FileLab: a label chosen to be used as part of the output files names. It is specified in double quotes. If unspecified, the default label "Temp" will be used.

- FilePath: the directory where the user wants the output Excel files to be stored. If unspecified, the default directory "C:\\Documents and Settings\\" will be used.

- age: to specify the name of a numeric variable containing age information; age must be in months not rounded. None of the three age-related z-scores will be calculated if age is missing.

- weight: the name of the numeric variable containing body weight information. Body weight must be in kilogram. Weight-related z-scores are not calculated if body weight is missing.

- height: the name of the numeric variable containing height information. Height must be in centimeters. Height-related z-scores are not calculated if height is missing.

- oedema: 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. Although it is highly recommended that this variable is provided by the survey, it is possible to run the analysis without specifying this variable. If unspecified, the default vector rep("n",dim(mydf)[1]) with values considered as non-oedema is used. Missing values will be treated as non-oedema. For oedema, weight related z-scores (zwfa and zbfa) are NOT calculated (set to missing), BUT if age is non-missing they are treated as being < -3 SD in the weight-related indicator prevalence estimation.

- sw: the name of a numeric variable containing sampling weight. If unspecified, the rep(1,dim(mydf)[1]) vector is used, i.e. where all records have equal sampling weights, and un-weighted analysis is performed. Negative values are not allowed.

Note: It is possible to miss one or more optional arguments. The corresponding default values will be used whenever an argument is missing.

EXPORTED FILES

This function does not provide any value; instead it creates, in directory FilePath, two Excel files: one containing z-score variables and other relevant created variables ("FileLab_z.csv"), and another displaying a table with under- and over-nutrition prevalence rates (and corresponding 95% confidence interval) and z-score means and standard deviations ("FileLab_prev.csv"). The latter summarizes results for each of the three indicators, by gender and age group.

The file FileLab_z_st.xls contains all variables in the input data frame mydf plus the following variables:

Variable	Description
cbmi	calculated BMI=weight / squared (height)
zwfa	Weight-for-age z-score
zhfa	Height-for-age z-score
zbfa	BMI-for-age z-score
fwfa	Flag for zwfa<-6 or zwfa>5
fhfa	Flag for zhfa<-6 or zhfa>6
fbfa	Flag for zbfa<-5 or zbfa>5

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.

EXAMPLE

Step 1: Before using who2007 function for the first time, remember to follow the 3 steps in the *Setup and run* given above.

Step 2: Import your data file into R. The following example is using data set survey_who2007.csv contained in the package who2007_R.

```
#Example: dataframe=survey.who2007
> survey.who2007<-
    read.csv("[dir]\\survey_who2007.csv",header=T,sep=",",skip=0,na.strings="")
```

Step 3:

```
#Standard analysis
> who2007(FilePath=" D:\\R who2007 example",
    FileLab="survey_who2007",mydf=survey.who2007,sex=sex,age=agemons,
    weight=weight,height=height,oedema=oedema,sw=SW)
```

Output: The files survey.who2007_z.csv and survey.who2007_prev.csv are created in the directory "D:/R who2007 example".