NDNS Years 1-3: weighting variables (core sample only)

The NDNS requires weights to correct for differences in sample selection and response. The weights adjust for differential selection probabilities of households and individuals, non-response to the individual questionnaire and non-response to the nurse visit. Non-response weights were generated using calibration weighting methods. The sample has been drawn and the weights designed so that all three years can be used together. However, there are some variables that are not available in all years so weighted for the individual years have also been included. The variables are described in Table 1.

Table 1 Description of the weights

Name	Description	Use for
Years 1-3 combined		
wti_ady123	Adult weight for individual and diary – combined Yr1-3	Any analyses of interview and food diary data for adults (19 years+) in the core sample combined Yr1-3 data
wti_chy123	Child weight for individual and diary – combined Yr1-3	Any analyses of interview and food diary data for children (18 months-18 years) in the core sample combined Yr1-3 data
wtn_ady123	Adult weight for nurse – combined Yr1-3	Any analyses of nurse data for adults (19 years+) in the core sample combined Yr1-3 data
wtn_chy123	Child weight for nurse – combined Yr1-3	Any analyses of nurse data for children (18 months-18 years) in the core sample combined Yr1-3 data
wti_ady12316	Adults 16+ weight for individual and diary – combined Yr1-3	Specific analysis of smoking behaviour/alcohol consumption for adults and older teenagers (16 years +) in the core sample combined Yr1-3 data
wtb_ady123	Adult weight for blood data – combined Yr1-3	Any analyses of blood sample data for adults (19 years+) in the core sample combined Yr1-3 data
wtb_chy123	Child weight for blood data – combined Yr1-3	Any analyses of blood sample data for children (18 months-18 years) in the core sample combined Yr1-3 data
Years 1-2 combined		
wti_ady12	Adult weight for individual and diary – combined Yr1&2	Any analyses of interview and food diary data for adults (19 years+) in the core sample combined Yr1 & 2 data
wti_chy12	Child weight for individual and diary – combined Yr1&2	Any analyses of interview and food diary data for children (18 months-18 years) in the core sample combined Yr1 & 2 data
wtn_ady12	Adult weight for nurse – combined Yr1&2	Any analyses of nurse data for adults (19 years+) in the core sample combined Yr1 & 2 data
wtn_chy12	Child weight for nurse – combined Yr1&2	Any analyses of nurse data for children (18 months-18 years) in the core sample combined Yr1 & 2 data
wti_ady1216	Adults 16+ weight for individual and diary –	Specific analysis of smoking behaviour/alcohol consumption for adults

	combined Yr1&2	and older teenagers (16 years +) in the core sample combined Yr1 & 2 data
Year 1 only		
wti_adulty1	Adult weight for individual and diary – Year 1 only	Any analyses of interview and food diary data for adults (19 years+) in the core sample Year 1 data only
wti_childy1	Child weight for individual and diary – Year 1 only	Any analyses of interview and food diary data for children (18 months-18 years) in the core sample Year 1 data only
wtn_adulty1	Adult weight for nurse – Year 1 only	Any analyses of nurse data for adults (19 years+) in the core sample Year 1 data only
wtn_childy1	Child weight for nurse – Year 1 only	Any analyses of nurse data for children (18 months-18 years) in the core sample Year 1 data only
wti_ady1_16	Adults 16+ weight for individual and diary – Year 1 only	Specific analysis of smoking behaviour/alcohol consumption for adults and older teenagers (16 years +) in the core sample Year 1 data only
Year 2 only		
wti_adulty2	Adult weight for individual and diary – Year 2 only	Any analyses of interview and food diary data for adults (19 years+) in the core sample Year 2 data only
wti_childy2	Child weight for individual and diary – Year 2 only	Any analyses of interview and food diary data for children (18 months-18 years) in the core sample Year 2 data only
wtn_adulty2	Adult weight for nurse – Year 2 only	Any analyses of nurse data for adults (19 years+) in the core sample Year 2 data only
wtn_childy2	Child weight for nurse – Year 2 only	Any analyses of nurse data for children (18 months-18 years) in the core sample Year 2 data only
wti_ady2_16	Adults 16+ weight for individual and diary – Year 2 only	Specific analysis of smoking behaviour/alcohol consumption for adults and older teenagers (16 years +) in the core sample Year 2 data only

Individual non-response weight

The individual non-response weight was generated for the analysis of fully responding individuals; individuals who responded to the individual interview and completed at least three food diary days. The individual non-response weight was generated using calibration methods. An iterative procedure was used to adjust a starting weight until the distribution of the (weighted) sample matched that of the population for a set of key variables. The adjustment keeps the values of the final weights as close as possible to those of the initial weights, which ensures the properties of the initial weights are retained in the final calibrated weights. The composite selection weights described above were used as the initial weights.

The key variables used to create the individual weight were age (grouped) by sex and Government Office Region (GOR). The population figures used were taken from the 2009 mid-year population estimates[1].

The aim of the calibration weighting was to reduce non-response bias resulting from differential non-response at the household and individual interview. The calibration weights generated were re-scaled so that the sum of the weights equalled the number of participating individuals; these are the final individual weights for the core sample. Thus the final individual weights adjust for dwelling unit, catering unit and individual selection, the Run In sample, and for the age/sex and regional profiles of participating individuals in the included survey years.

Nurse interview non-response weight

Participants who completed at least three food diary days were asked for their permission for a nurse to visit. Non-response weights were generated to adjust for differences between participants and non-participants to the nurse visit. These weights should be used for any analyses of nurse level data.

The nurse non-response weights were generated through a two-step process. Firstly, the data were used to model response behaviour[2]. Logistic regression was used to model the relationship between an outcome variable (response to the nurse interview) and a set of predictor variables (a set of socio-demographic respondent and household characteristics collected from the individual interview). The model generates the predicted probability each respondents would take part in the nurse interview, given the characteristics of the respondent and the household. These predicted probabilities are then used to generate a set of non-response weights; respondents with a low predicted probability get a larger weight, increasing their representation in the sample.

These non-response weights are then calibrated to get the final nurse weights. The nurse calibration used weighting totals based on (weighted) estimates from the individual questionnaire. This final step made respondents to the nurse visit match the population distribution in terms of age, sex and region and match the weighted respondents to the individual questionnaire in terms of household size, ethnicity of Main Food Provider, economic activity of the Household Reference Person (this was the economic activity of the respondent, if the respondent was an adult). The initial weights were the weights produced by the nurse non-response model.

---

[1] Produced for England and Wales by ONS, for Scotland by GROS and for Northern Ireland by NISRA.

[2] This step was not done on Y1 data. The small sample sizes resulted in weak non-response models and this stage was left out.

As before the calibration weights were re-scaled so that the sum of the weights equalled the number of participants who had a nurse visit. These final nurse weights adjust for unequal selection, non-response to the household and individual interview and non-response to the nurse visit.

Alcohol weights (Year 1, Year 2 and Years 1-2 combined)

An additional weight was required for a specific analysis of smoking behaviour and alcohol consumption of individuals aged 16 and over. This weight was generated for all core sample participants at general sample addresses[3], who were aged 16 and over and had competed an individual interview and at least three diary days. An additional nurse weight was not required for this specific sample. As before, calibration methods were used to generate the weights. The initial weight was the composite selection weight. This weight was adjusted using an iterative procedure to give a final weight that made the age, sex and regional and distribution of the weighted sample match that of the population aged 16 and over. The final weight makes the weighted sample of individual's representative of the UK population age 16 and over.

[3] It was more efficient to exclude the child boost sample than to include them but weight them down.