Public Health England

Protecting and improving the nation's health

Food Standards Agency
food.gov.uk

National Diet and Nutrition Survey

Years 7 and 8 (2014/15-2015/16)

User Guide for UK Data

NatCen Social Research

MRC Elsie Widdowson Laboratory

UKDA version

At **NatCen Social Research** we believe that social research has the power to make life better. By really understanding the complexity of people's lives and what they think about the issues that affect them, we give the public a powerful and influential role in shaping decisions and services that can make a difference to everyone. And as an independent, not for profit organisation we're able to put all our time and energy into delivering social research that works for society.

The Medical Research Council Elsie Widdowson Laboratory (MRC EWL), formerly called MRC Human Nutrition Research (HNR), conducts internationally-renowned research aimed at improving health through nutrition. Our mission is to conduct nutrition research and surveillance to improve the health of the population and to inform nutrition policy and practice.

NatCen Social Research
35 Northampton Square
London EC1V 0AX
www.natcen.ac.uk

A Company Limited by Guarantee
Registered in England No.4392418.
A Charity registered in England and Wales (1091768) and Scotland (SC038454)

Contents

Glossary 5

1. Introduction 8

1.1 The NDNS RP 8
1.2 NDNS RP reports 9

2. Survey design 10

2.1 Sample design and response 10
2.2 Survey structure 11

3. Archive documentation 12

4. Using the data 13

Years 7-8 datasets 13
4.2 UK Nutrient Databank 14
4.3 Variables in the datasets 14
4.4 Missing values conventions 15
4.5 Merging datasets 15
4.6 Serial number composition 16

5. Dietary, spot urine and blood data 18

5.1 Using the dietary data 18
5.2 Using spot urine iodine data 20
5.3 Using the blood data 21

6. Weighting variables 23

6.1 Description of weights 23
6.2 Single weights for all individuals 23
6.3 Individual non-response weight 24
6.4 Nurse interview non-response weight 24
6.5 Recent Physical Activity Questionnaire (RPAQ) non-response weight 25
6.6 Blood sample non-response weights 26
6.7 Selecting the appropriate weight variable 27

Appendix A UK Nutrient Databank 28

A.1 Introduction to the UK Nutrient Databank 28
A.2 Description of the Nutrient Databank fields and variables 29

A.2.1 Food name and food code 29
A.2.2 Food code markers 29
A.2.3 Sub food group name and code 29
A.2.4 Water and vitamin loss 30
A.2.5 Maximum weight 30
A.2.6 Base and unit (F codes only) 30
A.2.7 Dilution 30
A.2.8 Edible portion 30
A.2.9 Comments and descriptions 31
A.2.10 Nutrition data 31
A.3 Quality assurance 31
A.4 Nutrient fields and disaggregation categories in the Year 7 databank 32

Glossary

25-OHD	25-hydroxyvitamin D
5-A-Day	5 portions of fruit and vegetables per day (5 x 80g portions)
AOAC fibre	Intake of fibre defined using the AOAC method of analysis
Ca	Calcium
CAPI	Computer Assisted Personal Interview
Databank	UK Nutrient Databank
DH	Department of Health
DLW	Doubly Labelled Water
Fe	Iron
FM	Fat Mass
FQ	food quotient
FSA	Food Standards Agency
G	Gram
GOR	Government Office Region
HbA1c	Glycosylated Haemoglobin or Haemoglobin A1c
HoloTC	Holotranscobalamin
HRP	Household Reference Person
IMS	Intrinsic Milk Sugars
Kcal	kilocalorie
Kg	Kilogram
KJ	Kilojoule
L	Litre
LC-MS/MS	Liquid chromatography-tandem mass spectrometry
LIDNS	Low Income Diet and Nutrition Survey
LDL	Low Density Lipoprotein Cholesterol

LRNI	Lower Recommended Nutrient Intake
□g	Microgram
MAFF	Ministry of Agriculture, Fisheries and Farming
MG	Milligram
Mins	Minerals
MFP	Main Food Provider
MRC	Medical Research Council
MRC EWL	Medical Research Council Elsie Widdowson Laboratory
MRC HNR	Medical Research Council Human Nutrition Research
NatCen	NatCen Social Research
NDNS RP	National Diet and Nutrition Survey Rolling Programme
NISRA	Northern Ireland Statistics and Research Agency
NMES	Non-Milk Extrinsic Sugars
NSP	Non-Starch Polysaccharide
PABA	Para-aminobenzoic acid
PAF	Postcode Address File
PHE	Public Health England
PSU	Primary Sampling Unit
RNI	Reference Nutrient Intake
RPAQ	Recent Physical Activity Questionnaire
Se	Selenium
sTfR	Soluble Transferrin Receptors
Sub	Subsidiary
Trig	Triglyceride or Triacylglycerol
UCL	University College London
UK	United Kingdom
V25OHD_Std	Standardised 25-OHD variable

V25OHD	Unadjusted 25-OHD variable
VDSP	Vitamin D standardization Programme
Vit C	Vitamin C
Vits	Vitamins
Zn	Zinc

1. Introduction

1.1 The NDNS RP

The NDNS RP is designed to assess the diet, nutrient intake and nutritional status of the general population aged 1.5 years and over living in private households in the UK. The NDNS RP is jointly funded by Public Health England (PHE),[1] an executive agency of the Department of Health, and the UK Food Standards Agency (FSA)[2,3,4] and is carried out by a consortium comprising NatCen Social Research (NatCen)[5] and MRC Elsie Widdowson laboratory (MRC EWL), formerly known as MRC Human Nutrition Research (HNR).[6] Fieldwork in Northern Ireland is carried out by the Northern Ireland Statistics and Research Agency (NISRA).[7]

The NDNS RP provides the only source of high quality nationally representative data on the types and quantities of foods consumed by individuals, from which estimates of nutrient intake for the population are derived.[8] Results are used by Government to develop policy and monitor progress on diet and nutrition objectives of UK Health Departments, for example those set out in the Healthy Lives, Healthy People White Paper in England.[9] The food consumption data are also used by FSA to assess exposure to chemicals in food, as part of the risk assessment and communication process in response to a food emergency or to inform negotiations on setting regulatory limits for contaminants.

The NDNS programme began in 1992 as a series of cross-sectional surveys, each covering a different age group. Since 2008, the NDNS has been a RP covering adults and children aged 1.5 years and over. The NDNS collection also includes separate survey assessments of dietary sodium in adults.

The specific aims of the NDNS RP are to:

- provide quantitative data on the food and nutrient intakes, sources of nutrients and nutritional status of the UK population aged 1.5 years and above
- provide information on trends in food consumption, nutrient intake and nutritional status in different age groups

- describe the characteristics of individuals with intakes of specific nutrients above or below the national average
- produce a database of food consumption which will be used to calculate intakes of natural toxicants, contaminants, additives and other food chemicals
- measure blood and urine indices that provide evidence of nutritional status or dietary biomarkers, and to relate these to dietary, physiological and socio-demographic data
- provide height, weight and other anthropometric measurements and examine their relationship to socio-demographic, dietary, biochemical and health data
- monitor the diet of the population to establish the extent to which it is adequately nutritious and varied
- monitor the extent to which the diets of population sub-groups vary from expert recommendations
- assess total energy expenditure and physical activity levels and patterns in the study population

1.2 NDNS RP reports

Further information about the NDNS collection and the published reports can be found on the gov.uk site: https://www.gov.uk/government/collections/national-diet-and-nutrition-survey

Published tables for Years 7-8 can be found at:
https://www.gov.uk/government/statistics/ndns-results-from-years-7-and-8-combined

2. Survey design

2.1 Sample design and response

The NDNS RP is a survey of the food consumption, nutrient intakes and nutritional status of people aged 1.5 years and older living in private households. The survey is carried out in all four countries of the UK and is designed to be representative of the UK population. The survey aimed to collect data from a UK representative sample of 1000 people per fieldwork year, 500 adults (aged 19 years and older) and 500 children (aged 1.5 to 18 years). Recruitment in both Wales and Northern Ireland was boosted to 200 participants per year in order to achieve country-specific, representative dietary health data. Field work was conducted throughout the year (from April 2014 through August 2015 for Year 7 and April 2015 through August 2016 for Year 8) in order to take into account potential seasonal variations in food consumption.

The sample was drawn from the Postcode Address File (PAF), a list of all the addresses in the UK. In order to improve cost effectiveness the addresses were clustered into Primary Sampling Units (PSUs), small geographical areas based on postcode sectors, randomly selected from across the UK. A list of addresses was randomly selected from each PSU.

Overall for Years 7-8 a UK sample of 8,848 addresses was selected from 316 PSUs. In each PSU, 28 addresses were randomly selected. At each address, the interviewer established the number of households and, in cases where there were two or more, selected one household at random.

In order to achieve (as far as possible) equal numbers of adults (aged 19+ years) and children (aged 1.5-18 years) in the sample, at some addresses only children were selected to take part. The addresses were randomly allocated to one of two groups to determine whether an adult and a child, or a child only, was selected for interview.

At 10 of the selected addresses the interviewer selected one adult and, where present, one child for inclusion in the survey (“basic” addresses). The 18 remaining addresses were for a “young person/child boost” and the interviewer only carried out interviews in households with children. In households containing more than one eligible person (adult and/or child), interviewers selected the participant(s) using a random selection procedure.

In total 1417 adults and 1306 children gave fully productive interviews (consisting of three or four diary days). Of these (i.e. those who had completed a diary), 50% of adults (704) and 25% of children (329) went on to give a blood sample.

2.2 Survey structure

There are two main parts to the survey: an interviewer stage and a nurse visit.

Stage 1: Interviewer visit:

- Four-day food diary[10]
- Face-to-face Computer Assisted Personal interview (CAPI)[11]
- Height and weight measurements
- Smoking and drinking self-completion questionnaires (aged 8-17 years)
- Physical activity self-completion questionnaire (aged 16+ years)
- Collection of spot urine sample (aged 4+ years)

On successful completion of the interviewer stage (including three or four completed days of the food diary), each participant was invited to take part in the next stage, a visit from a nurse.

Stage 2: Nurse visit:

- Fasting blood sample (aged 4+ years)
- Non-fasting blood sample (aged 1.5-3 years)
- Physical measurements: waist and hip (aged 11+ years), demispan (aged 65+) and infant length (under 2 years)
- Blood pressure (aged 4+ years)
- Collection of information about prescribed medicines & dietary supplements

In addition, a subsample of Year 7 participants were recruited for a Doubly Labelled Water (DLW) sub-study to measure energy expenditure. DLW data is not included in the Years 6-7 dataset but is instead included in a separate Years 6-7 DLW dataset.

3. Archive documentation

The NDNS RP documentation has been organised into the following sections:

- **Survey documents**
  This contains the CAPI documentation for the interviewer visit and nurse schedules, self-completion questionnaires, showcards and consent forms. Separate documents are provided for Year 7 and Year 8 unless there was no change between the two fieldwork years in which case one document covering both years (and with Y7&8 in the title) is provided.
- **Data related documents**
  NDNS RP Yr7-8 Variable List - this contains a list of the variables on each dataset and the survey year to which it applies

  NDNS RP Yr7-8 Derived Variables – this contains the SPSS syntax specification for each of the derived variables included in the data.
- **Supporting documents**
  This contains details of food coding and instructions for office editing of the diaries and coding of the CAPI data and self-completions. Also, documents related to four-day food diary, interviewer and nurse project instructions and protocols.

Note that the questionnaire show the variable names used in the CAPI programme for Year 8. In some cases the variables in the data set have a different name where changes have been made to the question or routing between Years 7 and 8.

4. Using the data

Years 7-8 datasets

Data collected during the survey are contained in different data files described below.

Name of Dataset	No. of records	Description of Dataset
NDNS_Yr7-8a_indiv		Contains data for all fully productive individuals i.e. completed three/four food diary days. It contains information from the household questionnaire, main individual schedule, self-completions, physical measurements and nurse visit (where one occurred). It also includes blood sample results, and spot iodine data.
NDNS_Yr7-8a_hhold		Contains data on household composition, sex, age and marital status for all individuals in co-operating households.
NDNS_Yr7-8a_FoodLevelDietaryData		Diary data. Includes nutrient data and disaggregation at food level. Also, shows who else was present at the eating occasion, where the participant was located, whether the television was on and whether or not the participant was sitting at a table.
NDNS_Yr7-8a_DayLevelDietaryData_Foods		Daily food consumption data calculated using recipe main food groups and recipe sub food groups data.
NDNS_Yr7-8a_DayLevelDietaryData_Nutrients		Daily intakes of macronutrients, micronutrients and disaggregated foods.
NDNS_Yr7-8a_PersonLevelDietaryData		Mean intakes of nutrients, food consumption data calculated using recipe main food groups and recipe sub food groups data plus disaggregated food at the participant level. Also includes derived variables such as LRNI and RNI indicators and percentages.

4.2 UK Nutrient Databank

The UK Nutrient Databank used in the NDNS RP is saved as four different data files, one for each survey year. Details about the Nutrient Databank and its use can be found in Appendix A of this User Guide.

Name of Dataset	Description of Dataset
NDNSYr7NutrientDatabank_2016-05-06	Contains nutrient data assigned to foods and supplements for Year 7
NDNSYr8NutrientDatabank_2018-03-19	Contains nutrient data assigned to foods and supplements for Year 8

4.3 Variables in the datasets

The individual and household datasets contain questionnaire variables (excluding variables used for administrative purposes), demographic information including household composition, laboratory results and derived variables.

The dietary datasets contain variables coded from the diaries at food, day and person levels, plus dietary reference values and derived variables.

The variables included in all the datasets are detailed in the "**NDNS RP Yr7-8 UK Variable list**" document in the data section of the documentation. This document is the best place to look in order to plan your analysis. It includes:

- Major categories of variables (e.g. General Health, Blood Sample, Day Level Dietary Nutrients)
- Sub categories of variables (e.g. Longstanding illness (within General Health), Measurements from laboratory analysis (within the Blood Sample and Urine Collection sections), Nutrients including supplements (within Day Level Dietary Nutrients)
- Source of each variable (e.g. Individual questionnaire, Diary, Nurse visit, Self-completion booklet, Laboratory, Derived variable etc.)

Details of the question wording relating to a variable in the household and individual datasets is provided in the interview section documentation for the appropriate survey year. The "**NDNS RP Yr7-8 UK Derived Variables**" document provides information on how the variables were derived.

4.4 Missing values conventions

Missing value conventions are applied to most of the derived variables as well as the original questionnaire variables. The "**NDNS RP Yr7-8 UK Derived Variables**" specification should be consulted for details.

-1 Not applicable. This code is used to signify that a particular variable did not apply to a given participant because of internal routing (e.g. questions for children only) or because the participant did not participate in a particular element of the survey (e.g. refused a nurse visit).

-4 Question not applicable to survey year. This code indicates that that particular question was not asked/or element was not in a survey year.

-8 Don't know/Can't say

-9 No answer/Refusal

The above conventions apply to the majority of the variables on the data, however some variables have been attributed specific missing value codes, for example blood and urine results on the individual dataset, and "5 A Day" variables on the dietary data files. The description for each of the missing value codes are specified in the each variable value label.

For a full list of variables on the dataset and to which survey they apply see the "**NDNS RP Yr7-8 UK Variable list**".

4.5 Merging datasets

As various data are contained in different datasets, users may need to merge several datasets together for the purposes of their analysis. Individual serial number, survey year, age, sex and country variables are included in all the datasets for consistency.

4.6 Serial number composition

Serial numbers on the data consist of the following:

Variable name	Definition	Composition	File/files included
SERIALH	Household serial number	7 digits. The same number is allocated to each member of the same household. The first number corresponds to the survey year	Household and individual files
SERIALP	Individual identifier for each household member in a productive household	9 digits. SERIALH + PGRID	Household file only
SERIALI	Individual serial number for each productive individual (i.e. completed 3 or more diary days	8 digits SERIALH + ADCHILD ADCHILD is coded 1 for adult and 2 for child	Household, Individual and all Dietary files

The individual file also contains the person number of the Household Reference Person (HRP) and the Main Food Provider (MFP) (variables HRPNO and MFPNUM respectively). To create individual serial numbers for either the HRP or MFP, add HRPNO or MFPNUM to SERIALH.

Note that the HRP or MFP numbers correspond to the person number within each household. Therefore, due to the recoding of each productive individual to a 1 for the adult and 2 for the child, the HRP or MFP may not be the same individual although they may have the same serial number and vice versa. An example is shown below for a household consisting of 2 adults and 2 children.

	SERIALH	PGRID	SERIALP	Productive Y/N	SERIALI	Is HRP? Y/N	Is MFP? Y/N	HRP serial	MFP serial
Adult 1	7010101	1	701010101	No	N/A	No	Yes	N/A	70101011
Adult 2	7010101	2	701010102	Yes	70101011	Yes	No	70101012	N/A
Child 1	7010101	3	701010103	No	N/A	N/A	N/A	N/A	N/A

Child 2	7010101	4	701010104	Yes	70101012	N/A	N/A	N/A	N/A

In this example, the MFP and the individual are not the same person even though they have the same MFP and individual serial numbers. For clarification and cross checking, the age and sex of the HRP and MFP have been provided.

5. Dietary, spot urine and blood data

5.1 Using the dietary data

It is important to note the following when using the dietary data from Years 7-8 of the NDNS RP:

Days of the week

- The survey is designed so that all days of the week were evenly represented.[12]

Dietary coding

- Changes have been made to the dietary coding of homemade recipes and some purchased convenience foods in the NDNS RP, compared with previous standalone NDNS, which should be considered by those wishing to calculate food consumption data. In the RP all individual ingredients of a homemade recipe as reported in the food diary, or components of the purchased product as described on the food packaging, have been coded as their separate food codes and linked together under the appropriate Recipe Food Group, which highlights that those food codes were consumed together in one composite dish. The following variables should be used when calculating food consumption data:
    - RECIPEMAINFOODGROUPCODE
    - RECIPEMAINFOODGROUPDESC
    - RECIPESUBFOODGROUPCODE
    - RECIPESUBFOODGROUPDESC
- An example is provided here: A homemade dish of Thai chicken curry containing chicken, Thai curry sauce, and onion would appear in the RP Food Level dietary dataset as three entries with the food names; CHICKEN BOILED LIGHT MEAT ONLY, THAI CURRY SAUCE PURCHASED, and ONIONS BOILED, linked to the MAINFOODGROUPDESC of "*chicken and turkey dishes", "miscellaneous"* and *"vegetables not raw"*, respectively. As these three foods were consumed together in one composite dish they are assigned to the RECIPESUBFOODGROUPDESC of *Other chicken/turkey including homemade recipe dishes*. In previous NDNS datasets this homemade dish would have shown as one entry,

assigned as one food name, and the proportions of the composite dish consumed would not have been known.

- Recipes names are included in the Years 7-8 dataset (also included in the Years 1-4 and Years 5-6 datasets).
- To estimate absolute food consumption of one specific food type examine the FOODNAME and MAINFOODGROUPDESC variables, whilst examining disaggregation variables of any foods that are composites (NB disaggregation data is only provided for certain categories of meat, fish, fruit and vegetables). For example, to estimate absolute intakes of sausages from all sources you would need to include all the specific discrete portions of sausages, as well as calculate the percentage of sausages within all composite foods such as meat pies.
- All foods consumed have a base unit of grams that is, the amount consumed is described in grams. The exceptions are dietary supplements and artificial sweeteners. These have a base unit based on their form i.e. tablet, teaspoon. To avoid errors when calculating consumption, these have only been included in the food level dietary data file. When using this file, it should be noted that, for dietary supplements and artificial sweeteners, the value in the Total_Grams column is not a value in grams but a value in terms of the base unit, i.e. 0.5 for a granulated artificial sweetener would refer to 0.5 of a teaspoon not 0.5 grams.

5 A Day calculations

- Appendix A of the NDNS RP Years 7-8 report provides full details of:
    - the methods used for the disaggregation of meat, fish, fruit and vegetables
    - The variables/methodology used for 5 A Day calculations (section A.4.3)
    - all other dietary data methodologies used in the RP

Free sugars and AOAC fibre

Following the publication of the SACN Carbohydrates and Health report in 2015,[13] the recommendation is that the definition of free sugars and the AOAC definition of fibre should be adopted in the UK, replacing non-milk extrinsic sugars (NMES) and non-starch polysaccharides (NSP) respectively. Free sugars and AOAC fibre are included in the Years 7-8 dataset along with

NMES and NSP.[14] Further details of the methodology used to add free sugars and AOAC fibre is provided in Appendix AA of the NDNS RP Years 7-8 report.

Vitamin D

Following the publication of the SACN vitamin D report in 2016,[15] a RNI of 10μg vitamin D has been added to the Years 7-8 dataset for those aged 2 years and over and a safe intake of 10μg vitamin D has been added for those aged 1.5 to 2 years.[16]

Serum folate, red blood cell folate and unmetabolised "free" folic acid

Data for blood folate variables are included in the Years 7-8 dataset for all NDNS age groups.[17]

Comparing NDNS RP data to previous NDNS data

In comparing the NDNS RP data to past published reports of the NDNS, the differences in duration of assessment must be taken into account. Dietary assessment over a four-day period will provide similar mean intakes from assessment over a seven day period, but the variation in intakes will be different as will the percentage of participants consuming any given food over the recording period (percentage consumers), especially for foods that are consumed infrequently. Moreover, estimates of proportions of individuals above or below certain cut-off values, such as Lower Reference Nutrient Intakes (LRNIs) will be affected by assessments of different duration. For this reason, for the previous NDNS of young people aged 4 to 18 years and NDNS of adults aged 19 to 64 years reports, the seven day information from previous NDNS surveys was converted to four days using bootstrapping techniques. Details of the bootstrapping methodology used and descriptive statistics of bootstrapped data for NDNS foodgroups are provided in Appendix K of the NDNS RP Years 7-8 report (see section 1.2 of this User Guide for links to published reports).

5.2 Using spot urine iodine data

For Years 7-8, spot urinary iodine data have been supplied for fully continent participants over the age of four years. For those who provided a spot urine sample, but a urinary iodine result could not be obtained, the interview data have been included in the dataset but these individuals have been attributed a specific missing value code for each result variable and given a urine weighting of '0'.

5.3 Using the blood data

The complete blood results dataset have been supplied for those aged 1.5 years and over including both data for blood analytes published in the NDNS RP Years 7-8 tables as well as data for additional blood analytes measured but not reported. For those who provided a blood sample but results could not be obtained, the interview data has been included in the dataset but these respondents have been attributed a specific missing value code for each blood result variable and given a blood weight of 0.

Low Density Lipoproteins (LDL), Triglycerides (Trig) and Glucose (Glucose) result data are only provided for fasted blood samples. Glucose (Glucose), Selenium (Se) and Zinc (Zn) result data has been supplied for those aged 7 years and over.

Folate

Serum and red blood cell folate data free and (unmetabolised) folic acid data have been supplied for participants aged 1.5 years and over in Years 7-8.[18]

Fat soluble vitamins

From Year 5 onwards there was a change in the method used for the analysis of fat soluble vitamins, for more information see "**NDNS RP Yr5-6 User Guide**".

Vitamin D status (25-hydroxyvitamin D)

25-hydroxyvitamin D (25-OHD) was measured using the Diasorin Liaison method. These kits were later withdrawn and reformulated by the manufacturers because some samples gave inaccurately high results, greater than 100nmol/L, as a result of sample-specific interference.

Both Diasorin Liaison 25-OHD methods were standardised against an in-house MRC EWL (then called MRC HNR) liquid chromatography-tandem mass spectrometry (LC-MS/MS) method which gave results equivalent to the international gold-standard isotope-dilution-liquid chromatography-tandem mass spectrometry methods, as validated under the Vitamin D Standardization Program (VDSP). Concentrations after standardisation are approximately 3% higher than the results as

assayed. Standardised results (v25OHD_Std) are archived alongside the unadjusted data (v25OHD).

Holotranscobalamin (HoloTC)

Serum HoloTC data in addition to serum vitamin B_{12} have been supplied for participants aged 1.5 years and over as another measure of vitamin B_{12} status.

Soluble transferrin receptors (sTfR), Homocysteine

Analysis of these blood analytes ceased at the end of Year 5 of the NDNS RP, therefore sTfR data have not been supplied in the Years 7-8 dataset.

Retinyl palmitate

Analysis of this blood analyte ceased at the end of Year 4 of the NDNS RP, therefore retinyl palmitate data have not been supplied in the Years 7-8 dataset.

6. Weighting variables

6.1 Description of weights

The NDNS RP 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 and RPAQ questionnaires, non-response to the nurse visit and non-response to the blood sample. Non-response weights were generated using a mixture of non-response modelling and calibration weighting methods. Five weights were generated for the NDNS RP Years 7-8 datasets these are described in Table 1.

Table 1 Description of the NDNS RP Years 7-8 weights

Weight name	Description of weight	Use for
wti_Y78	Weight for non-response by individuals to the individual questionnaire and diary	Any analysis of individuals using data from the individual questionnaire or diary. Including analysis of Smoking & Drinking data (collected in self-completions & CAPI)
wtn_Y78	Weight for non-response by individuals to the nurse visit	Any analysis of individuals using data collected at the nurse visit
wtb_Y78	Weight for non-response by individuals to the blood sample	Any analysis of individuals using blood sample data
wtr_Y78	Weight for analysis of RPAQ (all individuals aged 16+)	Any analysis of RPAQ info for individuals aged 16+
wtsu_Y78	Weight for analysis of spot urinary iodine data (aged 4+)	Any analysis of individuals with spot urinary iodine data

6.2 Single weights for all individuals

There is a single weight for all individuals, rather than separate weights for adults and children. This means the sample needs to be filtered by age to ensure the correct ages are included. However, this means different age breaks to those presented in the NDNS RP Years 7-8 published tables can be used, i.e. 16 to 18 year olds can be combined with adults (19 years and over), which allows more flexibility in reporting.

6.3 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 weighting methods. An iterative procedure is used to adjust a starting weight until the distribution of the (weighted) sample matches 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 initial weights were a set of selection weights. These selection weights corrected for the unequal selection probabilities that were built into the sample design; for the random selection of individuals in households where more than one person was eligible and for the random selection of catering units and dwelling units at multi-unit addresses.

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 mid-year population estimates.[19] The average population of the two years was used; this was generated using the most recent years of population data (2014).

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 (wti_Y78). Thus the final individual weights adjust for dwelling unit, catering unit and individual selection, and for the age/sex and regional profiles of participating individuals. This weight should be used for any analyses of interview and food data in the Years 7-8 data.

6.4 Nurse interview non-response weight

Participants who completed three or four food diary days (i.e. those deemed fully productive) were asked to consent to a nurse visit. Around 70% of these participants went on to do a nurse interview. Non-response weights were generated to adjust for differences between participants and non-participants to the nurse visit.

There is a small difference in the approach used to create the nurse weights; the calibration step was dropped as the model produced weights that made the weighted profile of the nurse visit match the weighted profile of all individuals very closely. The modelling step, however, remained the same; a logistic regression was be used to model the relationship between response to the nurse interview (coded into a binary outcome variable) and a set of predictor variables (socio-demographic, participant and household/catering unit characteristics collected during the interview). Adults and children were modelled separately as response behaviour can vary between the two groups. The weights from the two models were then combined into a single weight.

The model generated a predicted probability for each participant. This is the probability the participant would take part in the nurse interview, given the characteristics of the individual and the household/catering unit. Participants with characteristics associated with non-response were under-represented in the nurse sample and therefore receive a low predicted probability. These predicted probabilities were then used to generate a set of non-response weights; participants with a low predicted probability got a larger weight, increasing their representation in the sample.

As before, the nurse weights were re-scaled so that the sum of the combined adult and child weights equalled the number of participants who had a nurse visit. These are the final nurse weights for the sample (wtn_Y78) and adjust for unequal selection, non-response to the household/MFP and individual interviews and non-response to the nurse visit. The nurse weights should be used for all analyses of nurse level data.

6.5 Recent Physical Activity Questionnaire (RPAQ) non-response weight

All individuals aged 16 years and over were asked to record their physical activity over the previous seven days in a self-completion booklet (the RPAQ).

Response behaviour was modelled using a logistic regression. The same set of predictor variables used to model non-response to the nurse visit was used to model non-response to RPAQ, namely, socio-demographic, participant and household/catering unit characteristics collected during the individual interview. The 16 to 18 year olds were modelled with the adult respondents as children under the age of 16 were not asked to fill in the self-completion booklet.

The RPAQ weights were re-scaled so that the sum of the combined adult and child weights equalled the number of participants who had completed RPAQ. These are the final RPAQ weights for the sample (wtr_Y78) and adjust for unequal selection, non-response to the household/MFP and individual interviews and non-response to the RPAQ.

Note that response to RPAQ and the nurse visit was not hierarchical; it was possible for a respondent to complete the RPAQ section but not the nurse visit, and vice versa.

6.6 Blood sample non-response weights

An additional set of weights was generated to correct for differential non-response to giving a blood sample. Non-response, whether due to refusal or inability to give a blood sample, will cause the blood data to be biased if there are systematic differences between individuals that provide a blood sample and individuals that do not. Blood samples were taken during the nurse visit. Only participants who fulfilled certain eligibility criteria were asked whether they would be prepared to give a blood sample. Participants were ineligible if they:

- had a clotting or bleeding disorder (e.g. conditions such as haemophilia and low platelets (thrombocytopenia))
- had **ever** had a fit
- were **currently** on anticoagulant drugs, e.g. Warfarin therapy
- had volunteered information that they are HIV or Hepatitis B or C positive

The blood weights were generated using a logistic regression model. Adults and children were modeled separately. The non-response weights from the model were combined with the final nurse weights to give the final blood weights for adults and children (wtb_Y78). These weights adjust for non-response to the blood sample, non-response to the nurse visit, non-response to the individual questionnaire and unequal selection probabilities. The weights were scaled, so the mean weight equalled one and the weighted sample size matched the unweighted sample size. These weights should be used for any analysis of blood sample data.

6.7 Selecting the appropriate weight variable

Any analyses that incorporate information from more than one stage should use the weight from the 'last' stage. For example, a cross tabulation using a variable from the nurse visit and a variable from the individual questionnaire should use the nurse weights (i.e. the weights from the latter stage of the survey), similarly, a cross tabulation of blood sample data and data collected at the nurse visit should use the blood weight.

Appendix A UK Nutrient Databank

A.1 Introduction to the UK Nutrient Databank

The food level dataset provides nutrients for the foods in amounts as consumed whilst the UK Nutrient Databank dataset provides nutrient information per 100g.

The UK Nutrient Databank (databank) contains extensive information on the nutrient content of foods commonly consumed in the United Kingdom (UK) and enables nutrient intakes to be calculated from consumption data. The databank was originally developed by the Ministry of Agriculture, Fisheries and Food (MAFF) for the 1990 Dietary and Nutritional Survey of British Adults and subsequently updated for the series of NDNS surveys between 1995 and 2000.[20] Ownership of the databank transferred to the FSA where it was updated for the NDNS of adults aged 19 to 64 years,[21] the Low Income Diet and Nutrition Survey (LIDNS),[22] and prior to commencing the NDNS RP in 2008.[23] In 2010 responsibility for the NDNS RP and ownership of the databank transferred to the Department of Health (DH). In April 2013, responsibility transferred to the Department's Executive Agency, Public Health England (PHE).

Data in the databank (managed by PHE and MRC EWL) is largely the same as that in the published UK food composition tables[24] but includes a larger range of processed foods and composite dishes and no missing values. The nutrient data assigned to foods originate from three main sources:

- Ongoing programme of nutrient analysis led by PHE
- Food manufacturers and information from food labels
- Recipe calculations for homemade dishes, and some manufactured products

In order to estimate nutrient intakes, values are assigned for all nutrients in all foods. Where reliable information is not available for some nutrients, data were obtained by extrapolating estimates from similar foods. All data were carefully evaluated before being incorporated into the databank.

In the NDNS RP it is essential that the databank is up-to-date and, as far as possible, reflects the nutrient composition of the food supply for each year reported. Hence a programme of updates and

revisions is a continuing aspect of the RP, with a yearly update carried out by PHE in conjunction with MRC EWL.[25] Each year of the RP is coded separately using a contemporaneous version of the databank. Updating of the databank includes the addition of new foods as well as revision of nutrient composition of existing foods, either at food group level following a programme of reanalysis, or extrapolating estimates by taking in to account the reformulations of manufacturers and changes in fortification practices. As changes in the databank are partly driven by the availability of new analytical data, new data are produced only occasionally so a gradual change in the nutrient content of the food supply may appear as a step change in the nutrient databank with the same foods having a different composition for some nutrients in one year of the RP compared to the next. It is important to utilise contemporaneous food composition data to produce accurate assessments of dietary intake.

A.2 Description of the Nutrient Databank fields and variables

A.2.1 Food name and food code

All foods are assigned a food code and a clear name is given to describe the code. The name may include specific qualities of the food, informing the user it is a fortified food, or from a specific brand. The food codes in the databank cannot be linked directly to foods in McCance and Widdowson's Composition of Foods.[26]

A.2.2 Food code markers

Each food code is assigned as an 'F' or 'R' code. 'F' indicates that the code exists as an atomic food and 'R' denotes a recipe code usually comprising 2 or more food codes. Details of the recipes held in the 'R' codes are not provided in this dataset.

A.2.3 Sub food group name and code

A full breakdown of the NDNS food groups is provided in Appendix R of the NDNS RP Years 7-8 report.[27] Each food code is assigned to a subsidiary (sub) food group, expressed as an integer with an alphabetical suffix, which is a food group level of greater detail than the main food groups. The databank displays the sub food group name as well as the code.

A.2.4 Water and vitamin loss

Recipe 'R' food codes may have water and vitamin loss assigned, which provides a correction to the nutrient data for associated cooking losses.

A.2.5 Maximum weight

This is an edit check field to assist users of the nutrient data to pick up any coding errors in portion sizes. Amending the maximum weight will not impact the nutrient values in the databank. Maximum weights are set at sub food group level and do not usually have to be adjusted for each code. Food codes that would not appear in dietary assessment records, such as raw meat, have a maximum weight set at 1.

A.2.6 Base and unit (F codes only)

These fields describe how the nutrient data are expressed. For most foods the nutrients are entered per 100 grams (base = 100, unit = grams). For vitamin and mineral supplements, nutrients are entered per tablet, capsule, teaspoon or drop (base = 1, unit = tablet, capsule, teaspoon or drop). The base and unit of a food code must match the form in which items are recorded in the dietary assessment method.

A.2.7 Dilution

The majority of food codes have a dilution of 1. Concentrated soft drinks and dried products that are made up have a dilution factor greater than 1. For example, if a concentrated drink is usually made up 1 part concentrate to 4 parts water the dilution factor would be 5. This provides the user with additional detail regarding the food in its unconcentrated form; however this is not used in calculating nutrient intakes.

A.2.8 Edible portion

The majority of food codes have an edible portion of 1; however this will be less than 1 for foods that include waste, such as meat weighed with bones.

A.2.9 Comments and descriptions

The comments box provides details on when the food code was created, including any updates and the data source. The description box provides details on the number and brand names of products included in calculating the nutrient information, or the food codes used in collating the recipe. Some food codes (particularly dietary supplements) contain nutritional compounds not measured in the nutrient databank (e.g. vitamin K or lutein).

A.2.10 Nutrition data

Each food code in the databank has a value assigned for 56 nutrients, including energy, provided in specific units of measurement (see table A.1). Additionally, to ensure accurate reporting of specific food types in NDNS, each food code present in the databank has been disaggregated into 28 specific food components (see table A.2).[28] Disaggregation values (g amount in per base unit, e.g. 75g 'fruit' in 100g 'fruit pie') are assigned into the databank. Nutrient values reported as 'trace' are assumed as zero in the nutrient databank. Some data will have been rounded into the appropriate number of decimal places.

A.3 Quality assurance

The databank is designed to perform automated quality assurance checks to nutrient data,[29] when food codes are added or updated:

Kcal	(protein x 4) + (fat x 9) + (carbohydrate x 3.75) + (alcohol x 7)
kJ	(protein x 17) + (fat x 37) + (carbohydrate x 16) + (alcohol x 29)
Carbohydrate	total sugars + starch
Total sugars	sum of all individual sugars
Total sugars	intrinsic + non milk extrinsic sugars
Total iron	haem iron + non haem iron
Total carotene	β-carotene + (α-carotene x 0.5) + (β-cryptoxanthin x 0.5)
Vitamin A	Retinol + (total carotene/6)
Total N	Protein/Nitrogen Conversion Factor (NCF)
Total fat	should be ≥ the sum of fatty acids
Fatty acids	0.6 X total fat is ≤ the sum of the fatty acids.

A.4 Nutrient fields and disaggregation categories in the Year 7 databank

Table A.1 Nutrient fields available in the databank

Nutrient (unit of measure)		
Water (g)	Lactose (g) ^	Vitamin B6 (mg)
Total Nitrogen (g)	Other sugars^ (g)	Vitamin B12 (µg)
Nitrogen conversion factor	Saturated fatty acids (g)	Folate (µg)
Protein (g)	Cis-monounsaturated fatty acids (g)	Pantothenic acid (mg)
Fat (g)	Cis-n3 fatty acids (g)	Biotin (µg)
Carbohydrate (g)^	Cis-n6 fatty acids (g)	Sodium (mg)
Energy (kcal)	Total trans fatty acids (g)	Potassium (mg)
Energy (kJ)	Cholesterol (mg)	Calcium (mg)
Alcohol (g)	Retinol (µg)+	Magnesium (mg)
Englyst Fibre (NSP) (g)*	Total carotene (µg)+	Phosphorus (mg)
AOAC fibre (g)**	Alpha-carotene (µg)+	Iron (mg)
Starch (g)^	Beta-carotene (µg)+	Haem iron^ (mg)
Total sugars (g)^	Beta cryptoxanthin (µg)+	Non-haem iron (mg)
Non-milk extrinsic sugars*** (NMES) (g)^	Vitamin A (retinol equivalents) (µg)+	Copper (mg)
Free sugars (g) ^ ****	Vitamin D (µg)	Zinc (mg)
Intrinsic and milk sugars***** (IMS) (g) ^	Thiamin (mg)	Chloride (mg)
Glucose (g) ^	Riboflavin (mg)	Iodine (µg)
Fructose (g) ^	Niacin equivalent (mg)	Manganese (mg)
Sucrose (g) ^	Vitamin C (mg)	Selenium (µg)
Maltose (g) ^	Vitamin E (mg)	

^ Carbohydrate is expressed as monosaccharide equivalents. Other sugars include oligosaccharides, where data is available on their levels. Haem iron is calculated as 40% of the iron in fish and meat.

* Includes total non-starch polysaccharides.
** Includes resistant starch, lignin and non-starch polysaccharides captured by AOAC method.

***Includes all sugars in fruit juices, table sugar, honey, sucrose, glucose and glucose syrups added to foods + 50% of the sugars in canned, stewed, dried or preserved fruits. Non-milk extrinsic sugars cannot be determined by chemical analysis so values are estimated based on available information on the types and sources of sugar in the food

**** Includes all monosaccharides and disaccharides added to foods by the manufacturer, cook or consumer, plus sugars naturally present in honey, syrups and unsweetened fruit juices (see Appendix AA).

*****Includes all sugars in fresh fruit and vegetables + 50% of the sugars in canned, stewed, dried or preserved fruits + lactose in milk and milk products.+ Vitamin A retinol equivalent is calculated as Retinol + (total carotene/6). Where, Total carotene is β-carotene + (½ α-carotene) + (½ β-cryptoxanthin). Fortified sources of vitamin A are entered as retinol in the nutrient databank.

Table A.2 Disaggregation categories used in the NDNS RP

Disaggregation category (g)	
Fruit (Fresh and canned fruit)	Other Red Meat
Dried Fruit	Burgers (Burgers and grill steaks)
Fruit Juice	Sausages
Smoothie Fruit	Offal
Tomatoes	Poultry (white meat)
Tomato puree	Processed Poultry
Brassicaceae	Game Birds
Yellow red Green (Yellow, Red & Dark Green Leafy Vegetables)	White Fish
Other Vegetables	Oily Fish
Beans (Beans and pulses)	Canned Tuna
Nuts	Shellfish
Beef (red meat)	Cottage cheese
Lamb (red meat)	Cheddar Cheese
Pork (red meat)	Other Cheese
Processed Red Meat	

[1] https://www.gov.uk/government/organisations/public-health-england

[2] https://www.food.gov.uk/

[3] Responsibility for nutrition policy in England and Wales transferred from FSA to Health Departments in 2010. Management of the NDNS RP also transferred to the Department of Health (DH) in England at that time. From 1 April 2013, responsibility for the survey transferred to DH's Executive Agency, Public Health England (PHE).

[4] Additional recruitment in the devolved countries is funded as required by Government bodies in Scotland, Wales and Northern Ireland.

[5] http://www.natcen.ac.uk/

[6] http://www.mrc-ewl.cam.ac.uk/

[7] http://www.nisra.gov.uk/

[8] Ashwell M, Barlow S, Gibson S, Harris C (2006) National Diet and Nutrition Surveys: the British experience. Public Health Nutrition 9(4) 523-530.

[9] Department of Health Healthy Lives, Healthy People: Our Strategy for public health in England White Paper http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_121941

[10] See the UK Years 7-8 report (https://www.gov.uk/government/statistics/ndns-results-from-years-7-and-8-combined) and section 5.1 of this user guide for more details.

[11] Collecting information on shopping and food preparation practices, cooking facilities in the household, eating habits and food avoidance. Also included a section for the Main Food Provider (MFP).

[12] This is also the case with the Years 5-6 data. However, in the Years 1-4 combined data, there remains a slightly higher proportion of weekend days than weekdays.

[13]https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/445503/SACN_Carbohydrates_and_Health.pdf

[14] Variables have also been added to the Years 1-4 and Years 5-6 datasets.

[15]https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/537616/SACN_Vitamin_D_and_Health_report.pdf

[16] It should be noted that the RNI has not been added retrospectively to the Years 1-4 and Years 5-6 datasets for those aged 4 to 64 years. However intakes as a % of the RNI can be calculated for Years 1-6 using the syntax provided in the Years 7-8 derived variables document.

[17] Following the re-publication of the Years 1 to 4 folate report, blood folate variables have also been added to the Years 1-4 and Years 5-6 datasets.

[18] In addition, data has also been added to the Years 1-4 and Years 5-6 datasets for serum and red blood cell folate and free (unmetabolised) folic acid.

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

[20] Smithers, G. (1993) MAFF's Nutrient Databank. Nutrition and Food Science; 2; 16-1.

[21] Henderson L, Gregory J, Swan G. National Diet and Nutrition Survey: adults aged 19 to 64 years. Volume 1: Types and quantities of food consumed. London: TSO, 2002.

[22] Nelson M, Erens B, Bates B, Church S, Boshier T. 2007. Low Income Diet and Nutrition Survey. Volume 2. Food Consumption. Nutrient intake. London, The Stationery Office.

[23] National Diet and Nutrition Survey: Results from Years 1 to 4 (combined) of the Rolling Programme2008/09-2011/12 [online] https://www.gov.uk/government/publications/national-diet-and-nutrition-survey-results-from-years-1-to-4-combined-of-the-rolling-programme-for-2008-and-2009-to-2011-and-2012.

[24] The summary edition and related supplements are referenced and located as follows: Composition of Foods Series, Sixth Edition and Supplements as cited in Bates B, Lennox A, Bates C & Swan G. National Diet and Nutrition Survey: Headline results from Years 1 and 2 (combined) of the Rolling Programme (2008/9 – 2009/10) http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsStatistics/DH_128166 (accessed 20/04/12).

[25] Previously FSA or DH when ownership of the Nutrient Databank lay with them.

[26] McCance and Widdowson's The Composition of Foods integrated dataset [online] http://tna.europarchive.org/20110116113217/http://www.food.gov.uk/science/dietarysurveys/dietsurveys/

[27] National Diet and Nutrition Survey: Appendix R Main and subsidiary food groups [online] https://www.gov.uk/government/statistics/ndns-results-from-years-7-and-8-combined
.

[28] Fitt E, Mak TN, Stephen AM, Prynne C, Roberts C, Swan G & Farron-Wilson M. (2010) Disaggregating composite food codes in the UK National Diet and Nutrition Survey food composition databank. Eur J Clin Nutr 64: S32-S36.

[29] Results from these calculations may not match the final values exactly (especially when comparing against analytical data) as a small margin of difference is acceptable in food composition.