Diet and Nutrition Survey of Infants and Young Children 2011

User Guide

MRC Human Nutrition Research
NatCen Social Research
Newcastle University

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.

NatCen Social Research
35 Northampton Square
London EC1V 0AX
T 020 7250 1866
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

1 Background.

The Diet and Nutrition Survey of Infants and Young Children (DNSIYC) was commissioned by the Department of Health (DH) and Food Standards Agency (FSA). Responsibility for nutrition policy in England and in Wales transferred from FSA to Health Departments in 2010, but remains with FSA in Scotland and Northern Ireland. Management of the DNSIYC contract also transferred to DH at this time; the core United Kingdom (UK) survey continues to be jointly funded by FSA and DH, with the additional recruitment in Scotland, Wales and Northern Ireland funded by organisations in those countries. The programme was carried out by a consortium of four organisations: MRC Human Nutrition Research (HNR), based in Cambridge, NatCen Social Research (NatCen)[i], based in London, MRC Epidemiology Unit, based in Cambridge and the Human Nutrition Research Centre at Newcastle University. Fieldwork in Northern Ireland was carried out by the Northern Ireland Statistics and Research Agency (NISRA).

The aims of the DNSIYC were to:

- provide detailed, quantitative information on the food and nutrient intakes, sources of nutrients, and nutritional status of a representative sample of infants and young children aged 4 to 18 months from the UK population, as a basis for developing government policy and measuring progress towards other government objectives
- provide detailed, quantitative information on breast milk and breast milk substitutes consumed by the population group under study
- describe the characteristics of participants with intakes and/or status of specific nutrients that are above and below national reference values, and evaluate the diet of this population compared to current national recommendations
- produce a database of food consumption to provide the basis for the calculation of likely dietary intakes of natural toxicants, contaminants, additives and other food chemicals for risk assessment
- provide length (height), weight and other body measurements and examine their relationship with dietary intake and status, and health and social factors
- examine the extent to which feeding practices adopted by carers of this population group differ from national policy for infant health

- provide some information on the dietary habits of the mother (and other key family members) and link this to the nutrient intakes and nutritional status of this population group
- carry out stable isotopically-labelled water assessment in sub-samples of the survey group in order to estimate breast milk intake and body composition in children consuming any breast milk, as well as fluid intake and body composition in non-breastfed children, and
- measure blood indices that give evidence of nutritional status and relate these to dietary, physiological and social data

2 Survey design.

DNSIYC is a survey of the food consumption, nutrient intakes and nutritional status of people aged 4 up to 18 months 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 sample was selected from Child Benefit (CB) records (estimated uptake of 98%) provided by Her Majesty's Revenue and Customs (HMRC)[ii]; the Healthy Start[iii] (HS) sample also used the HS recipient database (estimated to cover 22% of infants aged 4 to 18 months in the UK) provided by the Department of Health (DH). The CB and HS samples were stratified by Government Office Region, Index of Multiple Deprivation (IMD) scores and population density to ensure representativeness.

NatCen selected postcode sectors using CB claimant records provided by HMRC and information about HS recipients from DH. NatCen received postcode sector-level counts of children in receipt of CB and HS. These counts were used to create Primary Sampling Units (PSUs). PSUs were postcode sectors or groups of postcode sectors and clustering PSUs in this way helped improve the cost-effectiveness of the survey. Postcode sectors containing fewer than 50 eligible children were grouped with neighbouring sectors so that each PSU contained a minimum of 50 children in receipt of CB[iv].

The survey aimed to collect data from a UK representative sample of 1800 infants and young children. The PSUs for the core sample and HS boost were drawn together. The 178 PSUs were selected from the sorted PSU sampling frame and each PSU was to contain both core and HS children. Spreading the HS boost across 178 PSUs, rather than concentrating it within a small number of additional PSUs, reduced the effects of clustering. Thirty four additional PSUs were selected for the Scottish Boost sample. There was no overlap between the Scottish Boost and the HS boost.

A weighted selection approach was used to boost the number of HS children in the 178 PSUs. A weighting factor of 1.216 was applied to children on HS and all other children were given a weighting factor of 1 in order to obtain the required numbers of HS recipients. The weighted number of children was then generated for each PSU. Hence, a PSU containing 145 children in total, 120 of which were not on HS and 25 of which were on HS, got a weighted count of 120 + (25 X 1.216) = 150.4. The 178 core PSUs were selected with probability proportional to the weighted number of eligible children within them. This gave PSUs with a large number of HS children a higher chance of being selected.

From the entire core sample, a total of 2683 parents of children gave fully productive interviews (consisting of three or four diary days) and of these 973 children attended a clinic visit.

There were two main parts to the survey: an interviewer stage and a clinic visit with key components being:

- Face-to-face interviews, conducted using Computer Assisted Personal Interviewing (CAPI) with the parent most involved in feeding the sampled child.
- Dietary data collection, using a four-day estimated food diary
- Anthropometric measurements (maternal height and weight; child length, weight and head circumference; and skinfold thickness).
- Collection of a blood sample for estimation of iron and vitamin D status.

The interviewer stage included an extensive face-to-face Computer Assisted Personal interview (CAPI) with the parent of the sampled child; placement of a four-day food diary; and measurements of maternal height and weight and child length and weight. The interviewer also collected information on child development and health problems, feeding and food avoidance, plus information of the parents eating habits.

On successful completion of the interviewer stage (including three or four completed days of the food diary), the parent was invited to attend a local clinic where additional child and maternal anthropometric measurements and child blood samples were taken.

In addition, parents were asked to take part in a stable isotope dose protocol in order to measure the child's body composition, fluid intake and breast milk intake (if breastfed) through a urine collection protocol over 5 or 14 days.

Field work was conducted from January 2011 through to August 2011.

3 Documentation.

The documentation has been organised into the following sections

- **Survey documents**
  This contains the CAPI documentation for the interviewer visit, showcards, documents related to the four-day food diary and breast milk diary, consent forms, interviewer and clinic project instructions including measurement protocols and other survey documents.

- **Coding and editing instructions**
  These contain details of food and drink coding, and instructions for office editing and coding of the CAPI data.

- **Data related documents**
  This contains a list of the variables on each dataset and a derived variable specification.

4 Using the data.

4.1 Datasets

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

Name of Dataset	No. records	Description of Dataset
DNSIYC_Indiv_A	2,683	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, and clinic visit (where one occurred) including measurement and laboratory results.
DNSIYC_Hhold_A	10,279	Contains data on household composition, sex, age and marital status for all individuals in co-operating households
DNSIYC_FoodLevelDietaryData_A	218,779	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. 10/10/2013 data change – single day changed to a single food code
DNSIYC_DayLevelDietaryData_Foods_A	10,663	Food diary data. Daily food consumption data calculated using recipe main food groups and recipe sub food groups data. 10/10/2013 data change – Extraction of infant specific foods i.e. formula sub-categories and commercial infant food sub-categories from dietary databank when included in a recipe.
DNSIYC_DayLevelDietaryData_Nutrients_A	10,663	Food diary data. Daily intakes of macronutrients, micronutrients and disaggregated foods.
DNSIYCPersonLevelDietaryData_A	2,683	Food diary data. 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. 10/10/2013 data change – Changes made to only dietary food variables, corrected to 4 day average

Due to a number of post-reporting edits, the archived data may not always match the tables in the DNSIYC report on the DH website (see section 6).

4.2 Using dietary data

It is important to note the following when using the dietary data from DNSIYC food diaries:

- The DNSIYC data files have been prepared to the same format as those produced for the NDNS rolling programme[v], however there are some differences to food groups that should be considered if merging the different survey data. The food codes within two food infant-specific groups were further divided into more detailed categories than their sub-food Group. Infant formula contains the milk sub-categories; 'first milk', 'hungrier babies milk', Follow-on, 'growing up milk', soy-based milk, and 'other milk' (inc. hypoallergenic, goats, 'goodnight milk', extra hungry). Commercial toddler/infant foods contains the sub categories; 'meat and fish based products and dishes', 'other savoury based dishes' (not snacks), 'fruit based foods and dishes', 'dairy based foods and dishes', 'cereal based foods and dishes', and 'snacks' (sweet and savoury). The following variable should be used when calculating food consumption data of these sub-categories;
    - SUBCATEGORY
- Dietary coding of homemade recipes and some purchased convenience foods complements the method used in the NDNS rolling programme and should be considered by those wishing to calculate food consumption data. In DNSIYC 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 chicken curry containing chicken, Thai curry sauce, and onion would appear in the DNSIYC 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*.

- To estimate absolute food consumption of one specific food type examine the FOODNAME and MAINFOODGROUPDESC variables, whilst also ensuring disaggregation 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 eggs from all sources you would need to include all the specific discrete portions of egg, as well as calculate the percentage of egg within all composite foods such as mayonnaise, quiche, cakes, and pastries.

- 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, please note 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.
- For further details regarding the dietary data methodologies please refer to Appendix D of the main report.

- The food group structure is fully detailed in Appendix M of the main report.

- Breast milk intake was calculated based on the time for each feed, at 13.5g/min with a maximum of 135g per feed for those aged 4 to 7 months 10g/min with a maximum of 100g per feed for those aged 8 to 18 months. The reported daily consumption of breast milk from the food diary has been recorded in DNSIYC as an adjusted volume set to a maximum daily intake of 1200ml. This upper limit of daily breast milk intake was reviewed in line with current literature (refer to Appendix D of the main report).

- Breast milk intake was reported both in the food diary and also in a separate breast milk diary if the mother and child had opted to take part in the stable isotope Protocol 1. The breast milk diary recorded breast milk only for 14 days, whether expressed or from the breast. Breast milk intake from the breast milk diary is reported as the following four variables:

MeanBMIntake_kJ	Breast milk diary data. Mean daily intake of breast milk recorded in breast milk diaries - Energy (kJ)
MeanBMIntake_kcal	Breast milk diary data. Mean daily intake of breast milk recorded in breast milk diaries - Energy (kcals)
MeanBMIntake_g	Breast milk diary data. Mean daily intake of breast milk recorded in breast milk diaries (grams)
NumberOfBMDiaryDays	Breast milk diary data. Number of days recorded in breast milk diary

4.3 Using the physical measurement data

Physical measurements were collected at Stage 1 (interviewer) including infant length, (height), weight and head circumference, and at Stage 2 (clinic) including the same measurements as Stage 1 plus infant triceps and subscapular skinfold thickness. The mother's height and weight were also collected at both stages. If evaluating physical measurement data by age, then the correct child age variable should be selected for appropriate Stage that the child was measured.

Age variable Stage 1 – (DVAGEM)

Age variable Stage 2 – (CLINAGE)

Age and sex-adjusted Z-scores (also known as standard deviation scores (SDS)) have been calculated for physical measurements at each Stage by comparison to the UK-WHO[vi]/WHO 2006[vii] Growth Standards.

4.4 Using stable isotope data

Stable isotope data has been supplied for parents who attended a clinic visit and consented for their child to take part in the stable isotope component, except where one or more of the following criteria was met:

- Drinking less than 50% of the tracer water dose.
- Collecting fewer than three post-dose samples.
- Not recording dates and times of sample collection.
- Not collecting a pre-dose sample.
- The child only consuming small amounts of breast milk from the mother, therefore receiving only a small amount of the stable isotope from the mother (breastfed protocol only).

If one or more of these criteria was true then a negative result code (SIRes) was assigned and the participant excluded from the dataset for that reason. Where there is a missing value, a missing value code was assigned and where there is a valid result SIRes=1.

The child's age at Stage 2 (clinic) should be used when using this data. Stable isotope data are expressed with the assumption that 1ml = 1g of breast milk = 0.67kcals. Assumptions used to calculate breast milk intake using the dose-to-mother method can be found here in the following publication:

http://www-pub.iaea.org/MTCD/publications/PubDetails.asp?pubId=8168

Breast milk intake was calculated using the stable isotope method and was also recorded using a breast milk diary for comparison (refer to section 4.2).

4.5 Using blood data

Blood result data has been supplied for parents who attended a clinic visit (Stage 2) and consented for their child to provide a blood sample. The child's age at Stage 2 (clinic) should be used when using this data.

4.6 Derived variables

For a full list of the derived variables in the datasets please refer to the separate DNSIYC Derived Variable specification. This also contains the syntax used to derive these variables.

4.7 Missing value conventions

-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 mothers/carers only) or because the participant did not participate in a particular element of the survey (e.g. refused a clinic visit).

-8 Don't know

-9 No answer/ Refusal

These conventions have also been applied to most of the derived variables. The DNSIYC Derived Variable specification should be consulted for details.

For a full list of variables on the datasets please see the DNSIYC variable list.

5 Weighting variables.

The Diet and Nutrition Survey of Infants and Young Children (DNSIYC) required a set of weighting factors to adjust the sample for differences in sample selection and response. The weighting factors adjust for differential selection probabilities of boost sample members, non-response to the individual questionnaire, non-response to the clinic visit and non-response to providing blood and urine samples.

Four weights were generated for the DNSIYC data set - these are described in Table 1.

Table 1 Non-response weights available in the data set and their use

Name	Description	Use for
WT_INTERVIEW	Weight for individual and diary data	Any analyses of interview and food diary data
WT_CLINIC	Clinic weight	Any analysis of data collected during the clinic visit, with the exception of blood sample or urine sample data.
WT_BLOOD	Blood sample weight	Any analysis of blood sample data
WT_URINE	Urine sample weight	Any analysis of urine sample data

A brief description of how the weights were generated is given below. Please see Appendix B in the main DNSIYC report for more details.

5.1 Interview weighting factors

An interview weight was generated for the 2,683 parents who responded to the individual interview and completed three or four food diary days for their child. This weighting factor was generated using a combination of logistic regression modelling and calibration. The aim was to reduce bias resulting from sampling error and differential non-response.

The first step was to model response behaviour to the interview using logistic regression. The DNSIYC sample contained two boost samples; a boost of Healthy Start recipients and a boost sample of individuals living in Scotland. The modelling was therefore carried out separately for core sample from England, Wales and Northern

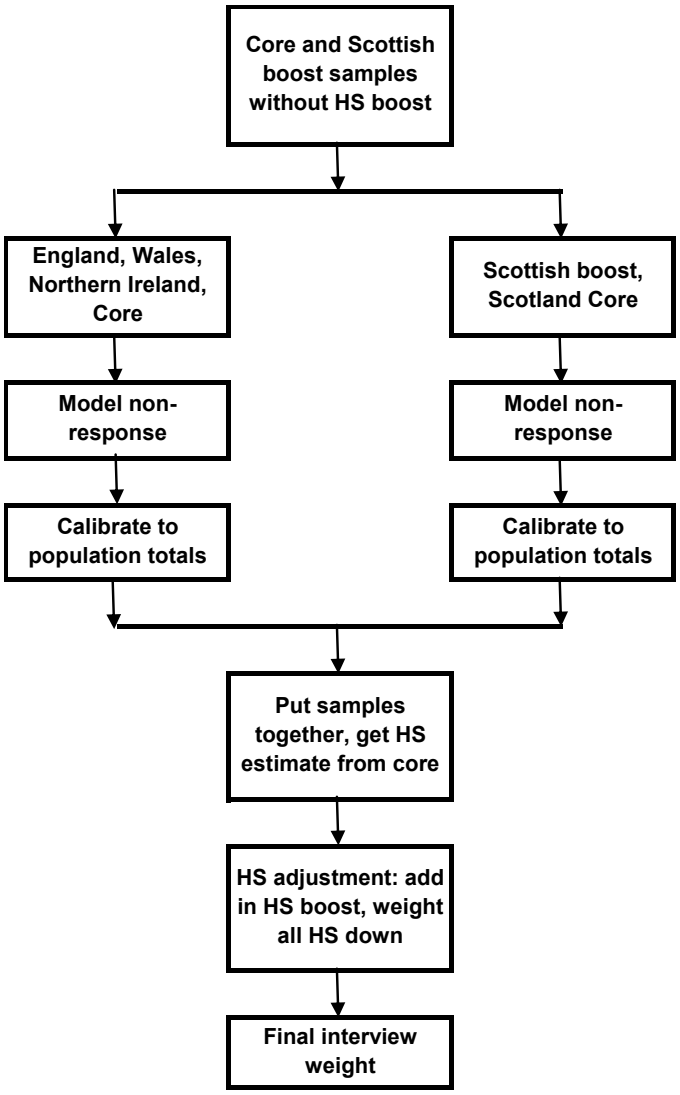
Ireland and for the core and boost sample from Scotland. The Healthy Start recipients were excluded at this step. The weighting factors from the Scottish non-response model fed into both the overall weighting factors and into a separate weighting factor for analysis of the Scottish sample only. Having a separate model for Scotland ensures any estimates for Scotland in the overall sample match those produced for the Scottish-only sample.

The logistic regression model was used to generate predicted probabilities of response. For each participant, this is the predicted probability of them taking part in the interview, given their individual characteristics and the characteristics of their household. The weights were generated as the inverse of the predicted probabilities, respondents with a low predicted response probability received a high non-response weighting factor, increasing their representation in the sample.

The next step was to calibrate the non-response weighting factors generated by the model. Again, this was carried out separately for core England, Wales and Northern Ireland sample and for the combined Scottish core and boost samples. An iterative procedure was used to adjust the non-response weighting factor until the distribution of the (weighted) sample matched that of the population for a set of key variables. This step made the profile of the sample match the population for child's age and gender, age of mother at the time of the child's birth and region. The population figures for calibration were taken from birth counts[1]. The calibrated weighting factors were combined into a single weighting factor and a final adjustment was made to incorporate the Healthy Start boost into the sample. This adjustment made the combined proportion of Healthy Start recipients (core sample plus boost) match that of the weighted core sample (i.e. the best population estimate available). An overview of the process is given in Figure 1.

[1] This means the weights also account for any non-take up of Child Benefit

Figure 1: Overview of methods used for the interview weighting factor



5.2 Clinic weighting factors

All core and Healthy Start parents were invited to attend a clinic where anthropometric measurements could be taken. Scottish boost sample members were not eligible. DNSIYC also required a weighting factor for non-response to the clinic visit.

The clinic weights were generated by modeling non-response to the clinic visit using a logistic regression. The outcome variable was whether or not an eligible individual attended a clinic and the predictor variables were taken from the DNSIYC interview. 2,212 participants were eligible for the clinic visit, 973 attended. The predicted probabilities of response produced by the model were used to generate clinic non-response weighting factors. The weighting factors were scaled, so the mean weighting

factor equaled one and the weighted sample size matched the unweighted sample size. The clinic weights should be used to analyse any data collected during the clinic visit, with the exception of blood sample or urine sample data, these data have separate weights due to the higher levels of non-response to the collection of biological samples.

5.3 Blood sample weighting factors

Participants who attended a clinic were asked if their child would give a blood sample. 513 usable samples were collected from participating children. As before, non-response to the blood sample was modelled using a logistic regression model. Information collected at the interview and clinic visit was used to model response to the blood sample. The non-response weighting factors from the model were combined with the final clinic weighting factors to give the final blood weighting factors. These weighting factors correct for non-response to the individual questionnaire, the clinic visit and the blood sample and should be used for any analysis of blood sample data. The weighting factors were scaled, so the mean weighting factor equalled one and the weighted sample size matched the unweighted sample size.

5.4 Urine sample weighting factors

As well as being asked for a blood sample, the participants who attended a clinic were also asked if a urine sample could be taken from their child. A total of 611 usable samples were collected from participating children. The methods used to create weighting factors to compensate for non-response to the urine sample were the same as those used to create weighting factors for non-response to the blood sample; logistic regression was used to model non-response behaviour using information collected at both the interview and clinic visit. The non-response weighting factors from the model were combined with the final clinic weighting factors to give the final urine weighting factors. These weighting factors correct for non-response to the individual questionnaire, the clinic visit and the urine sample and should be used for any analysis of urine sample data. The weighting factors were scaled, so the mean weighting factor equalled one and the weighted sample size matched the unweighted sample size.

6 DNSIYC report.

Further information about the Diet and Nutrition Survey of Infants and Young Children 2013 and the published report can be found on:

http://transparency.dh.gov.uk/2013/03/13/dnsiyc-2011/

References and endnotes

[ii] Previously called the National Centre for Social Research (NatCen)

[ii] HMRC supplied a sample of names, addresses of child benefit claimants under Paragraph 9, Schedule 5, Tax Credits Act 2002 which gives authority to supply information to other Departments for the purposes of provision of information for health purposes. CB records were used as a sampling frame and selected sample supplied to the Department of Health (DH) for the purpose of DNSIYC. The sample transfer between HMRC and DH was in line with Government security standards and with the agreement of the HMRC Data Guardian from the business area from which data is sourced. Data transfer was in adherence to the strict data transfer rules, and with the correct legal gateways in place.

[iii] Healthy Start is a Government scheme set up to offer a nutritional safety net for pregnant women, new mothers and children under 4 years of age in very low income families, and encourage them to eat a healthier diet. The scheme provides vouchers to put towards the cost of milk, fruit and vegetables or infant formula, and coupons are also given to exchange for free Healthy Start vitamin supplements (see Annexe A).

[iv] Minimum sample per PSU was based on CB counts since all HS children are assumed to also be claiming CB.

[v] The National Diet and Nutrition Survey (NDNS) is a UK survey of the food consumption, nutrient intakes and nutritional status of people aged 1.5 years and older living in private households. The NDNS is currently structured as a 'rolling programme' of continuous fieldwork. Headline results are published annually: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsStatistics/DH_128166

[vi] Royal College of Paediatrics and Child Health/World Health Organisation. The UK_WHO_Growth Charts: Early Years. London: RCPCH, 2009 Available online: http://www.rcpch.ac.uk/growthcharts

[vii] http://www.who.int/childgrowth/standards/en/