

Appendix A

Dietary data collection and processing in National Diet and Nutrition Survey 2019 to 2023

A.1 Introduction

For years 1 to 11 (2008 to 2019) of NDNS, dietary data was collected over 4 consecutive days using a paper food diary with estimated portion sizes. From year 12 (2019), the survey moved to collecting 24-hour dietary recalls on non-consecutive days using Intake24, an online dietary data collection tool. This change was made to provide a more automated data collection method, compatible with new technologies and with potential to improve data quality, reduce costs and enable the survey to scale in the future. The introduction of Intake24 in NDNS represents a significant transition for the survey. A formal evaluation has been undertaken within NDNS to assess the impact of the method change on the survey data. See the [stage 3 evaluation report](#) for details.

A.2 Food Diary

Details of the food diary methodology used up to year 11 can be found in appendix A of the [NDNS Years 9 to 11 report](#). In summary, participants were asked to keep a record in a paper food diary of everything eaten or drunk over 4 consecutive days. Food portion photographs were provided for a small number of frequently consumed foods but the majority of portion sizes were recorded by participants in household measures (for example 2 dessert spoons of baked beans, one Kit Kat (2 fingers) or, for packaged foods, the weight indicated on the packet). Foods and portions recorded in the diary were subsequently entered by trained coders into the dietary assessment system DINO (Diet In Nutrients Out). The linked food composition data came from respective years of NDNS Nutrient Databank (NDB) (see section A.5).

A.3 Intake24

Intake24 is an online automated 24-hour dietary recall tool based on the multiple pass method, primarily designed for self-completion (Rowland and others, 2018; Bradley and others, 2016). Intake24 was developed by Newcastle University, originally with funding from Food Standards Scotland and is licenced under the [Open Government Licence](#). The tool is maintained and developed in collaboration between Cambridge University, Monash University (Australia) and Newcastle University (UK). The version of Intake24 used for NDNS years 12 to 15 (UK Locale, System Version 3, 2019) was provided and adapted by the University of Cambridge, based on the original, with technical advisory input from Newcastle University. A [demo](#) version of the Intake24 system is available.¹

Participants completing Intake24 are asked to record everything they ate and drank the previous day (midnight to midnight). Participants are guided through webpages in the

¹ The demo version of Intake24 is updated regularly and so may not be exactly the same as the version used in years 12 to 15.

following order:

- meal based quick list (captures time, and lists foods)
- detail pass (foods selected, portion details provided)
- meal gap review
- last chance review

Pre-defined meal and snack times (for example breakfast at 08:00) are displayed in chronological order. However, the user can choose to complete entries in any order they wish, change the times and can delete or add more eating occasions. The multiple pass method means that foods and drinks are first entered as free text in the 'quick list' phase of the dietary recall to collect a list of all foods and drinks consumed during the previous day. The participant is then taken through the 'quick list' and for each item reported is asked to select from a list of potential matching food and drink names linked to a food composition code within Intake24. Participants can add, delete or change foods they have reported. Once a food or drink is selected, participants are presented with appropriate portion size estimation options. The majority of foods in Intake24 have a range of portion size photo images that the participant can review and select the image closest to the amount they consumed. If there is no photo, there is the option to report portion size as household measures such as individual items, different spoon sizes, or small, medium or large servings. Some foods offer both photos and household measures. For drinks, there are images of cups and glasses with a slider to indicate volume consumed or photos of bottles and cans.

Intake24 includes a custom spell checker which works to correct both phonetic misspellings and typing errors in the free text. The system also prompts participants for items commonly consumed together (for example milk in tea or coffee, spread on toast), and prompts for missing foods and drinks if there are long time gaps between reported eating occasions or if only a few foods have been reported.

Intake24 includes an embedded pre-coded food list. Participants are asked to select, as far as possible, their foods and drinks from the food list provided in Intake24. If they cannot find an exact match for the food consumed, the tool prompts them to choose the closest matching item where possible. If the participant still cannot find a suitable match for their food they can report this as a missing food. The missing food function asks the participant to provide information about the food for coding at a later stage (undertaken by the research staff based at the MRC Epidemiology Unit, University of Cambridge).

After reporting each eating occasion, participants are asked to record where they had bought or obtained most of the food for that meal or snack in order to differentiate food and drink from the household grocery shop from that obtained from out of home sources such as restaurants and takeaways and 'food on the go'. At the end of each dietary recall, questions prompt participants to record if their intake was typical for that day (and if not, the reason why), if they were following any special diet, the type of oil generally used for cooking (if any) and details of any dietary supplements taken.

For children aged 10 years or younger, a parent or carer is asked to complete the dietary recall with input from the child as appropriate. Children aged 11 years and older could complete the recall themselves with details confirmed with others where necessary. If the dietary recall was completed by a proxy (parent or carer) there may have been occasions when they were unsure of the details of a particular food or an entire meal (such as for a child attending school). At the end of each dietary recall there is a question asking if the recall had been completed by a proxy (that is on the selected participant's behalf) and, if yes, whether there were any difficulties. Any gaps in the recall could be noted here and taken into

account during the data checking process.

At the end of the fourth completed dietary recall, Intake24 provides participants with the option to view on-screen dietary feedback. The feedback provides an averaged daily personal intake for selected foods and nutrients compiled from an average of the participant's 4 dietary recalls combined (see [appendix J](#)). Participants who completed fewer than 4 recalls were not provided with feedback.

A.3.1 Intake24 assistance

During the initial interview (face-to-face or telephone), participants were asked to complete their first recall independently using Intake24. Interviewers then checked with participants if they had any problems completing their recall. In year 12 interviewers were asked to review the first recall before submission if possible. From year 13, this review moved to a check that face-to-face participants did not have any problems completing and submitting their recall. For those interviewed by telephone a follow-up phone call was scheduled. In the majority of cases, participants then went on to complete subsequent recalls independently with no further involvement from interviewers.

For participants who were unable to complete their first recall independently, for example due to internet access issues or lack of confidence with technology, face-to-face recall assistance was offered by interviewers during in-home interviews. If a participant indicated that they were unable to complete the second recall independently, a follow-up visit was arranged whereby the interviewer scheduled a date and time to return to the household and complete the recall with the participant. At the end of each visit the participant indicated whether they could complete the following recall independently (so that some participants may have had assistance with the second recall but not the third and fourth).

For remote interviews, or where internet access was poor in the area, telephone recall assistance was scheduled with research staff at the Epidemiology Unit. Participants were sent a printed food photograph atlas prior to the scheduled phone call to aid the estimation of portion sizes during completion of the recall. The researcher read out the instructions and prompts in Intake24 and the participant provided the information on their food and drink consumption. The researcher entered this into Intake24. At the end of the recall, the researcher arranged an appointment for the next dietary recall if assistance was still required.

If all 4 dietary recalls were completed with assistance, the participant was sent a copy of their dietary feedback via email or post if they requested it.

A.4 Dietary data processing

Foods and portion sizes are assigned codes within Intake24, allowing the system to generate nutrient output at the individual item level for all foods and drinks reported. As with the coding of the paper diaries in NDNS, all foods within Intake24 were assigned nutrient composition data using the NDNS NDB (see section A.5.1). Prior to analysis, the raw Intake24 nutrient output was imported into a bespoke database to facilitate data checks and to assign foods reported as missing to an appropriate food code and portion size, using the original free text search term and missing food details provided by the participant.

A.4.1 Quality checks

Recall data were monitored according to a series of indicators of recall completeness. A number of thresholds were adopted for consideration informed by the research team's

experience with processing and checking dietary data, and published studies using Intake24 and other similar dietary assessment methodology. This included monitoring the number of recalls:

- with 9 or fewer food items
- with 3 or fewer eating or drinking occasions (occasions when a participant reported consuming only a drink without food were included in the count)
- with completion time of under 3 minutes
- where the energy (calorie) intake was very low (less than 400kcal) or very high (more than 4000kcal).

These measures were not designed to classify recalls as complete or incomplete and recalls that met any of these criteria were not necessarily of poorer quality than recalls that did not. The purpose of the measures was to indicate where there may have been issues related to data collection that required further investigation and to provide a set of checks that could be used to assess the tool's performance across the age groups and against other UK surveys using Intake24. Therefore, no recalls were excluded from the dataset on the basis of these checks alone.

In addition to the basic counts, more detailed, manual checks were undertaken on a sub-set of 10% of recalls to identify instances of:

- multiple food items in the participant's search term (for example, 'toast cereal yoghurt') and only one food item was coded
- inconsistencies between the search term and the food code selected, for example searched for chicken stir-fry but selected prawn stir-fry
- 'orphan' foods (a reported food that appeared to have been eaten on its own, for example beef steak when it would commonly be eaten with other foods such as chips or salad)

These checks were carried out on the year 12 data (October 2019 to March 2020) to assess how the new dietary data collection method was working in the survey, and on the first 4 months of data collected in year 13 (October 2020 to February 2021) following the move to remote interviews (see [appendix B](#)). Undertaking and reporting on the above checks was part of the monitoring of Intake24 to identify potential issues and improvements for tool functionality and usability, for example additional food prompt questions, improved portion estimation and clarification around the naming of foods. These changes were implemented at intervals during NDNS years 12 to 15 data collection. No adjustments were made to the dietary data itself as a result of the 10% checks as it was not possible to apply the checks and adjustments systematically to the overall NDNS dataset. It was considered that bias may have been introduced if selected adjustments had been made.

In addition, when all missing foods had been coded, individual recalls were reviewed where the total energy intake was less than 400kcal and where the participant had not stated that they consumed 'less than usual', or that they were on a weight loss diet. As a result of this, 7 recalls were considered 'invalid' and excluded from the final years 12 to 15 dataset.

As with the QA and QC approach taken for NDNS years 1 to 11 dietary data, portion size boxplots were generated by NDNS standard sex and age groups (1.5 to 3 years, 4 to 10 years, 11 to 18 years and 19 years and over) within each food group. Boxplots (otherwise known as box and whisker plots) show the distribution of data and identify any extreme outliers. Extreme outliers were identified from the boxplots as individual data points separate from the box and whiskers since they were more than 3 x IQR (Inter-quartile range: 75th

percentile to 25th percentile) from the nearest quartile for that intake (either the 25th or 75th percentile). These were examined on a case-by-case basis and reviewed in the context of the participant’s overall reported consumption. Extreme portion sizes that were considered to be implausible, and likely to be the result of errors in portion size selection, were adjusted. If there was an issue with the food code itself, this was changed. Adjustments were carried out in the bespoke dietary database by changing the portion code at the individual recall level.

Finally, boxplots were generated by NDNS age group to identify any infeasible or extreme energy and nutrient intake values. Extreme intakes that were considered to be implausible and likely to be the result of errors in portion size selection were adjusted at the individual recall level. Some extreme intakes were due to food composition inaccuracies in the NDB and these were corrected.

As a result of the extreme outlier checks for portion size, energy and nutrients, 0.2% of all food and drink entries in the final data set were adjusted.

A.5 NDNS nutrient databank (NDB)

Dietary data for NDNS is linked to food composition values in the NDNS NDB to produce estimates of nutrient intakes. The NDB is compiled with information from the [UK Composition of Foods Integrated Dataset](#) (CoFID) supplemented by manufacturers’ data gathered through food labels and web information and from the [FSA Food Recipes Database](#). The NDB contains a range of food codes, including manufactured products, homemade recipe dishes and dietary supplements and is maintained by a programme of regular updates.

A.5.1 Rationalisation of the NDNS NDB for Intake24

Intake24 includes an embedded pre-coded food list linked to NDB food codes. Before Intake24 was introduced into NDNS in year 12 (2019), a major review and update of the food codes to be included within Intake24 was undertaken. The aim was to provide a food list for participants that was simple and usable while reflecting, as far as possible, the heterogeneity of the food supply in the UK. This included moving to more generic codes in the NDB to allow a single code to represent a range of similar foods. This review process also provided an opportunity to rationalise the NDB, which had grown unsystematically over many years to contain over 5,500 foods and drinks, to enable more efficient management of updates in future years of NDNS.

Details of the NDB review and rationalisation have been published elsewhere (Amoutzopoulos and others, 2022). In summary, dietary intake data from the first 10 years (2008 to 2018) of the survey were examined to identify reported consumption rates (frequency and amount) for all 6098 foods included in the 2018 NDB. During the review, other factors were also taken into consideration including whether the product was still available on the market and variations between different brands and food preparations. All foods were then allocated to one of the following categories:

RETAIN	Foods that are frequently consumed and/or foods which are not necessarily reported frequently but are essential as they are not represented by other retained foods, for example, fortified or ‘free-from’ foods or foods used as ingredients in recipes
EXCLUDE	Foods that have been coded only a few times or have not been used or are no longer available

REPRESENT	Foods that could be represented by another food: similar foods that could be represented by a retained code, for example, canned crab represented by boiled crab
-----------	--

Overall, the numbers of foods assigned to each category were:

- **RETAIN = 2331 (39%)**
- **EXCLUDE = 1541 (26%)**
- **REPRESENT = 2061 (35%)**

Following this approach, those food codes in the EXCLUDE and REPRESENT categories were decommissioned in the NDB (61% of codes). A number of new generic food codes (n=150) were created mainly for sandwiches and salads which during years 1 to 11 were coded as individual components with the paper diary method. This resulted in a final database of 2,481 foods with linked nutrient composition data. Following the rationalisation process, the updated NDB (version UK_NDB_pre1) was incorporated into Intake24 for the start of year 12 of NDNS in October 2019.

A.5.2 Updates to the NDB for years 12 to 15

For NDNS, a programme of ongoing updates and revisions are made to the NDB so that the databank is up-to-date and as far as possible, reflects the nutrient composition of the food supply for each survey year reported. In years 1 to 11 the NDB was updated and applied to the dietary data on an annual basis. With the move to Intake24 the programme of updates has moved to biennial. However, due to the work to update the NDB for the transition to Intake24, and impact of COVID-19 on survey fieldwork, there were 2 updates, initially for year 12 and then for years 13 to 15.

Updating of the NDB may include the addition of new foods as well as revision of nutrient composition of existing foods, either at individual or food group level following a programme of reanalysis, or to take account of reformulation reported by manufacturers and changes in fortification practices. Therefore, the same foods may have a different composition for some nutrients in one year of NDNS compared with another year. Table A.1 lists the key updates applied over years 12 to 15. It should be noted that updates applied in one year apply to future years unless further changes are made.

Table A.1 Updates to the NDB

NDB name	Fieldwork year	Analytical data updates	Reviews of manufacturers' or label data by food group	Other
UK_NDB_1.1	Year 12 (2019 to 2020)	Selected micronutrients in cows' milk (riboflavin,	Bacon and ham ^a Baked beans ^a Bread and rolls ^a Breakfast cereals	

		vitamin B12, iodine)	Butter and fat spreads Cheddar cheese ^a Savoury sauces, pickles, gravies and condiments ^a Crisps and savoury snacks Pizza Sausages Soft drinks, not low calorie or diet Soup ^a Biscuits Yoghurt, fromage frais and other dairy desserts	
UK_NDB_ 2	Years 13 to 15 (2020 to 2023)		Bread and rolls ^a Breakfast cereals Buns, cakes and pastries Burgers and kebabs (takeaway burgers only) Commercial toddlers foods Dietary supplements Other milk ^d Pork and dishes ^b Ready meals ^c Soft drinks, not low calorie or diet Soup ^a Yoghurt, fromage frais and other dairy desserts	Micronutrients in wheat flour (non-wholemeal)

^a Only sodium values updated.

^b Not manufactured pork products or dishes.

^c Not a specific food group, food codes were identified across all food groups based on name.

^d Updates made to some plant based drinks and lactose free milks, and flavoured milk drinks.

A.5.3 Update to energy value for fibre

The NDB uses the Atwater system to assign energy values to foods based on 4.0 kcal per gram for protein, 9.0 kcal per gram for fat, 3.75 kcal per gram for available carbohydrate (expressed as monosaccharide) and 7.0 kcal per gram for alcohol (Food Standards Agency 2002). Historically, in the Atwater system, dietary fibre (considered unavailable carbohydrate) did not contribute to energy intakes. Analysis of published data on energy availability from fibre (measured using AOAC method) recommended that the system be modified by using 2.0 kcal per gram for fibre (FDA, 2016; FAO2003). Therefore this factor was applied retrospectively to the dietary data for years 12 to 15. This change has been made following the [UK regulation allowing for voluntary fibre content declaration on the nutrition label](#) in alignment with [EU Regulation No. 1169/2011](#) which also include the contribution of fibre to energy content using this factor 2.0 kcal per gram. In future, energy from fibre will be calculated within the NDB.

A.6 Disaggregation of composite dishes

In order to determine total intakes of fruit, vegetables, meats and fish, all foods containing any of these food types have been systematically disaggregated into their components (Fitt and others 2010). To enable monitoring of intakes of these foods against government guidelines, it is important to quantify amounts eaten as part of composite dishes, as well as their discrete portions, to provide more accurate estimates of total amounts consumed at an individual food level. For example, carrots may be eaten as an accompaniment to a main meal, but they may also be consumed as an ingredient within a stew, together with additional vegetables such as onions and celery. Disaggregated data allows the automated estimation of total intakes of fruit, vegetables, meat and fish in NDNS, including the contribution from composite dishes.

A.6.1 Calculation of 5 A Day using disaggregated data

One of the estimates required for NDNS reporting is the intake of fruit and vegetables, specifically how participants' intake compare to the government recommendations to eat [5 A Day](#). Information on the fruit and vegetable content of each disaggregated food code was used to calculate the number of portions of fruit and vegetables consumed for each participant aged 11 years and over, using a portion weight of 80 grams (150 grams for fruit juice). In line with the 5 A Day criteria, fruit juice and smoothies combined and pulses (including baked beans) were included in the calculation up to a maximum of one portion per day each. The calculations included the fruit and vegetable content of foods such as meat, fish, pasta, rice and egg dishes, pizza, breakfast cereals, yogurts and dairy desserts and soups. Puddings and fruit pies were also included as these can contain significant amounts of fruit. Other foods that fall into the 'high fat or high sugars' segment of the [Eatwell Guide](#) were excluded from the calculation on the grounds that healthy eating advice is to reduce consumption of foods in this group. Therefore, the fruit and vegetable content of the following NDNS food groups was excluded from the estimates:

- soft drinks (35.1, 35.2, 35.3 and 36.1, 36.2, 36.3)
- confectionery (33.1, 33.3)
- buns, pastries, cakes (including fruit cake) and biscuits (5.1, 5.2, 5.3, 6.1)
- sugar, preserves (including jam) and sweet spreads (32.1, 32.2, 32.3)
- savoury snacks (29.1, 29.2)

- ice cream (11.1, 11.2, 11.3)

Main and subsidiary NDNS food groups are listed in [appendix R](#). The calculation of 5 A Day portions was carried out as follows:

- daily consumption of fruit juice (from 100% fruit juice) and smoothies was capped at 150g (1 portion)
- daily consumption of baked beans and other pulses was limited to 80g (1 portion)
- daily consumption (g) of dried fruit was multiplied by 3 to account for effects of drying²
- daily consumption (g) of tomato puree was multiplied by 5 to account for effects of concentration³
- total weight (g) of fruit and vegetables (including the weights of baked beans and other pulses, dried fruit and tomato puree, modified as above) was divided by 80 to arrive at the number of fruit and vegetable portions
- total weight of fruit juice (from 100% fruit juice) and smoothies was divided by 150 to arrive at the number of portions of fruit juice and smoothies
- the number of fruit and vegetable portions and the number of portions of fruit juice and smoothies were summed to give total 5 A Day portions consumed

5 A Day portions were not calculated for children aged 10 years and younger as there is not a quantified portion size recommendation for younger children.

References

Amoutzopoulos, B.; Steer, T.; Roberts, C.; Collins, D.; Trigg, K.; Barratt, R.; Abraham, S.; Cole, D.J.; Mulligan, A.; Foreman, J.; Farooq, A.; Page, P. Rationalisation of the UK Nutrient Databank for Incorporation in a Web-Based Dietary Recall for Implementation in the UK National Diet and Nutrition Survey Rolling Programme. *Nutrients* 2022, *14*, 4551.

<https://doi.org/10.3390/nu14214551>

Bradley J, Simpson E, Poliakov I and others. [Comparison of INTAKE24 \(an Online 24-h Dietary Recall Tool\) with Interviewer-Led 24-h Recall in 11-24 Year-Old](#). *Nutrients* 2016. 8(6): 358 (viewed 16 August 2021)

Fitt E, Mak TN, Stephen AM, Prynne C, Roberts C, Swan G and 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

Food and Drug Administration. [Revision of the Nutrition and Supplement Facts Label](#). US Federal Register 2016. Vol.81. No.103

² Dried fruit (g) is multiplied by 3 to ensure that it is comparable to non-dried fruit on the basis of their respective micronutrient contents

³ Tomato puree (g) is multiplied by 5 to ensure that it is comparable to canned tomatoes on the basis of their respective carotene contents. For example, 10g tomato puree and 50g canned tomatoes would provide approximately the same amount of carotenes (178µg and 181µg respectively)

National Diet and Nutrition Survey: 2019 to 2023 (years 12 to 15)

Food and Agricultural Organization of the United Nations. Food Energy – [Food energy – methods of analysis and conversion factors. Report of a Technical Workshop 2003](#). Rome: FAO Food and Nutrition Paper no. 77.

Food Standards Agency (2002) McCance and Widdowson's The Composition of Foods, Sixth summary edition. Cambridge: Royal Society of Chemistry

Rowland, MK, Adamson AJ, Poliakov I, Bradley J, Simpson E, Olivier P, Foster E. [Field Testing of the Use of Intake24 - An Online 24-Hour Dietary Recall system](#). Nutrients 2018. 10: 1690 (viewed 16 August 2021)