



Use of 24-Hour Recalls to Assess Adherence to Cancer Specific Dietary Guidelines: Experiences from the Advancing Survivorship Cancer Outcomes Trial (ASCOT)

Victoria Ireland, Helen Croker, Sara Esser, Phillippa Lally, Rebecca J Beeken, Abigail Fisher & Rana Conway

To cite this article: Victoria Ireland, Helen Croker, Sara Esser, Phillippa Lally, Rebecca J Beeken, Abigail Fisher & Rana Conway (2025) Use of 24-Hour Recalls to Assess Adherence to Cancer Specific Dietary Guidelines: Experiences from the Advancing Survivorship Cancer Outcomes Trial (ASCOT), *Nutrition and Cancer*, 77:7-8, 871-880, DOI: [10.1080/01635581.2025.2514783](https://doi.org/10.1080/01635581.2025.2514783)

To link to this article: <https://doi.org/10.1080/01635581.2025.2514783>



© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.



Published online: 06 Jun 2025.



Submit your article to this journal [↗](#)



Article views: 539



View related articles [↗](#)



View Crossmark data [↗](#)

Use of 24-Hour Recalls to Assess Adherence to Cancer Specific Dietary Guidelines: Experiences from the Advancing Survivorship Cancer Outcomes Trial (ASCOT)

Victoria Ireland^a, Helen Croker^{a,b}, Sara Esser^a, Phillippa Lally^{a,c}, Rebecca J Beeken^{a,d}, Abigail Fisher^a and Rana Conway^a

^aResearch Department of Behavioural Science and Health, University College London, London, UK; ^bWorld Cancer Research Fund International, London, UK; ^cSchool of Psychology, University of Surrey, Guildford, UK; ^dLeeds Institute of Health Sciences, University of Leeds, Leeds, UK

ABSTRACT

People living with and beyond cancer (LWBC) are advised to follow World Cancer Research Fund (WCRF) dietary guidelines. However, there is no established methodology to assess adherence. This study aimed to: (i) develop methodology to process dietary recalls into a format comparable to WCRF guidelines and (ii) evaluate the impact of additional data processing on estimates of dietary intake, in people LWBC. Advancing Survivorship Cancer Outcomes Trial (ASCOT) participants completed two 24-h dietary recalls at four timepoints using myfood24. Five WCRF recommendations (limiting consumption of energy dense foods, red meat and processed meat, and increasing consumption of fruit and vegetables and wholegrains and pulses) were operationalized (e.g. ≤ 500 g red meat per week). Quality control checks indicated the need for additional processing, including changing portion sizes and choosing alternative items from myfood24 to improve accuracy. Compared to myfood24 output, the processed dietary data indicated lower intake of fruit and vegetables, and higher intake of NSP and AOAC fiber (all p s < 0.001). Developing methodology to allow assessment of 24-h dietary recall data against WCRF guidelines was possible and necessary but resource intensive. Additional data processing impacted estimates of the key foods and nutrients consumed by trial participants in a meaningful way.

ARTICLE HISTORY

Received 11 March 2025
Accepted 28 May 2025

Introduction

Through earlier diagnosis and improvements in treatments, more people are surviving cancer, leading to a larger population in the UK living with and beyond cancer (LWBC) (1,2). According to the World Cancer Research Fund (WCRF), people LWBC should adhere to the WCRF's lifestyle recommendations for cancer prevention if they are able to (3). To reduce diet-related risk, the WCRF advises individuals LWBC to limit consumption of red and processed meat, processed foods high in fat, sugar and alcohol; and increase consumption of fruit, vegetables, wholegrains and beans (3). Many adults in the UK fail to meet national dietary guidelines and consume more free sugar and saturated fat and less fruit, vegetables and fiber than recommended and there is little evidence that a cancer diagnosis motivates health protective changes (4–9).

The WCRF has identified a need for individuals LWBC to receive guidance on nutrition, but it is unclear how to provide behavior change support that is simple yet effective (3). The process of developing and evaluating interventions to support dietary improvement inevitably requires an assessment of dietary intake in relation to WCRF guidelines. The use of 24-h dietary recalls is now the most common method of collecting detailed dietary intake data, as participant burden is relatively low compared to diet diaries. When recalls are self-administered *via* online systems this also reduces researcher burden and increases cost effectiveness, although issues with data quality remain, for example due to problems with memory, portion estimation, and social desirability bias (10–15). As WCRF guidelines extend beyond nutrients (e.g. fat) to foods (e.g. red meat), depending on the data output from online recall software, additional processing may be required before dietary data can be used to assess guideline adherence. For

CONTACT Rana Conway ✉ r.conway@ucl.ac.uk  Research Department of Behavioural Science and Health, University College London, Gower Street, London WC1E 6BT, UK.

This article has been corrected with minor changes. These changes do not impact the academic content of the article.

© 2025 The Author(s). Published with license by Taylor & Francis Group, LLC.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

example, if the output does not include “red meat,” assessing adherence to the recommendation to limit consumption becomes complex. Other considerations include whether dietary variables assessed by the chosen software are consistent with WCRF recommendations, for example, whether fruit juice is counted as “fruit.” The number of items in a database with missing nutrient composition data also needs to be considered. As a full nutrient profile is generated irrespective of human error or software problems, it is essential that research studies include quality control checks to ensure data integrity (16). In practice, a pragmatic approach may be required to balance a desire for accuracy with available resource for checks, taking into account the volume of data being collected.

The Advancing Survivorship Cancer Outcomes Trial (ASCOT) is a brief intervention using habit-based advice to improve people LWBC’s adherence to WCRF guidelines for diet, physical activity, alcohol and smoking (17). A randomized controlled trial was conducted to determine the effectiveness of the ASCOT intervention, which necessitated addressing the challenges set out above. Developing and performing additional data processing procedures to assess adherence to WCRF guidance presented a challenge which would be resource intensive if implemented at each stage of data collection. As it was unclear whether additional processing would result in a measurable impact on estimates of dietary intake it was considered necessary to carry out an evaluation when collecting baseline data. The challenges experienced for ASCOT dietary assessment are likely to be encountered by others evaluating adherence to WCRF guidelines and are therefore shared here.

The aims of this study were to (i) develop methodology to process 24-h dietary recalls obtained using myfood24 to facilitate estimation of adherence to five WCRF dietary guidelines and (ii) assess the impact of performing additional data processing on estimates of dietary intake.

Materials and Methods

Participants and Data Collection

Participants were adults aged 18 years and older taking part in ASCOT. Details of the trial can be found elsewhere (17). In brief, participants were recruited from ten hospital sites in London and Essex *via* a survey. Hospital sites were asked to send a questionnaire to patients diagnosed with breast, prostate, or colorectal cancer between 2012 and 2015. As hospitals did not always accurately identify patients, some

participants were diagnosed outside of these dates. Respondents interested in learning more about a trial of a lifestyle intervention who met eligibility criteria were asked to provide informed consent and complete additional measures of their diet and physical activity. Ethical approval was obtained through the National Research Ethics Service Committee South Central—Oxford B (reference number 14/SC/1369).

Participants were asked to complete two 24-h dietary recalls at four timepoints: baseline, 3 months, 6 months, and 2 years post-intervention. The myfood24 software was chosen, as it was the first online 24-h recall tool developed and tested with UK adults (18). It allows participants to self-administer 24-h recalls *via* the Internet as it enables participants to search for items and estimate portion sizes using pictures or standard household measures or weight (18).

On the day of the recall, participants were sent an email with a personalized link to the myfood24 website. Researchers completed recalls over the phone if a participant did not use email or struggled with online completion (19,20). Recalls completed online by participants were not routinely reviewed but if participants reported difficulties or researchers noticed unusual entries when downloading data, participants were contacted to resolve queries (19).

Assessing Adherence to WCRF Guidance

Adherence to dietary recommendations was assessed directly using WCRF criteria where possible (e.g. ≤ 500 g red meat/week) or national guidelines when WCRF guidance was not quantitative (e.g. reduce intake of energy-dense foods) (Table 1).

As data collection began in 2015, recommendations were taken from the 2007 WCRF report which was available at the time (21). This included assessing intake of non-milk extrinsic sugars (NMES), which are defined as sugars that are not incorporated into the cellular structure of foods, including those added to foods or that naturally occur, for example in fruit juice (5,6). Naturally occurring sugars in milk and milk products, despite being extrinsic, are not included (5,6). As alcohol intake is often irregularly consumed, intake was assessed *via* a questionnaire rather than using the 24-h recall so is not reported here (17,21).

Development of Dietary Quality Control Checks

Quality control checks involved identifying extreme outliers for portion sizes and daily energy and nutrient intake to identify possible errors in data entry or

Table 1. Operationalization of World Cancer Research Fund recommendations for ASCOT.

WCRF recommendation (21)	Criteria for assessing adherence to recommendation
Limit consumption of energy-dense foods and fast foods and avoid sugary drinks.	Fat intake: 33% of energy intake or less (6). Sugar intake: non-milk extrinsic sugars (NMES) make up 10% of energy intake or less (6) or free sugars makes up 5% of energy intake or less ¹ (5).
Eat at least five portions of fruit and non-starchy vegetables a day.	Fruit and vegetables (F&V): 5 or more portions (400g or more) a day (21).
Eat relatively unprocessed grains and/or pulses at each meal.	Fiber (as a measure of unprocessed grains and pulses): at least 18g non-starch polysaccharides (NSP) per day (6) or 30g AOAC fiber or more per day ¹ (3).
Limit intake of red meat.	Red meat: 500g per week or less (21) (71g per day or less).
Avoid processed meat.	Processed meat: 0g per day (21).

Abbreviation: g: grams.

¹The national recommendations for sugar and fiber intake changed during the study so participants were assessed against both recommendations, therefore both values had to be obtained from the data.

the nutrient database. The nutrients selected for checks were those needed to assess compliance with WCRF guidance. Changes to foods or portion sizes were made where two researchers with expertise in nutrition or dietetics agreed that it was implausible or extremely unlikely that the entry was a true reflection of consumption. Whilst alcohol intake was not assessed *via* the recalls, total alcohol intake was checked because portion size errors for alcoholic drinks could potentially impact other nutrients such as energy intake. Cutoffs for checks were developed pragmatically when exploring baseline data to ensure major errors were identified whilst keeping the number of recalls requiring a check manageable for researchers.

Assessment of the Impact of Additional Nutrient Data Processing

Data from the first recall collected at baseline were used to assess the impact of additional data processing on estimates of dietary intake. This allowed the team to assess whether additional processing was necessary and justified. Paired sample *t*-tests were used to compare myfood24 output with data generated after additional processing.

Results

A total of 1348 patients diagnosed with breast, prostate, or colorectal cancer were recruited to the trial between August 2015 and July 2019 and 1225 of these participants completed the first 24-h recall when requested.

Data Processing Procedure

The final procedure developed for creating the variables required to assess adherence to WCRF guidance from myfood24 output is outlined in Table 2. The process involved using two myfood24 outputs: (i) the list of individual food items and their nutrient values; (ii) estimates of total energy and nutrient intake for the day.

Initial examination of baseline data identified several issues, outlined below, which prompted the decision to develop standard operating procedures to ensure processing was consistent across researchers and timepoints.

Fat and Sugar

Fat intake (% energy) was calculated directly from myfood24 output for fat and energy assuming 9 kilocalories per gram of fat (see Table 2) (23).

Initially, to determine NMES and free sugars, food items within each recall were individually assigned values, but this was unfeasible for the large volume of data collected. Therefore, a method of approximating NMES was developed using total sugar and lactose values provided in myfood24 output. As NMES includes 50% of the sugar found in dried, stewed or canned fruit, dried fruit was estimated by summing the sugar values of each dried fruit item in a recall, and then halved (5). Stewed and canned fruit were not commonly consumed, and so the pragmatic decision was made not to adjust NMES estimates for these. Additional adjustment was necessary to subtract intrinsic sugars from fruit and vegetables. Using data from 200 recalls, sugar (g) from all fruit and vegetable items (F&V sugar) was calculated manually, and a regression equation was then derived as follows:

$$\begin{aligned} \text{F \& V sugar} = & (0.093 * (\text{weight fruit} - \text{weight dried fruit})) \\ & + (0.039 * \text{weight veg}) \end{aligned}$$

The amount of intrinsic sugar coming from other sources, for example, the sugar found in wholegrains was considered negligible and therefore not estimated. Free sugar was estimated similarly, but as the definition does not include any of the sugar in dried, stewed or canned fruit this was subtracted (see Table 2) (5). The metabolizable energy conversion factor of 3.75 kilocalories per gram was used to calculate NMES and free sugars as a percentage of energy intake (23).

Fruit and Vegetables

Several corrections were applied to transform the values for fruit (g/day) and vegetables (g/day) in the

Table 2. Data processing methodology.

Dietary component	Operationalized recommendation	Variable provided in myfood24 output	Additional variables calculated	Process for estimating intake of dietary component
Fat and total sugar	$\leq 33\%$ E NMES $\leq 10\%$ E Free sugar $\leq 5\%$ E	fat (g), E(kcal) total sugar (g), lactose (g), E (kcal)	None "Dried fruit sugar" (g) = sum of total sugar (g) in each occurrence of dried fruit ^a "F&V sugar" (g), estimated using a regression equation ^b	$\text{Fat (\%E)} = (\text{fat} \times 9 / \text{E}) \times 100$ $\text{NMES (g)} = \text{total sugar} - \text{lactose} - (\text{dried fruit sugar} \times 0.5) - \text{F\&V sugar}$ then $\text{NMES (\%E)} = (\text{NMES} \times 3.75 / \text{E}) \times 100$ $\text{Free sugar (g)} = \text{total sugar} - \text{lactose} - \text{dried fruit sugar} - \text{F\&V sugar}$ $\text{Free sugar (\% E)} = (\text{free sugar} \times 3.75 / \text{E}) \times 100$
Fruit and vegetables	≥ 5 portions per day. Includes fresh, canned, frozen and F&V content of composite foods. Excludes potatoes. 80g = one portion except: ≥ 80 g pulses and beans = one portion, ≥ 150 ml juice ^b = one portion, 30g dried fruit = one portion (22)	fruit (g), vegetables (g), baked beans (g), fruit juice (ml), smoothies (ml), dried fruit (g)	"Adjusted baked beans" (g) = baked beans (g) * 0.85 (sauce). If > 80 g corrected to 80g "Adjusted juice ^b " (g) = juice ^b ≥ 150 ml corrected to 80g. juice ^b < 150 ml corrected using equation: $\text{volume} / 150 \times 80$ "Adjusted dried fruit" (g) = dried fruit (g) * 3	$\text{Fruit and vegetables eligible for portions (g)} = \text{fruit} + \text{vegetables} + \text{adjusted baked beans} - \text{juice}^b + \text{adjusted juice}^b - \text{dried fruit} + \text{adjusted dried fruit}$ $\text{Portions of fruit and vegetables} = \text{fruit and vegetables eligible for portions} / 80$
Fiber	≥ 18 g NSP per day ≥ 30 g AOAC per day	NSP (g) for some food items AOAC (g) for some food items	Foods with 0g NSP: "NSP approx." (g) = AOAC/1.3 Foods with 0g AOAC: "AOAC approx." (g) = NSP*1.3 Fiber-containing foods showing 0g NSP and 0g AOAC replaced with similar food.	$\text{Total NSP (g)} = \text{sum of NSP values} + \text{sum of NSP approx.}$ $\text{Total AOAC (g)} = \text{sum of AOAC values} + \text{sum of AOAC approx}$
Red meat	≤ 71 g per day	None appropriate	"Red meat approx." (g) assigned to each item in recall ^{a,c}	$\text{Total red meat (g)} = \text{sum of red meat approx}$
Processed meat	0g per day	None appropriate	"Processed meat approx." (g) assigned to each item in recall ^{a,c}	$\text{Total processed meat (g)} = \text{sum of processed meat approx.}$

Abbreviations: E: energy; g: grams; kcal: kilocalories; NMES: non-milk extrinsic sugars; F&V: fruit and vegetables; ml: millilitres; NSP: non-starch polysaccharides.

^aCalculated or identified manually from the list of individual items in recall.

^bIncludes fruit juices and smoothies.

^cSee methods for details.

myfood24 output to assess adherence with guidance to consume at least five portions of fruit and vegetables per day (5-a-day) (21). In the 5-a-day guidelines, one portion of vegetables is 80 g, and pulses and beans count as a maximum of one portion per day (22). As pulses and beans were not included in the myfood24 estimation of vegetables (g/day), entries of pulses and beans had to be identified manually in each recall. The most commonly eaten pulses or beans were baked beans and intakes of others were infrequent, so due to time constraints only baked beans were included in estimated vegetable intake. Each entry of baked beans was identified, adjusted to account for sauce (85% beans: 15% sauce) and portions > 80 g were reduced to 80g to align with a single portion of vegetables (see Table 2).

In 5-a-day guidance, ≥ 150 ml of fruit juice or smoothie is considered one portion of fruit (22). Most but not all fruit juices and smoothies were included in myfood24 estimation of fruit (g/day) so each

occurrence of fruit juice and smoothies in each recall was identified to enable a correction to be applied. Volumes ≥ 150 ml were assigned a weight of 80g to align with a single portion of fruit, and volumes smaller than 150ml were adjusted to correspond with the specified 80g per portion (see Table 2), for example, 75 ml juice was converted to 40g fruit (half a portion). Finally, all entries of dried fruit in our recalls were identified and portion weights were multiplied by three, in line with Public Health England methodology (22,24).

Fiber

Initial analysis of baseline data, in preparation for targeted intervention delivery, identified lower than expected fiber intakes. After closer scrutiny of individual recalls, it was found that some common fiber-containing foods had been assigned a value of 0g for AOAC fiber, NSP fiber, or both. If a value was

available for either AOAC or NSP, then a conversion factor of 1.3 was applied to estimate the missing AOAC and NSP value (5). Commonly consumed foods which were known to contain fiber but where both AOAC and NSP fiber values appeared as 0g, including oatcakes, muesli and frozen mixed vegetables, were identified in each recall and replaced with similar foods so that a fiber value would be available.

Red and Processed Meat

Red and processed meat were not included in myfood24 output, therefore these were estimated from the food items listed in each 24-h recall. WCRF definitions of red and processed meat were used, alongside ingredient lists for composite dishes, to create mutually exclusive lists for red meat and processed meat, for example meatballs were counted as red meat and the meat component of pork pies as processed meat (25). A standard operating procedure was developed to ensure consistency when estimating meat content of composite foods and for cuts of meat recorded including bone.

Additional Data Cleaning Required

The criteria for quality control checks on individual food items or recall totals are described in Table 3. When examining data to develop these cutoffs it was discovered that some extremely high intakes were due to errors in the myfood24 nutrient database, rather than participants selecting implausible portion sizes. For example, a commonly selected entry for cranberry juice contained 34g sugar/100g according to myfood24 but 9.2g sugar/100g according to label information, therefore all entries for this juice were identified and replaced with an alternative.

Portion options, including the photos presented to participants, also appeared to be responsible for some very large portions identified in the dataset. Example 1: participants selecting “fresh egg spaghetti” were shown images of cooked spaghetti, but myfood24 software treated the portion weight as uncooked and

more than doubled it. Example 2: participants selecting some varieties of bacon were presented with information that 100g was an “average portion,” the interface did not inform participants that this was four rashers, and many participants appeared to assume it was a single rasher and select three of four multiples of this, meaning 400g portions of bacon were not uncommon in recalls. In other instances, portion options “as bought” were shown without further explanation as to whether this meant a single biscuit or a packet, or an individual or large yogurt pot. When these and similar problems were identified, entries for these items in other recalls were then checked.

It was also noticed some food descriptions may have been unclear to participants. Example 1: several recalls included large portions of concentrated orange juice, without the addition of water. Concentrated orange juice is not widely available or commonly consumed in the UK but when searching for “orange juice” in the myfood24 interface, the first results were “orange juice, freshly squeezed” followed by “orange juice, concentrated, unsweetened.” Given the tendency for participants to choose items closer to the top of lists when using recall software (26), it was considered likely that concentrated juice had been selected in place of juice reconstituted from concentrated juice, which is widely available in the UK and labeled “from concentrate” on the front of pack. Recalls including ≥100g concentrated juice without an entry for water, were therefore identified and replaced with reconstituted juice. Example 2: an item called “Cooked sandwich ham” with portion options of 14g and 140g was sometimes selected as the only item consumed for lunch. Entries such as this for 140g “Cooked sandwich ham” were replaced with a ham sandwich incorporating both slices of ham and bread.

Detailed descriptions of data processing procedures are provided in Standard Operating Procedures (<https://osf.io/8wkf6/files/osfstorage/682724b92646793920c4c015>, <https://osf.io/8wkf6/files/osfstorage/682724bc33e47add47cea0f7>, <https://osf.io/8wkf6/files/osfstorage/682724b9d3384237719c4154>, <https://osf.io/8wkf6/files/osfstorage/682724c533e47add47cea0f9>).

Table 3. Criteria for quality control checks.

Criteria for checks on a food entry	Criteria for checks on daily totals
Portion size of food (excluding drinks) <3g or >400g	Energy <500kcal or >3500kcal
AOAC fiber content of food portion consumed >15g	Fruit and vegetables > 15 portions
	AOAC fiber > 40g
	Free sugar > 125g
	Fat > 150g
	Alcohol > 80g (10 units)
	Red meat > 200g
	Processed meat > 150g

Abbreviations: kcal: kilocalories; g: grams.

Comparison of myfood24 Output with ASCOT Processed Data

Mean estimates of intake were significantly different after processing compared with original myfood24 output, see Table 4. For example, estimated fruit intake increased in 8.4% and decreased in 17.2% of

Table 4. Comparison of dietary data from myfood24 output and ASCOT dietary data after processing and cleaning (N=1225).

Nutrient	myfood24 output (g) Mean (SD)	Cleaned data (g) Mean (SD)	P (paired t-test)	Recalls changed N (%)	Recalls where estimated intake increased after processing			Recalls where estimated intake decreased after processing		
					N (%)	increase (g) Mean (SD)	Range (g)	N (%)	decrease (g) Mean (SD)	Range (g)
Sugar ¹	84.05 (38.57)	44.63 (30.57)	<0.001	1225 (100)	0 (0)	0	0	1225 (100)	39.41 (21.42)	1.42–153.56
Fruit	220.61 (192.39)	209.44 (180.77)	<0.001	314 (25.6)	103 (8.41)	70.83 (47.50)	2.00–240.00	211 (17.2)	99.40 (129.78)	0.93–920.00
Vegetables	191.50 (144.39)	189.83 (143.37)	<0.001	34 (2.78)	11 (0.90)	8.02 (2.72)	1.80–12.00	23 (1.9)	92.90 (78.19)	7.13–277.00
NSP	14.17 (6.12)	14.64 (6.26)	<0.001	558 (45.56)	558 (45.6)	1.03 (1.71)	0.00–24.76	0	0	0
AOAC	8.17 (5.38)	19.64 (8.52)	<0.001	1213 (99.02)	1213 (99.02)	11.57 (6.85)	0.13–47.75	0	0	0

Abbreviations: g: grams; NSP: non-starch polysaccharides.

¹The myfood24-generated sugar variable was compared to the estimated free sugar variable.

recalls, with mean (SD) changes of 70.8 (47.5)g and 99.4 (129.8)g, respectively.

Discussion

Additional data processing steps were developed to facilitate assessment of adherence to WCRF dietary guidelines using myfood24. The development and implementation of these steps was resource intensive but necessary – comparison of dietary data before and after additional processing found considerable differences. Key learnings from ASCOT demonstrate the need to closely interrogate data generated by automated recall software and highlight specific issues research teams may wish to examine when conducting dietary surveys.

Whilst it has been proposed that data collection *via* self-administered dietary recalls reduces researcher burden during data collection, additional time may be required to develop and implement methodology to process data and assess adherence to specific dietary guidelines, such as WCRF recommendations (10–13). Comparing the myfood24 output with cleaned data shows processing resulted in a mean increase or decrease of around a portion of fruit for a quarter of participants, which is a meaningful quantity when assessing intervention impact. This was primarily due to myfood24 counting fruit juice as multiple fruit portions, while dietary guidelines limit juice to one portion, regardless of quantity consumed due to its free sugar content and lack of fiber. Without additional data processing fruit intake would therefore have been overestimated for those drinking fruit juice or smoothies. Similarly, fruit intake would have been underestimated for participants consuming dried fruit, if a correction factor had not been applied to calculate the equivalent non-dried fruit weight in line with guidance (22).

There is great potential to reduce researcher burden involved in dietary data cleaning with

advancements in data processing and Artificial Intelligence (AI) technologies, for example by flagging unusual values in a dataset and learning from researchers' decisions. Much of the processing described here to generate estimates of fruit and vegetable intake that is comparable to NHS five-a-day advice could be incorporated in recall software, as is done in the Intake24 recall tool (27). A similar approach could be applied to generate red and processed meat values. Definitions for values provided would need to be included as red and processed meat definitions vary, indeed WCRF definitions potentially differ from the typical layperson understanding of these concepts. Intake24 assigns weights for beef, lamb, pork, burgers, sausages, processed red meat and other red meat to each food item, and the Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA24) takes a similar approach (27,28). However, sausages and hamburgers containing preservatives are defined by WCRF as processed meat whereas sausages and hamburgers not containing preservatives are defined as red meat which adds a level of complexity to this approach (25). More precise dietary measurement would allow researchers to assess guideline adherence with greater confidence as well as facilitating a more nuanced exploration of diet-disease relationships, potentially contributing to the development of cancer-specific dietary guidance.

The absence of AOAC values in the myfood24 nutrient database at the time of our study may be explained by missing data in the UK food composition tables alongside no requirement for “back of pack” nutrient labels, which are used in myfood24, to list fiber (23,29–31). The National Diet and Nutrition Survey (NDNS) Nutrient Databank compiles AOAC values from multiple sources including recipes and estimation from NSP to ensure all foods have a value for this nutrient (32). Without additional processing of myfood24 output, 99% of participants AOAC fiber would have been underestimated by an average of

11.6 g. This change is over a third of the daily recommended amount of 30 g or more (3).

Results highlighted the importance of examining how missing data is represented in dietary analysis software and the need to explore the scope of missing data. When missing nutrient composition values are substituted with zero, it is inevitably impossible to distinguish between foods that do not contain the nutrient and foods where the value is missing. For ASCOT, addressing this issue required skilled researchers trained in nutrition to identify and correct missing fiber values but it is probable that items were overlooked, potentially leading to an underestimation of fiber intake. Teams without nutrition expertise may find this approach more challenging or may not even be aware of the issue.

Free sugars are chemically indistinguishable from total sugar and therefore must be calculated, either through an estimation based on other sugars as was done here, or on a food-by-food basis from recipes as the NDNS Nutrient Databank has done (32). This study did not have the resources required for the second approach, which would have provided greater accuracy therefore NMES and free sugars were estimated using a regression equation for sugar from fruit and vegetables. It is worth noting that as the equation was derived from data from this sample, it may not apply to other populations. As mean free sugar intake was 39.4 g/day lower than total sugar intake and there was considerable variation between participants (SD 21.4), total sugar is not an adequate proxy for free sugar (5). Given the inclusion of both AOAC fiber and free sugar in UK nutrition guidelines, the issues we encountered are likely to also impact other groups assessing compliance (5,33).

Issues identified with food names and portion sizes presented in the participant interface highlight the need for dietary analysis software to include clear and unambiguous food and portion options, both to reduce participant burden and simplify data cleaning. A problem inherent in self-administered dietary data collection is that it is impossible to ascertain what participants truly intended to report, particularly with more unusual entries but this issue is exacerbated by ambiguity in food names or portion sizes. Adding examples to food names as others have done, offering images of food packaging, and portion size pictures for the majority of foods might provide clarity, though these would need to be frequently reviewed to ensure continued accuracy (11). Indeed, nutrient databases must be regularly maintained to ensure the accuracy of food names, portion sizes, availability of items, and nutrient data, as the creators of myfood24

acknowledge (16,29). Having fewer items could allow for more efficient maintenance which has the potential to improve accuracy; indeed Intake24 reduced the number of food items in their database with this end in mind (16). Any reduction in the number of items in a database however would have to be carefully considered to ensure that its ability to reflect the diversity of diets in the UK was not adversely affected.

To reduce the need for researchers to explore recalls identified by control checks as having exceptionally large or small portion sizes, more stringent range checks could be implemented during data collection to prompt participants to review portion sizes (34). Range checks on the total values of the recall could also be performed at the point of data collection, for example highlighting to participants where the total energy intake for the day is under 500 kilocalories as Intake24 does (11) or highlighting where the energy intake is over a given amount. This would have to be balanced against the risk of increasing social approval bias and social desirability bias (35–37).

Pitfalls in using automated dietary recall systems has implications for individuals who use such systems to self-monitor their diets. It seems probable many using the software for personal monitoring would assume that the values provided are ready for use and may potentially modify their diet based on incorrect estimates.

The rigorous quality control checks employed for ASCOT uncovered a number of issues, and only those pertaining to the baseline data have been presented. Additional issues were identified in follow-up recalls which were corrected in the individual recall where identified, but not corrected throughout the dataset as it was not feasible to reprocess data collected at earlier time points. Nor was it desirable to process post-intervention data using a different methodology to that employed at baseline. The issues highlighted here, while being relevant to other myfood24 users in particular, also serve to highlight the need for vigilance and close examination of dietary data output from other software. Output data were subject to more intense scrutiny than is usual, partly due to our desire to estimate specific cancer-relevant dietary variables, rather than simply macro and micronutrient intakes. It is unclear whether our experiences are unusual as trial papers often lack a detailed description of dietary data processing methods, although these would be valuable for other research teams.

The methodology developed had several limitations. For example, the estimation of fruit and vegetables to assess compliance with 5-a-day guidance, did not

account for consumption of juice at multiple time points during the day as corrections were applied at the food item level rather than participant level. Therefore, fruit intake was overestimated for the small number of participants with multiple entries for fruit juice. Further, not all pulses and beans were included in the estimation of vegetable intake. Lastly, some meat may have been incorrectly identified as red meat or processed meat, as it could not be ascertained from the data whether items such as sausages and hamburgers contained preservatives, so all sausages were categorized as processed meat, and all hamburgers as red meat.

Conclusion

Developing methodology to assess adherence to WCRF dietary guidelines using online dietary recall software, is resource intensive but necessary. Some of this additional work could be automated by dietary data collection software or emerging AI technologies. A pragmatic approach was required to balance resources with improved accuracy of estimates of dietary intake. The additional data processing had a meaningful impact on estimates of key foods and nutrients and provided confidence that the final data was valid and appropriate for use in ASCOT.

Acknowledgements

The authors thank Cancer Research UK for funding the Advancing Survivorship Cancer Outcomes Trial (ASCOT) and the NHS trusts who helped recruit participants, ASCOT staff who supported data collection and the participants for taking the time to complete the measures.

Author Contributions

AF and RB designed the ASCOT study. VI and RC planned and carried out the analyses presented in this paper. VI, HC, SE, PL, RB, AL and RC collected the data. IV and RC wrote the manuscript. All authors reviewed, edited and approved the final manuscript.

Ethical Approval

Ethical approval was obtained through the National Research Ethics Service Committee South Central—Oxford B (reference number 14/SC/1369).

Disclosure Statement

No potential conflict of interest was reported by the authors.

Funding

This work was funded by Cancer Research UK (grant number C43975/A27498). The charity had no influence on the data collection, interpretation or reporting.

Data Availability Statement

The data that support the findings of this study are available from the corresponding author, RC, upon reasonable request.

References

1. Quaresma M, Coleman MP, Rachet B. 40-year trends in an index of survival for all cancers combined and survival adjusted for age and sex for each cancer in England and Wales, 1971–2011: a population-based study. *Lancet*. 2015;385(9974):1206–18. DOI: [10.1016/S0140-6736\(14\)61396-9](https://doi.org/10.1016/S0140-6736(14)61396-9)
2. Cancer Research UK. Our research into preventing cancer [Internet]. [cited 2023 Sep 12]. <https://www.cancerresearchuk.org/our-research-by-cancer-topic/our-research-into-preventing-cancer>.
3. World Cancer Research Fund, American Institute for Cancer Research. Continuous Update Project Expert Report 2018. Recommendations and public health and policy implications. [Internet]. 2018 May [cited 2023 Jul 3]. <https://www.wcrf.org/wp-content/uploads/2021/01/Recommendations.pdf>.
4. Williams K, Steptoe A, Wardle J. Is a cancer diagnosis a trigger for health behaviour change? Findings from a prospective, population-based study. *Br J Cancer*. [Internet]. 2013;108(11):2407–12. DOI: [10.1038/bjc.2013.254](https://doi.org/10.1038/bjc.2013.254).
5. Scientific Advisory Committee on Nutrition. Carbohydrates and health [Internet]. London; 2015 [cited 2023 Jul 3]. <https://www.gov.uk/government/publications/sacn-carbohydrates-and-health-report>.
6. Committee on Medical Aspects of Food Policy. Dietary Reference Values for Food Energy and Nutrients for the United Kingdom. Report of the Panel on Dietary Reference Values of the Committee on Medical Aspects of Food Policy. Norwich: The Stationary Office; 1991.
7. Public Health England. Eatwell guide [Internet]; 2016 [cited 2023 Jul 17]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/528193/Eatwell_guide_colour.pdf.
8. Public Health England, Food Standards Agency. National diet and nutrition survey. Rolling programme Years 9 to 11 (2016/2017 to 2018/2019) [Internet]. 2020 Dec [cited 2023 Jul 13]. https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/943114/NDNS_UK_Y9-11_report.pdf.
9. Public Health England, Food Standards Agency. NDNS: results from years 9 to 11 (combined) - data tables. NDNS_Y9_11_DESCRIPTIVE_STATISTIC_TABLES_FINAL-2 [Internet]. 2020 Dec [cited 2023 Jul 13]. <https://www.gov.uk/government/statistics/ndns-results-from-years-9-to-11-2016-to-2017-and-2018-to-2019>.

10. Timon CM, van den Barg R, Blain RJ, Kehoe L, Evans K, Walton J, Flynn A, Gibney ER. A review of the design and validation of web- and computer-based 24-h dietary recall tools. *Nutr Res Rev.* 2016;29(2):268–80. DOI: [10.1017/S0954422416000172](https://doi.org/10.1017/S0954422416000172)
11. Simpson E, Bradley J, Poliakov I, Jackson D, Olivier P, Adamson AJ, Foster E. Iterative development of an online dietary recall tool: INTAKE24. *Nutrients.* 2017;9(2):118. DOI: [10.3390/nu9020118](https://doi.org/10.3390/nu9020118)
12. Foster E, Lee C, Imamura F, Hollidge SE, Westgate KL, Venables MC, Poliakov I, Rowland MK, Osadchiy T, Bradley JC, et al. Validity and reliability of an online self-report 24-h dietary recall method (Intake24): a doubly labelled water study and repeated-measures analysis. *J Nutr Sci.* 2019;8:e29. DOI: [10.1017/jns.2019.20](https://doi.org/10.1017/jns.2019.20)
13. Thompson FE, Dixit-Joshi S, Potischman N, Dodd KW, Kirkpatrick SI, Kushi LH, Alexander GL, Coleman LA, Zimmerman TP, Sundaram ME, et al. Comparison of interviewer-administered and automated self-administered 24-hour dietary recalls in 3 diverse integrated health systems. *Am J Epidemiol.* 2015;181(12):970–8. DOI: [10.1093/aje/kwu467](https://doi.org/10.1093/aje/kwu467)
14. Wark PA, Hardie LJ, Frost GS, Alwan NA, Carter M, Elliott P, Ford HE, Hancock N, Morris MA, Mulla UZ, et al. Validity of an online 24-h recall tool (myfood24) for dietary assessment in population studies: comparison with biomarkers and standard interviews. *BMC Med.* 2018;16(1):136. DOI: [10.1186/s12916-018-1113-8](https://doi.org/10.1186/s12916-018-1113-8)
15. Shim JS, Oh K, Kim HC. Dietary assessment methods in epidemiologic studies. *Epidemiol Health.* 2014;36:e2014009. DOI: [10.4178/epih/e2014009](https://doi.org/10.4178/epih/e2014009)
16. Public Health England. Evaluation of changes in the dietary methodology in the National Diet and Nutrition Survey Rolling Programme from Year 12 (2019 to 2020) Stage 1 [Internet]; 2021. Sep [cited 2023 Jul 19]. <https://www.gov.uk/government/publications/evaluation-of-change-in-dietary-methodology-in-ndns-rolling-programme-stage-1>.
17. Beeken RJ, Croker H, Heinrich M, Smith L, Williams K, Hackshaw A, Hines J, Machesney M, Krishnaswamy M, Cavanagh S, et al. Study protocol for a randomised controlled trial of brief, habit-based, lifestyle advice for cancer survivors: exploring behavioural outcomes for the Advancing Survivorship Cancer Outcomes Trial (ASCOT). *BMJ Open.* 2016;6(11):e011646. DOI: [10.1136/bmjopen-2016-011646](https://doi.org/10.1136/bmjopen-2016-011646)
18. Carter MC, Albar SA, Morris MA, Mulla UZ, Hancock N, Evans CE, Alwan NA, Greenwood DC, Hardie LJ, Frost GS, et al. Development of a UK Online 24-h Dietary Assessment Tool: myfood24. *Nutrients.* 2015;7(6):4016–32. DOI: [10.3390/nu7064016](https://doi.org/10.3390/nu7064016)
19. Conway R, Heuchan G, Croker H, Esser S, Ireland V, Lally P, Beeken R, Fisher A. Comparison between self-completed and interviewer-administered 24-hour dietary recalls in cancer survivors: sampling bias and differential reporting. *Nutrients.* 2022;14(24):5236. DOI: [10.3390/nu14245236](https://doi.org/10.3390/nu14245236)
20. Nelson M, Erens B, Bates B, Church S, Boshier T. Low income diet and nutrition survey. Appendices and supplementary tables [Internet]. London; 2007. [cited 2023 Oct 23]. <https://webarchive.nationalarchives.gov.uk/ukgwa/20110110172405/http://www.food.gov.uk/science/dietarysurveys/lidnsbranch/>.
21. World Cancer Research Fund, American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. Washington DC; 2007.
22. NHS. 5 A Day: what counts? [Internet]; 2022 [cited 2023 Jul 3]. <https://www.nhs.uk/live-well/eat-well/5-a-day/5-a-day-what-counts/>.
23. Food Standards Agency. McCance and Widdowson's The Composition of Foods. 6th ed. Cambridge: royal Society of Chemistry; 2002.
24. Public Health England, Food Standards Agency. National Diet and Nutrition Survey. Years 9 to 11 of the Rolling Programme (2016/17 - 2018/19). Appendix A Dietary data collection and editing for Year 10 and 11 of the NDNS RP [Internet]; 2020 Dec [cited 2023 Jul 11]. <https://www.gov.uk/government/statistics/ndns-result-s-from-years-9-to-11-2016-to-2017-and-2018-to-2019>.
25. World Cancer Research Fund, American Institute for Cancer Research. Meat, fish and dairy products and the risk of cancer 2018 [Internet]; 2018 [cited 2023 Jul 13]. <https://www.wcrf.org/wp-content/uploads/2021/02/Meat-fish-and-dairy-products>.
26. Kirkpatrick SI, Subar AF, Douglass D, Zimmerman TP, Thompson FE, Kahle LL, George SM, Dodd KW, Potischman N. Performance of the Automated Self-Administered 24-hour Recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. *Am J Clin Nutr.* 2014;100(1):233–40. DOI: [10.3945/ajcn.114.083238](https://doi.org/10.3945/ajcn.114.083238)
27. Food Standards Scotland, Newcastle University, University of Cambridge, Monash University. Intake24 Sample Output [Internet]. [cited 2023 Sep 13]. <https://intake24.co.uk/info/output#content>.
28. National Cancer Institute. Nutrient and Food Group Analysis ASA24 [Internet]; 2023 [cited 2023 Sep 4]. <https://epi.grants.cancer.gov/asa24/researcher/analysis.html#tab2-panel>.
29. Carter MC, Hancock N, Albar SA, Brown H, Greenwood DC, Hardie LJ, Frost GS, Wark PA, Cade JE. Development of a new branded UK food composition database for an online dietary assessment tool. *Nutrients.* 2016;8(8):480. DOI: [10.3390/nu8080480](https://doi.org/10.3390/nu8080480)
30. Public Health England. McCance and Widdowson's composition of foods integrated dataset [Internet]; 2021. [cited 2023 Jul 18]. <https://www.gov.uk/government/publications/composition-of-foods-integrated-dataset-cofid>.
31. British Nutrition Foundation. Looking at labels [Internet]; 2022 [cited 2023 Oct 24]. <https://www.nutrition.org.uk/putting-it-into-practice/food-labelling/looking-at-labels/#backofpack>.
32. Public Health England. NDNS Y9-11 Appendix AA Calculation of free sugars and AOAC fibre in the NDNS RP. [Internet]; 2020 Dec [cited 2023 Jul 18]. <https://www.gov.uk/government/statistics/ndns-result-s-from-years-9-to-11-2016-to-2017-and-2018-to-2019>
33. Public Health England. Government dietary recommendations. Government recommendations for energy and nutrients for males and females aged 1-18 years and 19+ years. [Internet]. London; 2016 Aug [cited 2023 Oct 24]. <https://www.gov.uk/government/publications/the-eatwell-guide>.
34. Arab L, Hahn H, Henry J, Chacko S, Winter A, Cambou MC. Using the web for recruitment, screen, tracking,

- data management, and quality control in a dietary assessment clinical validation trial. *Contemp Clin Trials*. 2010;31(2):138–46. DOI: [10.1016/j.cct.2009.11.005](https://doi.org/10.1016/j.cct.2009.11.005)
35. Meijboom S, van Houts-Streppel MT, Perenboom C, Siebelink E, van de Wiel AM, Geelen A, Feskens EJM, de Vries JHM. Evaluation of dietary intake assessed by the Dutch self-administered web-based dietary 24-h recall tool (Compl-eatTM) against interviewer-administered telephone-based 24-h recalls. *J Nutr Sci*. 2017;6:e49. DOI: [10.1017/jns.2017.45](https://doi.org/10.1017/jns.2017.45)
36. Miller TM, Abdel-Maksoud MF, Crane LA, Marcus AC, Byers TE. Effects of social approval bias on self-reported fruit and vegetable consumption: a randomized controlled trial. *Nutr J*. 2008;7(1):18. DOI: [10.1186/1475-2891-7-18](https://doi.org/10.1186/1475-2891-7-18)
37. Novotny JA, Rumpler WV, Riddick H, Hebert JR, Rhodes D, Judd JT, Baer DJ, McDowell M, Briefel R. Personality characteristics as predictors of underreporting of energy intake on 24-hour dietary recall interviews. *J Am Diet Assoc*. 2003;103(9):1146–51. DOI: [10.1016/s0002-8223\(03\)00975-1](https://doi.org/10.1016/s0002-8223(03)00975-1)