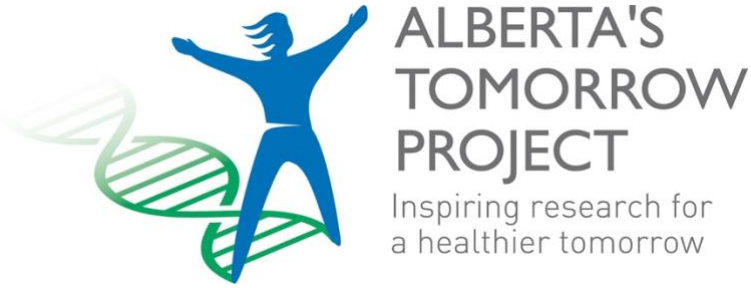

Considerations for Dietary Assessment in the Canadian Partnership for Tomorrow Project

April 15, 2018

Minor revisions July 8, 2019

Michelle L. Marcinow, Sharon I. Kirkpatrick, Jennifer E. Vena, Paula J. Robson

ACKNOWLEDGEMENTS

The authors are grateful to Vanessa DeClercq and Rachel Murphy for their comments on an earlier version of this document and to Miriam Price for editing and referencing assistance.

ABBREVIATIONS

24HR	24-hour recall
AMPM	Automated Multiple-Pass Method
ASA24	Automated Self-Administered 24-hour Dietary Assessment Tool
CDHQ	Canadian Diet History Questionnaire
CLSA	Canadian Longitudinal Study on Aging
CPTP	Canadian Partnership for Tomorrow Project
DHQ	Diet History Questionnaire
FFQ	Food frequency questionnaire
FR	Food record
R24W	Rappel 24h Web

TABLE OF CONTENTS

EXECUTIVE SUMMARY

Dietary factors are leading contributors to chronic disease and mortality globally and in Canada (1–3), and have been recognized as modifiable risk factors for certain cancers (4). However, much remains to be learned about how dietary factors interact with other modifiable and non-modifiable exposures and physiologic variables to influence disease risk in humans (5,6). Information collected from large prospective cohorts plays an important role in furthering our understanding of diet-disease relationships (7,8). To advance knowledge on how to promote health and prevent disease, it is critically important to use robust tools for collecting dietary information from participants in such cohorts (9). This guide is intended to be utilized by researchers designing nutritional epidemiological research and in particular, to guide the implementation of dietary assessment tools within the CPTP cohorts. The aim is to provide guidance on method selection, data collection, and analyses of dietary data, as well as stimulate discussions of harmonization of methods across cohorts to advance the evidence base.

Because objective measures such as biomarkers of diet are currently few, burdensome, costly, and limited in the information they provide about the types of foods and beverages people consume (5,6), researchers typically rely upon self-report tools. However, it has long been recognized that self-reported dietary data are affected by error, including systematic error or bias (9,10), leading some commentators to suggest that research should no longer rely on self-report approaches (11,12). However, much work has been conducted to better understand and address error in self-report dietary intake data (9,10). Such work has informed the development of novel technology-enabled tools to allow collection of the least-biased data possible, as well as the development of rigorous statistical approaches to mitigate the effects of error (13–16).

Based on what is known about sources and types of error in data captured using different types of tools, it has been recommended that a combination of tools may be the optimal way forward for cohort studies. Specifically, multiple 24-hour recalls (24HRs), administered in combination with a food frequency questionnaire (FFQ), may allow researchers to leverage the strengths of each instrument (10,14,17). Data from 24HRs provide comprehensive detail on intake and measure consumption with less bias than FFQs. On the other hand, FFQs measure intake over a longer period (e.g. past month or year) (18–20), meaning they are better able to capture intake of foods and beverages that may be consumed more episodically (e.g., whole grains, dark-green vegetables) but that may be important to diet-disease relationships. The availability of web- and mobile device-based dietary assessment tools for use in Canada and emerging statistical techniques to analyze the resulting data makes this multiple-tool scenario a realistic consideration for Alberta's Tomorrow Project (21), other cohorts within the Canadian Partnership for Tomorrow Project (CPTP) (22), and other health-related studies. With comprehensive and standardized measurement of dietary exposures across cohorts, the identification of promising strategies to reduce diet-related disease risk among Canadians can be furthered (9).

SECTION 1: INTRODUCTION

The Global Burden of Disease, Injuries, and Risk Factor study listed diet as one of the leading contributors to morbidity and mortality in Canada and globally (1–3). Particular to cancer, components or characteristics of the diet have been found to either reduce or increase risk of certain cancers (4). For example, intake of red and processed meats has been associated with an increased risk of advanced prostate cancer, while vegetarian and vegan diets have been associated with a slightly reduced risk of gastrointestinal cancers, cancers of female reproductive organs, and prostate cancer (4). As well, some types of vegetables, and fruits in general, have been identified as having a probable protective role against certain types of cancer (23,24). Despite the focus on singular foods and beverages, there is growing consensus on the importance of elucidating *overall* dietary patterns and the synergistic and potentially antagonistic effects of multiple dietary constituents in relation to disease risk (25–28).

Data on dietary patterns within Canadian populations are relatively sparse, posing a barrier to furthering evidence on diet-health relationships and the identification of strategies to promote health. The 2004 Canadian Community Health Survey collected comprehensive dietary intake data from a nationally-representative sample (using a single 24HR on the majority of the sample and two 24HRs on a subsample), with the data suggesting that current dietary patterns are not well aligned with recommendations for health (29,30). This survey was repeated in 2015 (31) but given the cross-sectional nature of the data, it is difficult to examine diet-related exposures in relation to disease endpoints or proximal markers of disease. However, cohort studies such as the Canadian Partnership for Tomorrow Project (CPTP) (22) and the Canadian Longitudinal Study on Aging (CLSA) (32) provide opportunities to address this knowledge gap within Canada. These studies are large and prospective in nature, offering unique opportunities to examine diet in relation to other risk factors, proximal markers of disease, and disease outcomes. With comprehensive and standardized measurement of dietary exposures within these cohorts, the identification of promising strategies to reduce diet-related disease risk among Canadians can be furthered (9).

To date, many of the cohorts within the CPTP, as well as the CLSA, have not collected comprehensive dietary intake data, in large part due to the lack of high-quality and easy-to-administer tools tailored to the Canadian context at the time the cohorts were initiated (9). However, recent developments have made web- and mobile device-based systems available for measuring dietary intake data in Canada, providing an opportunity to collect and eventually pool comprehensive dietary intake data across the cohorts. Thus, the existing CPTP (22) and other platforms can be leveraged to strengthen the evidence on diet-disease relationships, as well as to examine diet in relation to other factors, such as physical activity, sedentary

behaviour, and sleep, environmental exposures, and genetics. Collaborative efforts to facilitate data harmonization and pooling across studies will strengthen the evidence (33) and optimize the scientific impact of each individual cohort by enhancing statistical power and generalizability, as well as providing the potential to evaluate differences among the cohorts with respect to dietary exposures and disease-related outcomes (34).

The purpose of this guide is to outline considerations associated with measuring diet in epidemiologic research, particularly in the context of Canadian cohort studies. Included are a description of measurement error and its impact on estimating diet-disease relationships, as well as an overview of commonly-used dietary assessment measures and their characteristics related to measurement error. Also provided are considerations for the selection and implementation of dietary assessment tools within the CPTP cohorts, based on current thinking about best practices and evidence regarding feasibility of available tools. This user guide also includes an overview of practical considerations for collecting, processing, analyzing, interpreting, and reporting dietary data, directing users to additional resources that delve into these concepts in greater depth. While the document is not exhaustive given the breadth of considerations involved, it is intended to help researchers to get started and to identify issues that require further investigation to successfully implement comprehensive dietary assessment in epidemiologic research.

SECTION 2: MEASUREMENT ERROR AND ITS IMPLICATIONS FOR ESTIMATION OF DIET-DISEASE RELATIONSHIPS

A challenge related to dietary data and their use to examine diet-disease relationships is measurement error (10). Measurement error can cause attenuation of disease risk estimates and reduce statistical power (35), resulting in the need for large sample sizes to detect relationships between dietary exposures and disease-related outcomes (13). For example, in the Observing Protein and Energy Nutrition Study (36), attenuation due to error in estimation of protein intake was reported to cause a true relative risk of 2.0 to be reduced to 1.10-1.12 (13). It has long been recognized that there will likely always be error associated with self-report dietary assessment (37), but sources of error can be minimized to the extent possible and acknowledged to avoid misleading results (10).

Measurement error refers to the difference between the observed or measured value and the true value (10,18). In the context of dietary assessment, this reflects the difference between estimated intake of a dietary component based on self-report data and the true level of intake of that dietary component.

Measurement Error Webinar Series

The National Cancer Institute's webinar series (38) provides an in-depth overview of issues related to measurement error inherent in dietary intake data. Several webinars relate specifically to the estimation of diet-disease relationships, with one session dedicated to the use of 24HRs in combination with an FFQ in epidemiologic research.

The two main types of measurement error are random error and systematic error (also known as bias) (20,37). *Random error*, also referred to as classical measurement error, is independent of the true exposure (e.g., dietary component) (18,20). It can occur in either direction (under- or over-estimation) and does not affect the mean or median; thus, data affected by random error are not biased, but are not precise (13). A key source of random error in dietary data reflecting intake on a day or a few days is day-to-day variation in what individuals eat and drink (i.e., within-person variation) (20,38). Measurement error can also be *systematic (*referred to as bias), in that measurements consistently depart from the true value in the same direction (18,20). Systematic error may come about due to inaccurate recall, social desirability biases, and other sources of bias in reporting.

Recovery biomarkers, which provide estimates of true intake for energy, protein, potassium, and sodium, have been used in validation studies to characterize the types and extent of error affecting data collected using different self-report dietary tools. Such studies suggest that data from FFQs are affected by bias to a larger degree than data from 24HRs (36,39–41). Random

error, on the other hand, particularly affects data from short-term tools, such as 24HRs and food records (FRs), which are used to collect data for only a day or a few days. If a sufficient number of repeat data collections are available, averaging across days provides estimates of usual intake for a population or subpopulation such that the random errors cancel each other out (42). Investigators are usually unable to collect a sufficient number of recalls on a large sample (10), but smaller numbers of repeat measures (potentially for a sub-sample) and statistical modelling can be used to adjust for the effects of day-to-day variation (42). In contrast, systematic error, or bias, cannot be reduced or eliminated by repeat measures.

Importantly, although no self-report tool captures data that are free of bias (15,37), the findings of validation studies suggest that self-report data provide reasonable estimates for many nutrients and food groups (15,16). For example, estimates of potassium intake have been shown to be reasonable, suggesting the same may be true for estimates of fruit and vegetable. However, this is not the case for energy intake (15,16). Energy is contained in all foods and beverages (10); thus, error associated with each individual food or beverage accumulates (10,43). Thus, it is recommended that absolute energy estimates based on self-report not be relied upon as true indicators of caloric intake. Self-reported energy intake can be used for energy adjustment, which can partially mitigate the impact of bias on observed diet-disease relationships, as discussed in SECTION 4.These issues are also further detailed in Subar et al. (10).

SECTION 3: OVERVIEW OF DIETARY ASSESSMENT TOOLS

Choosing the most appropriate dietary assessment tool can be a challenge, particularly in studies with very large samples. Because objective measures of diet are few and burdensome (9,44), researchers are typically reliant upon self-report measures to capture diet. Generally, we are interested in using such tools to estimate usual intake (i.e. long-term average intake) (42,45,46), a consideration that must be borne in mind when determining the dietary data to be collected and how they will be analyzed.

Commonly-used self-report measures include 24HR, FR, FFQs, and brief tools or screeners, (9,20). Data on food, beverage, and supplement consumption collected using such tools can be linked to available nutrient and food group databases, such as the Canadian Nutrient File (9,18), to calculate estimated intake of foods, food groups, nutrients, and other dietary components, as well as patterns of intake. These estimates can then be used to model relationships between dietary exposures and health or disease outcomes (9,18).

The more you know

Understanding the advantages and limitations of different dietary assessment tools, and including the expertise of nutrition researchers, epidemiologists and statisticians, when choosing tools and identifying the corresponding statistical approaches to be used, can enhance the ability of a study to evaluate diet-disease relationships.

Self-report tools can be characterized based on whether they capture data on intake over the short or long term (18,19). Short-term tools, such as 24HR and FR, capture dietary intake for a single day or a few days. Data from short-term tools do not directly measure usual intake, but with repeat measures and the application of statistical modeling, distributions of usual intake can be estimated (42). Long-term tools, such as FFQs or screeners, are designed to collect data on usual intake over a long period, such as a month or a year (18,19). The main self-report tools, including the key sources of error, are described below, with a brief overview in Box 3.1.

Short-term Dietary Assessment Tools: 24-hour recalls (24HRs) and Food Records (FRs)

The 24HR consists of a structured interview to collect data on all foods and beverages consumed on a given day, typically the prior day from midnight-to-midnight (18–20). In interviewer-administered recalls, a trained interviewer uses open-ended questions to collect comprehensive details, including brand names, and how food was prepared (18–20). For example, a report of a sandwich for lunch is followed by probes for detailed information, such as the type (e.g., white or whole wheat) and brand name (e.g., Dempster's) of the bread, the items that were included on the sandwich (e.g., lunch meat, lettuce, tomato), and if condiments

were used. Food models, pictures, and other visual aids may be used to improve accuracy of portion size reporting (18–20). A recall administered by a trained interviewer usually requires approximately 20-60 minutes to complete (18,19). In addition to providing detailed dietary intake information, 24HRs can provide information about meal patterns (e.g., timing and composition of meals and snacks), consumption of foods at and away from home, and contextual factors (e.g., use of electronic devices during meals, eating with others) (18). Dietary supplement use can also be captured, which is important since supplements can make significant contributions to intake of certain nutrients (18–20). Standardized automated systems, such as the U.S. Department of Agriculture's Automated Multiple-Pass Method (AMPM), have been developed to facilitate the process, and improve accuracy, of interviewer-administered 24HRs (18,19). AMPM consists of: 1) a quick list of all foods and beverages consumed the prior day; 2) probes for forgotten foods; 3) inquiries into timing of consumption and eating occasions; 4) probing questions to gather details about the foods and beverages reported; and, 5) a final review, which probes the respondent about any other items not already reported (18–20).

Recall data are affected by random error, a key source of which is day-to-day variation (i.e., variation in dietary intake from one day to the next) (18,20). Random error can be accounted for with repeat measures (e.g., replicate 24HRs) and statistical modeling to account for the within-person variation (statistical modeling rather than averaging is usually required since the number of repeat recalls available is generally too low for averaging to sufficiently address random error) (42). Recall data are also affected by bias, sources of which include imperfect short-term memory or recall, inaccurate portion size reporting, and social desirability biases that may contribute to over-or underreporting of dietary components (19,20). However, 24HRs are recommended for use in various research initiatives, including epidemiological studies, because they have been shown to capture dietary intake with less bias compared to FFQs (13,15,16).

A key challenge in the use of 24HRs in large-scale research has been the prohibitive cost associated with interview-administered recalls (19), particularly if repeat administrations are required to account for day-to-day variation (42). As well, the need for manual coding by trained personnel contributes substantially to the cost and burden (18). To minimize errors and ensure consistency, a standardized protocol must be developed for coding each food/beverage item, coders must be trained extensively on the coding protocol, and quality control spot checks must be completed throughout the process (19,20). As a result, interview-based 24HRs have not generally been used in large-scale cohort studies.

In the past decade, technological advances in 24HRs have eliminated the need for trained interviewers and coders (47–51). For example, the Automated Self-Administered 24-hour (ASA24) Dietary Assessment Tool, developed by the US National Cancer Institute, is a web-based system developed to enable the collection of multiple 24HRs (as well as FRs) in large-scale research (51). The format and design of ASA24 are based on a modified version of the interviewer-administered AMPM. ASA24 integrates automated coding to eliminate the need for manual coding, which is laborious, time-consuming, and expensive (19).

Find out more about ASA24-Canada

Appendix A includes additional information about ASA24-Canada.

A demonstration version of ASA24 is available at: https://asa24.nci.nih.gov/demo/

ASA24-Canada (www.asa24.ca) is freely available and has been adapted to reflect the Canadian food supply, including differences between the US and Canada in brand names and fortification (51). It has been tested in subsamples of both the Alberta's Tomorrow Project (21,52) and the CLSA (32,53) cohorts. Findings from these feasibility studies suggest that participants are receptive to completing ASA24 on multiple occasions; however, the use of technology introduces new challenges, such as those associated with computer literacy, that must be taken into consideration (54). For example, it may be necessary to develop quick-start guides or video tutorials to help participants troubleshoot technological or other issues encountered. The National Cancer Institute website includes sample quick start guides that can be used or adapted for different types of studies and populations. In addition, researchers at Université Laval have developed the Rappel 24h Web (R24W), a web-based 24-hour recall for use with French-Canadian populations, based on the AMPM (55). Evidence based on a controlled feeding study supports the capacity of R24W to assess dietary intakes in a population of adults (56).

Food records, sometimes referred to as diaries, are similar to 24HRs in that they capture a detailed account of foods, beverages, and possibly, dietary supplements consumed over one or more days (e.g., 3-d FRs, 7-d FRs) (18–20). The difference between the tools is that 24HRs rely on short-term memory to *recall* dietary intake from the previous day (potentially eliciting recall biases), while with FRs, the intent is for participants to *record* information in real-time (potentially eliciting reactivity, elaborated upon below). With a written record, respondents are provided with a form and oral and/or written instructions prompting them to record relevant details for all foods and beverages consumed, such as brand name, preparation method, and where the food or beverages items were consumed. Portion size can be estimated using visual aids, such as food models or pictures, or measured using weighing scales or measurement cups/spoons. Food records may also capture intake of dietary supplements. Following completion, a trained interviewer may review the FRs and follow up with the respondent to fill in missing details (18–20).

Food records have similar advantages (e.g., collecting contextual factors) and challenges (e.g., inaccurate portion size estimation, social desirability bias) to 24HRs (18). In contrast to 24HRs (particularly those that are unannounced or not anticipated by respondents), data collected using FRs can be affected by reactivity bias in that respondents may alter their eating habits in response to monitoring (18–20). Respondents may also forget to record items in real-time and instead, rely on memory to fill in missing information. Last, there is a large burden placed on respondents to record intake (particularly if weighing of food is required), potentially impacting the quality of data, especially when many days of recording are required (18–20). Food records have been used in cohort studies to calibrate estimates derived from FFQs (18), but the large expense associated with manual coding and laborious data entry has made their use in large-scale cohorts impractical for the most part (18,19).

As with 24HRs, FRs have been enhanced by technological innovation (18). For example, ASA24-Canada, described above, provides the capacity to collect FRs in addition to recalls (57), providing promise for reducing burden as well as costs associated with coding and entering FR data. Mobile device-based applications have also been developed and may use cameras and image recognition for automated identification of foods and beverages and estimation of portion sizes (58,59). For example, the Technology Assisted Dietary Assessment (TADA) approach uses images obtained before and after eating occasions, along with fiducial markers to provide scale and perspective, to estimate the types and amounts of foods consumed (60). Work to automate food identification and portion size estimation for such tools is ongoing, but such tools are not yet ready for wide-scale implementation in large cohorts. Furthermore, to our knowledge, such tools have not yet been tailored for use in Canada, which is an important consideration given that the Canadian food supply, including foods and beverages available as well as fortification practices, differ from the food supply in other countries.

In addition to research-grade tools such as TADA, mobile device-based commercial applications for recording food intake for self-monitoring purposes (e.g., MyFitnessPal) have been made available (61). These commercial applications likely have similar limitations to FRs (e.g., reactivity) because they are designed to capture dietary intake in real-time. In some cases, reactivity may be a desired outcome of self-monitoring, but this source of bias is problematic for capturing usual dietary exposures in the context of epidemiologic research (61). For the most part, these applications have not been tested to assess the extent to which they accurately and reliably capture intake (61). There are also concerns with quality (e.g., limited Canadian content, currency) (62). Lastly, there may be issues relating to ownership and usage of data (e.g., potential to be shared with a third party) (63). Therefore, the use of commercial mobile device-based applications in large-scale cohort studies is not suggested, particularly with the availability of web-based dietary assessment tools that have been tailored specifically to Canada and are intended for use in research.

Long-term Dietary Assessment Tools: Food Frequency Questionnaires (FFQs) and Screeners

FFQs are designed to collect information regarding respondents' usual frequency of consumption of foods and beverages over a specified period, such as the previous month or year (18–20). Respondents are instructed to estimate how often a particular food or beverage or grouping is consumed, choosing from categories such as 'never' or 'less than once a month' or '6+ per day'. Many FFQs are designed to collect information on usual portion size or specify portion sizes as part of each question, sometimes with visual aids, such as images, to enhance reporting accuracy. An FFQ may also collect information regarding frequency of intake and dosage of commonly consumed dietary supplements (18–20). FFQs are typically self-administered and may require between 30-60 minutes to complete if the intent is to capture the total diet (18). Under certain circumstances, such as with populations with limited literacy, questionnaires can be interview-administered (18). Given that FFQs are comprised of a standard set of questions and food items, coding of the data is fairly straightforward. FFQs have traditionally been used in large-scale cohort studies because they have offered lower burden and cost in comparison to 24HRs or FRs (10).

A key challenge in the completion of FFQs is that they can be cognitively complex (18–20). Recalling long-term food and beverage intake and determining the average frequency with which items were consumed can be difficult (18–20). Additional sources of bias include inaccurate portion size estimation, the influence of recent diet on recalling long-term intake, and the use of a finite list of items that might not capture foods and drinks commonly consumed by particular populations (18–20). Biomarker-based studies indicate that FFQ data are affected by bias to a greater extent than those from 24HRs or FRs (15,16,36,39). It is thus suggested that estimates of nutrient intake derived from FFQs should be viewed as crude approximations, and they are not considered the optimal main tool for epidemiologic studies, particularly when used on their own without calibration with less-biased data (64). It should be noted however that since they query a longer period of time, FFQs may be better able to capture episodically-consumed foods, such as dark-green vegetables, compared to short-term tools (18–20). As a result, they may be of value in combination with short-term tools in the context of epidemiologic research.

Technological advances related to FFQs are largely limited to the development of online versions (65). For example, the Diet History Questionnaire III (DHQ-III), developed by the National Cancer Institute, and the Canadian adaptation (CDHQ-II) are offered in web-based versions (66,67).

Learn more about CDHQ

Appendix B includes additional information on the CDHQ.

A demonstration version is available at: https://epi.grants.cancer.gov/dhq2/webquest/demos.html

The CDHQ-II is freely available to researchers and clinicians, and an updated version, the CDHQ-III, is under development based on foods and beverages reported using 24HRs in the 2015 Canadian Community Health Survey and will be completed by the end of 2019. The CDHQ-III will be available in English and French (similar to previous versions) as well as past-year and past-month, with and without portion sizes. Similar to the DHQ-III, the CDHQ-III will be available as an online questionnaire only. Paper- and web-based versions of the CDHQ-II were evaluated for reliability, feasibility, and acceptability in a sub-sample of Alberta's Tomorrow Project participants (68); nutrient estimates from the two versions were comparable and the majority of participants indicated willingness to complete the web-based version in the future (68). These findings indicate that the CDHQ is amenable for use in large studies, though as noted above, the FFQ is not the ideal sole tool for use in epidemiologic research.

Screeners are brief tools similar in design to FFQs, with the exception that the former provide information on targeted aspects of the diet (18–20). Screeners can be used to obtain basic information regarding a limited number of foods and beverages, or provide information regarding dietary habits (18–20). Screeners typically have low participant burden and can be quick to complete (18). However, the lack of comprehensive dietary data limits analytic opportunities, including the consideration of overall diet quality and patterns. Further, sources of bias that affect data from FFQs also affect screeners (18). For prospective studies that used screeners at baseline, it may be possible to integrate more comprehensive tools such as recalls so that the screener data can be calibrated to reduce bias (13).

Box 3.1. Overview of Self-Report Dietary Assessment Tools

- **24-hour recalls (24HRs)**: Collect comprehensive detail regarding all foods and beverages (and possibly supplements) consumed the prior day or 24-hours (18–20). Data from 24HRs are mostly affected by random error versus bias (in contrast to FFQs and screeners) (15,16,36,39). A key source of random error is variation in dietary intake from day-to-day (18,20), while sources of bias include inaccurate memory recall of dietary components or portion sizes and social desirability biases (19,20). 24HRs capture intake with less bias than FFQs (13,15,16), but their use in large-scale cohort studies has been infrequent due to cost and burden (10). However, technological advances have eliminated the need for trained interviewers and coders (47–51), making recalls feasible for use in large cohort studies. Can also provide information on meal patterning, as well as contextual factors, such as where meals were eaten and with whom.

- **Food Records (FRs):** Collect a detailed real-time account of foods, drinks and possibly, dietary supplements consumed over periods usually ranging from one to seven days (18–20). Sources of measurement error in data from FRs are similar to those affecting 24HR data, but FR data are affected by reactivity (versus recall) bias. A large number of reporting days can result in less respondent cooperation (19). Technological innovation (18) has decreased the costs and burden associated with collecting and analyzing FR data (58,59). Can also provide information on meal patterning, as well as contextual factors, such as where meals were eaten and with whom.

- **Food Frequency Questionnaires (FFQs):** Retrospectively collect information regarding respondents' usual frequency of consumption of foods and beverages over a specified period, such as the previous month or year (18–20). Many FFQs incorporate portion size questions and FFQs may also query frequency of use of dietary supplements (18–20). FFQ data are affected by bias to a larger extent than are data from 24HRs (15,16,36,39), likely due to the fact that FFQs can be cognitively complex to complete (18–20). FFQs are not considered the optimal sole tool for large cohort studies. As technology has made it feasible to use 24HRs in large studies (47–51), there is an anticipated shift toward the use of 24HRs in combination with FFQ data to leverage the strengths of each method (17).

- **Screeners:** Retrospectively collect basic information regarding a limited number of foods and beverages, or to provide information regarding dietary habits (18–20). Screeners are affected by similar sources of bias as FFQs (18).

The Potential of Biomarkers in Nutritional Epidemiology

There is great interest in the potential use of biomarkers to advance nutritional epidemiology (69–71). Biomarkers provide an objective measure of dietary intake and, in some cases, can be used as a reference measure to evaluate self-report tools (69). Biomarker data may also improve the predictive power of self-report intake to more accurately associate nutritional status with disease risk (14,17). However, their use in cohort studies currently is likely to be limited by feasibility.

Recovery biomarkers are biologic products directly related to intake and not affected substantially by inter-individual differences in metabolism (72). Recovery biomarkers are considered to be measures of true intake because they provide nearly unbiased measures of intake (69). Because of this, recovery biomarkers are useful for quantifying measurement error in self-report data (15,16,36,39), as well as calibrating data from self-report tools to remove error (72). However, only a few recovery biomarkers have been identified, including doubly-labeled water for energy intake (73) and 24-hour urine for protein (74), potassium, and sodium intake (15,16). *Concentration biomarkers* reflect the concentration of specific chemicals or compounds in blood, urine, or tissues and can be used as an indirect measure of dietary intake (19). Unlike recovery biomarkers, these biomarkers do not represent a direct measure of true intake because they are affected by differences in metabolism or personal attributes, such as smoking status and body weight (14,72). Thus, concentration biomarkers cannot be used as measures of absolute intake or for assessing error in self-report dietary intake data (14). However, there have been efforts to combine concentration biomarkers, such as serum carotenoids, with self-report data to improve estimates of diet-disease relationships (14,17) (see SECTION 5). *Predictive biomarkers* have a relationship with intake that is stable, time-dependent, and sensitive to intake in a dose-response manner (75,76). Examples of candidate predictive biomarkers are urinary sugars (a biomarker of total sugars intake) (77) and plasma alkylresorcinols (a biomarker of whole grain wheat and rye intakes) (78). Similar to concentration biomarkers, predictive biomarkers do not represent a direct measure of true intake; however, they are postulated to improve estimation of associations between diet and disease (77).

Overall, while the discovery of new biomarkers to understand the relation between diet and health is a lively area of nutrition research, the use of those recognized as markers of true intake currently remains limited due to high cost and burden (79). Additionally, biomarkers do not provide detailed information about what respondents actually eat and drink or contextual factors that might be pertinent to understanding how eating patterns can impact health- and disease-related outcomes (10).

SECTION 4: CONSIDERATIONS FOR ASSESSING DIET IN CANADIAN COHORT STUDIES

The collection of dietary data to support examination of diet-disease relationships in large-scale cohort studies, such as CPTP (22) and CLSA (32), has been challenging due to the limited availability of dietary assessment tools that are tailored to the Canadian context, cost-effective, and appropriately tested (9). At the time of designing the baseline data collection strategies for cohorts within CPTP, it was not logistically feasible to collect comprehensive dietary intake data from thousands of participants due to cost and burden (9). Questions related to fruit and vegetable intake were included within CPTP (80) to provide face validity to participants who might have anticipated being queried about dietary risk factors. It was recognized, however, that the resulting data were not ideal for examining associations between diet, health and disease markers, and other end points (81). Additional dietary assessment has been undertaken within CPTP, but not consistently across the cohorts (Box 4.1). With technological advances in dietary assessment (51,65), it is now possible for CPTP and other large-scale research initiatives to collect comprehensive dietary data efficiently and affordably, opening many avenues of inquiry related to diet and health (9).

Box 4.1 Dietary Tools used in CPTP Cohorts

Alberta's Tomorrow Project (21)
- CDHQ-I (2000-2008) (23)
- Fruit and vegetable screener (2009-2015) (81)
- ASA24 & CDHQ-II (feasibility testing in a subsample) (68,52)

Atlantic PATH (109)
- Fruit and vegetable screener
- Multifactor screener and frequency questionnaire -- questions specific to the Atlantic population

BC Generations Project (110)
- Fruit and vegetable screener

CARTaGENE (111)
- CDHQ-II
- Fruit and vegetable screener

Ontario Health Study (112)
- Fruit and vegetable screener

The considerations outlined below are based on current evidence for reducing bias and estimating diet-disease relationships with as much precision as possible—their implementation relies upon the use of technology-enabled tools mentioned above and described in the Appendices. Implementation of these approaches will require the development of capacity not only in the collection of dietary data, but also their appropriate analysis using cutting-edge statistical techniques that continue to emerge.

Selecting Dietary Assessment Tools

Since there are few measures available that capture intake without bias (9,44), the aim in any given study should be to capture the least-biased data on dietary exposures (17,18); current evidence suggests this can be best achieved in cohort studies by collecting multiple 24HRs along with an FFQ (82).

Combining tools appears to be an ideal approach because it provides maximum analytical flexibility (i.e., strategically capitalizing on the strengths of each tool, while minimizing limitations) (18). In a simulation study, the combination of short- and long-term dietary data was shown to increase the precision of estimates as well as the power to detect relationships between dietary exposures and disease-related outcomes (17). Results of the simulation study suggest the use of four to six 24HRs over the course of a year, along with the administration of an FFQ (17), for epidemiologic research. A more recent validation study showed that multiple recalls performed better than an FFQ in terms of the accuracy of estimates of intake of energy, protein, potassium, and sodium (83). However, FFQ data are valuable for providing information on episodically-consumed foods and beverages (17). The FFQ assesses long-term intake with a single administration, while the number of repeat 24HRs that would be required to obtain similar information on episodically-consumed foods may be impractical (17).

If it is not possible to collect 24HRs from the full sample, the FFQ can be used as the main tool, with administration of a less-biased tool in a subsample to allow the use of statistical techniques to partially correct for error (40,84)**.** In this case, 24HRs (or biomarkers, if possible) can be used as the reference tool, allowing calibration to reduce bias in the FFQ data (13). The reference data are collected in a calibration sub-study, which can be either internal or external to the main study (see SECTION 5) (13). When using 24HRs, a repeat administration is needed to account for within-person variation.

In combining tools, the timing of administration of tools is important (85). Collections of 24HRs should be spread out over the study period since intake can vary from day to day and across seasons, while the FFQ should be administered towards the end of the study period (85). If investigators are interested in measuring changes over time, additional assessments should be conducted throughout the duration of the study (17,85).

Feasibility Testing in Alberta's Tomorrow Project

The feasibility of administering a combination of dietary assessment tools was examined in a study in a subset of Alberta's Tomorrow Project participants, who were asked to complete four administrations of the web-based ASA24 and one administration of the online DHQ II over a four month time period (52). More than half of participants completed two or more recalls, and those who completed a greater number of recalls were more likely to also complete the DHQ II (52). Using technology to administer 24HRs may introduce new issues, such as computer literacy (54) (see appendices for considerations to overcome

In addition, Freedman et al. (2010) have proposed combining self-report dietary data with concentration biomarker data for improving statistical power to detect diet-disease relationships in cohort studies (69). This may be a consideration for the future.

Analytic Considerations

Carefully considering the dietary assessment tool(s) to be used helps to lay the groundwork for statistical analyses to be conducted and associated study design considerations. For example, if the 24HR is used as the main dietary tool, investigators will ideally plan for multiple administrations and include an administration of a FFQ to all study participants to capture episodically-consumed foods and beverages (17). Alternatively, the use of regression calibration (described further below) to mitigate measurement error when a FFQ is used as the main dietary tool for data collection will require that a reference method (84), such as 24HRs or recovery biomarkers, be collected from a subsample of study participants (13). Strategic planning in collaboration with statisticians well in advance of study implementation is strongly recommended. Optimal analyses must be matched to the data and research question, but general approaches are briefly described below.

24-hour recalls as the main tool for assessing diet-disease relationships

As previously discussed, the use of a 24HR as the main dietary assessment tool in large cohort studies has been made possible by technological advancements (51,65). When data from both recalls and a FFQ are available for the full sample, FFQ data are not used as reference measures (as would be the case when a FFQ is the main tool and 24HR data are the reference) but instead are used as covariates to help predict true usual intake (14,82).

Several methods have been developed to model usual intake using 24HR data. Among these, the National Cancer Institute Method is widely used and can be applied to predict usual intake estimates for use in regression analyses (46,86), with inclusion of variables from the FFQ as covariates. Details and code to support the implementation of the National Cancer Institute Method are available at https://epi.grants.cancer.gov/diet/usualintakes/method.html.

Alternatively, regression calibration techniques can be implemented to combine 24HR and FFQ data (17), but the method is more complex in comparison to circumstances in which an FFQ is the main study tool. This approach is discussed in more depth in the National Cancer Institute's Measurement Error Webinar Series (https://epi.grants.cancer.gov/events/measurement-error/). It is expected that analytic approaches for cases in which 24HRs represent the main tool in the context of epidemiologic research will continue to develop.

Food frequency questionnaires as the main tool for assessing diet-disease relationships

As previously noted, FFQs have traditionally been used as the main dietary assessment tool in large-scale cohort studies (19). However, a reference measure should be used in tandem to reduce bias in the FFQ data (84). The most commonly-used approach to analyze data in this situation is regression calibration (13,18) (Box 4.2). To enable this approach, a calibration sub-study in which both the FFQ and the reference tool are administered is needed (13,39). The calibration sub-study can be internal or external to the main study. If the sub-study is external, then the study population and the main dietary assessment tool should be similar to those in the main study (18). The ideal reference tool used in the sub-study would capture dietary intake without bias (84). However, in the absence of recovery biomarkers or other unbiased measures of intake for most nutrients, less biased tools (mainly 24HRs) are used to provide reference data (18). Although imperfect, using a less biased measure has been shown to be preferable to ignoring measurement error (13). Collection of multiple administrations of 24HRs in at least a subsample is recommended to enable statistical approaches to address within-person variation in intake. The subsample should be representative of the overall sample. The size of the subsample needed is dependent on the dietary component of interest, with a larger number of recalls necessary to capture multiple instances of consumption of more episodically-consumed components(87). This is described in more detail elsewhere (87).

Energy adjustment

Energy adjustment is needed to control for confounding that may result if energy intake is associated with disease risk and can partially mitigate the effects of measurement error in self-report dietary data (39,43). Energy adjustment is a method in which dietary components are evaluated in relation to total energy intake (18). The most commonly used methods for energy adjustment are the nutrient density and residual methods (18). For example, nutrient density can be calculated by dividing nutrient and food group intakes by total energy intake and expressing values either as a proportion (e.g., % kcal from fat) or in units relevant to the particular nutrient (intake/1000 kcal) (18). Energy adjustment accounts for factors that affect energy requirements (e.g., metabolism, body composition, physical activity) (43) and can provide a measure of diet composition, while controlling for confounding that may result if total energy intake is associated with the disease outcome of interest (10,18,43). Further, because error in energy reporting is correlated with error in other dietary components (10,39,43), energy adjustment can be used to partially mitigate the effects of measurement error in self-report dietary data.

Box 4.2 Considerations Related to Regression Calibration

- Regression calibration is used to alleviate biased estimation of associations between diet and disease outcomes caused by measurement error in self-report tools (13,18).
    - The relationship between true intake and observed intake must be known and can be obtained by conducting a calibration sub-study (13).
    - The sub-study can be internal or external to the main study, but if external, then both the study population and the main dietary assessment tool should be similar to those in the main study (18).
    - A random sample of participants from the main study must complete the main and reference tools (13); using a less biased measure as the reference tool is better than ignoring measurement error (13).

- Regression calibration *does not* recover the loss of power to detect diet-disease relationships (13). The following approaches have been suggested to remedy the loss of power associated with measuring dietary exposures affected by measurement error (13):
    - Increase study sample size or conduct a meta-analysis of existing cohort studies (13); however, these approaches will not account for unmeasured confounders that may lead to inaccurate interpretation of findings (36,113).
    - Use the least-biased tool possible to reduce the loss of power in large-scale cohorts (13,17).

Interpreting and Reporting Study Findings

Cohort studies are valuable undertakings that can provide critical information regarding relationships between diet and health in populations and subpopulations (9). To make meaningful contributions to the literature and to facilitate syntheses of evidence, reports of findings from such studies should include detailed descriptions of the methods used to assess dietary exposures and to analyze the resulting data. The Strengthening Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut) checklist provides insights into the details that should be reported along with findings from cohort studies (88). To ensure that relevant attention is paid to critical details beginning at the stage of study planning, collaborations among nutritionists, epidemiologists, and statisticians are strongly recommended to aid in the selection of assessment tools, implementation, analyses, and interpretation.

SECTION 5: FUTURE DIRECTIONS

Integrating robust dietary assessment into cohorts will help lay the groundwork for leveraging novel technologies, biomarkers, and emerging analytic approaches in the future. Indeed, the field of dietary assessment is one in which there is currently much innovation.

In addition to technological advances related to self-report tools described earlier, there are emerging technologies, such as image recognition (58,59) and natural language processing, which may advance dietary assessment. Image-based food recording is promising for automatic recognition of food items and estimation of portion sizes (58,59,65), while natural language processing may be used for transcribing spoken dietary records (89). Although these technologies are not yet ready for wide-scale implementation, investment in infrastructure for capturing and analyzing dietary exposures within CPTP will position the cohorts to take advantage of emerging technologies in the future. Such efforts will likely need to be accompanied by complementary strategies to 'bridge' between existing and novel methods such that data collected at different time points can be compared.

The application of metabolomics to nutrition research also shows promise (90,91). Metabolomics is the study of small-molecule metabolites found in biological fluids (e.g., blood, saliva and urine) (71) that are produced through the metabolism of foods, medicine, and/or toxins (71,90). Metabolomics can be applied in nutrition research to identify dietary patterns that may be markers of disease risk (71,90,91). Further, the discovery of new dietary biomarkers has emerged as a potential application of metabolomics. For example, linking dietary intake with metabolomic profiles has led to the identification of potential markers of meat and vegetable intake (91). However, there are challenges with the implementation of metabolomics in large-scale cohort studies (90,91). These include barriers to accurate quantification of metabolites obtained from tissues (e.g., liver, gut, kidney) that may be ethically difficult to access in population-based studies, as well as the time required for sample and data processing and the cost of current analytical techniques (e.g., mass spectrometry) to identify metabolites (92). Recent advances in high-throughput technologies that can simultaneously examine thousands of metabolites have made the application of metabolomics at a population level more feasible (92). Challenges to be addressed include the development and standardization of analytical techniques that can be applied to large-scale datasets (92).

There also continues to be an evolution in methods of analyzing data from nutritional epidemiology studies. Traditionally, a reductionist approach (e.g., investigating associations between one nutrient or other dietary component and a disease outcome at a time) has been applied to the examination of diet-disease relationships (93); however, this approach does not

account for the complex nature of diet (93). The complexity of the synergistic and opposing effects of nutrients and bio-active components in the foods and beverages we consume cannot be captured by studying single dietary components (94,95). Furthermore, diet may change over time, with implications for health, as the multivariate and dynamic nature of eating patterns both over the lifespan and at critical points might be relevant to disease risk. There is thus growing interest in embracing the complexities of eating patterns when examining diet-disease relationships (9,96). To facilitate such research, attention has turned to approaches to address the multidimensional and dynamic nature of dietary patterns (97,98). While traditional approaches have included cluster and factor analyses as well as the use of diet quality indices (99), Reedy et al. (98) have highlighted potential approaches that could be used to answer questions such as what combination of foods explains variation in a set of intermediate health markers (e.g., reduced rank regression) (100) or minimizes cancer risk (neural networks) (101) or what aspects of the diet are most strongly associated with reducing cancer risk (classification and regression trees) (102). To address dynamism, time-varying approaches are being considered to address the complexity of dietary data collection at critical points over time (85).

In sum, there are several exciting developments in terms of our capacity to collect and analyze data to shed light on how diet influences health and disease. To support the generation of a cohesive body of evidence on the diet-disease nexus, standardized collection of comprehensive dietary data across studies is needed.

SECTION 6: SUMMARY

Barriers to the collection of robust dietary intake data in cohort studies in Canada have arisen due to the lack of cost-effective, easy-to-administer, and appropriately tested tools (9). Traditionally, in Canada and elsewhere, FFQs or screeners have been used as the main study tools to assess dietary exposures in large-scale cohort studies (10). However, validation studies addressing the structure of measurement error in self-report tools have shown that data from 24HRs are less biased in comparison to food frequency questionnaires (36,39–41). To capture dietary intake comprehensively and with the least bias possible (13), it is suggested that multiple 24HRs should be used as the main dietary assessment measure, in combination with an administration of an FFQ (17). As noted earlier, the administration of 4 to 6 recalls over a year-long period has been suggested.

In the past, 24HRs have been cost-prohibitive for use on a large scale, therefore limiting their utility in cohort studies. However, advances in technology (65), including the advent of web-based automated self-report recalls (51,57), are making the use of recalls as the main study tool feasible. Of particular salience to Canadian cohort studies, web-based 24HRs and FFQs have been tailored to the Canadian context (66,103), with work ongoing to update nutrient and food databases to better reflect Canadian dietary intake (9). Statistical and other (e.g., deep learning) techniques are also emerging for cases in which recalls are used as the main dietary assessment measure (14,17) in cohort studies, as well as for addressing the complexity of dietary patterns.

In conclusion, the progress that has been made in adapting web-based dietary assessment tools for the Canadian context (9) provides an exciting opportunity for CPTP and other Canadian cohorts. It is now possible to collect comprehensive dietary intake data within the cohorts and analytic strategies to maximize the value of these data are emerging. With such data, researchers can consider how different dietary components and patterns relate to cancer and chronic disease risk. The potential to unearth relationships between diet and disease is an exciting endeavour with the potential to shape future messaging for public health policies and programs.

SECTION 7: RECOMMENDED RESOURCES

The following are available resources that will guide investigators with method selection and provide recommendations regarding strategies for data collection and analyses of dietary data:

- Dietary Assessment Primer: https://dietassessmentprimer.cancer.gov/
- Measurement Error Webinar Series: https://epi.grants.cancer.gov/events/measurement-error/
- Diet and Physical Activity Measurement Toolkit: https://www.dapa-toolkit.mrc.ac.uk/
- Diet Assessment Methodology: https://epi.grants.cancer.gov/dietary-assessment/Chapter%201_Coulston.pdf

The following websites are dedicated to ASA24 and the DHQ/CDHQ, providing information, instructions, and access to other resources:

- Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA24):
    - Canadian version: http://asa24.ca/
    - U.S. Version: https://epi.grants.cancer.gov/asa24/
- Diet History Questionnaire (DHQ-II and CDHQ-II): https://epi.grants.cancer.gov/dhq2/

The following resources provide guidance related to the analysis of dietary data to examine relationships between intake and an outcome:

- The National Cancer Institute Method and SAS Macros https://epi.grants.cancer.gov/diet/usualintakes/macros.html
- Estimating Usual Dietary Intake From National Health and Nutrition Examination Survey Data Using the National Cancer Institute Method https://www.cdc.gov/nchs/data/series/sr_02/sr02_178.pdf

The following provides guidelines for investigators on reporting the findings of nutritional epidemiological studies:

- Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): http://www.strobe-nut.org/

REFERENCES

1. Forouzanfar MH, Afshin A, Alexander LT, Anderson HR, Bhutta ZA, Biryukov S, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. Lancet. 2016;388:1659–724.
2. Institute for Health Metrics and Evaluation. Global Burden of Diesase country profiles - Canada. 2015.
3. Afshin A, Sur PJ, Fay KA, Cornaby L, Ferrara G, Salama JS, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. Lancet. 2019;393(10184):1958–72.
4. World Cancer Research Fund International. Our cancer prevention recommendations. 2016.
5. Canadian Cancer Society. Research in cancer prevention. 2017.
6. American Association for Cancer Research. AACR Cancer Progress Report 2016. Clincal Cancer Res. 2016;22(Suppl 1):S1–137.
7. Doiron D, Raina P, Fortier I. Linking Canadian population health data: Maximizing the potential of cohort and administrative data. Can J Public Heal. 2014;104(3):e258–61.
8. Manolio T, Bailey-Wilson J, Collins F. Genes , environment and the value of prospective cohort studies. Nature. 2006;7:812–9.
9. Labonté M-È, Kirkpatrick SI, Bell RC, Boucher BA, Csizmadi I, Koushik A, et al. Dietary assessment is a critical element of health research – Perspective from the Partnership for Advancing Nutritional and Dietary Assessment in Canada. Appl Physiol Nutr Metab. 2016;41(10):1096–9.
10. Subar AF, Freedman LS, Tooze JA, Kirkpatrick SI, Boushey C, Neuhouser ML, et al. Addressing Current Criticism Regarding the Value of Self-Report Dietary Data. J Nutr. 2015;1–7.
11. Archer E. Opinion: A Wolf in Sheep's Clothing. 2013.
12. Archer E. Discussion of " Dietary assessment is a critical element of health research – Perspective from the Partnership for Advancing Nutritional and Dietary Assessment in Canada " − Misrepresentations distort the scientific record. Appl Physiol Nutr Metab. 2017;42(84):2017.
13. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing With Dietary Measurement Error in Nutritional Cohort Studies. J Natl Cancer Inst. 2011;103(14):1086–92.
14. Freedman LS, Midthune D, Carroll RJ, Tasevska N, Schatzkin A, Mares J, et al. Using regression calibration equations that combine self-reported intake and biomarker measures to obtain unbiased estimates and more powerful tests of dietary associations. Am J Epidemiol. 2011 Dec;174(11):1238–45.
15. Freedman LS, Commins JM, Moler JE, Willett W, Tinker LF, Subar AF, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. Am J Epidemiol. 2015;181(7):473–87.
16. Freedman LS, Commins JM, Moler JE, Willett W, Tinker LF, Subar AF, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for

potassium and sodium intake. Am J Epidemiol. 2015;181(7):473–87.
17. Carroll RJ, Midthune D, Subar AF, Shumakovich M, Freedman LS, Thompson FE, et al. Taking advantage of the strengths of 2 different dietary assessment instruments to improve intake estimates for nutritional epidemiology. Am J Epidemiol. 2012;175(4):340–7.
18. National Cancer Institute US National Institues of Health. Dietary Assessment Primer. 2015.
19. Thompson FE, Subar AF. Dietary assessment methodology. In: Coulston AM, Boushey CJ, Ferruzzi MG, editors. Nutrition in the Prevention and Treatment of Disease. 3rd ed. New York: Academic Press; 2013. p. 5–46.
20. Gibson R. Principles of Nutritional Assessment. 2nd ed. New York: Oxford University Press; 2005.
21. Alberta's Tomorrow Project.
22. Borugian MJ, Robson P, Fortier I, Parker L, McLaughlin J, Knoppers BM, et al. The Canadian Partnership for Tomorrow Project: Building a pan-Canadian research platform for disease prevention. Can Med Assoc J. 2010;182(11):1197–201.
23. Robson PJ, Solbak NM, Haig TR, Whelan HK, Vena JE, Akawung AK, et al. Design, methods and demographics from phase I of Alberta's Tomorrow Project cohort: a prospective cohort profile. Can Med Assoc J Open. 2016;4(3):E515–27.
24. World Cancer Research Fund / American Institute for Cancer Research. Food, Nutrition, Physical Activity, and the Prevention of Cancer: a Global Perspective. American Institute for Cancer Research. Washington, DC; 2007.
25. Krebs-Smith SM, Subar AF, Reedy J. Examining Dietary Patterns in Relation to Chronic Disease: Matching Measures and Methods to Questions of Interest. Circulation. 2015 Sep;132(9):790–3.
26. National Cancer Institute. Extending Methods in Dietary Patterns Research: Workshop. Rockville, MD; 2016.
27. Office of Disease Prevention and Health Promotion. 2015–2020 Dietary Guidelines for Americans. 2015. p. 144.
28. World Cancer Research Fund. Diet, Nutrition, Physical Activity and Cancer: a Global Perspective: A summary of the Third Expert Report [Internet]. American Institute for Cancer Research. 2018. 116 p. Available from: http://gco.iarc.fr/today%0Ahttps://dietandcancerreport.org
29. Garriguet D. Overview of Canadians' Eating Habits. 2004.
30. Jessri M, Nishi S, L'Abbé M. Assessing the Nutritional Quality of Diets of Canadian Adults Using the 2014 Health Canada Surveillance Tool Tier System. Nutrients. 2015;7:10447–68.
31. Canada S. Canadian Community Health Survey (CCHS) - 2015. 2016.
32. Raina, Parminder S Wolfson C, Kirkland SA, Griffith LE, Oremus M, Patterson, Christopher Tuokko H, Penning, Margaret Balion CM, et al. The Canadian Longitudinal Study on Aging. Can J Aging. 2009;28(3):221–9.
33. Fortier I, Doiron D, Burton P, Raina P. Invited Commentary Invited Commentary : Consolidating Data Harmonization — How to Obtain Quality and Applicability ? 2011;174(3):261–4.

34. Fortier I, Raina P, Heuvel ER Van Den, Griffith LE, Craig C, Saliba M, et al. Maelstrom Research guidelines for rigorous retrospective data harmonization. 2016;1–13.
35. Freudenheim J, Marshall J. The problem of profound mismeasurement and the power of epidemiological studies of diet and cancer. Nutr Cancer. 1988;11(4):243–50.
36. Subar AF, Kipnis V, Troiano, Richard P Midthune D, Schoeller DA, Bingham S, Sharbaugh CO, et al. Using Intake Biomarkers to Evaluate the Extent of Dietary Misreporting in a Large Sample of Adults: The OPEN Study. Am J Epidemiol. 2003;158(1–13).
37. Beaton GH, Burema J, Ritenbaugh C. Errors in the interpretation of dietary assessments. Am J Clin Nutr. 1997;65(Suppl):1100S–1107S.
38. National Cancer Institute US National Institues of Health. Measurement Error Webinar Series. 2016.
39. Kipnis V, Subar AF, Midthune D, Freedman LS, Ballard-Barbash R, Troiano RP, et al. Structure of dietary measurement error: Results of the OPEN biomarker study. Am J Epidemiol. 2003;158(1):14–21.
40. Neuhouser M, Tinker L, Shaw P, Schoeller D, Bingham S, Horn L, et al. Use of recovery biomarkers to calibrate nutrient consumption self-reports in the Women's Health Initiative. Am J Epidemiol. 2008;167(10):1247–59.
41. Schatzkin A, Kipnis V, Carroll RJ, Midthune D, Subar AF, Bingham S, et al. A comparison of a food frequency questionnaire with a 24-hour recall for use in an epidemiological cohort study: results from the biomarker-based Observing Protein and Energy Nutrition (OPEN) study. Int J Epidemiol. 2003;32:1054–62.
42. Dodd KW, Guenther PM, Freedman LS, Subar AF, Kipnis V, Midthune D, et al. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. J Am Diet Assoc. 2006 Oct;106(10):1640–50.
43. Willett WC, Howe GR, Kushi L. Adjustment for total energy intake in epidemiologic studies. Am J Clin Nutr. 1997;65(Suppl):1220S–1228S.
44. Jenab M, Slimani N, Bictash M, Ferrari P, Bingham S. Biomarkers in nutritional epidemiology: applications, needs and new horizons. Hum Genet. 2009;125(5):507–25.
45. Freedman LS, Midthune D, Carroll RJ, Krebs-smith S, Subar AF, Troiano RP, et al. Adjustments to improve the estimation of usual dietary intake distributions in the population. Nutr Methodol. 2004;134:1836–44.
46. Tooze JA, Kipnis V, Buckman DW, Carroll RJ, Freedman LS, Guenther PM, et al. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: The NCI method. Stat Med. 2010;29(27):2857–68.
47. Rhodes DG, Murayi T, Clemens JC, Baer DJ, Sebastian RS, Moshfegh AJ. The USDA Automated Multiple-Pass Method accurately assesses population sodium intakes. Am J Clin Nutr. 2013;97(5):958–64.
48. Moshfegh AJ, Rhodes DG, Baer DJ, Murayi T, Clemens JC, Rumpler W V, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. 2008;(88):324–32.
49. Blanton CA, Moshfegh AJ, Baer DJ, Kretsch MJ. The USDA Automated Multiple-Pass Method accurately estimates group total energy and nutrient intake. J Nutr. 2006;136(10):2594–9.
50. NCC: Nutrition Coordinating Center. NDSR Software. 2016.

51. Subar AF, Kirkpatrick SI, Mittl B, Zimmerman TP, Thompson FE, Bingley C, et al. The Automated Self-Administered 24-Hour Dietary Recall (ASA24): A Resource for Researchers, Clinicians, and Educators from the National Cancer Institute. J Acad Nutr Diet. 2012;112(8):1134–7.
52. Solbak NM, Lo Siou G, Paek S, Al Rajabi A, Vena JE, Kirkpatrick SI, et al. Evaluating the feasibility of administering a combination of online dietary assessment tools in a cohort of adults in Alberta, Canada. 2017.
53. Gilsing A, Mayhew AJ, Payette H, Shatenstein B, Kirkpatrick SI, Amog K, et al. Validity and reliability of a short diet questionnaire to estimate dietary intake in older adults in a subsample of the Canadian longitudinal study on aging. Nutrients. 2018;
54. Kirkpatrick SI, Gilsing AM, Hobin E, Solbak NM, Wallace A, Haines J, et al. Lessons from studies to evaluate an online 24-hour recall for use with children and adults in Canada. Nutrients. 2017;9(100):1–13.
55. Jacques S, Lemieux S, Lamarche B, Laramée C, Corneau L, Lapointe A, et al. Development of a Web-Based 24-h Dietary Recall for a French-Canadian Population. Nutrients. 2016;8(724):1–12.
56. Lafrenière J, Lamarche B, Laramée C, Robitaille J, Lemieux S. Validation of a newly automated web- based 24-hour dietary recall using fully controlled feeding studies. BMC Nutr. 2017;3(34):1–10.
57. Subar A, Mittl B, Zimmerman, TP, Kirkpatrick S, Schap T, Miller A, Wilson M, et al. The Automated Self-Administered 24-Hour (ASA24) is now mobile and can collect both 24-Hour recalls and food Records. FASEB J. 2016;30(1).
58. Probst Y, Nguyen D, Tran M, Li W. Dietary Assessment on a Mobile Phone Using Image Processing and Pattern Recognition Techniques: Algorithm Design and System Prototyping. Nutrients. 2015;7(8):6128–38.
59. Zhang W, Yu Q, Siddiquie B, Divakaran A, Sawhney H. “Snap-n-Eat”: Food Recognition and Nutrition Estimation on a Smartphone. J Diabetes Sci Technol. 2015;9(3):525–33.
60. Khanna N, Boushey CJ, Kerr D, Okos M, Ebert DS, Delp EJ. An Overview of The Technology Assisted Dietary Assessment Project at Purdue University. IEEE. 2010;290–5.
61. Franco RZ, Fallaize R, Lovegrove JA, Hwang F. Popular Nutrition-Related Mobile Apps: A Feature Assessment. JMIR mhealth uhealth. 2016;4(3):385.
62. Lieffers JR, Vance VA, Hanning RM. Use of mobile device applications in Canadian dietetic practice. Can J Diet Pract Res. 2014;75(1):41–7.
63. Cook VE, Ellis AK, Hildebrand KJ, Hpte M. Mobile health applications in clinical practice: Pearls, pitfalls, and key considerations. Ann Allergy, Asthma Immunol. 2018;117(2):143–9.
64. Briefel R, Flegal K, Winn D, Loria C, Johnson C, Sempos C. Assessing the nation’s diet: limitations of the food frequency questionnaire. J Am Diet Assoc. 1992;92(8):959–62.
65. Stumbo PJ. New technology in dietary assessment: a review of digital methods in improving food record accuracy. Proc Nutr Soc. 2013;72(1):70–6.
66. National Cancer Institute US National Institutes of Health. Diet History Questionnaire II and Canadian Diet History Questionnaire II (CDHQ II).
67. Csizmadi I, Boucher BA, Siou G Lo, Massarelli I, Rondeau I, Garriguet D, et al. Using national dietary intake data to evaluate and adapt the US Diet History Questionnaire: the

stepwise tailoring of an FFQ for Canadian use. Public Health Nutr. 2016;1–9.
68. Siou G Lo, Csizmadi I, Boucher BA, Akawung AK, Whelan HK, Sharma M, et al. The Comparative Reliability and Feasibility of the Past-Year Canadian Diet History Questionnaire II : Comparison of the Paper and Web Versions. 2017;9:133.
69. Freedman LS, Kipnis V, Schatzkin A, Tasevska N, Potischman N. Can we use biomarkers in combination with self-reports to strengthen the analysis of nutritional epidemiologic studies? Epidemiol Perspect Innov. 2010;7(2):1–9.
70. Potischman N, Freudenheim JL. Biomarkers of Nutritional Exposure and Nutritional Status: An Overview. J Nutr. 2003;133:873S–874S.
71. Hedrick VE, Dietrich AM, Estabrooks PA, Savla J, Serrano E, Davy BM. Dietary biomarkers: advances, limitations and future directions. Nutr J. 2012;11(109):1–14.
72. Kaaks R, Ferrari P, Ciampi A, Plummer M, Riboli E. Uses and limitations of statistical accounting for random error correlations, in the validation of dietary questionnaire assessments. Public Health Nutr. 2002;5(6A):969–76.
73. Livingstone MBE, Black AE. Biomarkers of Nutritional Exposure and Nutritional Status. J Nutr. 2003;133(3):895S–920S.
74. Bingham SA. Urine Nitrogen as a Biomarker for the Validation of Dietary Protein Intake. J Nutr. 2003;133:921S–924S.
75. Tasevska N, Midthune D, Potischman N, Subar AF, Cross AJ, Bingham SA, et al. Use of the predictive sugars biomarker to evaluate self-reported total sugars intake in the Observing Protein and Energy Nutrition (OPEN) study. Cancer Epidemiol Biomarkers Prev. 2011 Mar;20(3):490–500.
76. Tasevska N, Runswick SA, McTaggart A, Bingham SA. Urinary sucrose and fructose as biomarkers for sugar consumption. Cancer Epidemiol Biomarkers Prev. 2005;14(5):1287–94.
77. Tasevska N. Urinary sugars—a biomarker of total sugars intake. Nutrients. 2015 Jul;7(7):5816–33.
78. Landberg R, Aman P, Friberg LE, Vessby B, Adlercreutz H, Kamal-Eldin A. Dose response of whole-grain biomarkers: alkylresorcinols in human plasma and their metabolites in urine in relation to intake. Am J Clin Nutr. 2009 Jan;89(1):290–6.
79. Thompson FE, Subar AF, Loria CM, Reedy JL, Baranowski T. Need for Technological Innovation in Dietary Assessment. J Am Diet Assoc. 2010;110(1):48–51.
80. Canadian Partnership for Tomorrow Project. Health and Risk Factor Questionnaire.
81. Csizmadi I, Kahle L, Ullman R, Dawe U, Zimmerman TP, Friedenreich CM, et al. Adaptation and evaluation of the National Cancer Institute's Diet History Questionnaire and nutrient database for Canadian populations. Public Health Nutr. 2007;10(1):88–96.
82. Kipnis V, Midthune D, Buckman DW, Dodd KW, Patricia M, Krebs-smith SM, et al. Modeling Data with Excess Zeros and Measurement Error: Application to Evaluating Relationships between Episodically Consumed Foods and Health Outcomes. Biometrics. 2009;65(4):1003–10.
83. Park Y, Dodd KW, Kipnis V, Thompson FE, Potischman N, Schoeller DA, et al. Comparison of self-reported dietary intakes from the Automated Self-Administered 24-h recall, 4-d food records, and food-frequency questionnaires against recovery biomarkers. Am J Clin Nutr. 2018 Jan;107(1):80–93.

84. Kipnis V, Carroll RJ, Freedman LS, Li L. Implications of a new dietary measurement error model for estimation of relative risk: application to four calibration studies. Am J Epidemiol. 1999;150(6):642–51.
85. Freedman LS, Midthune D, Dodd KW, Carroll RJ, Kipnis V. A statistical model for measurement error that incorporates variation over time in the target measure, with application to nutritional epidemiology. Stat Med. 2015;34:3590–605.
86. Tooze JA, Midthune D, Dodd KW, Freedman LS, Krebs-Smith SM, Subar AF, et al. A new method for estimating the usual intake of episodically- consumed foods with application to their distribution. J Am Diet Assoc. 2006;106(10):1575–87.
87. Kirkpatrick SI, Subar AF, Tooze JA. Statistical Approaches to Mitigate Measurement Error in Dietary Intake Data Collected Using 24-hour Recalls and Food Records/Diaries. In: Schoeller DA, Westerterp M, editors. Advances in the Assessment of Dietary Ingtake. Taylor & Francis Group; 2017. p. 25.
88. Lachat C, Hawwash D, Ocké MC, Berg C, Forsum E, Hörnell A, et al. Strengthening the Reporting of Observational Studies in Epidemiology — Nutritional Epidemiology (STROBE-nut): An Extension of the STROBE Statement. PLOS Med. 2016;13(6):1–15.
89. Lacson R, Long W. Natural language processing of spoken diet records (SDRs). In: AMIA 2006 Symposium Proceedings. 2006. p. 454–8.
90. Astarita G, Langridge J. An emerging role for metabolomics in nutrition science. J Nutrigenet Nutrigenomics. 2013;6(4–5):181–200.
91. O'Sullivan A, Gibney MJ, Brennan L. Dietary intake patterns are reflected in metabolomic profiles: Potential role in dietary assessment studies. Am J Clin Nutr. 2011;93(3):314–21.
92. Fearnley LG, Inouye M. Metabolomics in epidemiology: from metabolite concentrations to integrative reaction networks. Int J Epidemiol. 2016;1319–28.
93. Ocké MC. Evaluation of methodologies for assessing the overall diet: dietary quality scores and dietary pattern analysis. Proc Nutr Soc. 2013;72:191–9.
94. Newby PK, Tucker KL. Empirically Derived Eating Patterns Using Factor or Cluster Analysis: A Review. Nutr Rev. 2004;62(5):177–203.
95. Schulze MB, Hoffmann K. Methodological approaches to study dietary patterns in relation to risk of coronary heart disease and stroke. Br J Nutr. 2006;95:860–9.
96. Dietary Guidelines Advisory Committee. Part D. Chapter 2: Dietary patterns, foods and nutrients, and health outcomes. Scientific Report of the 2015 Dietary Guidelines Advisory Committee. 2015.
97. Liese AD, Krebs-Smith SM, Subar AF, George SM, Harmon BE, Neuhouser ML, et al. The Dietary Patterns Methods Project: Synthesis of Findings across Cohorts and Relevance to Dietary Guidance. J Nutr. 2015;145(3):393–402.
98. Reedy J, Wirfält E, Flood A, Mitrou PN, Krebs-Smith SM, Kipnis V, et al. Comparing 3 dietary pattern methods-cluster analysis, factor analysis, and index analysis-with colorectal cancer risk. Am J Epidemiol. 2010;171(4):479–87.
99. Moeller SM, Reedy J, Millen AE, Dixon B, Newby P, Tucker KL, et al. Dietary Patterns: challenges and opportunities in dietary patterns research. J Am Diet Assoc. 2007;107(7):1233–9.
100. Hoffmann K, Schulze MB, Schienkiewitz A, Nöthlings U, Boeing H. Application of a New Statistical Method to Derive Dietary Patterns in Nutritional Epidemiology. Am J

Epidemiol. 2004;159(10):935–44.
101. Rothney MP, Neumann M, Béziat A, Chen KY. An artificial neural network model of energy expenditure using nonintegrated acceleration signals. J Appl Physiol. 2007;103(4):1419–1427.
102. Camp NJ, Slattery ML. Classification tree analysis: a statistical tool to investigate risk factor interactions with an example for colon cancer (United States). Cancer Causes Control. 2002;13:813–23.
103. ASA24-Canada.
104. ASA24.
105. Zimmerman TP, Potischman N, Douglass D, Dixit-joshi S, Kirkpatrick SI, Subar AF, et al. The Effect of Editing Open-Ended Text Responses on Nutrient and Food Group Estimates from the Automated Self-Administered 24- Hour Dietary Recall (ASA24). Procedia Food Sci. 2015;4:160–72.
106. Bryant H, Robson PJ, Ullman R, Friedenreich, Christine, Dawe U. Population-based cohort development in Alberta, Canada: A feasibility study. Chronic Dis Can. 2006;27(2):51–9.
107. Subar AF, Ziegler RG, Thompson FE, Johnson CC, Weissfeld JL, Reding D, et al. Is shorter always better? Relative importance of questionnaire length and cognitive ease on response rates and data quality for two dietary questionnaires. Am J Epidemiol. 2001;153(4):404–9.
108. Subar A, Thompson F, Kipnis V, Midthune D, Hurwitz P, McNutt S, et al. Comparative validation of the Block, Willett, and National Cancer Institute food frequency questionnaires: the Eating at America's Table Study. Am J Epidemiol. 2001;154(12):1089–99.
109. Atlantic Path Study.
110. BC Generations Project.
111. CARTaGENE: a new era of research.
112. Ontario Health Study.
113. Egger M, Schneider M, Davey-Smith G. Spurious precision? Meta-analysis of observational studies. Br Med J. 1998;3(16 (7125)):140–4.

APPENDIX A

Automated Self-Administered 24-Hour Dietary Assessment Tool-Canada (ASA24-Canada)
ASA24-Canada website: http://asa24.ca/
National Cancer Institute website with general information on ASA24:
https://epi.grants.cancer.gov/asa24/

Background: ASA24 was developed by the U.S. National Cancer Institute to make it possible to collect 24-hour recalls (24HRs) in large-scale research (51); it has recently been enhanced to enable the collection of food records (FRs) as well. ASA24 provides the ability to complete recalls and records on both laptops/desktops and mobile devices (smartphones or tablets).

A Canadian adaptation (ASA24-Canada) has been freely available since 2014 and uses an adapted database reflective of the Canadian food supply (based on the Canadian Nutrient File) (103). Food and beverage items unique to Canada have been added and items not available in Canada have been removed (103). Many changes made to develop the Canadian version reflect differences in brand name and restaurant items between Canada and the U.S. Additionally, metric measurements were added to aid with portion size estimation (see figure A.1) (103). ASA24-Canada includes a French language version.

The level of detail captured using ASA24, along with the robust and comprehensive database, is unmatched by commercial mobile dietary intake tracking applications. Additional information is available on the ASA24-Canada website and the National Cancer Institute's ASA24 website.

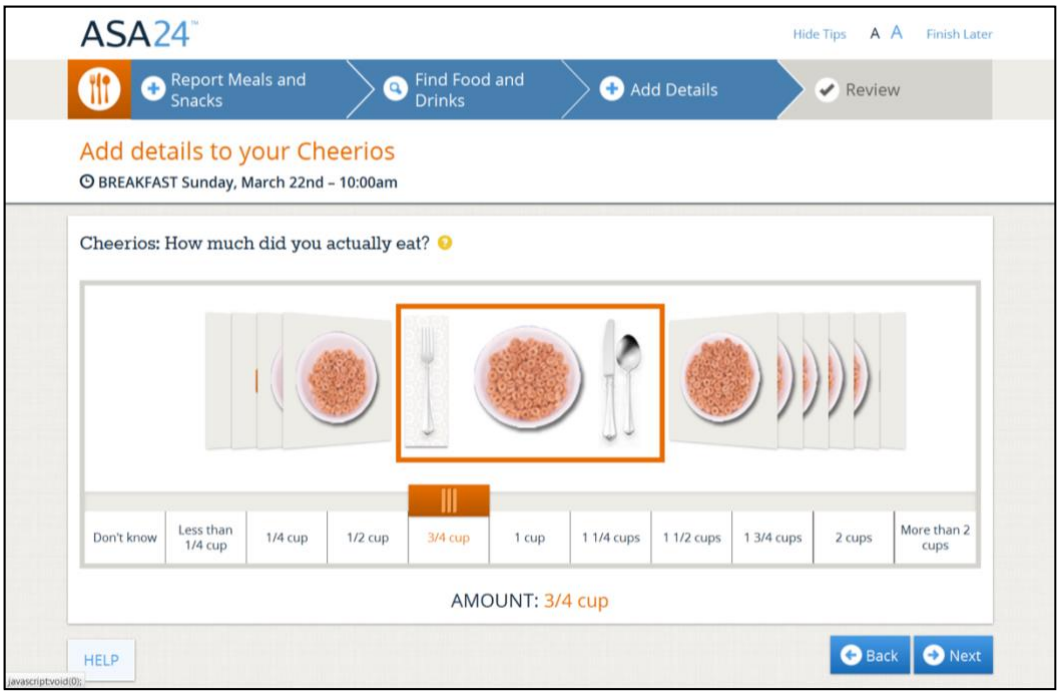

Figure A.1: Sample set of portion size options available in ASA24 with an associated visual representation.

Feasibility Testing: Alberta's Tomorrow Project (21) and the Canadian Longitudinal Study on Aging (CLSA) have completed feasibility studies with adult Canadian populations to inform the use of ASA24 in cohort studies (32,52,54). A random sample (n=331) of participants (ages 36-82) completed up to four 24HRs over four months, followed by the completion of the DHQ-II (52,54). Items adapted from the System Usability Scale were used to inquire into participants' experiences using the ASA24 system. In general, participants were receptive to completing ASA24 and indicated willingness to complete multiple recalls (54), but there are challenges to be considered. Using technology to administer 24HRs introduces new issues that investigators must consider during data collection (54), such as computer literacy, cognitive skills, user experiences, preference for type of devise used, and costs associated with training study staff to support participants when assistance is needed to complete ASA24 and to coordinate various aspects of data collection (e.g., recruitment, data cleaning) (54).

Practical Considerations for Administering ASA24-Canada within CPTP cohorts: The lessons learned from these studies and others informed practical considerations for administering ASA24-Canada within CPTP cohorts (54):

- Pilot test protocols in the specific population and setting to assess potential challenges with ASA24-Canada administration and the need for training of staff to provide support for participants to complete recalls.
- Though there are no costs associated with accessing ASA24-Canada, resources are needed to support participants and monitor ASA24 completion (e.g., training staff to

ensure they are familiar with ASA24-Canada, maintaining email accounts to provide help to participants).

- Provide participants with an opportunity to complete a demonstration of ASA24-Canada, if possible. This may alleviate challenges with navigating the system by allowing participants to become familiar with the steps involved in completing a 24HR.
- Ensure participants have realistic expectations regarding the time required to complete the recall. Consider which optional modules to include based on time available and respondent burden. For example, asking about the source of each food or beverage may add substantial time to a recall administration.
- If the recall is completed at home, multiple attempts should be allotted to increase the proportion of participants that will complete the recall.
- Take into consideration the study population and characteristics that may affect completion of ASA24-Canada (e.g., older adults may prefer the use of mobile/tablet devices instead of a desktop/laptop).
- Consider telephone administered recalls to assist participants in situations where they are unable to complete the recall independently (e.g., vision problems, computer illiteracy).
- Unscheduled recalls (i.e., participants can log in to system at any time) may increase the proportion of participants who complete recalls. *Please note:* Investigators must consider the increased risk of reactivity, such as participants' modifying recalls for the specified recall period at a later date.
- If feasible, be flexible in the availability of study staff to support participants in completion of ASA24-Canada in studies involving scheduled recalls (i.e., participants can only log in to the system on pre-determined dates).
- Provide sign-in credentials to participants close in time to completion of the recalls as possible to minimize login related issues. Consider either a) e-mailing sign-in information to participants to enable copy and pasting or b) mailing the information on bright colored paper or study-specific magnets.
- Ensure recalls are monitored by study staff on an ongoing basis, with particular attention paid to incomplete recalls.

Recommendations for Reviewing and Cleaning the Data: Investigators using ASA24-Canada should review data to determine if the data files contain missing data, outliers, or text entries with open-ended text (i.e., "match not found" or "other" are chosen by respondents) (103,104). The following document provides general guidelines regarding reviewing and cleaning the data: https://epi.grants.cancer.gov/asa24/resources/cleaning.html.

Outliers can occur when individuals report very high or low dietary intakes. Outliers may also indicate errors in reporting, coding, or databases used to estimate dietary intake (18). However, investigators should not necessarily eliminate outliers from the data set without examining them carefully. Applying cut-offs for the inclusion or exclusion of dietary data can potentially result in the loss of individual data that contain no more measurement error than those recalls

within the cut-off range (18). Disregarding data will also lead to a smaller sample size, resulting in the loss of study power to detect diet-disease relationships (18). In regard to text entries, it may not be feasible to check through each open-ended text entry. However, Zimmerman et al. (2015) suggest that this level of cleaning and recoding may not be necessary in a large study sample (105).

Any missing data, incorrect matches or outliers may require recoding and reanalysis of the data, which must be completed outside of the ASA24-Canada system (103,104). Investigators should consider the time and effort needed to clean the data based on the size of the study sample and the research question(s) of interest (103,104).

Database considerations: ASA24-Canada uses an adapted version of the U.S. ASA24 database that is reflective of the Canadian food supply (103). The database was adapted using the Canadian Nutrient File (CNF 2015) and a Canadian recipe database used for surveillance. For items within the CNF and the recipe database that were without complete nutrient profiles, appropriate Food and Nutrient Database for Dietary Studies (FNDDS) values were used instead. The dietary supplement database is based on Health Canada's Licensed Natural Health Products Database and is supplemented by the NHANES Dietary Supplement Database. The most recent version also links to the U.S. Department of Agriculture's Food Patterns Equivalents Database and can provide amounts of food groups and other dietary components, such as oils or added sugars (103).

APPENDIX B

Canadian Diet History Questionnaire II (CDHQ-II)
http://epi.grants.cancer.gov/dhq2/

Background: The original DHQ is a food frequency questionnaire that was developed by the National Cancer Institute. The DHQ-II was subsequently developed based on updated dietary data (66), with a Canadian adaptation (CDHQ-II) also made freely available to investigators. The Canadian questionnaire was modified to better reflect Canadian diets, which have been observed to be modestly different from those in the U.S. due to differences in food availability and fortification practices (67,81).

An earlier Canadian adaptation, the CDHQ-I (81), was used in Alberta's Tomorrow Project cohort to examine reported intake using the original DHQ database and a Canadian-modified database (106). Findings revealed significant differences in mean estimates for certain nutrients, such as folate, vitamins A, D (women only), iron and B_6. Although dietary habits may be similar between the U.S. and Canada, these findings highlighted the importance of using a tailored FFQ (and associated database) to accurately capture the dietary intake of Canadians (81).

The CDHQ-II food list was based on analyses of 24HRs completed by adults surveyed in the 2004 Canadian Community Health Survey (67). Most decisions pertaining to the inclusion and exclusion of food and beverage items were based on analyses of CCHS data; however, foods recently added to the Canadian market (e.g., bottled water with added vitamins and minerals) were also evaluated and where appropriate, added to the CDHQ-II (67). As well, nutrition experts and researchers were consulted to advise on the inclusion of foods that may be relevant to emerging new hypotheses in nutritional epidemiology and chronic disease research (67). The CDHQ-II consists of 165 questions in total, of which 153 are food questions (figure B.1), 10 are related to nutritional supplement use (vitamins, minerals and herbal supplements), with one question about the cooking of meat and another inquiring about vegetarian diets. One hundred and forty-three food questions inquire about portion sizes. The CDHQ-II is available in paper- or web-based versions in two formats, including past year and month, with portion size (66). Additional information is available on the National Cancer Institute's DHQ and CDHQ website. Work is underway to develop an online version of the CDHQ-III (expected in late 2019) using data on food consumption from the 2015 Canadian Community Health Survey. The comparability of data from the CDHQ-II and CDHQ-III is expected to be reasonable since they were developed using similar procedures and Canadian intake data.

Figure B.1: Example of a sample set of questions related to soft drink or pop consumption in the last 12 months. (66)

Reliability and Feasibility Testing: Pilot research examining the response rates of approximately 400 participants (aged 55-74) and 1000 (aged 20-70) for the original paper-based DHQ in two different studies showed that response rates varied from 70-85%, and these rates were not statistically different from rates reported by shorter FFQs used in other epidemiological studies (107,108). Technological advances have led to web-based versions of the DHQ; thus, Alberta's Tomorrow Project investigated the comparative reliability and feasibility of the paper-and-pencil and web-based versions of the CDHQ-II (68). Participants were randomly assigned to one of two groups: 1) paper, web, paper; and, 2) web, paper, web over a six-week time period. An overall response rate across all collection points and questionnaire groups was found to be 89.7% and 72.7% for the paper-and-pencil and web-based versions, respectively (68). Many participants indicated that they would be willing to complete the web-based version in the future, and 59% preferred completing the web-based version instead of the paper-and-pencil version. The study also demonstrated that the two versions of the CDHQ-II provided comparable nutrient estimates, indicating the flexibility for questionnaire administration in either paper or online formats (68). Both versions are described below (66).

Paper-Based vs. Web-Based Forms of the CDHQ-II (66)

- The **paper-and-pencil** format of the CDHQ-II asks about dietary intake in the past year and includes questions about portion size (66). Sample forms are provided and can be found at the following site: https://canadiandhqii.com/survey_resources.html. Please note that the sample forms *should not* be copied and completed by research subjects unless the researcher intends to input the data by hand. The codebooks for the paper forms are currently available in English; however, the codes and column values may be used with the French forms. The expenses associated with using this format are related to printing and scanning of the questionnaire. Typically, investigators contract commercial vendors to create custom forms and provide scanning services or use in-house IT resources. Researchers have the option to modify the CDHQ-II to suit their research needs if the following guidelines are met: https://epi.grants.cancer.gov/dhq2/forms/coding/ (66).

- The **web-based (DHQ*Web)** CDHQ-II is freely available and is identical in content compared to the paper-based format (66). The DHQ*Web was designed to prompt respondents to follow automated skip patterns. This supports the collection of high-quality dietary data because respondents are unable to complete the questionnaire if there are missing or inconsistent responses. The questionnaire is designed to be self-administered, but as with ASA24-Canada, staff support should be trained and ready to provide assistance for participants who may encounter challenges completing the questionnaire (66).

Recommendations for Coding and Analysis: The Diet*Calc analysis software is freely available for analysis of CDHQ-II data (66). An ASCII text file containing data from completed questionnaires must be created from either paper-based or web-based formats. If using paper-based forms, this text file can be created by a scanner or a data entry system, while the data file is created automatically when using the DHQ*Web. The format of each line of data are specified in the questionnaire's codebook and data dictionary (used by Diet*Calc) (66). The codebook and data dictionary are available at: https://epi.grants.cancer.gov/dhq2/forms/.

Guidelines for coding DHQ data have been developed: https://epi.grants.cancer.gov/dhq2/forms/coding/. Investigators should review questionnaire data for missing codes or error characters. A missing code indicates that the respondent did not answer a required question, while an error character indicates

Download Diet*Calc

The software is available for download at: https://epi.grants.cancer.gov/DHQ/dietcalc/download.html

that a respondent marked two or more responses to a question for which only one answer was required.

Diet*Calc will interpret CDHQ-II data by using the food frequency data collected and the food and nutrient database (a database that has the nutrient content for each food by sex and serving size) to provide these nutrient and food group intake estimates (66). The CDHQ-II nutrient and food database was created from the 2004 Canadian Community Health Survey (CCHS; Cycle 2.2) 24HR data. Similar foods reported in CCHS were grouped into categories to determine their contribution to the overall Canadian diet, and for inclusion into the food list (66).