Advancing methods to capture and analyze dietary patterns

by

Joy Hutchinson

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Doctor of Philosophy
in
Public Health Sciences

Waterloo, Ontario, Canada, 2024

© Joy Hutchinson 2024

Examining Committee Membership

The following served on the Examining Committee for this thesis. The decision of the Examining Committee is by majority vote.

External examiner	**DR. MATTHIAS SCHULZE** Professor Department of Molecular Epidemiology German Institute of Human Nutrition Potsdam-Rehbrücke
Supervisor	**DR. SHARON KIRKPATRICK** Professor School of Public Health Sciences Faculty of Health, University of Waterloo
Internal Members	**DR. MICHAEL WALLACE** Associate Professor Department of Statistics and Actuarial Science Faculty of Mathematics, University of Waterloo
	DR. BENOÎT LAMARCHE Professor École de nutrition Faculté des sciences de l'agriculture et de l'alimentation, Université Laval
Internal-External Member	**DR. HEATHER KELLER** Professor Department of Kinesiology and Health Sciences Faculty of Health, University of Waterloo

Author's Declaration

This thesis consists of material all of which I authored or co-authored: see Statement of Contributions included in the thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Statement of Contributions

Joy Hutchinson was the sole author for Chapters 1-3 and 8, which were written under the supervision of Dr. Sharon Kirkpatrick and were not written for publication.

This thesis is comprised in part by four manuscripts that have been published or have been prepared for publication.

The exceptions to sole authorship within this thesis are included below. Joy Hutchinson led or co-led each of the following manuscripts:

Research presented in Chapter 4:

Joy Hutchinson and Tabitha Williams are co-first authors on the manuscript developed from this research. Development of the initial screener was led by Joy Hutchinson and Sharon Kirkpatrick. To develop the screener, Joy Hutchinson reviewed existing screeners, mapped the 2019 Canada's Food Guide healthy food choices recommendations, and developed draft screener questions and response options. Tabitha Williams led cognitive interviews conducted in English, with Joy Hutchinson and Ailish Westaway conducting note taking during interviews. Alexandra Bédard led cognitive interviews conducted in French, with Camille Pitre conducting note taking during interviews. Health Canada led the face and content validity testing of the screener. Joy Hutchinson, Tabitha Williams, and Sharon Kirkpatrick consulted with Health Canada and expert advisors throughout the study process. They applied insights from expert advisors, cognitive testing, and face and content validity testing to revise questions to arrive at the final screener. Joy Hutchinson, Tabitha Williams, and Sharon Kirkpatrick led the drafting and revising of the manuscript. Expert advisors and collaborators, including Simone Lemieux, Kevin Dodd, Benoît Lamarche, Patricia Guenther, Jess Haines, Angela Wallace, Alicia Martin, Maria Laura da Costa Louzada, Mahsa Jessri, Dana Lee Olstad, Rachel Prowse, Janis Randall

Simpson, and Jennifer Vena, contributed to the conception and development of the screener and provided feedback throughout.

Citation:

Hutchinson JM*, Williams TE*, Westaway AM, Bedard A, Pitre C, Lemieux S, Dodd KW, Lamarche B, Guenther PM, Haines J, Wallace A, Martin A, da Costa Louzada ML, Jessri M, Olstad DL, Prowse R, Randall Simpson J, Vena JE, Kirkpatrick SI. 2023. Development of the Canadian Food Intake Screener to assess alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations. *Applied Physiology, Nutrition, and Metabolism*, 48(8), 603-619. doi:10.1139/apnm-2023-0019. *Co-first authors.

Research presented in Chapter 5:

Sharon Kirkpatrick, Joy Hutchinson, Kevin Dodd, Patricia Guenther, Benoît Lamarche, Jess Haines, and Angela Wallace conceived and designed this study. Data collection was led by Joy Hutchinson and Angela Wallace. Joy Hutchinson led the data cleaning, analysis of data, and designed the scoring system for the screener, with support for cleaning, analysis, scoring, and interpretation provided by Sharon Kirkpatrick, Kevin Dodd, Patricia Guenther, and Benoît Lamarche. Joy Hutchinson and Sharon Kirkpatrick drafted and revised the manuscript for publication. Expert advisors and collaborators, including Kevin Dodd, Patricia Guenther, Benoît Lamarche, Jess Haines, Angela Wallace, Maude Perreault, Tabitha Williams, Maria Laura da Costa Louzada, Mahsa Jessri, Simone Lemieux, Dana Lee Olstad, Rachel Prowse, Janis Randall Simpson, Jennifer Vena, and Kathleen Szajbely provided feedback and support throughout.

Citation:

Hutchinson JM, Dodd KW, Guenther PM, Lamarche B, Haines J, Wallace A, Perrault M, Williams TE, da Costa Louzada ML, Jessri M, Lemieux S, Olstad DL, Prowse R, Randall Simpson J, Vena JE, Szajbely K, Kirkpatrick SI. 2023. The Canadian Food Intake Screener for

assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: Scoring system and construct validity. *Applied Physiology, Nutrition, and Metabolism*, 48(8), 620-633. doi:10.1139/apnm-2023-0018

Research presented in Chapter 6:

Sharon Kirkpatrick conceived of the review and planned it with the full team of co-authors. Amanda Raffoul, Alexandra Pepetone, Sanaa Hussain, and Sharon Kirkpatrick conducted the search and screening. Joy Hutchinson and Tabitha Williams conducted extraction and Lesley Andrade conducted verification. Joy Hutchinson synthesized results and led the writing of the manuscript. Alexandra Peptone contributed to writing the methods section of the manuscript. All co-authors provided critical input to the manuscript and approved the final manuscript.

Citation:

Hutchinson JM, Raffoul A, Pepetone A, Andrade L, Williams TE, McNaughton S, Leech R, Reedy J, Shams-White MM, Vena JE, Dodd KW, Bodnar L, Lamarche B, Wallace MP, Deitchler M, Hussain S, Kirkpatrick SI. Advances in methods for characterizing dietary patterns: A scoping review. *medRxiv*. Preprint: 10.1101/2024.06.20.24309251 (Submitted to BJN).

Research presented in Chapter 7:

Joy Hutchinson designed this study. Joy Hutchinson led analyses with support for interpretation provided by Sharon Kirkpatrick, Dylan Spicker, Benoît Lamarche, Michael Wallace, Abel Torres-Espín, and Mélina Côté. Joy Hutchinson drafted the final manuscript with critical revisions provided by co-authors.

Citation:

Hutchinson JM, Spicker D, Lamarche B, Wallace M, Côté M, Torres-Espín A, Kirkpatrick SI. Multidimensional dietary patterns and their joint associations with intersecting sociodemographic characteristics among adults in Canada.

Abstract

Background: Suboptimal diet quality is linked to poor health outcomes and is associated with many sociodemographic characteristics, including several that are indicators of inequities. Assessments of dietary intake have shifted over time from investigations of single foods and nutrients towards dietary patterns. This change has led to developments in methods to capture and analyze dietary patterns, from short tools that quickly assess the overall diet to novel analytic methods. These methodological advances present opportunities to better understand dietary patterns in Canada and globally.

Research objectives: The objectives of this research were to **(1)** develop a brief dietary screener to assess alignment of dietary intakes with the 2019 Canada's Food Guide healthy food choices recommendations; **(2)** develop a scoring system for the screener and assess the construct validity of the screener; **(3)** identify novel methods used to characterize dietary patterns through a scoping review of the literature; and **(4)** explore the capacity of probabilistic graphical models to expand our understanding of the joint relationships between multidimensional dietary patterns and intersecting sociodemographic characteristics.

Methods and results: The first study in this dissertation (**Chapter 4**) discusses the process to develop the Canadian Food Intake Screener. This was achieved by mapping the dietary guidance in the 2019 Canada's Food Guide and reviewing existing tools to develop a draft screener, which was reviewed by Health Canada and external collaborators (n=15). The screener was revised iteratively based on feedback from cognitive testing conducted among adults aged 18-65 years in English (n=17) and French (n=16) and from face and content validity testing conducted with experts (n=16). The screener was well understood overall and testing informed refinement to finalize the Canadian Food Intake Screener, which includes 16 questions to rapidly assess alignment of adults' intake with the 2019 Canada's Food Guide healthy food choices recommendations.

The second study in the dissertation (**Chapter 5**) was conducted to develop a scoring system for the screener and evaluate the screener's construct validity among adults aged 18 to 65 years. Analysis of variance (ANOVA) was used to compare screener scores among subgroups with known differences in diet quality. The correlation between scores on the screener and the Healthy Eating Food Index-2019 (HEFI-2019), which also assesses alignment of intake with the 2019 Canada's Food Guide healthy food choices recommendations, was assessed. Adults aged 18-65 years (n=154) completed the screener, answered a range of questions about their health and sociodemographic characteristics, and completed up to two 24-hour dietary recalls. The mean screener score was 35 points (SD = 4.7; maximum 65), ranging from 25 (1st percentile) to 45 (99th percentile). Meaningful differences in screener scores were observed in hypothesized directions by gender identity ($p = 0.06$), perceived income adequacy ($p = 0.07$), education ($p = 0.02$), and smoking status ($p = 0.003$). The correlation between screener and HEFI-2019 scores was 0.53 (SE = 0.12). The screener demonstrated moderate construct validity, indicating that it is appropriate for use for rapid assessment of alignment of adults' intake with the healthy food choices recommendations when comprehensive dietary assessment is not possible.

In **Chapter 6**, novel methods used to characterize dietary patterns in peer-reviewed literature were summarized using a scoping review. The databases MEDLINE, CINAHL, and Scopus were searched using keywords such as such as machine learning, latent class analysis, and least absolute shrinkage and selection operator (LASSO) to identify novel methods used to describe dietary patterns. Of 5274 records identified, 24 met the inclusion criteria. Twelve of 24 articles were published since 2020. A range of methods was applied to identify dietary patterns, with nine studies using approaches that have applications in machine learning to characterize dietary patterns, and the remaining 15 using other novel methods such as latent class analysis, LASSO, or treelet transform. Future work to guide the application, interpretation, and comparability of these methods is necessary to enable synthesis of the literature to inform policies and programs.

The final study (**Chapter 7**) in this dissertation examined the ability of probabilistic graphical models to explore the joint relationships between dietary patterns and sociodemographic characteristics. While prior research has established relationships between dietary patterns and sociodemographic characteristics, it has rarely considered the multidimensional relationships between dietary components or possible intersecting relationships among sociodemographic characteristics. Mixed graphical models, a network method, enable explorations of these complex joint relationships, which have largely been unexplored in the Canadian context. We conducted a secondary analysis of first 24-hour dietary recalls collected from adults aged 18 years and above who participated in the 2015 Canadian Community Health Survey Nutrition (n=14 097). Mixed graphical models were used to identify joint relationships between amounts consumed in grams of thirty log-transformed food groups and age, sex, education, income, household food security status, geographic region, employment status, and smoking status. Sociodemographic characteristics formed a network, with several pairwise relationships. Several dietary components also formed networks, often patterning by food group. Age and sex were the sociodemographic characteristics most strongly connected to dietary components. This research applied mixed graphical models to provide deeper insights into the internal structures of the dietary patterns of adults in Canada, and how sociodemographic characteristics are jointly related with dietary patterns. Probabilistic graphical models offer promise to complement existing methods to characterize dietary patterns, such as indices.

Conclusions: This dissertation makes contributions to dietary patterns research with respect to both collecting data reflective of overall dietary patterns and analytic methods to capture their complexity. The advances from these studies can be applied to inform targeted research and policies promoting public health nutrition in Canada and beyond.

Acknowledgments

I am grateful to have received a Vanier Canada Graduate Scholarship (Social Sciences and Humanities Research Council) to fund my PhD studies. I am also grateful for additional research and travel funding from the University of Waterloo to support my learning, including the Provost Doctoral Entrance Award for Women.

Thank you to my committee members and examiners, Dr. Sharon Kirkpatrick, Dr. Benoît Lamarche, Dr. Michael Wallace, Dr. Heather Keller, and Dr. Matthias Schulze. I am grateful to have learned from and worked with such thoughtful and attentive scholars throughout my PhD.

Thank you first and especially, to Dr. Sharon Kirkpatrick, for modeling excellence in everything that you do. When I reflect on the growth that I have been able to achieve throughout this degree, so much of what I have learned is directly attributable to your wisdom. I am grateful for the many opportunities for collaboration that you have provided, for how you have supported me to grow as a researcher, and for how you helped me to cultivate curiosity. Thank you for supporting me to pursue my interests and for guiding me towards avenues that I may not have considered. I have most enjoyed our one-on-one chats, when we both get excited about a niche area of diet assessment and dream up a new project that we inevitably don't have time for (yet).

Thank you to the mentors who have added layers of richness to my degree. Particularly, I would like to thank my committee members, Dr. Benoît Lamarche and Dr. Michael Wallace, for the many insights you have provided throughout my studies. It has been a privilege to learn from you and to have been supported by such generous collaborators who are always willing to weigh in, push me to consider different perspectives, and spark excitement for new ideas. I am also grateful for the mentorship of Dr. Kevin Dodd throughout my degree. Some of my favorite memories during my PhD were spent on Zoom troubleshooting bugs in code together. Thank you to the advisors for the Canadian Food Intake Screener and the diet/AI working group for the

encouragement, mentorship, and wisdom you provided to strengthen the research in this dissertation and for supporting my development as a researcher.

Thank you to my peers – Alex, Alyssa, Lesley, Martin, Navreet, Lisa, Tabi, Mona, Carolyn, and Kirsten – for your support throughout this journey. Starting a PhD during the height of the COVID pandemic certainly had its challenges, but I always felt supported by this team. It is a gift to work with so many talented and kind researchers and I look forward to maintaining friendships and collaborations. Particularly, I feel lucky to say that I have looked forward to Monday mornings because of weekly writing groups, which have helped me to accomplish writing-related goals and have provided support throughout the PhD journey. I am especially grateful to my office buddy, Alyssa, for always being open to taking a trip to Grad House or a café and for all the long pauses you took from your work to chat about research or other, less focused topics.

Thank you to the other exceptional past and present students within the School of Public Health Sciences and the University of Waterloo broadly. I have thoroughly enjoyed working with such collaborative students, especially through the SPHS peer mentorship program, the SPHS Conference, and the diet/AI working group.

To my friends, I appreciate you. Thank you for being available, and for the many ways you have shaped my life. To my family, I love you and am grateful for your continuing support and belief.

It is a privilege to do this research. I feel grateful to be able to spend my time working with people and on projects that are deeply interesting and feel meaningful. I could not have reached this point without the support of mentors, friends, and family, and I am exceptionally grateful for the role that each of you have played to help me follow this path to completion.

Most of all, thank you to Caleb. I could not have done this without you. Thank you for always reminding me that there is life beyond work. Here's to our future.

Table of Contents

Examining Committee Membership ii

Author's Declaration iii

Statement of Contributions iv

Abstract vii

Acknowledgments x

List of Figures xv

List of Tables xvi

List of Abbreviations xvii

Statement of Positionality xviii

Chapter 1: Introduction 1

1.1. Overview and scope 1
1.2. Dissertation organization 2

Chapter 2: Literature Review 4

2.1. An overview of dietary assessment methods 4
2.2. Methods to capture self-reported dietary intake 5
2.3. Methods to capture dietary intake in the Canadian context 9
2.4. Methods to analyze dietary intake data 11
2.5. Shifts in assessment: From single components to dietary complexity 13
2.6. Associations between dietary intake and sociodemographic characteristics 23
2.7. Summary and rationale 25

Chapter 3: Research Objectives 28

Chapter 4: Development of the Canadian Food Intake Screener to assess alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations 30

4.1. Overview 32
4.2. Introduction 33
4.3. Methods 35
4.4. Results 51
4.5. Discussion 64
4.6. Conclusion 68
4.7. References 71

Chapter 5: The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: Scoring system and construct validity 78

5.1. Overview 80
5.2. Introduction 81
5.3. Methods 82
5.4. Results 96
5.5. Discussion 101
5.6. Conclusion 108
5.7. References 110

Chapter 6: Advances in methods for characterizing dietary patterns: A scoping review 117

6.1. Overview 119
6.2. Introduction 120
6.3. Methods 122
6.4. Results 124
6.5. Discussion 146
6.6. Conclusion 149
6.7. References 152

Chapter 7: Multidimensional dietary patterns and their joint associations with intersecting sociodemographic characteristics among adults in Canada 160

7.1. Overview 161
7.2. Introduction 162
7.3. Methods 165
7.4. Results 168
7.5. Discussion 178
7.6. Conclusion 183
7.7. References 185

Chapter 8: General Discussion 192

8.1. Overview 192
8.2. Summary of key findings 193
8.3. Contributions of the dissertation 196
8.4. Overall strengths and limitations of the dissertation 200
8.5. Implications for public health policy and research 204
8.6. Conclusion 207

References 209

Appendices 221

Appendix A: Advisors who provided input on the development of the Canadian Food Intake Screener 221

Appendix B: Initial questions in the Canadian Food Intake Screener/ Questionnaire court canadien sur les apports alimentaires to assess alignment of adults' intake with the Canada Food Guide-2019 healthy food choices recommendations 222

Appendix C: Canadian Food Intake Screener in English and French 233

Appendix D: Details about the Automated Self-Administered Dietary Assessment Tool (ASA24-Canada-2018) and cleaning of 24-hour dietary recall data 248

Appendix E: Canadian Food Intake Screener/Questionnaire court canadien sur les apports alimentaires: Scoring System READ ME File 252

Appendix F: Sample Code to Score the Canadian Food Intake Screener 254

Appendix G: Deriving Healthy Eating Food Index-2019 variables 257

Appendix H: Hypothesized differences in Canadian Food Intake Screener scores across subgroups based on known differences in diet quality 259

Appendix I: Distribution of total Healthy Eating Food Index-2019 scores[1] among adults aged 18-65 years in a study to examine the construct validity of the Canadian Food Intake Screener 260

Appendix J: Search strategies for MEDLINE, CINAHL, and SCOPUS 261

Appendix K: Key extraction fields for review of novel analytic methods to characterize dietary patterns 269

Appendix L: Directed acyclic graph of relationships between exposures of interest and dietary patterns 270

Appendix M: Bureau of Nutritional Science Condensed Variables for analysis of dietary patterns in the 2015 CCHS Nutrition 271

Appendix N: Weighted adjacency matrices 277

Appendix O: Bootstrapped confidence intervals of edge weights for the full network including sociodemographic characteristics and dietary components among adults ≥18 y in the 2015 Canadian Community Health Survey Nutrition 284

List of Figures

Figure 1. Process for development of the Canadian Food Intake Screener to assess alignment of adults' intake with the healthy food choices recommendations 36

Figure 2. Guiding principles for the development and evaluation of the Canadian Food Intake Screener 37

Figure 3. Flowchart describing participation in a study to assess the construct validity of the Canadian Food Intake Screener after applying inclusion and exclusion criteria 86

Figure 4. Distribution of Canadian Food Intake Screener scores among adults aged 18-65 years (n=154) in a study to examine the construct validity of the screener 99

Figure 5. PRISMA diagram illustrating the screening process for a scoping review exploring innovative methods for the analysis of dietary intake data and characterization of dietary patterns 125

Figure 6. Network of sociodemographic characteristics among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition 174

Figure 7. Network of dietary components among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition 175

Figure 8. Full network of dietary components and sociodemographic characteristics among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition 176

Figure 9. Centrality indices of variables included in the full network with sociodemographic and dietary components among adults ≥18 in the 2015 Canadian Community Health Survey Nutrition 177

List of Tables

Table 1: Summary of studies applying probabilistic graphical models to identify dietary patterns or joint relationships between dietary components 17

Table 2: Dietary guidance mapped to final questions within the Canadian Food Intake Screener, in English and French, to assess alignment of intake with the 2019 Canada's Food Guide healthy food choices recommendations among adults aged 18-65 years 41

Table 3: Characteristics of participants (n=33) in cognitive interviews in English and French to evaluate the comprehension and face validity of the Canadian Food Intake Screener 52

Table 4: Examples of issues identified in cognitive interviews in English and French to evaluate the comprehension and face validity of the Canadian Food Intake Screener[1] 57

Table 5: Scoring system for the Canadian Food Intake Screener 89

Table 6: Sociodemographic and health characteristics of adults (n=154) participating in a study to evaluate the construct validity of the Canadian Food Intake Screener 96

Table 7: Comparison of mean Canadian Food Intake Screener scores, of a possible 65 points, among subgroups known to differ in dietary quality among adults 18-65 years (n=154) 100

Table 8: Study characteristics across included studies applying novel methods to characterize dietary patterns 127

Table 9: Characteristics of studies (n=24) identified in a scoping review of novel analytic methods to characterize dietary patterns 129

Table 10: Description of dietary patterns (n=24) identified in a scoping review of novel analytic methods to characterize dietary patterns 133

Table 11: Novel methods applied to identify dietary patterns across included studies 144

Table 12: Characteristics of adults in the 2015 Canadian Community Health Survey Nutrition 169

Table 13: Mean one-day intakes of 30 food groupings used to identify dietary pattern networks among adults in the 2015 Canadian Community Health Survey Nutrition 170

List of Abbreviations

24HDR	24-hour dietary recall
ANOVA	Analysis of variance
ASA24	Automated Self-Administered 24-hour Dietary Assessment Tool
CCHS	Canadian Community Health Survey
CÉGEP	Collège d'enseignement général et professionnel
CFG-2019	2019 Canada's Food Guide
CI	Confidence interval
EBIC	Extended Bayesian Information Criterion
FFQ	Food frequency questionnaire
HEFI-2019	Healthy Eating Food Index-2019
LASSO	Least absolute shrinkage and selection operator
MCMC	Markov Chain Monte Carlo
NCI	National Cancer Institute
PUMF	Public Use Microdata Files

Statement of Positionality

Given that it is not possible to fully disentangle my social position from my work as a researcher, it is worth considering how my positionality may have influenced this dissertation research. During the development and evaluation of the screener in **Chapters 4** and **5**, my own preconceived ideas, biases, and expectations for the development and evaluation of a screener may have influenced study outcomes. This was the first time that I worked to develop and test a dietary assessment instrument, and there were many lessons that I learned throughout the process that I could apply to the development and evaluation of future tools. A majority of our study team including myself identified as White and women, and I recognize that this social position is generally overrepresented within the field of nutrition and dietetics, and that this overrepresentation could mean that we missed some perspectives in the development and evaluation of the screener.

During the scoping review, described in **Chapter 6**, we assessed dietary patterns characterized using novel methods. This review primarily sought to identify novel methods being used to characterize dietary patterns. However, we also extracted information on the extent to which studies considered sociodemographic characteristics that could be indicators of equity and described how these were incorporated into analyses of dietary patterns. We took care to incorporate sociodemographic characteristics as a consideration within this review after reading critiques noting the potential for novel methods like machine learning to widen inequities.(1–3) We limited our extraction related to equity to how sociodemographic characteristics were incorporated into analyses to assess relationships with dietary patterns, but it is important to consider that methods for data collection and sample selection also inform potential bias in a study.(1) As a researcher beginning to engage with these methods, I may not have fully captured or considered all potential ways in which analyses may have been less equitable than they could have been.

For the analysis of the 2015 Canadian Community Health Survey Nutrition (**Chapter 7**), I reflect that I expected to find joint relationships between dietary patterns and sociodemographic characteristics. I am a researcher who is still early in my journey of engaging with and learning about novel methods, and I acknowledge my responsibility to learn about the potential for these methods to widen inequities and the importance of developing models responsibly to explore questions of public health relevance that actively promote equity. I acknowledge the privilege that comes with being a researcher and the responsibility that I must use evidence from this research to advocate for reducing inequities where possible. It is not possible to conduct research that is totally objective, and I sought to honestly reflect on the factors that may have influenced my work.

Chapter 1: Introduction

1.1. Overview and scope

Dietary patterns are inextricably linked to health and wellbeing.(4–7) Additionally, several sociodemographic characteristics, such as lower income or educational attainment or living in a food-insecure household, are associated with poorer diet quality.(8–11) Several methods have been developed to capture and analyze dietary intake data.(12,13) Methods to capture these data include food frequency questionnaires, 24-hour dietary recalls (24HDRs), food records, and dietary screeners.(12) Each of these self-report assessment methods has associated strengths and limitations that make them amenable to specific research contexts. In some cases, shorter assessment tools may be desirable for feasibility, whereas in other contexts, nuanced dietary intake data collected using more detailed assessment tools can enable the application of complex methods to better understand facets of dietary intake. For example, brief screeners may be more desirable for situations in which resources are limited or when researchers wish to minimize participant burden, whereas a 24-hour dietary recall requires more resources but provides greater detail and has less systematic measurement error.(14)

When analyzing dietary intake data, statistical approaches have traditionally been applied to investigate differences and changes in dietary exposures. However, such approaches are limited in their ability to capture potential synergistic or antagonistic relationships that form dietary patterns.(15,16) Novel methods, including machine learning, are starting to be applied to address nutrition questions to explore dietary patterns and relationships with health and other factors.(13,17,18) This has introduced several possibilities for dietary assessment, with opportunities to explore dietary patterns in ways that better reflect their totality. However, novel methods do not always provide added benefit compared to traditional approaches,(18) and desired complexity must be carefully weighed with ease of interpretability.(19)

Robust dietary assessment relies on both appropriately captured data and well-matched analytic approaches, and it is thus important to consider both how dietary data are captured and analyzed to assess dietary patterns. This dissertation seeks to advance diet assessment methods at the population-level through the development and evaluation of a brief dietary screener assessing alignment with national dietary guidelines, a scoping review of the literature to identify novel methods being used to characterize dietary patterns, and a study exploring the potential of a network method to expand our understanding of the relationship between multidimensional dietary components and intersecting sociodemographic characteristics. Much of this dissertation is rooted in the Canadian context; however, the findings provide insights for dietary assessment broadly. Overall, the goal of this dissertation is to progress methods that capture and analyze diet to advance our understanding of dietary patterns that reflect the heterogeneity within the population.

1.2. Dissertation organization

Chapter 2 provides an overview of the literature, with consideration of approaches for dietary assessment, methods to capture and analyze dietary intake data, shifts in assessment from single components to dietary patterns analysis, dietary assessment in the Canadian context, and associations between dietary intake and sociodemographic characteristics. This chapter is concluded with a summary and rationale for the dissertation.

This is followed by **Chapter 3**, which provides the objectives of the four studies in this dissertation, described in **Chapters 4** to **7**.

Chapters 4 to **7** include manuscripts that have been published, submitted, or are prepared for publication. **Chapter 4** describes the development of a brief dietary screener to assess alignment with the 2019 Canada's Food Guide healthy food choices recommendations

among adults. **Chapter 5** describes the evaluation of this dietary screener and its scoring system among adults. **Chapter 6** describes a scoping review of studies that characterized dietary patterns using novel methods. **Chapter 7** describes a study that uses a novel method of assessment, probabilistic graphical models, to explore joint relationships between dietary components and sociodemographic characteristics. The four manuscript chapters aimed to: (1) develop a brief dietary screener assessing alignment with the 2019 Canada's Food Guide healthy food choices recommendations; (2) score and conduct construct validity testing[1] of this screener; (3) identify novel methods used to characterize dietary patterns; and (4) explore how dietary patterns are associated with sociodemographic characteristics among adults in Canada using probabilistic graphical models.

Findings from this dissertation are synthesized in **Chapter 8**, with an overview of the results from each manuscript, contributions of the dissertation, strengths and limitations, implications for public health policy and research, and conclusions. This dissertation advances our understanding of dietary patterns and explores different approaches to capture and analyze this complex construct. This research also demonstrates the need for clear principles for the application of dietary assessment tools and analytic approaches such that methods can be readily and appropriately applied. Finally, this dissertation highlights the importance of including researchers with expertise in dietary assessment methods in nutrition research, so that questions match methodological approaches.

[1] Construct validity testing is used to determine whether a tool is appropriate for its desired purpose based on whether it performs as expected.(20,21)

Chapter 2: Literature Review

2.1. An overview of dietary assessment methods

Poor diet is well-established as a key risk factor for chronic disease.(4,6,7) To understand dietary intake and how it relates to health and wellbeing, methods are required to capture this construct with reasonable accuracy. Diet assessment methods are foundational to nutritional sciences. Without carefully evaluated methods to assess how people eat, we would not be able to form conclusions about diet-health relationships, such as the relationship between saturated fat intake and cardiovascular disease.(22) Self-report methods like food frequency questionnaires have been integral to learning about this relationship.(22)

Methods to assess dietary intakes range from capturing specific dietary components to overall dietary patterns.(23,24) Different types of tools to assess intake include objective measures, such as biomarkers, and self-reported methods, such as brief dietary screeners.(23,24) Although objective measures represent an important component of diet assessment methods, this dissertation focuses on self-reported methods of assessment. Self-report methods are generally used to assess dietary intake because there are a limited number of nutritional biomarkers that can capture associations that reflect dietary intake.(23,24) An even smaller number of biomarkers, known as recovery biomarkers, are able to accurately capture the relationship between what was eaten at a given time and its excretion.(24) These markers cannot capture overall dietary intake and instead provide an indicator of true intake for single components, such as energy and protein. Thus, self-reported methods are currently essential for researchers wishing to understand total dietary patterns.

As the only feasible option to evaluate total dietary intake and across a large scale, the significant limitations of self-reported methods of assessment must be noted. Self-report methods are subject to various forms of measurement error related to reporting and recall

errors, which result in both random and systematic measurement error, the major types of error affecting self-report dietary assessment tools.(14,25) The National Cancer Institute's Dietary Assessment Primer provides an authoritative overview for researchers conducting diet assessment using self-reported dietary intakes.(12,14) Random error, otherwise known as within-person variation, affects shorter term assessment tools to a greater degree than longer term tools.(14) Random error represents the difference between the reported intake on any given day and an individual's usual intake.(14) Systematic error is typically present to a greater degree in longer-term assessment tools, and results in estimates that are biased in a consistent direction away from true usual intake.(14) While random error can be mitigated through repeat measures, systematic error cannot be addressed in this way.(26) The magnitude of error and strategies to mitigate it vary across self-reported instruments, but consistently, all self-report methods are subject to error. Nevertheless, nutrition research progresses recognizing the limitations of available methods and seeking to improve methods and analytic techniques to better understand dietary patterns and their relationships with health and characteristics that are associated with wellbeing.

2.2. Methods to capture self-reported dietary intake

To assess self-reported dietary intake, whether the dietary components of interest are nutrients, foods, or overall dietary patterns, researchers need appropriate methods. A variety of tools exist, including 24HDRs, food frequency questionnaires, food records, and dietary screeners.(12,14,23) Researchers are typically interested in understanding a population's usual or long-term intake, rather than their intake on a given day.(26) This is because it is long-term intake that is most relevant to health, and dietary guidance typically pertains to habitual intake rather than what is consumed each and every day.(26–28) While some tools explicitly capture long-term intake by asking participants to recall their intake over a given period, other tools may

capture intake over shorter periods of time, such as a single day. For these tools, additional steps may need to be taken to estimate usual dietary intake distributions using repeated measures and statistical techniques to account for within-person variation.(26,29) Self-report dietary assessment tools each have benefits and drawbacks that vary depending on the research setting and study population and the appropriate tool should be carefully selected depending on the research question.

Food records may be used to conduct population-level assessments of dietary intake or to examine diet-health relationships.(14) Food records are a short-term tool, and though averaging over several days could begin to approximate usual intake, statistical modeling using multiple administrations of n-day food records to represent long-term usual intake can estimate distributions that better approximate usual intake(26) Food records require participants to record their food and beverage intake in real time over a set number of days, often four to seven.(23) Records are typically self-administered and often require that participants record the meal, food items, brand name, preparation method, and portion size using visual guides or weighing.(14) Strengths of food records include that they do not rely upon recall, which may be beneficial for individuals who have not yet fully developed their cognitive abilities or who experience cognitive challenges. Depending on how they are administered, their open format allows participants to enter foods and beverages that could be missed in other assessment methods.(23) Limitations include that food records require participant literacy levels to be sufficient to independently complete records and they have relatively high participant burden that increases with the number of days requested and could lead to attrition or selection bias. Further, it is possible that participants may modify what they eat and drink due to the fact that they are completing a food record, known as reactivity.(12,14,23) The primary source of measurement error in food records is random error compared to systematic error, although systematic error also affects record data.(14)

A less burdensome method that is commonly used in surveillance studies and can be used to examine relationships with health outcomes is the 24HDR.(14) Recalls provide detailed data about dietary intake during the previous 24 hours or the prior day from midnight to midnight. Although a single recall does not represent usual dietary intake, this can be approximated through the group mean, repeat recalls, or statistical techniques.(14,26,29) 24HDRs rely on participant memory and are typically conducted by an interviewer,(14,23) but computer-assisted versions are increasingly used and are cheaper to apply.(30,31) A notable advancement to 24HDRs was the development of the Automated Multiple-Pass Method, which probes for commonly missed foods and beverages and helps to minimize reporting error, by the United States Department of Agriculture.(32,33) Strengths of 24HDRs include that participants completing interviewer-assisted recalls do not need to be able to read or write to complete a recall;(23) however, this does not apply to self-administered online 24HDRs. Other strengths include that while 24HDRs rely on memory, participants are accessing immediate memory, which is easier to retrieve compared to remembering over longer periods of time.(23) Limitations of 24HDRs include that a single recall cannot be used to estimate distributions of usual dietary intake for a group, though this can be overcome with multiple recalls and statistical modeling.(26) 24HDRs are also limited by the costs of implementing interviewer-administered recall(s) and potential attrition and possible selection bias due to respondent burden associated with the depth and detail.(14,23) 24HDRs are primarily subject to random error, though systematic error is also present, and varies depending on the dietary component of interest.(14,26)

A commonly used dietary assessment method in epidemiological research is the food frequency questionnaire.(14) Food frequency questionnaires require participants to recall dietary intakes over a period of time, often a month or a year, and therefore inherently query usual intake.(14) The number of questions included typically ranges from 80-120,(12) requiring approximately 30-60 minutes to complete.(14,23) Strengths of food frequency questionnaires

include that they are relatively inexpensive to implement, capture long-term intake, and are typically self-administered.(14,23) These tools contain less random error than shorter-term instruments; however, they are limited by substantial systematic error that arises because of the cognitive challenges related to estimating and averaging intake over long periods of time, and potentially, missing details on preparation methods or foods not included in pre-specified food lists.(14,23) Given the systematic error present in food frequency questionnaires, if possible, it is useful to include 24HDRs to calibrate food frequency questionnaires used as the primary instrument in epidemiological research.(12,14)

Dietary screeners are another long-term tool that aim to capture usual dietary intake and are often used to quickly assess intake in settings where resources may be constrained.(14) These brief instruments ask a limited number of questions to obtain a rapid assessment of intake of a dietary component or a dietary pattern.(14) Dietary components evaluated may include particular nutrients such as sodium,(34) or food components like fruits and vegetables.(35) Evaluation of dietary patterns using screeners can include assessing alignment with dietary guidance, such as through the short Healthy Eating Index.(36) Screeners typically refer to the past month or year, are frequency-based, and many do not query portion sizes of foods or beverages.(12,14) Strengths of screeners include that they are quick and inexpensive to apply. However, they are limited in regards to providing detailed dietary intake data when compared to more in-depth, longer methods.(14) Like food frequency questionnaires, dietary screeners are subject to substantial systematic error, though they contain less random error than shorter-term instruments.(12,14) They are not a replacement for more comprehensive methods of assessment; however, they can be useful for providing high-level indications of dietary intake.(12,14)

Each self-report method to assess dietary intake has associated strengths and limitations that should be carefully considered when selecting a method for capturing intake in a study.

Measurement error affects all types of self-report assessment methods, and its implications should be carefully considered when evaluating study results.(14) When possible, researchers should attempt to mitigate the effects of measurement error during study design and analysis.(25)

2.3. Methods to capture dietary intake in the Canadian context

In Canada, suboptimal dietary intake is the fifth highest risk factor contributing to morbidity and mortality(37) and is estimated to result in nearly $14 billion in healthcare costs annually.(38) Despite the strong links between diet and health, there are relatively little data on the dietary intakes of Canadians. There have only been three "national" assessments collecting comprehensive data on the dietary intakes of Canadians, once in the 1970s(39) and again in 2004 and 2015 through the nutrition-focused cycles of the Canadian Community Health Survey (CCHS).(40,41) These surveys were not truly national assessments because they each excluded the territories, among other key population subgroups.(39–41) The 2004 and 2015 CCHS Nutrition are cross-sectional surveys that enable investigations of the dietary intakes of individuals living in the provinces using 24HDRS, with a subsample of individuals participating in each survey having completed a second recall.(40,41)

Dietary intake is routinely assessed in Canada through a series of screeners included in the cross-sectional Canadian Health Measures Survey, which has been implemented every two years since 2007.(42,43) Additionally, the annual cross-sectional waves of CCHS include the Behavioural Risk Factor Surveillance System fruit and vegetable module ,(42,44) which consists of six questions about fruit and vegetable intake over the past 30 days.(35,44) Beyond these national surveys, there have been several small-scale studies measuring dietary intake in Canada,(45–49) though the population, method of assessment, and duration of studies vary.

This paucity of dietary data presents barriers to effectively developing and evaluating nutrition policy in Canada. Healthy eating is a policy priority in Canada, evidenced by the development of the Healthy Eating Strategy,(50) a multi-pronged strategy that led to the 2019 update to Canada's Food Guide,(51) targets to reduce sodium consumption,(52) and changes to nutrition labels.(53) The 2019 Canada's Food Guide represented a shift from previous iterations of dietary guidance, moving away from recommended numbers of servings of food groups to broader guidance on overall patterns of eating and proportions of intake.(51) Within the 2019 Canada's Food Guide, the healthy food choices recommendations are primarily represented by a plate image that is half comprised of fruits and vegetables, one quarter comprised of whole grains, and one quarter comprised of protein foods.(54) This updated guidance also provided recommendations for eating practices, such as cooking with others and enjoying your food, recognizing that healthy eating is encompassed by more than the foods that we eat but also consists of the social and cultural act of eating.(51,55)

Health Canada expressed a desire to assess alignment with the 2019 Canada's Food Guide healthy food choices recommendations, leading to the development and evaluation of the Healthy Eating Food Index-2019 (HEFI-2019).(56,57) This tool demonstrated reasonable validity in a national dataset and allows for the investigation of dietary patterns in Canada broadly, with opportunities to examine total and component scores.(56) The HEFI-2019 was tested using 24HDRs but can theoretically be applied to data from other comprehensive measures of assessment, like food frequency questionnaires or food records. A brief measure to quickly assess alignment with the healthy food choices recommendations in the Food Guide would complement the HEFI-2019 by providing a tool that could be implemented in contexts in which time or financial resources are limited. This could enable collection of dietary intake data across a greater range of contexts to explore dietary patterns and their relationships with sociodemographic characteristics or health among individuals in Canada.

2.4. Methods to analyze dietary intake data

Using data generated by self-report methods, several approaches have been developed to analyze dietary intake data. Broadly, these methods will be conceptualized as traditional or classical statistical approaches and novel methods to analyze dietary data, although there is no clear delineation between what is traditional and what is novel.(58) For the purposes of this dissertation, traditional approaches include methods to estimate usual intake distributions by accounting for within-person variation in data from short-term assessment methods, hypothesis-driven models to estimate relationships between dietary intake and characteristics of interest, and data-driven approaches that have been commonly applied to characterize of dietary patterns, such as factor analysis.(58–60) Methods to estimate distributions of usual intake include software implementations like the National Cancer Institute (NCI) method, which can be applied to data from 24HDRs when at least a subsample of the population has completed a second recall.(29,61) Other statistical approaches to account for measurement error in data from self-report methods of assessment include regression calibration to mitigate the effects of systematic error by calibrating data from the primary method of assessment to a more accurate instrument.(14,62)

Hypothesis-driven traditional statistical approaches have the benefit of producing results that are directly relevant to the research question and many approaches are simple enough for individuals with applied backgrounds to apply and interpret. Traditional statistical approaches can be used in a variety of contexts, ranging from exploratory analyses to confirmatory approaches to determine causal links. In the case of dietary patterns analysis, traditional methods have included both *a priori* and *a posteriori* approaches, allowing insights into how dietary patterns differ across populations and subpopulations and how these patterns are associated with health.(4,63) However, limitations of traditional statistical approaches include

that they typically require researcher assumptions, which could hinder discoveries that humans may not have conceived.(58) Further, when there is a high degree of complexity in the research question or data source, traditional statistical methods may not be the optimal approach.(13)

As the field of nutrition has progressed, so too have the types of research questions that individuals have sought to answer. These questions move beyond assessing relationships between single nutrients and health outcomes to overall dietary patterns that allow for consideration of synergistic relationships among dietary components and how these patterns may relate more broadly to the food environment, sociodemographic characteristics, and health outcomes.(16,64) As more complex questions have emerged, so have more complex approaches to analyze data. Novel methods, such as machine learning, are gaining popularity with a range of methods being applied to tackle complex problems including precision nutrition, image-based dietary assessment, and dietary patterns analysis.(13,18,65,66)

Machine learning is a subset of the field of artificial intelligence.(13) It encompasses a wide range of algorithms that can be used to learn from data for the purposes of prediction, classification, and identifying patterns.(13,18) Machine learning includes supervised and unsupervised approaches, whereby researchers may provide some or no input to models.(13,17) This capacity for models to learn from data rather than requiring pre-specified hypotheses opens possibilities to allow dietary patterns to emerge from data. While machine learning continues to grow in popularity, the distinctions between machine learning and traditional statistical approaches are often vague. Machine learning is typically cited as having greater emphasis on prediction whereas traditional statistical approaches have greater emphasis on inference.(67) However, machine learning is rooted in statistical approaches, and there is great overlap between machine learning and traditional approaches and the types of questions that they seek to answer.(67,68) Within this dissertation, machine learning is defined as algorithms that learn without specific instructions, often focusing on prediction.(13,68)

As the application of machine learning algorithms have gained popularity, an emphasis on developing explainable algorithms has grown.(13,69) To achieve explainability, both interpretability and fidelity are required, meaning that models need to both be comprehensible and have sufficient precision to describe the task.(69) Depending on the model, interpretability or complexity may be emphasized as they trade off, leading to black box models as complexity increases.(19) Within nutrition, explainability and interpretability are often desirable, as understanding how the model arrived at its results may be more important than a complex model; for example, when considering relationships with health outcomes with the aim of informing dietary guidance and recommendations to reduce chronic disease risk.(13,18)

2.5. Shifts in assessment: From single components to dietary complexity

Dietary assessment traditionally focused on the evaluation of single nutrients when preventing deficiency was a primary public health nutrition concern.(70) As concerns about deficiencies decreased across high income countries over the twentieth century, attention shifted to the prevention of chronic disease and interest likewise expanded to include foods and food groups and eventually, dietary patterns.(16,70) Each dietary assessment method has the ability to capture dietary patterns in different capacities.(14) Food records, 24HDRs, and food frequency questionnaires each aim to capture total dietary intake and the resulting data are amenable to the application of *a priori* diet quality indices and *a posteriori* methods applied to investigate dietary patterns. While screeners may be multifactorial and can assess a dietary pattern at a high level (such as alignment with a set of recommendations), their brevity does not allow them to assess total dietary intake.(14)

To account for the fact that people typically do not consume foods or nutrients in isolation, researchers developed diet quality indices and scores to assess overall dietary patterns.(16) The Healthy Eating Index, the most well-known of these indices, was developed to measure alignment with the Dietary Guidelines for Americans.(71–73) Indices of these types are based on *a priori* expert knowledge, with researchers acknowledging the shortcomings of these types of tools, including that they compress a multidimensional construct into a unidimensional score and that weights for scores are often assigned somewhat arbitrarily.(16,73) Other approaches to characterize dietary patterns have included *a posteriori* methods, such as principal component analysis or factor analysis.(74) In these methods, statistical modeling is used to generate dietary patterns rather than relying on expert knowledge.(75) However, researchers must still select food groups and detected associations are not adjusted for indirect effects of other variables.(76)

Interest in understanding dietary patterns continues to grow, as evidenced by their identification as a priority for research in the 2015 update to the Dietary Guidelines for Americans.(77) New developments in dietary patterns research are shifting to the use of novel methods to characterize dietary patterns, with the goal of better considering multidimensionality and synergy.(15,16) While these methods are being widely adopted for dietary pattern analysis, there has not been a review of the literature to determine which methods are being applied and how these methods are being used to characterize dietary patterns.

A promising method that has been applied to identify dietary patterns is probabilistic graphical models.(78–82) Probabilistic graphical models account for dietary complexity by jointly modeling distributions between variables.(83) Joint relationships are visually represented through graphical networks, with relationships between variables, or nodes, indicated through edges, demonstrating conditional dependencies.(84–86) Among undirected

graphical models, Gaussian graphical models provide the opportunity to explore joint relationships among continuous variables while Ising models have been used among discrete variables.(83,85) However, there are many cases in which researchers may wish to incorporate both continuous and categorical variables in their analysis. Mixed graphical models offer a solution to this problem, allowing the inclusion of both continuous and categorical variables within models.(85,87) These models allow researchers to investigate pairwise relationships conditioned on all other nodes in the network and are useful when individuals wish to understand the direct and indirect relationships among complex information inputs, such as dietary intake.(85)

Mixed graphical models offer the opportunity to visualize relationships between dietary patterns and sociodemographic characteristics,(78) considering how certain characteristics may demonstrate dependencies with each other and dietary components. Models are thus able to capture dietary synergies or interdependencies that are often missed when using *a priori* summary scores or cluster methods that categorize people within a dataset into a particular dietary pattern based on similar characteristics. Additionally, probabilistic graphical models offer the benefit of being easily interpretable through graphical representations. While models can be made more complex through the addition of variables, researchers ought to weigh a desire for accuracy and complexity with interpretability and should carefully select relevant inputs.(86) In this sense, probabilistic graphical models offer a balance between expert knowledge to determine the appropriate variables that should be included in models and model design with data-driven techniques. Several studies have applied probabilistic graphical models to characterize dietary patterns or joint relationships among dietary components (**Table 1**).(76,78,79,81,82,88–95) These studies explored the interrelationships between dietary components and some estimated scores from the patterns identified through networks to examine relationships with health.(78,89–93) These studies demonstrate the potential to apply

probabilistic graphical models to identify dietary patterns, and particularly their capability to directly incorporate sociodemographic characteristics.

Table 1: Summary of studies applying probabilistic graphical models to identify dietary patterns or joint relationships between dietary components

Authors	**Year**	**Study description**	**Graphical model**	**Measure of dietary intake**	**Other factors included in model** (i.e., health, sociodemographic characteristics)	**Overview**
Iqbal et al.(80)	2016	27 120 adults in Germany in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam cohort. Objective: Explore the interrelation between food intakes by sex among adults in Germany.	Gaussian graphical model; Sensitivity analysis with semiparametric Gaussian copula model	148-item FFQ; condensed to 49 food groups	N/A	Two networks of dietary patterns were identified by sex.
Schwedhelm et al.(82)	2018	814 adults in the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam cohort in Germany. Objective: Describe meal and habitual dietary patterns in German adults.	Semiparametric Gaussian copula model	Up to 3 24HDRs (averaged); condensed to 39 food groups	N/A	Five networks of dietary patterns were identified. Four networks split intake by meals (breakfast, lunch, dinner, snacks) and the fifth network provided a total habitual dietary pattern.
Behrouzi et al.(76)	2019	662 older adults from four studies (Nu-Age, ProMuscle, ProMO, V-Fit) in the Netherlands. Objective: Identify conditional independence relationships between muscle function	Copula graphical model	2-7 day food records	Indicators of muscle function (i.e., hand grip), age, sex, smoking status, anthropometric measures	A network of nutrient intake, indicators of muscle function, sociodemographic characteristics, and anthropometric measures was identified.

		indicators, nutrients, and other variables.				
Fereidani et al.(88)	2020	401 females hospitalized in Iran (134 cases of breast cancer; 267 controls). Objective: Compare dietary patterns between patients with and without breast cancer.	Gaussian graphical model	168-item FFQ; condensed to 28 food groups	N/A	Two networks of dietary patterns were estimated by cases and controls.
Gunathilake et al.(89)	2020	1245 adults in Korea (415 cases of gastric cancer; 830 controls). Objective: Identify dietary pattern networks associated with gastric cancer.	Gaussian graphical model	106-item FFQ	N/A	Three dietary pattern networks were identified; one total, and two that were stratified by sex. Dietary pattern scores were then used and categorized into tertiles. Logistic regression was used to estimate the odds of gastric cancer.
Hoang et al.(78)	2021a	7423 adults in South Korea in the Cancer Screening Examination Cohort. Objective: Explore the interactions between demographics, dietary behaviors, and comorbidities.	Mixed graphical model	106-item FFQ	Demographic characteristics (age, sex, marriage status, education, employment, income, smoking status, drinking status), physical activity, body mass index, metabolic markers (blood pressure, total cholesterol, fasting glucose, glomerular filtration rate)	Three networks were identified. The first network included only dietary components. The second network included a dietary pattern score, demographic characteristics, and health characteristics. The dietary pattern score was estimated from amount consumed multiplied by eigenvector centrality. The third network included all dietary components,

						demographic characteristics, and health characteristics.
Hoang et al.(91)	2021b	10 777 adults in South Korea participating in cross-sectional screenings at the National Cancer Center. Objective: Identify dietary patterns using graphical models, identify dietary pattern scores, and assess differences in cancer prevalence by diet pattern scores.	Gaussian graphical model	FFQ	N/A	Three networks were estimated; one total network, and two that were stratified by sex. Dietary pattern scores were then estimated using Eigenvector centrality and logistic regression was used to calculate the odds of cancer.
Jayedi et al.(92)	2021	850 adults in Iran participating in a cross-sectional study. Objective: Identify dietary networks among adults in Iran and explore associations with adiposity.	Gaussian graphical model	168-item FFQ; condensed to 30 foods, grouped into six communities	N/A	A dietary pattern network was identified. Dietary pattern scores were estimated by taking standardized intakes of food and multiplying by factor loading scores from principal component analysis. ANOVA and Logistic regression were used to test associations between dietary pattern scores and adiposity.
Schwedhelm et al.(79)	2021	365 pregnant women in the United States participating in the Pregnancy Eating Attributes Study. Objective: Compare meal-specific dietary patterns among pregnant women with	Gaussian graphical models	Three 24HDRs (each trimester); condensed to 40 food groups	N/A	Dietary pattern networks were estimated for each meal (breakfast, lunch, dinner, snack) for the low and high diet quality tertile separately.

		high and low diet quality.				
Gunathilake et al.(90)	2022	7874 adults in South Korea participating in a cohort study. Objective: Identify dietary patterns and how they contribute to the development of cancer and compare these patterns with reduced rank regression and principal component analysis.	Gaussian graphical model	106-item FFQ; condensed to 35 food groups; grouped into 17 communities	N/A	Two networks of dietary patterns were identified; one for all participants and one for participants matched by age and sex. Dietary pattern scores were estimated based on node centrality and standardized intake. Cox proportional hazard ratios were used to estimate associations with cancer risk.
Jahanmiri et al.(93)	2022	850 adults in Iran participating in a cross-sectional study. Objective: Explore centrality of nodes in dietary patterns and explore associations with metabolic syndrome.	Gaussian graphical model	168-item FFQ; condensed to 39 food groups	N/A	A dietary pattern was identified, and dietary pattern scores were calculated using standardized intake multiplied by factor loading score from principal component analysis.
Landaeta-Diaz et al.(94)	2023	1242 adults in Chile with anhedonia symptoms. Objective: Identify dietary patterns among adults in Chile among those with and without anhedonia symptoms.	Gaussian graphical model	13-item FFQ	N/A	Three networks of dietary patterns were estimated. The first was estimated for the total sample, the second among those with anhedonia symptoms, and the third among those with food-related anhedonia.
Slurink et al.(95)	2023	74 132 adults in the Lifelines prospective cohort study in the Netherlands.	Mixed graphical model	110-item baseline Flower-FFQ (1 main questionnaire on energy and macronutrient	Incident prediabetes, baseline sociodemographic characteristics,	Two networks were estimated. The first estimated joint relationships between dairy products and food

		Objective: Investigate prospective associations of dairy intake with prediabetes.		intake, 4 complementary food questionnaires).	lifestyle risk factors, clinical markers	groups. The second network included dairy products, food groups, health characteristics, and sociodemographic characteristics.

In Canada, machine learning algorithms have been applied to investigate a range of nutrition questions. A study using data from three 24HDRs and a food frequency questionnaire found that statistical approaches, including logistic regression and least absolute shrinkage and selection operator (LASSO) penalized regression, and machine learning algorithms predicted adequate consumption of fruits and vegetables with similar accuracy.(96) This study noted that while machine learning presents opportunities for the field of nutrition, it is not a silver bullet analytic approach and should be carefully matched to appropriate problems.(96) Another Canadian study used machine learning algorithms to predict cardiovascular disease outcomes using a single 24HDR from the 2015 CCHS Nutrition and found that supplements, caffeine, and alcohol were the most important components for predicting cardiovascular disease.(97) The use of a single 24HDR with unmitigated random error to investigate nutrients' relationships to a long-term health outcome may partially explain the weak relationships observed,(97) though there is no established method to account for measurement error within studies applying machine learning to dietary data. Another study used k-means cluster analysis and LASSO regression to investigate the relationship between diet quality and sociodemographic characteristics in the Canadian context. The results suggested that females, non-smokers, and immigrants were more likely to have higher diet quality compared to other groups.(98) This study also used a single 24HDR to investigate the relationship between diet quality and sociodemographic characteristics.(98) Finally, natural language processing has been applied in Canada in a range of instances to explore food processing, food labels, and nutritional habits through social media.(99–101) While these studies did not seek to identify dietary patterns, they provide glimpses into how machine learning methods can be applied to examine questions related to dietary intake and public health nutrition in Canada.

2.6. Associations between dietary intake and sociodemographic characteristics

Examinations of dietary intake, using a range of methods to capture and analyze data, have shown that several sociodemographic characteristics, including lower income, lower educational attainment, living in food-insecure households, and smoking, are associated with poorer diet quality.(8–11,102–104) Age and sex are also associated with diet quality, with older adults and females typically having better quality diets compared to younger adults and males.(56,104–107)

In Canada, a study investigating the dietary intakes of individuals who participated in the 2004 CCHS Nutrition found that prevalence of inadequacy for several nutrients was higher among adults with the lowest levels of income adequacy or educational attainment.(8) Likewise, having higher household income adequacy or educational attainment was associated with higher intakes of dairy, fruit, and vegetables among adults and children.(8) Another study investigating differences in diet quality among adults by income, educational attainment, and neighborhood deprivation in Canada found that gaps in diet quality have either remained stable or worsened between 2004 and 2015.(10)

Living in a food-insecure household is similarly associated with suboptimal dietary intake among adults. (9,102,108–110) A study investigating the diet quality of individuals who participated in the 2015 CCHS by severity of household food security status found that after adjusting for income, education, and geographic region, diet quality did not differ by food security status.(9) However, there was a significant trend of increasing mean percent energy from ultra-processed foods with worsening severity of food insecurity across age sex groups after adjusting for the covariates.(9) Other research focused on individuals in Canada in relation to food security status has also demonstrated compromises in intake of nutrients, foods, and

food groups.(102,108–111) Although individuals in Canada overall have poor diet quality,(56,112) inequities are apparent with respect to low income, low educational attainment, and living in food-insecure households.

While investigations of relationships between single sociodemographic characteristics and dietary intake are informative, it is increasingly relevant to consider how the numerous characteristics that form people's identities may relate to dietary patterns. An investigation of diet quality by income adequacy and racial identity among adults found that income adequacy was associated with diet quality scores but racial identity was not independently associated.(113) Once models incorporated both income adequacy and racial identity, the interaction was associated with diet quality scores, indicating that both constructs may contribute to diet quality in conjunction.(113) Because of diet's importance to health, it is essential to consider how dietary patterns may differ by sociodemographic characteristics, particularly those that have known associations with compromises in dietary quality or health. These markers of inequities may intersect with each other to create compounding effects on dietary intake.

Current methods for characterizing dietary patterns make it challenging to consider how numerous sociodemographic characteristics may be jointly related to dietary patterns. While evidence indicates that dietary intakes differ among several population subgroups in Canada, the dietary patterns of these subgroups are not well understood beyond the use of *a priori* methods. However, these *a priori* methods are limited in their ability to capture dietary synergy and may be limited by pre-existing hypotheses, which may not consider all potential differences or intersections among population subgroups. Data-driven approaches such as machine learning algorithms may be useful to learn about multidimensional dietary patterns among populations and population subgroups and how they are jointly related with sociodemographic characteristics, such as age, sex, income, education, and household food security status. By allowing dietary patterns to emerge from the data rather than forcing a particular structure onto

the sample, these approaches could complement existing methods to identify dietary patterns. It is important to understand the heterogeneity of dietary patterns within populations and how they relate to health, which can be considered through novel approaches.(114) Using a range of approaches to better understand dietary patterns could help to consider diversity, informing nutrition policy that meets the needs of all people in Canada.

2.7. Summary and rationale

Dietary intake is a key contributor to health and disease risk.(6,7) Better understanding dietary patterns and how these differ among population subgroups requires attention to both how dietary intake data are captured or collected and how they are analyzed. In some settings, it is most appropriate to use in-depth methods that may be more burdensome but yield more detailed data. In other instances, it is informative to use brief measures that can be quickly and inexpensively administered and have low participant and researcher burden, allowing for the examination of dietary patterns in broader population subgroups and contexts that may otherwise be unexplored. Likewise, when analyzing diet, there are cases in which it is useful and appropriate to apply hypothesis-led, classical approaches, whereas in other instances, novel, data-driven approaches can be helpful to learn relationships that may otherwise not have been considered. As researchers seek to maximize available tools to evaluate dietary intakes, applying a flexible range of tools to capture and analyze dietary patterns can further extend the possibilities for discovery. This dissertation addresses key gaps to capture dietary patterns in Canada and explores novel methods to analyze multidimensional dietary patterns, considering the influence of sociodemographic characteristics on dietary patterns.

Recent updates to dietary guidance through the 2019 Canada's Food Guide have led to a desire to understand how the dietary patterns of individuals in Canada align with this food-based guidance. The HEFI-2019 provides opportunities to comprehensively assess alignment

with the healthy food choices recommendations,(57) though comprehensive measures of dietary intake may not always be possible. Thus, it is desirable to have a shorter tool to quickly assess alignment with the healthy food choices recommendations among adults to allow for research and surveillance that offers less detail but with lower burden. To address this gap, Health Canada contracted the development of a brief dietary screener.

Traditional approaches for the analysis of dietary patterns, such as their identification through *a priori* approaches such as diet quality indices(57,63,115) or through *a posteriori* approaches such as principal component analysis have allowed for an improved understanding of diet-health relationships beyond simply investigating single dietary components.(63,74,116) However, these methods may be limited in their capacity to capture the full complexity of dietary patterns, such as the synergistic and antagonistic relationships among dietary components.(16) In such cases, methods that have not traditionally been applied within the field of nutrition could provide unique insights. There are already a wide range of questions that have been explored using methods that are novel within the field of nutrition.(13,17,18,65,66) However, there has not been a substantive exploration of how these methods are being applied to characterize dietary patterns.

While there is evidence that dietary intakes of individuals in Canada differ by sociodemographic characteristics, such as age, sex, income, education, and household food security status,(8–11,56,102) prior investigations have often considered single dietary components or used *a priori* indices to assess diet quality and typically do not consider potential interrelationships among sociodemographic characteristics. In reality, dietary patterns are complex, with likely synergistic and antagonistic effects from consuming dietary components in combination.(16,64) Likewise, dietary patterns are unlikely to be shaped by single sociodemographic characteristics, but rather by the multiple identities that people hold. To account for this complexity, novel analytic approaches may be appropriate to learn previously

unconsidered joint relationships between dietary components and sociodemographic characteristics. Though novel methods have been applied to nutrition questions in Canada, these methods have not yet been applied to conduct a detailed exploration of dietary patterns. Through the application of probabilistic graphical models to characterize dietary patterns in relation to sociodemographic characteristics, we will explore the capacity of these models to advance our understanding of the interrelationships between these two complex constructs.

Collectively, this research will advance dietary assessment methods, enabling a better understanding of dietary patterns to provide evidence that can inform nutrition policies and programs to support healthy eating.

Chapter 3: Research Objectives

This thesis consists of four studies investigating methods to capture and analyze dietary patterns. The overall goal is to advance our capacity to characterize diet quality relative to dietary guidance and to provide insights into how dietary patterns can be characterized using novel methods. This work highlights different techniques that can be used to generate new data and leverage existing dietary data in Canada and broader contexts. The goal of **Chapter 4** and **5** was to develop and evaluate a brief dietary screener to assess alignment between dietary intake and the 2019 Canada's Food Guide healthy food choices recommendations among adults. **Chapter 6** sought to understand the range of novel methods being used to identify dietary patterns. **Chapter 7** used secondary data from the nutrition-focused 2015 CCHS to explore the potential of a novel analytic approach to improve our understanding of the joint relationships between multidimensional dietary patterns and intersecting sociodemographic characteristics among adults.

Specific Objectives

Chapter 4: Development and cognitive testing of a brief dietary screener

1. To develop a brief screener to assess overall alignment of dietary intake of adults aged 18-65 years, with marginal and higher health literacy, with the 2019 Canada's Food Guide healthy food choices recommendations.

Chapter 5: Scoring and assessment of construct validity of a brief dietary screener

2. To assess the construct validity of a dietary screener assessing alignment with the 2019 Canada's Food Guide by examining variability in screener scores, whether the screener produces scores that differ between subgroups with known differences in dietary quality,

and the correlation coefficient between screener scores and total HEFI-2019 scores among a sample of adults.

Chapter 6: A scoping review of novel methods to characterize dietary patterns

3. To synthesize peer-reviewed literature applying novel methods to characterize dietary patterns.

Chapter 7: Identification of dietary patterns considering sociodemographic characteristics

4. To examine the capacity of probabilistic graphical models, a novel method for dietary pattern analysis, to expand our understanding of the nexus between dietary patterns and sociodemographic characteristics through a layered analysis.

Chapter 4: Development of the Canadian Food Intake Screener to assess alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations

Status: Accepted version. Published in *Applied Physiology, Nutrition, and Metabolism*, 2023; 48(8): 603-619. Doi: 10.1139/apnm-2023-0019.

Authors: Joy M. Hutchinson*[1], Tabitha E. Williams*[1], Ailish M. Westaway[1], Alexandra Bédard[2], Camille Pitre[2,3], Simone Lemieux[2,3], Kevin W. Dodd[4], Benoît Lamarche[2,3], Patricia M. Guenther[5], Jess Haines[6], Angela Wallace[6], Alicia Martin[7], Maria Laura da Costa Louzada[8], Mahsa Jessri[9], Dana Lee Olstad[10], Rachel Prowse[11], Janis Randall Simpson[6], Jennifer E. Vena[12], Sharon I. Kirkpatrick[1]

1. School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada
2. Centre Nutrition, santé et société (NUTRISS), Institut sur la nutrition et les aliments fonctionnels (INAF), Université Laval, QC, Canada
3. École de nutrition, Faculté des sciences de l'agriculture et de l'alimentation, Université Laval, Québec, QC, Canada
4. Division of Cancer Prevention, National Cancer Institute, Bethesda, MD, USA
5. Department of Nutrition and Integrative Physiology, University of Utah, Salt Lake City, UT, USA
6. Family Relations and Applied Nutrition, University of Guelph, Guelph, ON, Canada
7. Department of Geography, Environment and Geomatics, University of Guelph, Guelph, ON, Canada
8. Department of Nutrition, School of Public Health; Center for Epidemiological Research in Nutrition and Health, University of São Paulo, São Paulo, Brazil
9. Food, Nutrition and Health Program, Faculty of Land and Food Systems; Centre for Health Services and Policy Research, Faculty of Medicine, University of British Columbia, BC, Canada
10. Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
11. Division of Community Health and Humanities, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada

12. Alberta's Tomorrow Project, Cancer Research and Analytics, Cancer Care Alberta, Alberta Health Services, Calgary, AB, Canada

Acknowledgments: The authors are grateful to the participants in the cognitive testing and the experts who participated in the face and content validity testing. The authors are also grateful to Kimberley J. Hernandez, Lisa-Anne Elvidge Munene, Sylvie St-Pierre, and other Health Canada employees who participated in and provided input on various aspects of the project; Dr. Hannah Tait Neufeld of the University of Waterloo who reviewed the screener for appropriateness for use with Indigenous populations and provided related advice; Dr. Gordon B. Willis of the United States National Cancer Institute for feedback on the cognitive interviewing protocol; and Sanaa Hussain of the University of Waterloo for assistance with referencing.

4.1. Overview

The objective of this project was to develop a brief self-administered dietary screener, in English and French, to rapidly assess alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations. In consultation with Health Canada and external advisors (n=15), guiding principles were defined. Existing screeners were scanned, and the healthy food choices recommendations were mapped to inform questions and response options. Cognitive interviews were conducted in English (n=17) and French (n=16) with adults aged 18-65 years from April to June 2021 to assess understanding of questions and face validity; recruitment emphasized variation in sociodemographic characteristics. Face and content validity were assessed with experts in nutrition, surveillance, and public health (n=13 English, 3 French) from April to May 2021. The testing indicated that the screener was well-understood overall but informed refinements to improve comprehension of the questions and their alignment with the healthy food choices recommendations. The resulting Canadian Food Intake Screener/Questionnaire court canadien sur les apports alimentaires includes 16 questions to rapidly assess alignment of intake with the 2019 Canada's Food Guide healthy food choices recommendations, including healthy foods and foods to limit, in situations in which comprehensive dietary assessment is not feasible.

4.2. Introduction

Suboptimal dietary patterns are a key risk factor for noncommunicable chronic diseases in Canada and globally (Afshin et al. 2019; Lim et al. 2012; Lozano et al. 2012). Diets low in whole grains, fruits, vegetables, nuts and seeds, and omega-3 fatty acids and high in sodium have been shown to be responsible for more deaths globally than any other risk factor (Afshin et al. 2019). To promote healthy eating and reduce diet-related chronic disease risk, many countries publish food-based dietary guidelines, presenting "specific, culturally appropriate, and actionable recommendations" (Food and Agriculture Organization 2022; Herforth et al. 2019). In 2019, Health Canada released an updated Canada's Food Guide (CFG-2019), with recommendations on healthy food choices and healthy eating habits (Health Canada 2021, 2022a). In a shift from prior iterations, CFG-2019 does not provide recommendations on the number of servings per day or serving sizes for food groups based on age and sex. Instead, through the healthy food choices recommendations, CFG-2019 recommends eating a variety of healthy foods each day, including fruits and vegetables, whole-grain foods, and protein foods, emphasizing more frequent consumption of plant-based protein foods. The CFG-2019 plate provides a visualization of the desirable proportions of foods from these categories in relation to one another (Health Canada 2022a). CFG-2019 aims to promote healthy eating and overall nutritional well-being and to support improvements to the food environment in Canada (Health Canada 2022b).

For research and surveillance purposes, it is of interest to assess alignment of eating patterns and practices with CFG-2019, for example, to inform targeted interventions to address disparities in alignment with the guidance among population subgroups. The Healthy Eating Food Index-2019 (HEFI-2019) facilitates assessment of the alignment of dietary intake with CFG-2019 healthy food choices recommendations in research and surveillance contexts in which comprehensive dietary intake data, such as from 24-hour dietary recalls, are available (Brassard

et al. 2022a, 2022b). Data from dietary recalls are recommended for characterizing the dietary intake of populations and subgroups due to their comprehensiveness, as well as their greater accuracy relative to frequency-based instruments (Kirkpatrick et al. 2022a; National Cancer Institute 2015). However, while online self-administered recall platforms have eased researcher and respondent burden (Lafrenière et al. 2017; Subar et al. 2012), recalls can be time consuming to collect and may not be amenable to all settings. Further, appropriate use of recall data requires substantial expertise and extensive cleaning and analytic efforts (Kirkpatrick et al. 2022b).

In contrast, brief dietary questionnaires, informally called "screeners," can be used for rapid assessment of food and beverage intake over a given period, such as the past month or year (National Cancer Institute 2015). Screeners often focus on specific dietary components (e.g., fruits and vegetables, fiber) (Centers for Disease Control and Prevention n.d.; Tangney et al. 2019) but may be multifactorial (Colby et al. 2020; de Rijk et al. 2021; Gnagnarella et al. 2018; Lafrenière et al. 2019). Prior multi-factorial screeners have aimed to assess dietary intake relative to food-based dietary guidance (Colby et al. 2020; de Rijk et al. 2021; Gabe and Jaime 2019).

The objective of this study was to develop a brief screener to assess overall alignment of adults' dietary intake with the CFG-2019 healthy food choices recommendations. The screener was developed for use with adults, aged 18-65 years, with marginal and higher health literacy, and is intended for self-administration in English and French. The current paper describes the development process, including cognitive testing to assess whether screener questions were understood as intended, and face and content validity with a panel of experts. The screener's scoring system and construct validity are described in the accompanying paper (Hutchinson et al. 2022). A separate brief questionnaire, the Canadian Eating Practices Screener, developed to

assess adults' alignment with the CFG-2019 healthy eating habits recommendations, is described elsewhere (Haines et al. 2022).

4.3. Methods

Development of the screener

Screener development and evaluation were undertaken in collaboration with Health Canada and guided by a team of external expert advisors, including nutrition researchers and practitioners (**Appendix A**); this group included English- and French-speaking individuals. Many of the advisors were involved in the development of the HEFI-2019 (Brassard et al. 2022a, 2022b), supporting consistent interpretation of the underlying dietary guidance and alignment of the screener with the HEFI-2019.

The development of the screener drew upon the messaging related to healthy food choices in CFG-2019, including the plate depicting the recommended proportions of the food categories (Health Canada 2022a). Also considered were the food choice components within the Healthy Eating Recommendations, which provide simple and actionable messages for consumers (Health Canada 2020), and the Dietary Guidelines, which are intended primarily for health professionals and policymakers (Health Canada 2022b). In addition to the guidance, the components included in the HEFI-2019 and their construction (e.g., inclusion or exclusion of particular foods) (Brassard et al. 2022a, 2022b) informed the screener questions to some extent. We also drew upon available information on the dietary intake of Canadians (e.g., key sources of food categories and nutrients such as saturated fats) (Harrison et al. 2019; Kirkpatrick et al. 2019a; Tugault-Lafleur and Black 2019).

The steps in screener development are illustrated in **Figure 1** and described below, though the process was iterative in terms of consultation with the advisors and refinement of the screener.

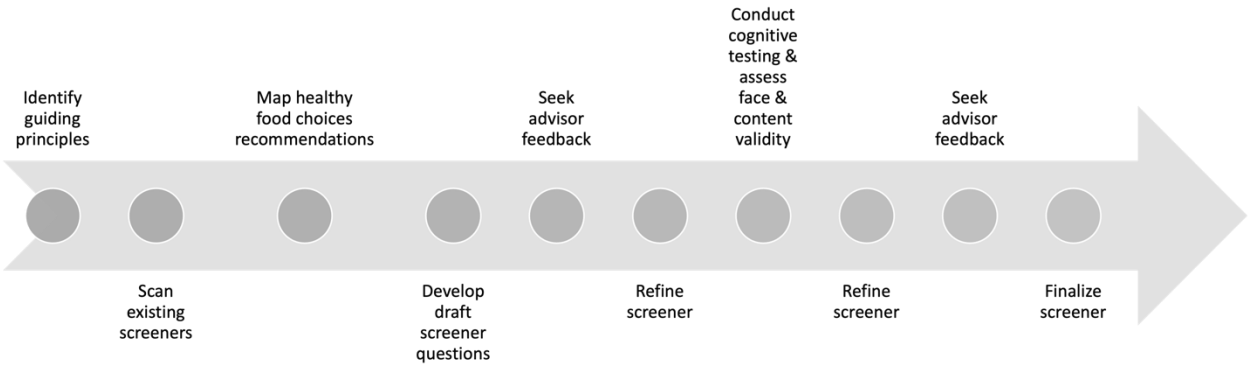


Figure 1. Process for development of the Canadian Food Intake Screener to assess alignment of adults' intake with the healthy food choices recommendations

Defining guiding principles for the screener

The development of the screener was informed by guiding principles, defined in collaboration with Health Canada *a priori* (**Figure 2**). These principles related to the development of a brief screener that is simple to use and score, assesses alignment with the healthy food choices recommendations overall, and considers the numeracy and literacy levels of the target population. Equivalence or comparability (Frongillo et al. 2019) in capturing alignment with the CFG-2019 healthy food choices recommendations across population subgroups was also considered. Additionally, the screener should demonstrate reasonable construct validity, described in the accompanying paper (Hutchinson et al. 2022). These principles were discussed with the advisors and informed initial decisions about the format and content of the screener. These decisions included assessing frequency of consumption versus

proportions, capturing foods and beverages versus nutrients, and confirming assessment of alignment with the healthy food choices recommendations overall.

- Simple to use and score.
- Brief (<10 minutes).
- Assess alignment with the 2019 Canada's Food Guide healthy food choices recommendations overall, not specific recommendations.
- Consider the numeracy and literacy levels of the target population.
- Consider equivalence (i.e., comparability) in capturing the construct across subgroups of the target population.
- Demonstrate reasonable validity for capturing the construct in the target population.

Figure 2. Guiding principles for the development and evaluation of the Canadian Food Intake Screener

First, although the CFG-2019 plate identifies the recommended proportions to be contributed by vegetables and fruit, whole-grain foods, and protein foods, the screener does not focus on proportionality. This is because screeners do not capture total dietary intake so a denominator, which would be needed to calculate proportions allocated to different types of foods, is unavailable. To capture usual consumption, participants could be asked to average proportions across eating occasions over some period, such as a month. However, this approach was hypothesized to be cognitively challenging, as well as difficult to score. Alternatively, repeat administrations of a screener focused on proportions on a given day could be used to capture usual proportions, but this would add burden for researchers and participants, undermining the goal of a brief screener. Furthermore, not all CFG-2019 recommendations related to healthy food choices are expressed using proportionality, such as the recommendations to limit intake of highly processed foods. Given the guiding principles related to literacy and numeracy demands, a frequency-based screener was thus developed, with the hypothesis that patterns of frequency of intake of different foods and beverages would provide an indication of the degree of alignment with the CFG-2019 healthy food choices recommendations. The past month, which is the typical period queried by screeners (Centers for Disease Control and Prevention n.d.;

National Cancer Institute 2021; Wijnhoven et al. 2018), was selected as the time frame of interest. A focus on the past month provides an indication of longer-term intake (versus intake on a given day) and may reduce error compared to recalling and averaging frequency of consumption over a longer period, such as a year (National Cancer Institute 2015). The screener does not query portion sizes, which is common for brief instruments (National Cancer Institute 2015).

Second, the screener focuses on frequency of intake of foods and beverages and not nutrients. The guidance is primarily food-based, though the Dietary Guidelines include recommendations related to intake of free sugars, saturated fats, and sodium (Health Canada 2022b). A screener specifically focused on one of these nutrients would likely include many questions and even then, may not accurately estimate nutrient intake (Tangney et al. 2019). Per the guidance and examinations of dietary intake among the population (Harrison et al. 2019; Kirkpatrick et al. 2019a), highly processed foods account for high proportions of intake of sugars, saturated fats, and sodium; thus, questions on highly processed foods were expected to provide a moderately strong signal in terms of the extent of alignment of dietary intake with the healthy food choices recommendations.

Finally, it was confirmed with the advisors that the screener and its scoring should focus on alignment with the recommendations overall, given that a brief multi-factorial screener cannot, by design, provide accurate estimates of intake of particular food categories.

Scanning existing screeners, mapping the healthy food choices recommendations, developing screener questions, and soliciting feedback from advisors

Screener development and evaluation were informed by existing screeners that have undergone validation (Centers for Disease Control and Prevention n.d.; Colby et al. 2020;

Gnagnarella et al. 2018; National Cancer Institute 2019; Tangney et al. 2019). These screeners provided insights into format and possible questions for inclusion, as well as response options. CFG-2019 recommendations pertaining to healthy food choices were then mapped (**Table 2**) to provide a roadmap against which to develop screener questions to ensure content validity. Initial screener questions were developed and revised iteratively based on feedback from Health Canada and the advisors. The screener was developed in English and translated to French. Translations were conducted by a professional firm and reviewed by bilingual researchers at Université Laval.

Given the guiding principles, a key emphasis in seeking advisor feedback on the draft questions was weighing trade-offs between a nuanced screener mapped closely to the recommendations versus a simple screener. It was expected that more detail would result in higher cognitive load and accordingly, more reporting error (Natarajan et al. 2010). This reporting error may be differential between individuals with different characteristics, for example, with respect to literacy (Choi and Cawley 2018; Keogh et al. 2020), resulting in less utility of the screener for use with diverse populations and potentially masking differences in alignment with the guidance among population subgroups. There was consensus among the advisors that the screener did not need to capture every nuance of the guidance (e.g., differentiating fruits canned in syrup from those not canned in syrup, capturing specific sources of unsaturated fats such as avocado). Such nuances can be more adequately addressed using more comprehensive methods, such as 24-hour dietary recalls.

The iterative feedback from the advisors led to consensus on wording questions as simply as possible, avoiding technical terms (e.g., fortified); querying foods of interest using colloquial terms (e.g., plant-based milks); including examples of commonly consumed foods and relevant exclusions but avoiding lengthy, exhaustive lists; using consistent question structure and response options; and ordering questions such that earlier questions cue responses to later

questions. We sought to avoid combining different types of foods (e.g., meat, cheese, and milk) in a single question to the extent possible, while also aiming for a short screener. This approach was deemed useful for minimizing cognitive load and ensuring clarity in the foods to consider in responding to each question. This approach also provided flexibility to account for emphases of the guidance, for example, on plant- versus animal-based protein foods, in the screener's scoring system, described in the accompanying paper (Hutchinson et al. 2022). The specific wording used in the screener questions differs in some instances from language used in the guidance and the use of wording such as 'sources of' does not imply that the foods or beverages must meet the nutrient criteria for the claim "source of", as defined in the Food and Drug Regulations (Government of Canada 2022).

The version of the screener evaluated in the first round of cognitive testing and in face and content validity testing consisted of 15 questions (**Appendix B**). Response options were adapted from the Dietary Screener Questionnaire and the Diet History Questionnaire (National Cancer Institute 2021; National Cancer Institute 2022; Millen et al. 2006) and ranged from never to 6 or more times per day.

Table 2: Dietary guidance mapped to final questions within the Canadian Food Intake Screener, in English and French, to assess alignment of intake with the 2019 Canada's Food Guide healthy food choices recommendations among adults aged 18-65 years (NOTE: questions in the table are not in order, refer to Appendix C for the final screeners in English and French)

Dietary Guideline	Healthy Eating Recommendation and/or other guidance	Final English screener questions	Final French screener questions
Vegetables, fruit, whole grains, and protein foods should be consumed regularly	Eat plenty of vegetables and fruit	Over the past month, how often did you eat potatoes, including baked, boiled, or mashed potatoes, or sweet potatoes? **Do not include** french fries, poutine, home fries, or hash browns.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des pommes de terre, y compris des pommes de terre au four, bouillies ou en purée, ou des patates douces? **N'incluez pas** les frites, les frites maison, la poutine, les pommes de terre rissolées et les galettes de pommes de terre (hash browns).
		Over the past month, how often did you eat fresh, cooked, frozen, or canned vegetables? **Do not include** potatoes, french fries, poutine, or other deep-fried vegetables, or vegetable juices and drinks.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des légumes crus, cuits, congelés ou en conserve? **N'incluez pas** les pommes de terre, les frites ou autres légumes frits, la poutine, les jus de légumes et les boissons de légumes.
		Over the past month, how often did you eat fresh, frozen, canned, or dried fruit? **Do not include** fruit juices and drinks.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des fruits frais, congelés, en conserve ou séchés? **N'incluez pas** les jus de fruits et les boissons aux fruits.

	Eat whole grain foods	Over the past month, how often did you eat **whole wheat or whole grain** breads, bagels, pasta, noodles, quinoa, oats, brown or wild rice, breakfast cereals, or other whole wheat or whole grain foods? **Do not include** white breads, bagels, pasta, noodles, rice, or refined breakfast cereals.	Au cours du dernier mois, à quelle fréquence avez-vous consommé du pain, des bagels, des pâtes, des nouilles, du quinoa, du gruau, du riz brun ou sauvage, des céréales à déjeuner ou tout autre **aliment fait de blé entier ou de grains entiers**? **N'incluez pas** le pain blanc, les bagels blancs, les pâtes ou les nouilles blanches, le riz blanc et les céréales à déjeuner raffinées.
	Eat protein foods	Over the past month, how often did you eat eggs, beef, pork, wild meat, chicken or other poultry, fish, shellfish, or other animal-based sources of protein? Include canned fish and canned poultry. **Do not include** fast food, hot dogs, sausages, beef jerky, bacon, ham, or other deli or luncheon meats.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des œufs, du bœuf, du porc, du gibier, du poulet ou autre volaille, du poisson, des fruits de mer ou d'autres sources de protéines d'origine animale? Incluez le poisson en conserve et le poulet ou autre volaille en conserve. **N'incluez pas** le fast-food, les hot-dogs, les saucisses, le bœuf séché (*beef jerky*), le bacon, le jambon et autres viandes de charcuterie ou froides.
		Over the past month, how often did you eat nuts, seeds, tofu, beans and lentils, peanut butter or other nut butters, or other plant-based sources of protein? **Do not include** green beans or packaged veggie burgers and plant-based meats.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des arachides, des noix, des graines, du tofu, des lentilles ou autres légumineuses, du beurre d'arachides ou autres beurres de noix, ou d'autres sources de protéines d'origine végétale? **N'incluez pas** les haricots verts et les hamburgers végétariens prépréparés du commerce et les simili-viandes d'origine végétale prépréparées du commerce.
		Over the past month, how often did you eat yogurt, kefir, or cheese?	Au cours du dernier mois, à quelle fréquence avez-vous consommé du yogourt, du kéfir ou du fromage?
Among protein foods, consume plant-based more often	Choose protein foods that come from plants more often	Over the past month, how often did you eat nuts, seeds, tofu, beans and lentils, peanut butter or other nut butters, or other plant-based sources of protein? **Do not include** green beans or packaged veggie burgers and plant-based meats.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des arachides, des noix, des graines, du tofu, des lentilles ou autres légumineuses, du beurre d'arachides ou autres beurres de noix, ou d'autres sources de protéines d'origine végétale? **N'incluez pas** les haricots verts et les hamburgers végétariens prépréparés du commerce et les

			simili-viandes d'origine végétale prépréparées du commerce.
Water should be the beverage of choice	Make water your drink of choice • Replace sugary drinks with water	The screener does not assess water consumption.	Le questionnaire court ne demande pas la consommation d'eau.
	Healthy drink options other than water can include: ● White milk (0% and 1% milk) ● Unsweetened fortified plant-based beverages such as soy beverage or almond beverage ● Unsweetened coffee and tea	Over the past month, how often did you have **white** cows' milk or **unsweetened** plant-based beverages (e.g., soy, almond, or oat milk)? **Do not include** small amounts in coffee or tea, or chocolate and other sweetened milk.	Au cours du dernier mois, à quelle fréquence avez-vous consommé du lait de vache ou des boissons d'origine végétale **non sucrées** (par exemple, du lait de soya, du lait aux amandes, du lait d'avoine)? **N'incluez pas** les petites quantités dans le café ou le thé, le lait au chocolat et autres laits aromatisés sucrés.
		Over the past month, how often did you have chocolate milk or other **flavoured** milk or **sweetened** plant-based beverages (e.g., soy, almond, or oat milk)? **Do not include** small amounts in coffee or tea, or diet/artificially sweetened or sugar-free beverages.	Au cours du dernier mois, à quelle fréquence avez-vous consommé du lait au chocolat ou autres laits **aromatisés sucrés** ou des boissons d'origine végétale **sucrées** (par exemple, du lait de soya, du lait aux amandes, du lait d'avoine). **N'incluez pas** les petites quantités dans le café ou le thé, les boissons diètes, les boissons sucrées avec des édulcorants et les boissons sans sucre.
		Over the past month, how often did you drink fruit juice, fruit-flavoured drinks, soda or pop, **sweetened** sports or energy drinks, **sweetened** hot or iced coffee or tea, or **sweetened** waters? **Do not include** diet/artificially sweetened or sugar-free beverages, such as diet soda.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des jus de fruits, des boissons aromatisées aux fruits, des boissons gazeuses **contenant du sucre**, des boissons sportives **sucrées**, des boissons énergisantes **sucrées**, des cafés **sucrés** ou thés **sucrés,** chauds ou glacés, ou des eaux **sucrées**? **N'incluez pas** les boissons diètes, les boissons sucrées avec des édulcorants et les boissons sans sucre, comme les boissons gazeuses diètes.

Foods that contain mostly unsaturated fat should replace foods that contain mostly saturated fat	Choose foods with healthy fats instead of saturated fats	Over the past month, how often did you have margarine or vegetable oils (e.g., olive, canola, or sunflower oil)? **Do not include** lard, coconut oil, palm oil, or butter.	Au cours du dernier mois, à quelle fréquence avez-vous consommé de la margarine ou des huiles végétales (par exemple, l'huile d'olive, l'huile de canola, l'huile de tournesol)? **N'incluez pas** le saindoux, l'huile de coco, l'huile de palme et le beurre.
Processed or prepared foods and beverages that contribute to excess sodium, free sugars, or saturated fat undermine healthy eating and should not be consumed regularly	Limit highly processed foods. If you choose these foods, eat them less often and in small amounts. ● Prepare meals and snacks using ingredients that have little to no added sodium, sugars or saturated fat ● Choose healthier menu options when eating out	Over the past month, how often did you eat hot dogs, sausages, beef jerky, bacon, ham or other deli or luncheon meats? **Do not include** fast food, canned fish, canned poultry, or packaged veggie burgers and plant-based meats.	Au cours du dernier mois, à quelle fréquence avez-vous consommé des hot-dogs, des saucisses, du bœuf séché (*beef jerky*), du bacon, du jambon ou autres viandes de charcuterie ou froides? **N'incluez pas** le fast-food, le poisson en conserve, le poulet ou autre volaille en conserve, les hamburgers végétariens prépréparés du commerce et les simili-viandes d'origine végétale prépréparées du commerce.
		Over the past month, how often did you eat food from fast food restaurants, such as burgers, french fries, poutine, pizza, submarine sandwiches, fried chicken, burritos, or tacos?	Au cours du dernier mois, à quelle fréquence avez-vous consommé de la nourriture provenant d'un restaurant de type *fast-food*, comme des hamburgers, des frites, de la poutine, de la pizza, des sous-marins, du poulet frit, des burritos ou des tacos?
		Over the past month, how often did you eat cookies, cakes, muffins, pastries, granola bars, protein bars, ice cream, candy, chocolate, sugary breakfast cereals, or other sugary foods?	Au cours du dernier mois, à quelle fréquence avez-vous consommé des biscuits, des gâteaux, des muffins, des pâtisseries, des barres tendres, des barres protéinées, de la crème glacée, des bonbons, du chocolat, des céréales à déjeuner sucrées ou d'autres aliments sucrés?
		Over the past month, how often did you eat crackers, chips, pretzels, popcorn, or other salty snacks?	Au cours du dernier mois, à quelle fréquence avez-vous consommé des craquelins, des croustilles, des bretzels, du maïs soufflé ou autres collations salées?

		Over the past month, how often did you eat **white** breads, bagels, rice, pasta, noodles, or other refined grains, such as breakfast cereals? **Do not include** whole wheat or whole grain foods.	Au cours du dernier mois, à quelle fréquence avez-vous consommé du pain **blanc**, des bagels **blancs**, du riz **blanc**, des pâtes et des nouilles **blanches** ou d'autres grains **raffinés**, comme les céréales à déjeuner raffinées? **N'incluez pas** les aliments à base de blé entier ou de grains entiers.

Cognitive testing

Cognitive testing is a qualitative, psychologically oriented method to investigate the ways in which research participants interpret and respond to survey questions, typically via individual interviews (Foddy 1996; Willis 2005; Willis and Artino 2013; Willis and Miller 2011). The goal is to determine whether each question is understood consistently in the way researchers intend (Collins 2003; Beatty and Willis 2007; Willis and Artino 2013). Cognitive testing draws upon insights from psychology regarding the cognitive processes involved in responding to survey items. Tourangeau's four-stage model describes these processes, which include 1) comprehension, 2) retrieval of information, 3) judgment or estimation, and 4) selection of a response to the question (National Research Council 1984). Most cognitive testing procedures target the level of comprehension because this is where problems most often occur (Foddy 1996; Willis et al. 2013). A cognitive interviewing reporting framework proposed by Boeije and Willis (2013) was used to guide reporting of this aspect of the process.

Data collection

Cognitive interviews were conducted from April to June 2021. The interviews in English were conducted by researchers at the University of Waterloo and those in French by researchers at Université Laval. Interviews in each language were led by researchers (TEW and AB) with training in qualitative methods. Ethics review and approval was obtained from the University of Waterloo Office of Research Ethics (ORE #42994), the Université Laval Research Ethics Board (REB #2021-088), and the Health Canada and Public Health Agency of Canada Research Ethics Board (REB #2020-044H).

While cognitive testing is often conducted informally, research suggests small sample sizes may fail to detect problems with survey questions, including those that may introduce

measurement error (Blair et al. 2006). Recommendations for sample size range from 10 to 30 total participants, or five to 15 participants per round for two to three rounds (Beatty and Willis 2007; Willis and Artino 2013; Willis and Miller 2011). To adequately probe for potential issues with comprehension of the screener with individuals with varied sociodemographic characteristics, a sample of approximately 32 participants, to take part in 16 interviews in English and 16 in French, was sought. Interviews were conducted in three rounds in each language, with four to eight participants per round, to allow for iterative refinement of the screener (Beatty and Willis 2007).

Potential participants to complete interviews in English were recruited through community organizations and social media and invited to complete an eligibility questionnaire hosted on Qualtrics (Qualtrics, Provo, UT). Potential participants to complete interviews in French were recruited through community organizations and a database of potential research participants and invited to complete an eligibility questionnaire and return it to the research coordinator via email. Eligible individuals were aged 18-65 years, lived in Canada, and were able to read the screener and complete a 45–60-minute interview in English or French using online teleconferencing software. The eligibility questionnaire also captured information on age, gender identity, racial/ethnic identity, educational attainment, and perceived income adequacy. Quota sampling was used to seek a balance of participants with varying educational attainment. Specifically, the aim was for half of participants in the study to have not completed post-secondary education, as a proxy for lower literacy levels (Laramee et al., 2007; Schillinger et al., 2006). An approximate balance between women and men was sought, with a desire to include some individuals identifying as non-binary. Purposive sampling was used to maximize variation in other sociodemographic characteristics, including age, racial/ethnic identity (Black, East/Southeast Asian, Indigenous, Latino, Middle Eastern, South Asian, White), and perceived income adequacy (“Thinking about your total monthly income, how difficult is it for you to make

ends meet?", with response options including very difficult, difficult, neither easy nor difficult, easy, and very easy) (Litwin and Sapir 2009).

Eligibility questionnaire data were reviewed relative to the quotas and purposive sampling criteria on an ongoing basis, and eligible individuals were invited by email to participate in an interview. Those who agreed were sent an information letter and an informed consent form in advance. These documents were reviewed at the beginning of the interview, at which time the participant was asked to provide verbal informed consent. Interviews in English were conducted using Zoom (Zoom Video Communications, San Jose, CA), except for one conducted by telephone because the participant did not have Internet access, and interviews in French were conducted using Microsoft Teams (Microsoft Corporation, Redmond, WA). A note-taker captured details of participants' responses during the interview. Most interviews were audio-recorded, with participant consent, to allow researchers to review recordings as needed. Participants who completed an interview received a $20 CAD honorarium via Interac e-transfer in appreciation of their time. Following each interview, the interviewer and note-taker completed a debriefing form to note overall impressions and reflections on the interview process. By the end of the third round in each language, diminishing returns (Beatty and Willis 2007) were noted, in that few new problems were being identified, and recruitment and data collection were concluded.

Cognitive interview guide

Because the screener is intended for self-administration, participants were asked to complete it independently before they reviewed each question and their response process with the interviewer using a think-aloud approach (e.g., "can you walk me through how you arrived at that number?") (Beatty and Willis 2007; Lenzner et al. 2016). The interviewer then used open-

ended verbal probes to gauge understanding and thought processes, using a semi-structured interview guide developed in English and translated to French. The probes were aligned with the cognitive stages of processing. For example, processing of the question, “In the past month, how often did you consume fresh, frozen, and canned fruit?”, requires the respondent to understand and interpret keywords and phrases, including “how often”, “consume”, and “fruit”; to recall the correct response by thinking about how frequently they consumed fruit in the past month and to make a judgment about what number to report; and finally, to provide a response that matches the options available within the screener (e.g., “2 times per week”) . The corresponding probes asked, “are there any words or ideas in this question that were difficult to understand?”, targeting comprehension, and “how sure or unsure are you that the number you provided is accurate?”, targeting retrieval of information and judgment. Sorting of foods and beverages across questions and assessment of face and content validity were integrated by asking respondents what kinds of fruits they thought of and those they excluded when answering each question. This process was repeated for each screener question.

To ensure comparability of testing approaches (Willis and Miller 2011), the English- and French-speaking teams used consistent interview guides, and the lead interviewer from the French-speaking team observed pilot interviews conducted in English with graduate students not involved in this research.

Data analysis

As is common in cognitive testing (Willis 2005; Beatty and Willis 2007; Willis and Artino 2013; Willis and Miller 2011), the interview notes and recordings were informally coded based on the suggestion that formal coding may add limited value when the focus is on understanding potential issues with survey questions (Willis 2005). After each round of

interviews in each language, issues that may have required changes to ensure that questions were understood as intended were identified and summarized (Lenzner et al. 2016; Willis and Miller 2011). The two teams debriefed between rounds and following the final round. At each point, the screener was modified to address issues identified in each language. To ensure translational equivalency, a decentering approach that recognized that problems identified in one language may require changes to the screener in both languages was applied (Brislin 1970; Willis et al. 2008). Prior to the final round of testing in English and the second round of testing in French, issues that had arisen were discussed with the advisors, who provided feedback and suggested modifications to the screener for the next round of interviews.

For reporting purposes, the issues identified were subsequently grouped into problem categories, or themes related to cognitive processes (Bobrovitz et al. 2015; Lenzner et al 2016; Thompson et al. 2022; Willis and Miller 2011).

Face and content validity testing

Face and content validity testing was conducted to examine whether the screener was well-constructed and grounded in an understanding of the underlying phenomenon of interest (Frongillo, et al. 2019; Kirkpatrick et al. 2019b). This testing was completed by Health Canada from April to May 2021. Experts in nutrition, surveillance, and public health were identified by Health Canada based on existing contacts, including individuals who participated in the process to update the CFG-2019 and/or other relevant projects, and invited by email to participate. Advisors involved in the development of the screener (Appendix A) were not invited to participate in this phase. Ethics approval was obtained from the Health Canada and Public Health Agency of Canada Research Ethics Board (REB# 2020-044H). According to the associated policies for ethical research conduct, the face and content validity experts were not

considered participants since they were not themselves the focus of the research; therefore, informed consent was not required.

Content experts who agreed to participate were sent, via email, the version of the screener tested in the first round of cognitive testing (Appendix B) and a Microsoft Excel (Microsoft Corporation, Redmond, WA) spreadsheet prompting them to comment on whether each screener question reflected the guidance it was intended to capture and was easy to understand, with space for comments. The spreadsheet was returned to Health Canada staff via email and the results summarized, including identifying questions that did not perform well according to multiple experts, as well as any global feedback on the screener. These results were shared with the cognitive testing teams and informed modifications to the screener in advance of the final rounds of cognitive interviews in each language.

4.4. Results

Cognitive testing

For the interviews in English, 193 potential participants completed the eligibility questionnaire, of whom 136 met the eligibility criteria. A total of 22 individuals were contacted, and 17 completed an interview. For recruitment of French-speaking participants, 101 potential participants completed the eligibility questionnaire, of whom 97 were eligible, and 16 were contacted and completed an interview. A total of 33 interviews (17 in English and 16 in French) were conducted. In total, 21 participants identified as women and 12 identified as men (**Table 3**). No participants identified their gender as non-binary. Participants represented a mix of racial identities, though the majority (n=19) identified as White. About half of the participants (n=13) had less than post-secondary education.

Table 3: Characteristics of participants (n=33) in cognitive interviews in English and French to evaluate the comprehension and face validity of the Canadian Food Intake Screener

Characteristic	English	French	Total
	n		
Total	17	16	33
Age (years)			
18-24	3	3	6
25-34	5	3	8
35-44	2	3	5
45-54	3	3	6
55-65	4	4	8
Gender identity			
Man	5	7	12
Woman	12	9	21
Racial identity			
White	9	10	19
Indigenous	0	4	4
East/Southeast Asian	4	0	4
Black	2	0	2
Middle Eastern	0	1	1
South Asian	1	0	1
Latino	0	1	1
Prefer not to answer	1	0	1
Perceived income adequacy[1]			
Very easy	2	2	4
Easy	5	4	9
Neither easy nor difficult	6	8	14
Difficult	3	2	5
Very difficult	1	0	1
Educational attainment[2]			
High school graduate	2	7	9
Some college	1	0	1
Some university	4	0	4
College graduate	2	3	5
University graduate	6	6	12
Postgraduate training or degree	2	0	2

[1] Perceived income adequacy was based on a question that asked, "Thinking about your total monthly income, how difficult is it for you to make ends meet?" (Litwin and Sapir 2009).

[2]The French eligibility questionnaire did not include "Some college", "Some university", or "Postgraduate training or degree" as response options.

Issues with the screener questions identified during cognitive testing generally fell into one of four themes, mainly related to comprehension, described below**.**

Lack of clarity about what to include or exclude

Lack of clarity about what to include or exclude in responses to screener questions related to groupings of foods and forms of foods that can be consumed in different ways. With respect to food groupings, in some cases, participants were uncertain about the types of foods that should be reported in response to a given question. Such problems most often occurred due to a lack of detail or examples in a question (**Table 4, example 1**). Contrarily, confusion also occurred when examples were too specific, as participants were unsure whether to "think outside the box" or report only the foods listed (Table 4, **examples 2 and 3**). These ambiguities were addressed by adding detail to existing questions or adding new questions to the screener, as well as through formatting and ordering of questions.

Many foods can be consumed in multiple ways. For example, milk can be consumed as a beverage or used in a sauce or added to a bowl of cereal. Similarly, oil can be used in cooking or as part of a salad dressing. In cognitive testing, some participants were unsure whether certain foods should be reported only if consumed by a particular method and may have overlooked other methods (Table 4, **example 4**). To address this ambiguity, general terms were used to encompass multiple methods of consumption (e.g., *have* instead of *drink* milk).

Keyword confusion

Some questions in the initial screener included keywords that were unclear or vague to participants, particularly in the testing in French, creating opportunities for misinterpretation (Table 4, **examples 5 and 6**)**.** Lack of clarity was addressed by using more specific keywords

(e.g., “lait de vache” instead of lait and “cow’s milk” instead of milk) to describe the foods that should be included in responding to each question.

Readability

Readability issues occurred when the structure of a question hindered participants’ comprehension. When encountering lengthy lists of examples, participants tended to miss details and consequently provided inaccurate responses (Table 4, **examples 7 and 8**). Including parentheses within a question to provide additional examples of particular food categories hindered readability, as participants felt the parentheses cued them to stop reading (Table 4, **example 9**). Such problems were alleviated through formatting changes and by re-wording and simplifying phrasing, for example, by reducing the number or changing the order of examples or breaking a single question into multiple questions.

Response option errors

There were a few cases in which participants reported quantity, rather than frequency of consumption, referencing serving sizes detailed by prior versions of CFG (Table 4, **example 10**). Because the screener instructs participants to report frequency in both the preamble and each question, no changes were made to address this issue.

Face and content validity testing

For the testing in English, 21 content experts were invited and 13 accepted and for the testing in French, five content experts were invited and three accepted. These experts were academics, dietitians/nutritionists, and federal employees. Sociodemographic information was

not collected from the content experts, but there was perceived variation in sex and race/ethnicity.

Overall, the content experts generally agreed that the initial screener questions reflected the guidance and were easy to understand. Issues raised were often consistent with those arising in the cognitive interviews. For example, the experts noted that the question assessing highly processed foods was overly complicated, with too many examples, such that it would not be understood by those with lower literacy levels. The experts' feedback supported simplification of this question and others. Some experts suggested aligning the screener questions and structure more closely with the guidance, for instance, by ordering questions in a manner consistent with the guidance and including more detail, such as specifying additional example foods noted in the guidance.

Refinements to the screener

Although questions were generally well-understood in both languages, changes were made to the order of the screener's questions as well as examples within each question to improve clarity and make the screener more intuitive for respondents (Table 4, **examples 7 and 8**). Rearranging questions and examples appeared to improve readability and comprehension by cueing respondents on which foods to include or exclude and emphasizing aspects of the question that were otherwise overlooked. In alignment with the guiding principle to develop a screener that is simple to use and the cognitive testing findings, changes were not made to address suggestions from the experts to order questions consistent with the guidance itself and to include more example foods.

More substantial modifications included the addition of a question to assess frequency of potato consumption (Table 4, **example 1**) because cognitive testing revealed that some

participants tended not to include potatoes when asked about their vegetable consumption. Instead, they viewed potatoes as a starchy food that might belong with grain foods. The question regarding highly processed foods was simplified to clarify what foods should and should not be included in responses.

In some cases, issues identified in one language entailed changes to the screener in both languages. For example, the expression “viande rouge maigre” (“lean red meat”) was unclear to participants in the cognitive testing in French, although the issue did not arise in English (Table 4, **example 6**). To ensure translational equivalency, “viande rouge maigre” (“lean red meat”) was replaced with “boeuf et porc” (“beef and pork”) in both versions of the screener. In other cases, changes were required in only one language. For example, in the cognitive testing in English, the term “have” (i.e., “how often did you have”) improved comprehension of the screener compared to terms like “drink”, which were overly specific and tended to limit respondents’ thought processes (Table 4, **example 4**). However, a direct translation of the term “have” was not available in French, and the term “consommer” (to consume) did not elicit the same issues in the French cognitive testing. Comprehensibility was prioritized over translational equivalency in this case, retaining “consommer” in French and “have” in English.

Formatting was used strategically, including line spacing and judicious use of bolding for emphasis; however, underlining and italics were avoided in the final screener to improve accessibility (City of Peterborough 2014; Kovac 2018).

Table 4: Examples of issues identified in cognitive interviews in English and French to evaluate the comprehension and face validity of the Canadian Food Intake Screener[1]

Example	Problem category	Focus of question	Cognitive issue	Modifications
1	Lack of clarity about what to include or exclude/Food groupings	Vegetables	Several participants did not consider potatoes in their response related to frequency of consuming vegetables. Participants explained they view potatoes as a starch rather than a fresh vegetable, and some felt potatoes should be reported with grain foods instead.	A question probing potato consumption, prior to the vegetable question, was added. This modification was tested in the final rounds of cognitive testing in English and French and appeared to help participants understand where to include potato consumption.
2	Lack of clarity about what to include or exclude/Food groupings	Pre-made and ready-to-eat meals	The initial version of the screener asked, "How often did you consume deep-fried foods and ready-to-heat or ready-to-eat dishes?" (in French, "Au cours du dernier mois, à quelle fréquence avez-vous consommé des aliments frits et des plats prêts à réchauffer ou prêts à manger?"). Several participants expressed difficulty understanding which foods to include in their responses to this question. This question also did not perform well in face and content validity testing with experts.	"Ready-to-heat or ready-to-eat" was simplified to "pre-made meals". However, in the next rounds of interviews, participants were uncertain whether all take-out foods, or only deep-fried and fast foods, should be reported. This question was then replaced with a more specific one probing frequency of consumption of fast foods, such as pizza, burgers, and French fries, to clarify inclusion criteria, while aligning with examples of highly processed foods within CFG-2019 healthy food choices guidance.

Example	Problem category	Focus of question	Cognitive issue	Modifications
				The updated question was tested in subsequent rounds of cognitive interviews in English and French and found to be easier to understand compared to previous iterations.
3	Lack of clarity about what to include or exclude/Food groupings	Sugary foods	The initial version of this question included example foods such as sugary breakfast cereals, cookies, and cakes (in French, des céréales sucrées pour le déjeuner, des biscuits, des gâteaux). In cognitive testing in French, some participants were uncertain whether the list of examples was exhaustive, or whether sugary snacks not specifically mentioned should be included.	To indicate that the list of examples was not exhaustive, "or other sugary foods" ("ou d'autres aliments sucrés) was added after the list of examples. This modification was tested in subsequent rounds of cognitive interviews in English and French and appeared to help participants to "think outside the box" and consider foods not specifically mentioned in the question.
4	Lack of clarity about what to include or exclude/ Foods consumed in multiple ways	Unsweetened milks Oils	The screener question probing unsweetened milk consumption originally asked, "how often did you drink milk and unsweetened plant-based beverages"? (in French, "à quelle fréquence avez-vous bu du lait et des boissons d'origine végétale non	The phrases "drink" ("boire") and "cook with or add" ("cuisiner ou ajouter") were changed to the more general keyword, "consume" ("consommer") to indicate that the foods could be consumed in multiple ways. Though using "consommer" appeared to resolve

Example	Problem category	Focus of question	Cognitive issue	Modifications
			sucrées"). In cognitive testing in both English and French, participants expressed uncertainty about whether to report only milk drunk in a glass or if it was appropriate to report milk used as an ingredient or component of a dish. Similarly, the question probing oil consumption asked respondents, "how often did you cook with or add vegetable oils or soft margarines to your foods?" (in French, "combien de fois avez-vous cuisiné avec des margarines molles ou des huiles végétales[…] ou en avez-vous ajoutées à vos aliments?"). In cognitive testing in both English and French, several participants did not consider oils or margarines used outside of cooking, such as in salad dressing or spread on toast, until prompted by the interviewer.	the issue in French, three participants in the second round of testing in English reported confusion with the word "consume". In the English version of the screener, the question was rephrased using the word "have" (e.g., "how often did you have milk and unsweetened plant-based beverages?"). A direct translation of the term "have" was not available in French, and the term "consommer" (to consume) did not elicit the same issues in the cognitive testing in French. Comprehensibility was prioritized over translational equivalency in this case, retaining "consommer" ("à quelle fréquence avez-vous consommé du lait de vache ou des boissons d'origine végétale non sucrées") in French. The modifications were tested in subsequent rounds of cognitive interviews in English and French, and the problems did not recur.

Example	Problem category	Focus of question	Cognitive issue	Modifications
5	Keyword confusion	Unsweetened milks	In cognitive testing in French, several participants were uncertain whether "lait" ("milk") referred only to cow's milk or other types of milk, such as plant-based beverages or milk from other animals. This issue did not occur in cognitive testing in English.	To clarify that "lait" refers to cow's milk, the phrase was replaced with "lait de vache" ("cow's milk") in the French version of the screener. To maintain translational equivalency, "milk" was replaced with "cow's milk" in the English version of the screener. The modifications were tested in the final rounds of cognitive interviews in English and French, and found to perform well. This specification (inclusion of cow's milk/lait de vache) was not included in the question on sweetened flavored milks because the issue did not arise for that question, and so adding another keyword to the question was not warranted.

Example	Problem category	Focus of question	Cognitive issue	Modifications
6	Keyword confusion	Animal-based proteins	The animal-based protein question originally queried consumption of "lean red meat" (in French, "viande rouge maigre"). In cognitive testing in French, several participants were uncertain which types of meats would be considered "lean" ("maigre"). This issue did not occur in cognitive testing in English.	The phrase "lean red meat" ("viande rouge maigre") was replaced with the more specific keywords, "beef and pork" ("boeuf et porc"). Changes were made to both versions of the screener to maintain translational equivalency. This wording was found to perform well in the final rounds of cognitive interviews in English and French.
7	Readability	Animal-based proteins	This question initially placed eggs at the end of a list of animal-based proteins including red meat, poultry, and shellfish. A participant in the first round of cognitive interviews in English explained they only skimmed the question after reading the first few examples and assumed only meat products should be included. They therefore did not consider eggs until prompted by the interviewer.	Since the other example foods listed were meat (and shellfish) products, eggs were moved to the front of the list of example food items. This modification was tested in the second round of cognitive interviews in English and French, and the problem did not recur.

Example	Problem category	Focus of question	Cognitive issue	Modifications
8	Readability	Salty snacks	Crackers were initially placed at the end of a list of salty snacks such as chips and pretzels. Several participants in the English interviews did not consider crackers in their responses, as they assumed the question was asking about foods conventionally viewed as "junk" foods. In contrast, crackers are generally considered "healthier".	Crackers were moved to the front of the list of example food items. This modification was tested in the final rounds of cognitive interviews in English and French, and the problem did not recur.
9	Readability	Plant-based proteins Oils Sweetened milks	Initial drafts of the screener included parentheses within some questions containing examples of specific foods that should be considered in responding. In cognitive interviews in French, several participants expressed that the parentheses made the question difficult to read and understand. They noted that they stopped reading when they encountered parentheses, causing foods listed within the parentheses to be overlooked.	In both the English and French versions, parentheses were removed or moved to the end of the question to enhance readability. These modifications were tested in the second rounds of cognitive interviews in English and French and the problem did not recur.

Example	Problem category	Focus of question	Cognitive issue	Modifications
10	Response errors	Vegetables Whole grains	Two participants in the English interviews reported their responses as quantities rather than frequencies for the vegetable question, and one of those participants did the same for whole grains. Both referenced "servings" as specified by prior iterations of Canada's Food Guide.	No changes were made to the screener, as respondents are instructed to report frequency in both the preamble and in each individual screener question.

[1]The versions of the screener, in English and French, tested in the initial round of cognitive interviews, as well as in face and content validity testing with experts, are available in Appendix B.

Final screener

The final version of the screener in each language includes 16 questions (Table 2). Nine assess consumption of healthy foods to "eat each day", including fruit; vegetables; potatoes; animal-based protein foods; plant-based protein foods; yogurt, kefir, and cheese; unsweetened cow's milk and plant-based beverages; whole-grain foods; and margarine and vegetable oils. Seven questions assess "foods to limit", including processed meat, fast food, sweetened cow's milk and plant-based beverages, other sugary beverages, sugary snacks, salty snacks, and refined grains. The final screener is available, in both English and French, in **Appendix C**.

4.5. Discussion

The Canadian Food Intake Screener/Questionnaire court canadien sur les apports alimentaires, available in English and French, rapidly assesses overall alignment of dietary intake with the healthy food choices recommendations in CFG-2019. The screener is intended for use in research with adults, aged 18-65 years, with marginal and higher health literacy. The screener was developed through an iterative process that included three rounds of cognitive interviews in each language along with ongoing feedback from advisors, as well as formal face and content validity testing with a separate panel of content experts. Results suggested that the screener was well understood in both languages, and informed refinements to question wording and screener structure to improve comprehension and minimize cognitive load. The screener can be completed in approximately five minutes.

Cognitive testing is a valuable method for identifying and correcting problems within a survey or screener and goes beyond conventional pre-testing to comprehensively examine respondents' understanding of each item (Beatty and Willis 2007; Foddy 1996). Previous cognitive testing studies have revealed issues with ambiguous language and keyword

misinterpretation (Bobrovitz et al. 2015; Eland et al. 2022), which was also observed in the present study. Small details, including individual words, can change a question's meaning; thus, the cognitive interviews were valuable to ensure the screener questions were understood as intended. Seemingly minor tweaks to the language and structure, such as changing "drink" to "have" (in the English version of the screener) and reordering examples, improved comprehension, helping to address the guiding principle related to an easy-to-use screener. Cognitive interviews also exposed the challenge of designing questions that are specific enough to cue respondents on what to include and exclude, while not being so overly specific that they limit respondents' thinking. Thompson et al. (2022) found a similar issue in the development of a food literacy questionnaire, wherein participants thought too narrowly about a particular Context if the frame of reference was not well-defined. To address this issue in the current screener, example lists were kept as short and simple as possible, and in some cases, reference to "other" foods (e.g., other plant-based protein foods, other salty snacks) indicates that the list is not exhaustive. Providing exclusion criteria also appeared to help guide respondents on what should be included when responding to each question.

Potential uses of the screener are discussed in the accompanying paper (Hutchinson et al. 2022). The screener assesses intake over the past month, consistent with other screeners (Centers for Disease Control and Prevention n.d.; National Cancer Institute 2021; Wijnhoven et al. 2018). Adaptation to the past week or year is possible, but these time frames have not been evaluated. The screener captures the main elements of the healthy food choices recommendations from CFG-2019; however, given its brevity, it is not comprehensive. For example, it does not query all examples of highly processed foods (e.g., frozen entrées, sauces), though a range of foods noted as highly processed within the guidance are included (Health Canada 2022a). Given challenges in accurately measuring water consumption due to

consumption throughout the day that is not structured around meals (Gandy 2015), the screener does not query frequency of water intake.

The screener was developed and evaluated for use with adults aged 18-65 years. It may be amenable to self-administration by older children without substantial modification. However, recalling and reporting frequency of intake over the past month is likely to be cognitively challenging for younger children. It has been suggested that children can begin to conceptualize time at around ages seven to eight years and to self-report their own intake using a frequency-based instrument starting at around 10 years (Livingstone et al. 2004). Future research could evaluate the administration of the screener to caregivers as proxy reporters for younger children, similar to the implementation of 24-hour dietary recalls in national surveillance (Health Canada 2006; Health Canada 2017). Future research could also evaluate the screener for use with older adults. Evaluation of the screener used in the Canadian Longitudinal Study on Aging suggested differential performance among younger versus older adults (Gilsing et al. 2018), suggesting unique considerations related to age. Within that evaluation, considerations related to digital literacy among older adults were raised (Gilsing et al. 2018), but such concerns may be lessening in the digital era. Data from Statistics Canada indicate that the proportion of adults aged 65 years and older who accessed the Internet for personal use in the last three months increased from 48% in 2012 to 71% in 2018 (Statistics Canada 2019).

A guiding principle for the development and evaluation of the screener was to consider equivalence across population subgroups in capturing alignment with the healthy food choices recommendations. Equivalence relates to comparability (Boer et al. 2018; Frongillo et al., 2019; He and van de Vijver 2012) and can be threatened by construct bias, such that the construct intended to be measured is not the same across groups, as well as item bias, such that items have different meanings across groups (Boer et al. 2018; He and van de Vijver 2012). With

respect to language, French-speaking advisors were involved throughout the process and versions in English and French were tested and modified simultaneously to maximize translational equivalency (Hebestreit et al. 2017; Kwon et al. 2020; Vieira et al. 2020). Nonetheless, English-speaking individuals were more heavily represented among the advisors and experts.

The screener queries a range of foods consumed by diverse populations. For example, the question on animal-based protein foods queries beef, pork, wild meat, chicken, and shellfish; and various types of grains, including rice, pasta, noodles, and breads are listed. Further, reasonable variation in sociodemographic characteristics among cognitive testing participants was achieved, supporting relatively broad perspectives, including on the specific foods queried. However, the sample was skewed toward women, no participants identified as non-binary, and few identified as Black or Indigenous. The development process was informed by input from advisors who included researchers and practitioners from Canada, the USA, and Brazil, including experts in food-based dietary guidance and dietary assessment. Nonetheless, the field of nutrition and dietetics in Canada and elsewhere is relatively homogeneous (McBurney 2022) and structural barriers, including racism and heteronormativity, uphold this homogeneity (White 2018; Carter 2020; Burt et al. 2021; Joy and McSweeney-Flaherty 2022). Given the lack of representation of individuals with diverse and intersecting gender, racial/ethnic, and other identities, important perspectives related to dietary intake and its measurement among subgroups of the population may have been overlooked. Further evaluation of the screener with specific subgroups may thus be warranted to assess whether interpretation of the questions and their face and content validity are consistent. Moreover, it is critical to improve diversity and representation in the field to ensure that heterogeneity of the population is appropriately considered, as well as to heighten consideration of cross-context equivalence within dietary assessment.

Additional considerations are salient to the development of the screener. The sample size was consistent with recommendations for cognitive testing (Beatty and Willis 2007; Willis and Artino 2013; Willis and Miller 2011). However, some cognitive processes are difficult to verbalize, and it is possible that interviews may identify problems that would not occur when the screener is administered in the field or fail to identify issues that would emerge in the field (Beatty and Willis 2007). Informal coding has been used successfully in cognitive testing research, especially when time and resource constraints limit the ability to conduct full transcription of interview recordings (Willis 2005; Beatty and Willis 2007; Willis and Artino 2013; Willis and Miller 2011). To assess the potential impact of more intensive coding, the recordings of the interviews in English were subsequently transcribed and coded using the framework analysis method (Gale et al. 2017; Ritchie and Spencer 1994). Although additional instances of issues with comprehension of specific questions were noted, the issues themselves were identified by the informal coding, and no new issues likely to have prompted refinements to the screener were identified (Williams 2022).

4.6. Conclusion

The Canadian Food Intake Screener allows for rapid assessment of the overall alignment of adults' dietary intake with the healthy food choices recommendations within CFG-2019. Collaboration with a range of advisors, along with cognitive interviews and face and content validity testing, facilitated development of a simple screener in both English and French intended for use with adults with marginal and higher health literacy. The screener requires about five minutes to complete and is amenable to research and surveillance contexts in which it is not possible to conduct comprehensive dietary assessment. As is the case with the HEFI-2019 (Brassard et al. 2022a, 2022b), appropriate use of the screener can promote consistent assessment of alignment of adults' dietary intakes with CFG-2019 healthy food choices

recommendations. This is critical to creating an evidence base that can be synthesized to inform policies and programs to narrow the gap between current dietary intake and the guidance.

Competing interests: This project was funded by and conducted in collaboration with Health Canada, through a contract to SIK. SIK has received funding from Agriculture and Agri-Food Canada, AI for Good, the Canadian Institutes of Health Research (CIHR), the Canadian Foundation for Dietetic Research, Health Canada, the National Institutes of Health, the Ontario Ministry of Research and Innovation, and the Social Sciences and Humanities Research Council of Canada. SIK is a member of the Health Canada Nutrition Science Advisory Committee and the CIHR Institute of Nutrition, Metabolism, and Diabetes Institute Advisory board. SL has received funding from CIHR. BL has received funding from CIHR (ongoing), the Fonds de recherche du Québec – Santé (FRQS) (ongoing), Fonds de recherche du Québec – Nature et technologies (NT) (ongoing), the Ministère de la santé et des services sociaux (MSSS) du Québec (ongoing), Health Canada (completed in 2021) and Atrium Innovations (completed in 2019). BL is a member of the Canadian Nutrition Society Advisory Board. JH has received funding from the Canadian Foundation for Dietetic Research, CIHR, Danone Institute International, Danone Institute North America, Health Canada, and the National Institutes of Health. The remaining authors have no competing interests to disclose.

Funding: This project was funded by Health Canada through a contract with SIK. JMH received funding from a Social Sciences and Humanities Research Council (SSHRC) Vanier Canada Graduate Scholarship, and TEW received funding from a SSHRC Canada Graduate Scholarship – Master's.

Data availability: Due to the nature of the data, they cannot be made openly available.

4.7. References

Afshin, A., Sur, P.J., Fay, K.A., Cornaby, L., Ferrara, G., Salama, J.S., et al. 2019. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. The Lancet. **393**(10184): 1958–1972. Doi:10.1016/S0140-6736(19)30041-8. PMID:30954305.

Beatty, P.C., and Willis, G.B. 2007. Research synthesis: The practice of cognitive interviewing. Public Opinion Quarterly. **71**(2): 287–311. Doi:10.1093/poq/nfm006.

Blair, J., Conrad, F., Ackermann, A., and Claxton, G. 2006. The effect of sample size on cognitive interview findings. Proceedings of the American Statistical Association: 4041–4046.

Bobrovitz, N., Santana, M.J., Kline, T., Kortbeek, J. and Stelfox, H.T. 2015. The use of cognitive interviews to revise the Quality of Trauma Care Patient-Reported Experience Measure (QTAC-PREM). Quality of Life Research. **24**(8): 1911–1919. Doi:10.1007/s11136-015-0919-5. PMID:25589232.

Boeije, H., and Willis, G. 2013. The Cognitive Interviewing Reporting Framework (CIRF). European Journal of Research Methods for the Behavioral and Social Sciences. **9**(3): 87–95. Doi:10.1027/1614-2241/A000075.

Boer, D., Hanke, K., and He, J. 2018. On detecting systematic measurement error in cross-cultural research: a review and critical reflection on equivalence and invariance tests. Journal of Cross-Cultural Psychology. **49**(5): 713–734. Doi:10.1177/0022022117749042.

Brassard, D., Elvidge Munene, L.-A., St-Pierre, S., Gonzalez, A., Guenther, P.M., Jessri, M., et al. 2022b. Evaluation of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Applied Physiology, Nutrition, and Metabolism. **47**(5): 582-594. Doi:10.1139/APNM-2021-0416. PMID:35030069.

Brassard, D., Elvidge Munene, L.-A., St-Pierre, S., Guenther, P.M., Kirkpatrick, S.I., Slater, J., et al. 2022a. Development of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Applied Physiology, Nutrition, and Metabolism. **47**(5): 595–610. Doi:10.1139/APNM-2021-0415. PMID:35030038.

Brislin, R.W. 1970. Back-translation for cross-cultural research. Journal of Cross-Cultural Psychology. **1**(3): 185–216. Doi:10.1177/135910457000100301.

Burt, K.G., Lopez, R., Landaverde, M., Paniagua, A., and Avalos, E. 2021. Systemic and institutionalized racism, not achievement gap factors, limit the success of Black, Indigenous, and People of Color in dietetics education and credentialing. Journal of Critical Dietetics. **6**(1).

Carter, S. 2020. Systemic racism and implicit bias in nutrition and dietetics. Food & Nutrition Magazine. Available from https://foodandnutrition.org/blogs/stone-soup/systemic-racism-implicit-bias-nutrition-dietetics/ [Accessed 3 August 2022].

Centers for Disease Control and Prevention. (n.d.). Surveillance of fruit and vegetable intake using the Behavioral Risk Factor Surveillance System. Centers for Disease Control and

Prevention, Atlanta, GA. Available from https://www.cdc.gov/brfss/pdf/fruits_vegetables.pdf [Accessed 3 August 2022].

Choi, A., and Cawley, J. 2018. Health disparities across education: the role of differential reporting error. Health Economics. **27**(3): e1–e29. Doi:10.1002/HEC.3609. PMID:29210133.

City of Peterborough. Guide to Accessible Documents - AODA. 2014. Version 3. City of Peterborough, Peterborough, ON. Available from https://iwscc.ca/wp-content/uploads/2017/06/Guide-to-Accessible-Documents.pdf [Accessed 3 August 2022].

Colby, S., Zhou, W., Allison, C., Mathews, A.E., Olfert, M.D., Morrell, J.S., et al. 2020. Development and validation of the short healthy eating index survey with a college population to assess dietary quality and intake. Nutrients. **12**(9): 1–24. Doi:10.3390/nu12092611. PMID:32867172.

Collins D. 2003. Pretesting survey instruments: an overview of cognitive methods. Quality of life research. **12**(3): 229–38. Doi:10.1023/a:1023254226592. PMID:12769135.

De Rijk, M.G., Slotegraaf, A.I., Brouwer-Brolsma, E.M., Perenboom, C.W.M., Feskens, E.J.M., and de Vries, J.H.M. 2021. Development and evaluation of a diet quality screener to assess adherence to the Dutch food-based dietary guidelines. British Journal of Nutrition. 1–11. Doi:10.1017/S0007114521004499. PMID:34776025.

Eland, N.D., Strand, L.I., Ostelo, R.W., Kvåle, A. and Magnussen, L.H. 2022. How do physiotherapists understand and interpret the "Pain Attitudes and Beliefs Scale"? A cognitive interview study. Physiotherapy Theory and Practice. **38**(4): 513–527. Doi:10.1080/09593985.2020.1774949. PMID:32520655.

Foddy, W. 1996. The in-depth testing of survey questions: a critical appraisal of methods. Quality and Quantity. **30**(4): 361–370. Doi:10.1007/BF00170142.

Food and Agriculture Organization. 2022. Food-based dietary guidelines. Food and Agriculture Organization, Rome, Italy. Available from https://www.fao.org/nutrition/education/food-dietary-guidelines/background/en/ [Accessed 3 August 2022].

Frongillo, E.A., Baranowski, T., Subar, A.F., Tooze, J.A., and Kirkpatrick, S.I. 2019. Establishing validity and cross-context equivalence of measures and indicators. Journal of the Academy of Nutrition & Dietetics. **119**(11): 1817-1830. Doi:10.1016/j.jand.2018.09.005. PMID:30470590.

Gabe, K., and Jaime, P. 2019. Development and testing of a scale to evaluate diet according to the recommendations of the Dietary Guidelines for the Brazilian Population. Public Health Nutrition. **22**(5): 785–796. Doi:10.1017/S1368980018004123. PMID:30744711.

Gandy J. 2015. Water intake: validity of population assessment and recommendations. European Journal of Nutrition. **54**(2): 11-6. Doi:10.1007/s00394-015-0944-8. PMID:26048039.

Government of Canada. 2022. Food and Drug Regulations (C.R.C., c. 870). Government of Canada, Ottawa, ON. Available from https://laws-lois.justice.gc.ca/eng/regulations/C.R.C.,_c._870/index.html [Accessed 11 November 2022].

Gnagnarella, P., Dragà, D., Misotti, A.M., Sieri, S., Spaggiari, L., Cassano, E., et al. 2018. Validation of a short questionnaire to record adherence to the Mediterranean diet: an Italian experience. Nutrition, Metabolism and Cardiovascular Diseases. **28**(11): 1140–1147. Doi:10.1016/J.NUMECD.2018.06.006. PMID:3007749.

Haines J, Martin A, Perreault M. 2022. Questionnaire (screener) to measure eating practices based on Canada's Food Guide eating habits recommendations. Canadian Nutrition Society Webinar: How to use newly developed research tools to measure food choices and eating behaviours based on the 2019 Canada's Food Guide.

Harrison, S., Brassard, D., Lemieux, S., and Lamarche, B. 2019. Consumption and sources of saturated fatty acids according to the 2019 Canada Food Guide: data from the 2015 Canadian Community Health Survey. Nutrients. **11**(9): 1964. Doi:10.3390/NU11091964. PMID:31438574.

He, J., and van de Vijver, F. 2012. Bias and equivalence in cross-cultural research. Online Readings in Psychology and Culture. **2**(2). Doi:10.9707/2307-0919.1111.

Health Canada. 2006. Canadian Community Health Survey. Cycle 2.2, Nutrition (2004): A guide to accessing and interpreting the data. Health Canada, Ottawa, ON. Available from https://www.hc-sc.gc.ca/fn-an/alt_formats/hpfb-dgpsa/pdf/surveill/cchs-guide-escc-eng.pdf [Accessed 3 August 2022].

Health Canada. 2017. Reference Guide to Understanding and Using the Data 2015 Canadian Community Health Survey—Nutrition. Health Canada, Ottawa, ON. Available from https://www.canada.ca/content/dam/hc-sc/documents/services/food-nutrition/food-nutrition-surveillance/ReferenceGuide2015CCHS-Nutr_Eng_Final_06192017.pdf [Accessed 3 August 2022].

Health Canada. 2020. Healthy Eating Recommendations. Health Canada, Ottawa, ON. Available from https://food-guide.canada.ca/en/healthy-eating-recommendations/ [Accessed 3 August 2022].

Health Canada. 2021. Revision process for Canada's Food Guide. Health Canada, Ottawa, ON. Available from https://www.canada.ca/en/health-canada/services/canada-food-guide/about/revision-process.html [Accessed 3 August 2022].

Health Canada. 2022b. Canada's Dietary Guidelines for Health Professionals and Policy Makers. Health Canada, Ottawa, ON. https://food-guide.canada.ca/en/guidelines/ [Accessed 3 August 2022].

Health Canada. 2022a. Canada's Food Guide. Health Canada, Ottawa, ON. Available from https://food-guide.canada.ca/en/ [Accessed 3 August 2022].

Hebestreit K., Yahiaoui-Doktor M., Engel C., Vetter W., Siniatchkin M., Erickson N., et al. 2017. Validation of the German version of the Mediterranean Diet Adherence Screener (MEDAS) questionnaire. BMC Cancer. **17**(1): 341. Doi:10.1186/s12885-017-3337-y. PMID:28521737.

Herforth, A., Arimond, M., Álvarez-Sánchez, C., Coates, J., Christianson, K., and Muehlhoff, E. 2019. A global review of food-based dietary guidelines. Advances in Nutrition. **10**(4): 590. Doi:10.1093/ADVANCES/NMY130. PMID:31041447.

Hutchinson, J.M., Dodd, K.W., Guenther, P.M., Lamarche, B., Haines, J., Wallace, A., Perreault, M., Lemieux, S., Olstad, D.L., Prowse, R., Randall Simpson, J., Vena, J.E., Szajbely, K., and Kirkpatrick, S.I. 2022 The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: Scoring system and construct validity. MedRxiv. Preprint. Doi: 10.1101/2022.12.19.22283575.

Joy, P., and McSweeney-Flaherty, J.M. 2022. Moving dietetics forward with queer pedagogy: a post-structural qualitative study exploring the education and training experiences of Canadian dietitians for LGBTQ care. Journal Academy of Nutrition and Dietetics. Doi:10.1016/J.JAND.2022.02.011. PMID:35217245.

Keogh, R.H., Shaw, P.A., Gustafson, P., Carroll, R.J., Deffner, V., Dodd, K.W., et al. 2020. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1 – basic theory and simple methods of adjustment. Statistics in Medicine. **39**(16): 2197. Doi:10.1002/SIM.8532. PMID:32246539.

Kirkpatrick, S.I., Baranowski, T., Subar, A.F., Tooze, J.A., and Frongillo, E.A. 2019b. Best practices for conducting and interpreting studies to validate self-report dietary assessment methods. Journal Academy of Nutrition and Dietetics. **119**(11): 1801–1816. Doi:10.1016/j.jand.2019.06.010. PMID:31521583.

Kirkpatrick, S.I., Raffoul, A., Lee, K.M., and Jones, A.C. 2019a. Top dietary sources of energy, sodium, sugars, and saturated fats among Canadians: insights from the 2015 Canadian Community Health Survey. Applied Physiology, Nutrition and Metabolism. **44**(6): 650–658. Doi:10.1139/APNM-2018-0532. PMID:30951373.

Kirkpatrick, S.I., Guenther, P.M., Subar, A.F., Krebs-Smith, S.M., Herrick, K.A., Freedman, L.S., et al. 2022b. Using short-term dietary intake data to address research questions related to usual dietary intake among populations and subpopulations: assumptions, statistical techniques, and considerations. Journal Academy of Nutrition and Dietetics. **122**(7): 1246-1262. Doi:10.1016/J.JAND.2022.03.010. PMID:35283362.

Kirkpatrick, S.I., Troiano, R.P., Barrett, B., Cunningham, C., Subar, A.F., Park, Y., et al. 2022a. Measurement error affecting web- and paper-based dietary assessment instruments: insights from the multi-cohort Eating and Activity Study for Understanding Reporting Error. American Journal of Epidemiology. **191**(6): 1125–1139. Doi:10.1093/AJE/KWAC026. PMID:35136928.

Kovac L. 2018. Accessible Writing Style. Accessibility for Ontarians with Disabilities Act. Accessibility for Ontarians with Disabilities Act. Available from https://aoda.ca/accessible-writing-style/ [Accessed 3 August 2022].

Kwon, Y.-J., Lee, H., Yoon, Y., Kim, H.M., Chu, S.H., and Lee, J.-W. 2020. Development and validation of a questionnaire to measure adherence to the Mediterranean diet in Korean adults. Nutrients. **12**(4): 1102. Doi:10.3390/nu12041102. PMID:32316107.

Lafrenière, J., Lamarche, B., Laramée, C., Robitaille, J., and Lemieux, S. 2017. Validation of a newly automated web- based 24-hour dietary recall using fully controlled feeding studies. BMC Nutrition. **3**(34): 1–10. Doi:10.1186/s40795-017-0153-3. PMID: 32153814

Lafrenière, J., Harrison, S., Laurin, D., Brisson, C., Talbot, D., Couture, P., et al. 2019. Development and validation of a Brief Diet Quality Assessment Tool in the French-speaking adults from Quebec. International Journal of Behavioral Nutrition and Physical Activity. **16**(1): 1-0. Doi:10.1186/s12966-019-0821-6. PMID:31387609.

Laramee, A.S., Morris, N., and Littenberg, B. 2007. Relationship of literacy and heart failure in adults with diabetes. BMC Health Serv Res. **7**(98). Doi: 10.1186/1472-6963-7-98. PMID: 17605784.

Lenzner, T., Neuert, C., and Otto, W. 2016. GESIS Survey Guidelines: Cognitive Pretesting. Mannheim, Germany. Doi:10.15465/gesis-sg_en_010.

Lim, S.S., Vos, T., Flaxman, A.D., Danaei, G., Shibuya, K., Adair-Rohani, H., et al. 2012. A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. **380**(9859): 2224–60. Doi:10.1016/S0140-6736(12)61766-8. PMID:23245609.

Litwin, H. and Sapir, E.V., 2009. Perceived income adequacy among older adults in 12 countries: findings from the survey of health, ageing, and retirement in Europe. The Gerontologist. **49**(3): 397-406. Doi:10.1093/geront/gnp036. PMID:19386829.

Livingstone, M.B.E., Robson, P.J., and Wallace, J.M.W. 2004. Issues in dietary intake assessment of children and adolescents. British Journal of Nutrition. **92**(S2): S213-22. Doi:10.1079/BJN20041169. PMID:15522159.

Lozano, R., Naghavi, M., Foreman, K., Lim, S., Shibuya, K., Aboyans, V., et al. 2012. Global and regional mortality from 235 causes of death for 20 age groups in 1990 and 2010: a systematic analysis for the Global Burden of Disease Study 2010. The Lancet. **380**(9859): 2095–128. Doi:10.1016/S0140-6736(12)61728-0. PMID:23245604.

McBurney, M.I. 2022. Are professors of human nutrition faculty at Canadian universities representative with respect to common social constructs of gender and race? Applied Physiology, Nutrition and Metabolism. **47**(5): 517–520. Doi:10.1139/apnm-2021-0771. PMID:35138933.

Millen, A.E., Midthune, D. Thompson, F.E., Kipnis, V., and Subar, A.F. 2006. The National Cancer Institute Diet History Questionnaire: validation of pyramid food servings. American Journal of Epidemiology. **163**(3): 279-288. Doi:10.1093/aje/kwj031. PMID:16339051.

Natarajan, L., Pu, M., Fan, J., Levine, R., A, Patterson, R.E., Thomson, C.A., et al. 2010. Measurement error of dietary self-report in intervention trials. American Journal of Epidemiology. **172**(7): 819-27. Doi:10.1093/aje/kwq216. PMID:20720101.

National Cancer Institute. 2015. Dietary Assessment Primer. National Cancer Institute, Bethesda, MD. Available from https://dietassessmentprimer.cancer.gov/ [Accessed 3 August 2022].

National Cancer Institute. 2019. Dietary Screener in the 2009 California Health Interview Survey. National Cancer Institute, Bethesda, MD. Available from https://epi.grants.cancer.gov/diet/screeners/CHIS2009.pdf [Accessed 3 August 2022].

National Cancer Institute. 2021. Dietary Screener Questionnaire in the NHANES 2009-10: Background. National Cancer Institute, Bethesda, MD. Available from https://epi.grants.cancer.gov/nhanes/dietscreen/ [Accessed 3 August 2022].

National Cancer Institute. 2022. Diet History Questionnaire III. Available from https://epi.grants.cancer.gov/dhq3/ [Accessed 3 August 2022].

National Research Council. 1984. Cognitive aspects of survey methodology: building a bridge between disciplines. National Academies Press, Washington, D.C. doi:10.17226/930.

Ritchie, J., Spencer, L. 1994. Qualitative data analysis for applied policy research. *In* Analyzing Qualitative Data. Routledge, London. Pp. 173–94.

Schillinger, D., Barton, L.R., Karter, A.J., Wang, F., and Adler, F. 2006. Does literacy mediate the relationship between education and health outcomes? A study of a low-income population with diabetes. Public Health Reports. **121**(3): 245-254. Doi: 10.1177/003335490612100305. PMID: 16640146.

Subar, A.F., Kirkpatrick, S.I., Mittl, B., Zimmerman, T.P., Thompson, F.E., Bingley, C., et al. 2012. The Automated Self-Administered 24-hour dietary recall (ASA24): a resource for researchers, clinicians, and educators from the National Cancer Institute. Journal of the Academy of Nutrition and Dietetics. **112**(8): 1134–1137. Doi:10.1016/j.jand.2012.04.016. PMID: 22704899.

Tangney, C., Rasmussen, H., Richards, C., Li, M., and Appelhans, B. 2019. Evaluation of a brief sodium screener in two samples. Nutrients. **11**(1). Doi:10.3390/NU11010166. PMID:30646541.

Thompson C, Adams J, Vidgen HA. 2022. Progressing the development of a food literacy questionnaire using cognitive interviews. Public Health Nutrition. **25**(7):1968-78. Doi:10.1017/S1368980021004560. PMID:34743775.

Tugault-Lafleur, C.N., and Black, J.L. 2019. Differences in the quantity and types of foods and beverages consumed by Canadians between 2004 and 2015. Nutrients. **11**(3): 526. Doi:10.3390/NU11030526. PMID:30823448.

Vieira, L.M., Gottschall, C.B.A., Vinholes, D.B., Martinez-Gonzalez, M.A., Marcadenti, A. 2020. Translation and cross-cultural adaptation of 14-item Mediterranean Diet Adherence Screener and low-fat diet adherence questionnaire. Clinical Nutrition ESPEN. **39**: 180-189. Doi:10.1016/j.clnesp.2020.06.018. PMID:32859314.

White, J. 2018. Barriers to becoming registered dietitians identified by African American students and practitioners. Journal of Critical Dietetics. **4**(1): 41-46. Available from https://criticaldieteticsblog.files.wordpress.com/2018/07/41-46-white.pdf

Wijnhoven, H., Elstgeest, L., de Vet, H., Nicolaou, M., Snijder, M., and Visser, M. 2018. Development and validation of a short food questionnaire to screen for low protein intake in

community-dwelling older adults: The Protein Screener 55+ (Pro55+). Public Library of Science ONE. **13**(5). Doi:10.1371/JOURNAL.PONE.0196406. PMID:29791454.

Williams, T. 2022. Cognitive testing of a brief dietary screener: A comparison of analysis methods. UWSpace. Available from *http://hdl.handle.net/10012/18766* [Accessed 29 September 2022].

Willis, G.B. 2005. Cognitive Interviewing. SAGE Publications, Inc. doi:10.4135/9781412983655.

Willis, G.B., and Artino, A.R. 2013. What do our respondents think we're asking? Using cognitive interviewing to improve medical education surveys. Journal of Graduate Medical Education. **5**(3): 353–356. Doi:10.4300/jgme-d-13-00154.1. PMID:24404294.

Willis, G.B., and Miller, K. 2011. Cross-cultural cognitive interviewing: seeking comparability and enhancing understanding. Field Methods. **23**(4): 331-341. Doi:10.1177/1525822X11416092.

Willis, G.B., Lawrence, D., Hartman, A., Stapleton Kudela, M., Levin, K., and Forsyth, B. 2008. Translation of a tobacco survey into Spanish and Asian languages: the Tobacco Use Supplement to the Current Population Survey. Nicotine Tobacco Research. 10(6): 1075–84. Doi:10.1080/14622200802087572. PMID:18584471.

Chapter 5: The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: Scoring system and construct validity

Status: Accepted version. Published in *Applied Physiology, Nutrition, and Metabolism*, 2023; 48(8): 620-633. Doi: 10.1139/apnm-2023-0018.

Authors: Joy M. Hutchinson[1], Kevin W. Dodd[2], Patricia M. Guenther[3], Benoit Lamarche[4,5], Jess Haines[6], Angela Wallace[6], Maude Perreault[6], Tabitha E. Williams[1], Maria Laura da Costa Louzada[7], Mahsa Jessri[8], Simone Lemieux[4,5], Dana Lee Olstad[9], Rachel Prowse[10], Janis Randall Simpson[6], Jennifer E. Vena[11], Kathleen Szajbely[1], Sharon I. Kirkpatrick[1]

1. School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada
2. Division of Cancer Prevention, National Cancer Institute, Bethesda, MD, USA
3. Department of Nutrition and Integrative Physiology, University of Utah, Salt Lake City, UT, USA
4. Centre Nutrition, santé et société (NUTRISS), Institut sur la nutrition et les aliments fonctionnels (INAF), Université Laval, QC, Canada
5. École de nutrition, Faculté des sciences de l'agriculture et de l'alimentation, Université Laval, Québec, QC, Canada
6. Department of Family Relations & Applied Nutrition, University of Guelph, Guelph, Ontario, Canada
7. Department of Nutrition, School of Public Health; Center for Epidemiological Research in Nutrition and Health, University of São Paulo, São Paulo, Brazil
8. Food, Nutrition and Health Program, Faculty of Land and Food Systems; Centre for Health Services and Policy Research, Faculty of Medicine, University of British Columbia, BC, Canada

9. Department of Community Health Sciences, Cumming School of Medicine, University of Calgary, Calgary, AB, Canada
10. Division of Community Health and Humanities, Faculty of Medicine, Memorial University of Newfoundland, St. John's, NL, Canada
11. Cancer Research & Analytics, Cancer Care Alberta, Alberta Health Services, Calgary, AB, Canada

Acknowledgements: The authors are grateful to the participants in the construct validity study, and to Kimberley J. Hernandez, Isabelle Rondeau, Lisa-Anne Elvidge Munene, Sylvie St-Pierre, and other Health Canada employees who participated in and provided input on various aspects of the project. The authors are also grateful to Dr. Didier Brassard from McGill University (formerly of Université Laval) for providing advice on the application of HEFI-2019 and Sanaa Hussain from the University of Waterloo for support with data cleaning and creation of the code book, as well as referencing.

5.1. Overview

The Canadian Food Intake Screener/Questionnaire court canadien sur les apports alimentaires was developed to rapidly assess alignment of adults' dietary intake over the past month with the 2019 Canada's Food Guide's healthy food choices recommendations. From July to December 2021, adults (n=154) aged 18-65 years completed the screener and up to two 24-hour dietary recalls. The screener scoring system was aligned with the Healthy Eating Food Index-2019 (HEFI-2019), to the extent possible. ANOVA compared screener scores among subgroups with known differences in diet quality. Using the recall data, the National Cancer Institute multivariate method was used to model HEFI-2019 components, with the screener score as a covariate, and the correlation coefficient between screener and total HEFI-2019 scores was estimated. The mean screener score was 35 points (SD=4.7; maximum, 65), ranging from 25 (1st percentile) to 45 (99th percentile). Differences in scores in hypothesized directions were evident by gender identity (p=0.06), perceived income adequacy (p=0.07), education (p=0.02), and smoking status (p=0.003). The correlation between screener and HEFI-2019 scores was 0.53 (SE=0.12). The screener's moderate construct validity supports its use for rapid assessment of alignment of adults' intake with the healthy food choices recommendations when comprehensive dietary assessment is not possible.

5.2. Introduction

In 2019, Health Canada released an updated Canada's Food Guide (CFG-2019), with recommendations on healthy food choices and healthy eating habits for Canadians aged two years and older (Health Canada 2022). The healthy food choices recommendations contain advice on the types of foods and beverages to eat each day and to limit (Health Canada 2022). Researchers and practitioners are often interested in assessing alignment between dietary intake and food-based dietary guidance. The Healthy Eating Food Index-2019 (HEFI-2019) was developed to assess alignment of intakes with the CFG-2019 healthy food choices recommendations using 24HDR data (Brassard et al. 2022a, 2022b). When it is not possible to administer 24HDR or other comprehensive dietary assessment instruments, brief dietary questionnaires, informally known as "screeners", can provide rapid insights into dietary intake (National Cancer Institute 2015; Thompson et al. 2015). For example, the Behavioral Risk Factor Surveillance System (Centers for Disease Control and Prevention n.d.) fruit and vegetable module has been administered in the Canadian Community Health Survey (CCHS) as an indicator of the healthfulness of dietary intake (Statistics Canada 2021).

The Canadian Food Intake Screener/Questionnaire court canadien sur les apports alimentaires (Appendix C) was developed to assess alignment of adult's dietary intake over the past month with the CFG-2019 healthy food choices recommendations (Health Canada 2022). The 16-question screener is intended for self-administration by adults aged 18-65 years with marginal and higher health literacy (Hutchinson et al. 2022). In collaboration with Health Canada and external expert advisors (Appendix A), guiding principles were defined to inform the screener's development and evaluation. The final versions in English and French were informed by three rounds of cognitive testing with adults aged 18 to 65 years and face and content validity testing with a separate panel of experts, as described elsewhere (Hutchinson et al. 2022). Although the CFG-2019 plate depicts the suggested proportionality of different food

categories (Health Canada 2022), the screener is frequency-based in keeping with the guiding principle of a simple screener and because not all recommendations are expressed using proportionality. Given evidence of gaps between the dietary intake of individuals aged two years and older and the healthy food choices recommendations (Brassard et al. 2022b), a focus on frequency of intake of healthy foods and foods to limit (Health Canada 2022) was deemed likely to differentiate higher from lower alignment with the recommendations.

Construct validity indicates whether a tool is suitable for its intended purpose, as determined by whether it performs as expected based on the theory underlying its development (Frongillo et al. 2019; Kirkpatrick et al. 2019). The objective of this study was to assess the construct validity of the Canadian Food Intake Screener by examining variability in screener scores, whether the screener produces scores that differ between subgroups with known differences in dietary quality, and the correlation coefficient between screener scores and total HEFI-2019 scores among a sample of adults. The screener's scoring system is also described.

5.3. Methods

Data collection was conducted from July to December 2021. The study was reviewed by and received ethics clearance through a University of Waterloo Research Ethics Committee (ORE #42973), the University of Guelph Research Ethics Board (REB #21-04-009), and the Health Canada and Public Health Agency of Canada Research Ethics Board (REB #2020-044H). Reporting follows the *Strengthening the Reporting of Observational Studies in Epidemiology-Nutrition* (STROBE-nut) guidelines (Lachat et al. 2016).

Reference measure and sample size calculation

In the absence of an unbiased measure of usual diet quality relative to the healthy food choices recommendations, total HEFI-2019 scores, based on repeat 24HDRs, were selected as the reference measure for examining the screener's construct validity. The HEFI-2019 has been shown to detect sufficient variation in scores reflecting alignment with the healthy food choices recommendations and to differentiate among subgroups with known differences in diet quality (Brassard et al. 2022b). The evaluation of the HEFI-2019 drew upon 24HDRs (Brassard et al. 2022b), which have been shown to capture intake with less bias than other self-report dietary assessment methods, such as food frequency questionnaires (Freedman et al. 2014, 2015). The collection of repeat non-consecutive 24HDRs enables application of methods to estimate HEFI-2019 scores based on the simulated distribution of usual intakes (Zhang et al. 2011; Brassard et al. 2022b).

The sample size calculation was based on the expected correlation coefficient between screener scores and total HEFI-2019 scores, estimated using repeat non-consecutive 24HDRs. A correlation coefficient >0.4 is often considered to indicate an acceptable association, whereas coefficients of 0.5-0.7 are often observed for frequency-based dietary questionnaires (Cade et al. 2007, Willett 2012) and are considered to indicate a reasonably good association (de Rijk et al. 2021; Gicevic et al. 2021; Gilsing et al. 2018; Schröder et al. 2011). Based on a hypothesized correlation coefficient of at least 0.4 (null hypothesis of 0.2, with an alpha of 0.05 (two-sided), 80% power), a sample size of 163 participants was needed. This hypothesized coefficient was based on the anticipated performance of a multidimensional screener, as well as observed correlation coefficients between other dietary screeners and various reference instruments (Colby et al. 2020; Gicevic et al. 2021; Schröder et al. 2011, 2012).

Sampling and data collection

A diverse sample of adults aged 18-65 years living in Canada who could complete the study in English or French was sought. Quota sampling was used to seek one-third of the sample completing the study in French, a planned oversample given that French is the first official language spoken among 21% of individuals in Canada (Government of Canada 2022). Quotas were also applied for gender identity, seeking at least 30% of participants identifying as men, given that nutrition research is often skewed toward women (Colby et al. 2020; de Rijk et al. 2021); racial/ethnic identity, seeking at least 25% identifying as a racial/ethnic identity other than White; and educational attainment, seeking at least 25% who had completed college or less. Exclusion criteria included having advanced training in nutrition, following a dietary pattern related to a health condition (e.g., gluten-free, low protein, enteral or parenteral nutrition), and self-identifying as currently undergoing treatment for a condition that may alter dietary intake (e.g., disordered eating, cancer).

In the absence of a sampling frame or resources to contract a survey firm, social media was used for recruitment to avoid a convenience sample, for example, of students. Postings on Twitter, Facebook, and Instagram described the study and invited prospective participants to visit an online questionnaire to assess their eligibility. The eligibility questionnaire was hosted on Qualtrics (Qualtrics, Provo, UT) and collected information on age, whether the potential participant resided in Canada, whether they could read and respond to questions in English or French, and characteristics associated with the exclusion criteria and quotas, noted above. Questions related to gender identity, racial/ethnic identity, and educational attainment were based on prior surveys (Hammond et al. 2022; Statistics Canada 2022). Participants who met the eligibility criteria were invited to complete informed consent and the survey if quotas were not yet filled, while those filtered out due to quotas having been achieved were thanked for their interest.

The survey, hosted on Qualtrics (Qualtrics, Provo, UT), included the Canadian Food Intake Screener, as well as a screener to assess adults' alignment with the CFG-2019 healthy eating habits recommendations, described elsewhere (Haines et al., 2022). The survey queried sociodemographic and health characteristics, using questions informed by prior surveys (Hammond et al. 2022; Statistics Canada 2022). Participants who completed the survey received a $15 Interac e-transfer or grocery store gift card. Data quality checks were based on survey completion time (not less than 5 minutes), completing the survey only once, and correct responses to at least one of three data integrity questions, as well as peaks in survey completion and mismatches between names and email addresses that were indicators of likely bogus participants (Pei et al. 2020; Storozuk et al. 2020).

Participants passing data quality checks were invited by email to complete two 24HDRs using the Canadian adaptation of the Automated Self-Administered 24-hour Dietary Assessment Tool (ASA24-Canada-2018), which was developed by the National Cancer Institute (NCI) and can be completed in English or French (National Cancer Institute 2022a, 2022b). ASA24 was adapted to the Canadian food supply by Health Canada (National Cancer Institute 2022b). Reported foods and beverages are auto-coded using the 2015 Canadian Nutrient File (Government of Canada 2021) and a recipe database used for surveillance by Health Canada (Health Canada 2017; National Cancer Institute 2022b). When foods in ASA24 could not be matched to codes in the Canadian database, US foods were used and when necessary, nutrient values were adjusted to account for fortification differences between the countries. The assignment of food codes enables estimation of energy and nutrient values per food and per day. Additional details on ASA24 are provided in **Appendix D**. The two 24HDRs were spread over approximately one month to align with the period captured by the screener, as well as across days of the week at the group level to permit accounting for potential weekend day vs. weekday

effects. Participants were provided with a $10 Interac e-transfer or grocery store gift card for completion of each of the 24HDRs.

There were a total of 1484 completions of the eligibility questionnaire (**Figure 3**). Of these, 845 were determined to be bogus participants. Of the remaining 639, 485 were ineligible based on the eligibility criteria and quotas, and 154 eligible participants completed the survey. A total of 95 participants completed the survey in English and 59 in French. Cleaning procedures and sample sizes for the 24HDRs are detailed below and in Appendix D.

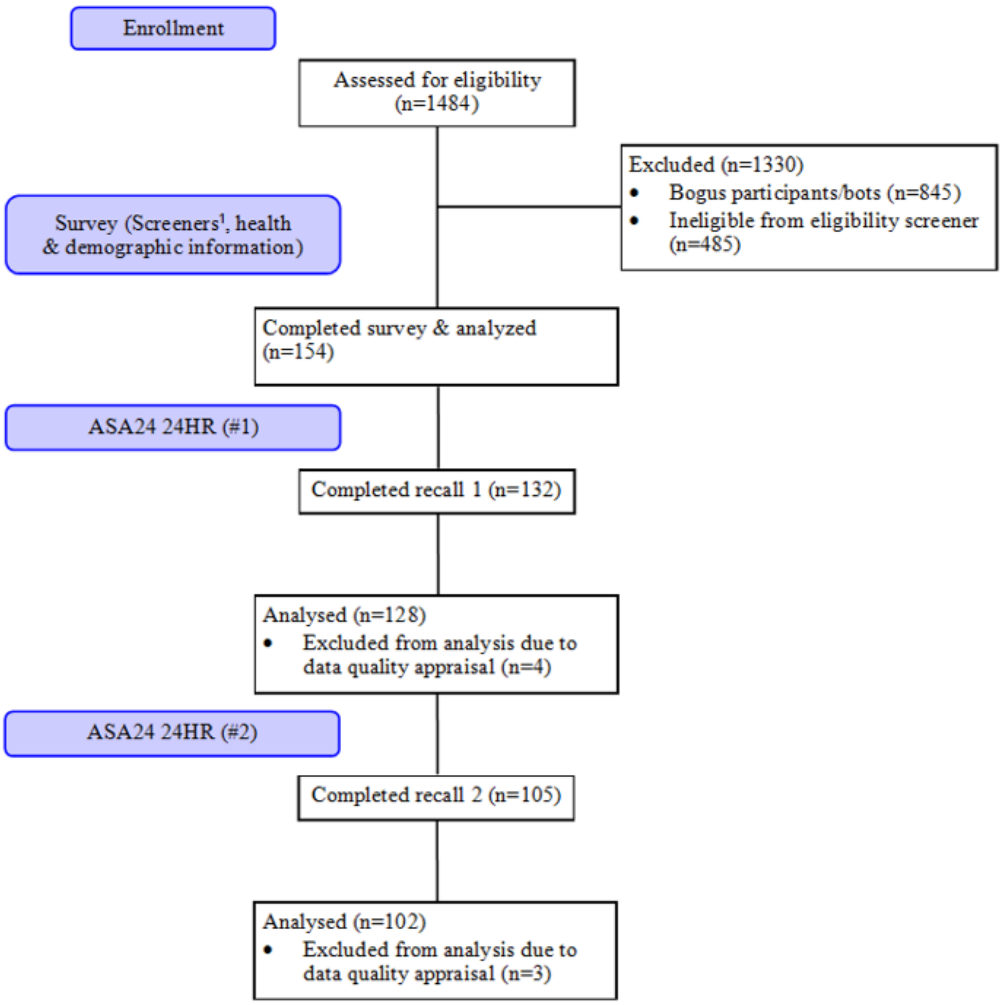


Figure 3. Flowchart describing participation in a study to assess the construct validity of the Canadian Food Intake Screener after applying inclusion and exclusion criteria

[1]Screeners included in the survey were the Canadian Food Intake Screener and a screener to assess adults' alignment with the CFG-2019 recommendations on healthy eating habits (Haines et al., 2022)

Canadian Food Intake Screener scoring system

The screener assesses frequency of intake over the past month (i.e., "Over the past month, how often did you eat…") of healthy foods and foods to limit (Health Canada 2022). Respondents choose from 10 response options, adapted from the Dietary Screener Questionnaire and the Diet History Questionnaire (National Cancer Institute 2021; National Cancer Institute 2022c; Millen et al. 2006), ranging from never to six or more times per day (Hutchinson et al. 2022). As a first step in scoring, the response options for each question were scored ordinally, using a scale of zero to nine (Gicevic et al. 2021). In keeping with the scoring approaches for the HEFI-2019 (Brassard et al. 2022a, 2022b), responses to questions that capture healthy foods were scored positively (i.e., more frequent consumption was assigned a higher score), while responses to questions that capture foods to limit were reverse scored (i.e., less frequent consumption was assigned a higher score).

Next, prior to data analyses, a scoring system was developed (**Table 5**), aligned as closely as possible with the HEFI-2019 scoring algorithm (Brassard et al. 2022a, 2022b). In the development of the screener, an effort was made to avoid combining different types of foods (e.g., meat, cheese, and milk; foods and beverages) in a single question to minimize cognitive load (Hutchinson et al. 2022). Further, based on cognitive testing, the screener includes a question specific to potatoes to help respondents report frequency of intake of both vegetables and grain foods more accurately (Hutchinson et al. 2022). Accordingly, in four cases, scores assigned to separate questions were summed to create "components", reflecting intake of total vegetables (including potatoes) and fruits, total protein foods, foods and beverages high in sugar, and foods and beverages high in sodium/saturated fat (Table 5). The inclusion of separate questions for animal- and plant-based proteins and for whole grains and refined grains enabled

derivation of ratios to reflect emphases within the CFG-2019 healthy food choices recommendations (Health Canada 2022). Specifically, ratios of scores assigned to frequencies were used for protein foods (plant-based protein foods/total protein foods) and grain foods (whole-grain foods/total grain foods). Scores based on frequency of intake of whole-grain foods and unsaturated oils round out the scoring system, for a total of eight components.

Weighting of the components was informed by the HEFI-2019 scoring algorithm (Brassard et al. 2022a, 2022b) (Table 5). For example, the vegetables and fruits component is scored up to 20 points in both the screener and HEFI-2019 scoring algorithms. The weights of the vegetables and fruits, grain foods, and protein foods components are consistent with the

Table 5: Scoring system for the Canadian Food Intake Screener

Component[1]	**Screener questions**	**Scoring approach**	**Maximum points**	**Standard for minimum score (zero)**	**Standard for maximum score**
Vegetables and fruits	Vegetables + potato + fruits	Responses scored 0-9 points[4] and summed, /27	20	Never consuming vegetables, potatoes, or fruit	Consuming each item 6 or more times per day
Whole-grain foods	Whole-grain foods	Response scored 0-9 points	5	Never consuming whole-grain foods	Consuming whole-grain foods 6 or more times per day
Grain foods ratio	Whole-grain foods/Total grain foods[2]	Ratio of frequencies	5	Never consuming whole-grain foods	Ratio of 1.0
Protein foods	Plant-based protein foods + yogurt, cheese, kefir + animal-based protein foods + unsweetened milk	Responses scored 0-9 points and summed, /36	5	Never consuming protein foods	Consuming each item 6 or more times per day
Plant-based protein foods	Plant-based protein foods/Total protein foods[3]	Ratio of frequencies	5	Never consuming plant-based protein foods	Ratio of 1.0
Unsaturated oils	Margarine and vegetable oils	Response scored 0-9 points	5	Never consuming margarine and vegetable oils	Consuming margarine and vegetable oils 6 or more times per day
Foods and beverages high in sugars	Sweetened milk + sweetened beverages + sugary snacks	Responses reverse-scored 0-9 points and summed, /27	10	Consuming each item 6 or more times per day	Never consuming sweetened milk, sweetened beverages, or sugary snacks

Foods high in sodium/saturated fat	Fast food + processed meat + salty snacks	Responses reverse-scored 0-9 points and summed, /27	10	Consuming each item 6 or more times per day	Never consuming fast food, processed meat, or salty snacks
Total maximum score			65		

[1]Components are derived only for calculating total screener scores. Individual component scores should not be used on their own as the Canadian Food Intake Screener is intended to provide rapid insights into *overall* alignment with the Canada's Food Guide-2019 healthy food choices recommendations.
[2] Total grain foods = whole-grain foods + refined grain foods
[3] Total protein foods = animal-based protein foods + plant-based protein foods + yogurt, cheese, kefir + unsweetened milk
[4] Respondents to the screener choose from 10 response options for each question, ranging from never consuming to consuming the given foods or beverages six or more times per day. As a first step in scoring, the response options for each question are scored continuously, using a scale of zero to nine points, with reverse-scoring for responses to questions that capture foods to limit (i.e., less frequent consumption is assigned a higher score).

CFG-2019 plate proportions of 50%, 25%, and 25%, respectively (Health Canada 2022). Weighting of components capturing foods and beverages to limit was aligned (20/65 possible points, 31%) with the HEFI-2019 scoring (25/80 possible points, 31%) (Brassard et al. 2022a, 2022b). Consistent with the guidance (Health Canada 2022), no upper limit on frequency was applied in determining the standards for maximum scores for healthy foods. The standard for minimum scores for healthy foods is zero frequency of consumption, whereas frequencies between the minimum and maximum are scored proportionately. For sources of added sugars, saturated fat, and sodium, the standard for full points is zero frequency of consumption, whereas frequencies between the minimum and maximum frequencies receive fewer points, consistent with the recommendations to limit these foods and beverages (Health Canada 2022). The scores for each component are summed to arrive at a total score, to a maximum of 65 points. Code for applying the scoring system is available in **Appendix E & F**.

Although the scoring process results in scores for screener "components", these are used solely to calculate total scores. Given its brevity, the screener is designed to provide a single score indicative of higher or lower alignment with the CFG-2019 healthy food choices recommendations overall.

ASA24-Canada-2018 data and variables to calculate HEFI-2019 scores

Of the 154 participants in the analytic sample, 132 completed the first 24HDR and 105 completed the second. The median number of days between the initial survey and the first recall was nine (interquartile range, 8-11 days) and between the two recalls was nine (interquartile range, 8-10 days). The coded dietary intake data from ASA24-Canada-2018 were reviewed by a registered dietitian (JMH), following protocols outlined by NCI (National Cancer Institute 2022d) and with input from a second registered dietitian (SIK) and a Health Canada nutritionist

experienced in coding and cleaning 24HDR data. Based on this review, four first 24HDRs and three second 24HDRs were removed (Appendix D), resulting in 128 first recalls (83% of the analytic sample) and 102 second recalls (66%) included in the analyses.

To arrive at HEFI-2019 component and total scores (Brassard et al. 2022a, 2022b), foods and beverages reported must be classified and amounts in appropriate units determined, as described in **Appendix G**.

Sociodemographic and health characteristics

Sociodemographic and health characteristics of interest were based on known differences in diet quality observed in other studies (**Appendix H**) (Alkerwi et al. 2017; Brassard et al. 2022b; Guenther et al. 2014; Hiza et al. 2013; Kuczmarski et al. 2016; Moubarac et al. 2017; Reedy et al. 2018; Tarasuk et al. 2010). Gender identity categories included men, women, and non-binary. Age was captured as a five-level categorical variable. Perceived income adequacy was assessed by asking participants, "Thinking about your total monthly income, how difficult is it for you to make ends meet?", with response options including very difficult, difficult, neither easy nor difficult, easy, and very easy (Litwin and Sapir, 2009). For examination of known differences, response categories were condensed to a three-level variable, including difficult or very difficult, neither easy nor difficult, and easy or very easy, based on the observed frequencies. Educational attainment was captured as high school graduate or less; some college; college graduate/Collège d'enseignement général et professionnel (CÉGEP), apprenticeship training, and/or trade programs; some university; university graduate; or postgraduate training or professional degree. For examination of known differences, these categories were collapsed to two, college graduate/CÉGEP/apprenticeship training/trade program or less versus higher educational attainment. Health literacy was assessed using the

Canadian adaptation of the Newest Vital Sign©, which asks six questions referencing a nutrition label (Mansfield et al. 2018). The original tool demonstrated reasonable criterion validity and reliability (Weiss et al. 2005), and the Canadian adaptation demonstrated face validity with Health Canada experts (Mansfield et al. 2018). Each question was scored as correct (1) or incorrect (0), and scores were summed to a possible six points (Weiss et al. 2005). For examination of known differences, a binary variable was created to indicate low or limited health literacy (score of 0-3) or adequate health literacy (score of 4-6) (Weiss et al. 2005). Finally, smoking status was assessed by asking participants whether they currently smoked tobacco (Statistics Canada 2021), and a binary variable was created to indicate those who currently smoked versus those who did not for examination of known differences.

Statistical analyses

Statistical analyses were conducted using SAS, version 9.4 (SAS Institute, Cary, NC, 2013). Descriptive statistics were generated to describe the sample. The assessment of construct validity included examining the variability in screener scores, comparing screener scores among subgroups with known differences in diet quality, and estimating the correlation coefficient between screener scores and total HEFI-2019 scores. The mean screener score and standard deviation were computed. ANOVA was used to compare screener scores among subgroups defined by characteristics known to differentiate dietary quality. Levene's test was used to assess whether the assumption of homogeneity of variances was not met, and Welch's ANOVA was used when it was not. Individuals with non-binary gender identities and those with missing responses to the gender identity question were not included in the assessment of known differences by gender identity due to small cell sizes (<5).

The approach used in the evaluation of the HEFI-2019 (Brassard et al. 2022b) was extended to incorporate the screener score as a covariate, enabling examination of relationships between screener scores and total HEFI-2019 scores. The multivariate Markov Chain Monte Carlo (MCMC) extension (Zhang et al. 2011) to the NCI method allows joint modeling of the numerators and denominators used to arrive at HEFI-2019 component and total scores, including dietary components consumed nearly daily by most persons and those consumed episodically (Zhang et al. 2011). The NCI method generates simulated true usual intakes, which allows estimation of the distribution of usual intakes, conditional on individual-level covariates while maintaining the joint distribution of these covariates as observed in the sample. Whole and non-whole-grain foods, animal- and plant-based protein foods, beverages, and "other" foods were modeled as episodically consumed based on ≥5% of recalls not including the given category (Krebs-Smith et al. 2010). All other dietary components were considered to be non-episodically consumed. In addition to the screener score, covariates included gender identity and age, as well as 24HDR sequence (first versus second recall) and whether the 24HDR was completed for a weekday versus weekend day (Nusser et al. 1996).

The output from the MCMC modeling was used to calculate the Pearson correlation coefficient between screener scores and total HEFI-2019 scores, computed as the square root of the R-squared of a linear regression model, with screener scores as the independent variable and total HEFI-2019 scores as the dependent variable. The square of the correlation coefficient, or the coefficient of determination (Kasuya 2018), is the proportion of variance in the dependent variable explained by the linear relationship with the independent variable. To assess whether higher screener scores were driven by higher overall intake, the correlation between screener scores and simulated usual energy intake based on the 24HDRs was assessed using the output from the MCMC modeling, computed as the square root of the R-squared of a linear regression

model, with screener scores as the independent variable and energy intake as the dependent variable.

For comparison, the R-squared is provided for the linear regression model that included the scores for the components generated in the process of scoring the screener, rather than the total screener score. Additionally, the implications of relaxing the standards for maximum scores for healthy foods and minimum scores for foods to limit were considered by comparing the correlation coefficients and R-squared estimates between models based on total scores derived from the original scoring system, described above, and based on an altered system that used truncation in the process of scoring. For example, for assigning minimum scores for foods to limit, such as processed meats, the response options were reverse scored continuously from zero to seven rather than zero to nine, with a score of seven applied to the three highest frequency categories. When truncation was applied to the plant-based protein component, scored as a ratio to total protein, the standard for the maximum score was 0.5, consistent with the HEFI-2019 thresholds (Brassard et al. 2022a, 2022b), whereas without truncation, the standard for the maximum score was 1.0. Further, to assess whether the *a priori* scoring system added value in aligning screener scores with total HEFI-2019 scores, the estimated correlation coefficient and R-squared were compared to those estimated using screener scores based on totaling the scores from zero to nine points applied to each question, without creating components. Finally, the NCI method was rerun without the screener score included to estimate HEFI-2019 scores in the sample based on simulated usual intakes.

Standard errors for statistics based on the NCI method were obtained using the bootstrap procedure (Tooze et al. 2010, Bland and Altman, 2015), with 200 replicates. P-values and standard deviations/errors are reported along with an interpretation of the level of evidence (Muff et al. 2022).

5.4. Results

Of the 154 participants who completed the screener, 74% identified as women and one quarter identified as men (**Table 6**). Participants were fairly evenly distributed by age categories, with the smallest proportion (10%) in the 55-65-year-old age category. Over half of participants (60%) identified as White, 13% as South Asian, 11% as East/Southeast Asian, and 6% as Black. Most participants stated it was neither easy or difficult (31%), easy (26%), or very easy (23%) to make ends meet. Most participants had completed some post-secondary education, with 16% graduating from college/CÉGEP or an apprenticeship training program or trade program, 14% and 34% completing some or graduating from university, respectively, and 24% completing postgraduate training or a professional degree. The vast majority (91%) reported they did not currently smoke. Characteristics of participants who completed 24HDRs versus those who did not were similar (data not shown).

Table 6: Sociodemographic and health characteristics of adults (n=154) participating in a study to evaluate the construct validity of the Canadian Food Intake Screener

Characteristic	Total n=154	English n=95	French n=59
		n (%)	
Gender identity			
Woman	114 (74.0)	68 (71.6)	46 (78.0)
Man	38 (24.7)	26 (27.4)	12 (20.3)
Gender-fluid, non-binary, and-or two-spirt	1 (0.7)	0 (0)	1 (1.7)
Not stated	1 (0.7)	1 (1.1)	0 (0)
Age (years)			
18-24	26 (16.9)	22 (23.2)	4 (6.8)
25-34	45 (29.2)	31 (32.6)	14 (23.7)

35-44	46 (29.9)	21 (22.1)	25 (42.4)
45-54	22 (14.3)	16 (16.8)	6 (10.2)
55-65	15 (9.7)	5 (5.3)	10 (17.0)
Racial/ethnic identity			
White	92 (59.7)	44 (46.3)	48 (81.4)
South Asian	20 (13.0)	19 (20.0)	1 (1.7)
East/Southeast Asian	17 (11.0)	14 (14.7)	3 (5.1)
Black	9 (5.8)	9 (9.5)	0 (0)
Indigenous	4 (2.6)	3 (3.2)	1 (1.7)
Middle Eastern	3 (1.9)	2 (2.1)	1 (1.7)
Latino	2 (1.3)	1 (1.1)	1 (1.7)
Other	6 (3.9)	2 (2.1)	4 (6.8)
Not stated	1 (0.7)	1 (1.1)	0 (0)
Perceived income adequacy			
Very difficult	3 (2.0)	3 (3.2)	0 (0)
Difficult	18 (11.7)	14 (14.7)	4 (6.8)
Neither easy nor difficult	48 (31.2)	34 (35.8)	14 (23.7)
Easy	40 (25.6)	21 (22.1)	19 (32.2)
Very easy	36 (23.4)	18 (19.0)	18 (30.5)
Don't know/prefer not to respond	9 (5.9)	5 (5.3)	4 (6.8)
Educational attainment			
High school graduate or less	14 (9.1)	11 (11.6)	3 (5.1)
Some college	5 (3.3)	5 (5.3)	0 (0)
College graduate/CÉGEP[1] or apprenticeship training and/or trade program	24 (15.6)	20 (21.1)	4 (6.8)
Some university	21 (13.6)	18 (19.0)	3 (5.1)

University graduate	53 (34.4)	29 (30.5)	24 (40.7)
Postgraduate training or professional degree (medical, law, etc.)	37 (24.0)	12 (12.6)	25 (42.4)
Health Literacy			
Low or limited health literacy[2]	21 (13.6)	19 (20.0)	2 (3.4)
Adequate health literacy[2]	133 (86.4)	76 (80.0)	57 (96.6)
Smoking status			
Daily or less than daily	13 (8.4)	10 (10.5)	3 (5.1)
Not at all	140 (90.9)	84 (88.4)	56 (94.9)
Not stated	1 (0.6)	1 (1.0)	0 (0)

[1]Collège d'enseignement général et professionnel
[2]Low or limited health literacy was based on a score of 0-3 and adequate health literacy was based on a score of 4-6 on the Newest Vital Sign.

The mean screener score was 35.0 (SD=4.7) of 65 possible points, ranging from 25.1 at the 1st percentile to 45.2 at the 99th percentile (**Figure 4**). There was moderate to strong evidence of differences in screener scores among subgroups in hypothesized directions (**Table 7**) by gender identity (p=0.06), perceived income adequacy (p=0.07), educational attainment (p=0.02), and smoking status (p=0.003). Differences were not observed by age group or health literacy.

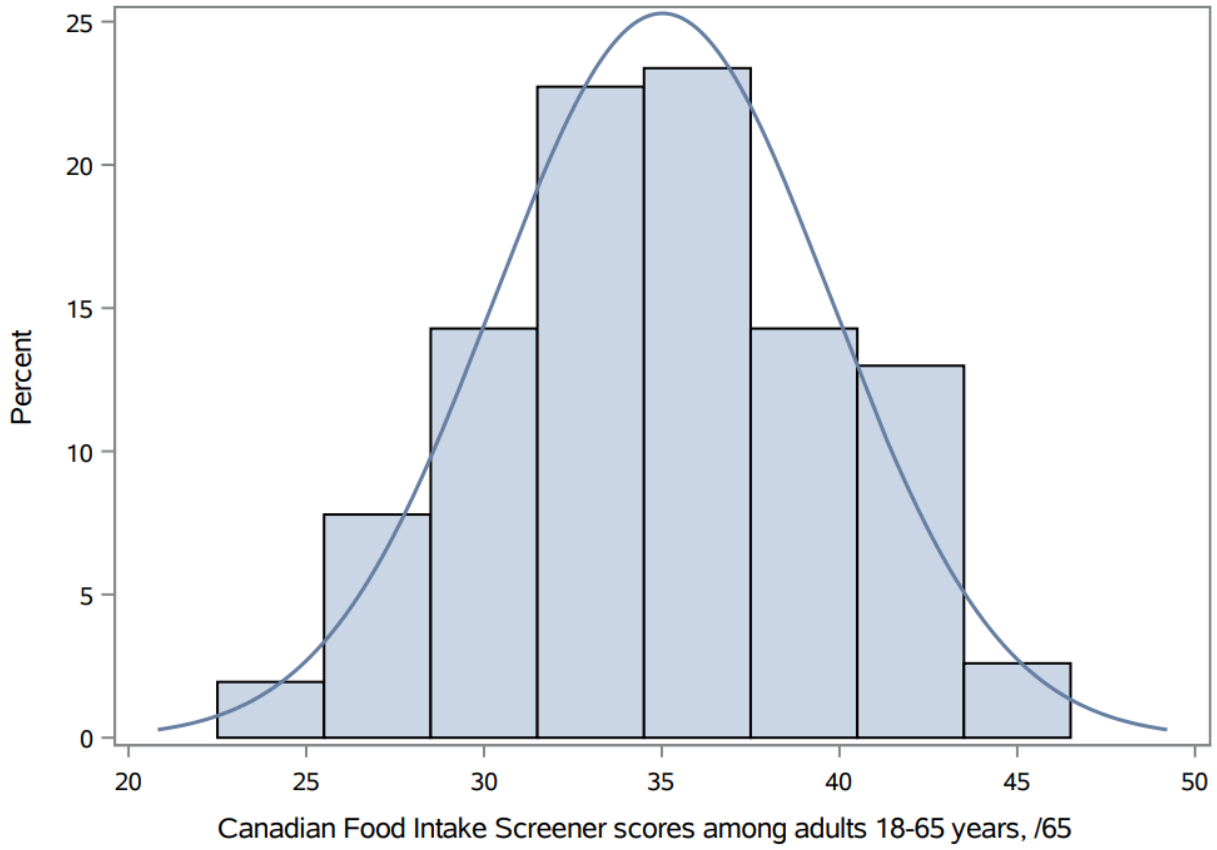


Figure 4. Distribution of Canadian Food Intake Screener scores among adults aged 18-65 years (n=154) in a study to examine the construct validity of the screener

The mean HEFI-2019 score based on simulated usual intakes was 40.9 (SE=1.2) of 80 possible points, ranging from 25.9 at the 1st percentile to 56.0 at the 99th percentile (**Appendix I**). A moderately strong Pearson correlation coefficient of 0.53 (SE=0.12) was observed between screener scores and total HEFI-2019 scores. Total screener scores explained 28% (SE=12%) of the variability in total HEFI-2019 scores. The correlation between screener scores and simulated usual energy intake was 0.08 (SE=0.09). In a model that included the scores for the screener components (generated to arrive at the total screener score) in place of the total score, they collectively explained 53% (SE=9.9%) of the variability in total HEFI-2019 scores.

Table 7: Comparison of mean Canadian Food Intake Screener scores, of a possible 65 points, among subgroups known to differ in dietary quality among adults 18-65 years (n=154)

Characteristic	Mean Screener Score	Standard Deviation	P-value[1]
Age			
18-24 (n=26)	34.8	4.8	0.88
25-34 (n=45)	35.1		
35-44 (n=46)	34.5		
45-54 (n=22)	35.3		
55-65 (n=15)	36.2		
Gender[2]			
Men (n=38)	33.8	4.7	0.06
Women (n=114)	35.4		
Perceived income adequacy[3]			
Difficult or very difficult (n=21)	33.0	4.7	0.07
Neither easy nor difficult (n=48)	34.7		
Easy or very easy (n=76)	35.5		
Educational attainment			
College graduate/CÉGEP[4] or apprenticeship training and/or trade program (n=43)	33.6	4.7	0.02

Some university or greater (n=111)	35.6		
Health literacy[5]			
Low or limited health literacy (0-3) (n=21)	33.9	4.7	0.22
Adequate health literacy (4-6) (n=133)	35.2		
Smoking status[6]			
Daily or less than daily (n=13)	30.5	4.6	0.003
Not at all (n=140)	35.4		

[1]P-values were derived using one-way ANOVA and applying Levene's test for homogeneity to determine whether a Welch's ANOVA was appropriate in the cases in which the assumption of homogeneity of variances was not met.
[2]Two individuals were not included because they identified as another gender identity or did not provide details on their gender identity.
[3]Nine individuals were not included because they had missing data on perceived income adequacy.
[4]Collège d'enseignement général et professionnel
[5]Low or limited health literacy was defined as a score of 0-3 and adequate health literacy was defined as a score of 4-6 on the Newest Vital Sign tool (Weiss et al. 2005).
[6]One individual was not included because they did not state their smoking status.

When truncation was applied in generating screener scores, the correlation coefficient and proportion of variability in total HEFI-2019 scores explained were slightly lower (data not shown). Similarly, when a scoring approach that did not incorporate summing across questions or deriving ratios to create components was used, the correlation coefficient and proportion of variability explained were lower compared to those based on the application of the scoring system (data not shown).

5.5. Discussion

The findings of this study suggest moderate construct validity of the Canadian Food Intake Screener, supporting its use to provide rapid insights into overall alignment of intake

over the past month with the CFG-2019 healthy food choices recommendations. The mean screener score was 35.0 of 65 possible points and reasonable variation was observed. The 99th percentile was 45.2, indicating considerable room for improvement with respect to alignment of intakes with the recommendations. Similarly, the 99th percentile for HEFI-2019 scores was 63 of 80 possible points based on analysis of the 2015 CCHS-Nutrition data (Brassard et al. 2022b) and 56 in the current sample.

The largest difference in screener scores observed between subgroups was five points, between those who did and did not report smoking. Similarly, in the evaluation of the HEFI-2019 using data from the 2015 CCHS-Nutrition, the corresponding difference in total HEFI-2019 scores was seven points (Brassard et al., 2022b). Differences were not observed in screener scores by age, possibly due to the categorical nature of the age variable and the small cell sizes for some categories. The sample size and skew toward higher perceived income adequacy and higher health literacy may underlie the lack of meaningful differences by these characteristics.

The correlation coefficient of 0.53 between screener and total HEFI-2019 scores is similar to that observed for other brief screeners relative to data from more comprehensive dietary assessment instruments. For example, correlation coefficients between Behavioral Risk Factor Surveillance System fruit and vegetable module data and data from food frequency questionnaires, 24HDRs, and food records range from 0.29 to 0.63 (Centers for Disease Control and Prevention n.d.). An evaluation of a 40-item screener to assess alignment with the Dutch food-based dietary guidelines among 751 adults found a Kendall's tau-b coefficient of 0.51 based on comparison to Dutch Healthy Diet index 2015 scores derived from a food frequency questionnaire (de Rijk et al. 2021). Similarly, an evaluation of the 14-item Mediterranean Diet Adherence Screener among a sample of 7146 adults found a Pearson's correlation coefficient of 0.52 with scores indicating adherence to the Mediterranean diet based on a food frequency questionnaire (Schröder et al. 2011). Among a sample of 50 college students, a Pearson's

correlation coefficient of 0.79 was observed between scores on the 22-item Short Healthy Eating Index and Healthy Eating Index-2015 scores, estimated using repeat 24HDRs (Colby et al. 2020). This relatively high correlation for the latter screener may relate, in part, to the more homogenous sample and the underlying quantitative recommendations, including recommended amounts to consume daily.

The HEFI-2019 (Brassard et al. 2022a, 2022b) is density-based, helping to decouple diet quality from quantity. In contrast, the focus on frequency of consumption within the screener and its scoring system raises the possibility that those with higher energy requirements and intakes, particularly of healthy foods, may receive higher screener scores. Within this sample, there was no meaningful correlation between screener scores and energy intake based on the 24HDRs. However, the measurement error in estimated energy intake based on self-report methods (Freedman et al. 2014; Kirkpatrick et al. 2022b) makes it challenging to detect an association between screener scores and true usual energy intake. Because the CFG-2019 does not provide quantitative recommendations for food categories, no upper limit on frequency was applied in determining the standards for maximum scores for healthy foods and minimum scores for foods to limit. Truncation was considered to relax the standards but did not improve the correlation coefficient with total HEFI-2019 scores, perhaps because it introduces its own problems, such as floor or ceiling effects (Ricart et al. 2022). Alternative assessment approaches, such as images of plates captured using mobile food records (Ho et al. 2020) and analyzed using machine learning, may allow for more direct consideration of relative intake of different food categories in the future, consistent with the focus of the CFG-2019 plate on proportionality (Health Canada, 2022). However, like other methods of dietary assessment, image-based technology is subject to measurement error (Ho et al. 2020).

While dietary screeners are not a replacement for comprehensive dietary assessment, their simplicity makes them a popular choice among researchers and practitioners (Kirkpatrick

et al. 2014, 2017). The Canadian Food Intake Screener can be self-administered in approximately five minutes and was designed to be as simple as possible, though it requires recalling and averaging frequency of intake over the past month, which can be cognitively challenging (National Cancer Institute 2015). In contrast, 24HDRs integrate multiple passes to prompt respondents to report their intake for the prior day in as much detail as possible. Although web-based self-administered systems, such as ASA24 (Subar et al. 2012) and Rappel 24h Web (R24W) (Jacques et al. 2016; Lafrenière et al. 2017), make it possible to collect 24HDRs in a range of contexts, a recall takes 20 to 30 minutes or more to complete, and non-consecutive repeat administrations are needed on at least a subsample to enable the estimation of distributions of usual intake (Dodd et al. 2006). Further, cleaning the recalls and preparing the variables to calculate HEFI-2019 scores in the current study required >75 hours of a registered dietitian's time. These steps are less burdensome when using secondary data that can be readily linked with information on reference amounts and free sugars. Nonetheless, the screener scoring system is, by design, simpler than the HEFI-2019 scoring algorithm.

Decisions about whether to use the screener versus the HEFI-2019 should be informed not only by considerations of resources and burden, but more importantly, by the research question. Scores based on the screener do not consider all components assessed by the HEFI-2019 (Brassard et al. 2022a, 2022b). Because the screener does not assess water consumption (Hutchinson et al. 2022), a beverages component is not included in the scoring system. Furthermore, the screener assesses frequency of consumption of key sources of added sugars (Mela and Woolner, 2018), saturated fat, and sodium, whereas HEFI-2019 scoring involves quantification of intake of free sugars, saturated and unsaturated fat, and sodium (Brassard et al. 2022b). Component scores used to calculate total screener scores should not be interpreted individually because they are based on one or a few questions that do not provide sufficient information to quantify intake of a specific group of foods, such as fruits and vegetables. Rather,

the total screener score is intended to signal the degree of overall alignment with the CFG-2019 healthy food choices recommendations. In contrast, given the underlying detailed intake data, the HEFI-2019 enables consideration of individual components, with interpretation of component scores recommended to contextualize total scores (Brassard et al. 2022a, 2022b).

The screener could be administered in national surveillance to provide high-level insights on the alignment of adults' intake with the CFG-2019 healthy food choices recommendations. However, it should not replace the collection of 24HDR data in nutrition-focused cycles of the CCHS (Health Canada 2004, 2015). Compared to the screener, 24HDR data are more accurate and allow for richer analyses (Dodd et al. 2006; Freedman et al. 2014, 2015). In addition to the calculation of HEFI-2019 scores (Brassard et al. 2022a, 2022b), such analyses may entail comparisons to the Dietary Reference Intakes (Institute of Medicine 2006), identify top sources of food categories and nutrients of public health concern, and characterize meal patterns and contextual factors.

The use of the screener to evaluate interventions requires care since exposure to an intervention can change reporting of intake (Neuhouser et al. 2008; Natarajan et al. 2010), though this may be less of a concern for interventions focused on environmental factors versus individual practices (National Cancer Institute 2015). As well, the level of change that the screener can detect has not been assessed. Relatedly, consideration of what constitutes meaningful differences in screener scores over time or among groups is needed. There is no threshold at which screener scores are indicative of meeting the CFG-2019 healthy food choices recommendations or of a 'healthy' diet. For knowledge translation purposes, scores can be categorized to describe the degree of alignment with the guidance (e.g., higher or lower than the mean or median in a given sample). However, scores should not be categorized for analysis as this results in loss of information and potentially, misclassification of values near the thresholds (Kirkpatrick et al. 2018).

In the current sample, the correlation between scores based on the screener scoring system and total HEFI-2019 scores was somewhat higher compared to that for an approach that did not involve the creation of components and weighting, indicating that the scoring system adds value in aligning screener scores with the healthy food choices recommendations. However, when scores assigned to separate screener components were considered in place of total screener scores, a larger proportion of variability in total HEFI-2019 scores was explained. Though the addition of independent variables to a model results in a higher R-squared estimate by definition, this finding suggests that the scoring algorithm could potentially be further optimized. Administration of the screener, potentially along with 24HDRs, in larger samples may enable refinement of the scoring system using data-driven approaches. Drawing upon 24HDR data for approximately 7600 individuals participating in the 2009-2010 National Health and Nutrition Examination Survey (86% of whom completed two recalls), Thompson et al. (2017) developed scoring algorithms to align estimates from the Dietary Screener Questionnaire more closely with those from 24HDRs. However, Gilsing et al. (2018) explored the development of scoring algorithms for the Canadian Longitudinal Study on Aging screener, using data from 200 older adults who completed four 24HDRs, and found the algorithms did not notably improve mean nutrient intake estimates. Additionally, data-driven approaches based on one dataset may not be portable to other datasets (Thompson et al. 2017), and there is a ceiling on the extent to which screeners and their scoring systems can be improved.

To improve the consistency of periods captured by the screener and the HEFI-2019 scores, modeling was used to estimate total HEFI-2019 scores based on the simulated distribution of usual intakes (Dodd et al. 2006, Brassard et al. 2022b). A similar approach has been used in other validation research (Colby et al. 2020, Gilsing et al. 2018; Thompson et al. 2017). Other methods, such as Bland-Altman analysis (Giavarina, 2015) to assess agreement between estimates from two error-prone instruments (de Rijk et al. 2021; Schröder et al. 2011,

2012), were not appropriate because the modeling used to produce HEFI-2019 scores does not produce an individual-level analogue to the screener score (Kirkpatrick et al. 2018, 2022a). While the screener was evaluated among adults aged 18-65 years, it could be evaluated and/or adapted for use with other populations, such as youth and older adults (Hutchinson et al. 2022).

Several additional considerations are relevant to our findings' interpretation. The collection of two non-consecutive 24HDRs enabled comparison of screener scores to total HEFI-2019 scores based on the simulated distribution of usual intakes, which represented the strongest available reference measure for assessment of the screener's construct validity since no objective marker of usual diet quality is known. Multivariate usual intake modeling accounted for within-person random error, as well as the multiple correlated dietary components considered in scoring the HEFI-2019 (Brassard et al. 2022b). However, 24HDR data are affected by bias (Freedman et al. 2014, 2015; Kirkpatrick et al. 2022b), and the correlation between screener and HEFI-2019 scores may be overestimated due to correlated errors between the screener and 24HDR data (Kirkpatrick et al. 2019). The sample of 154 was close to the desired size; and quotas for language, race/ethnicity, and educational attainment were achieved. However, the sample was skewed towards women and 24HDRs were available for a subset of participants. The smaller than intended sample with recall data limits the precision of the observed correlation between screener and HEFI-2019 scores.

Bogus participants are adaptive to quotas and other attempts at dissuasion (Shamon and Berning 2019; Pei et al. 2020; Wang et al. 2021). Intensive data quality appraisal ensured data from bogus participants were not included in the analyses. Finally, testing of the screener was conducted online, and effects related to mode of administration are possible (Bowling 2005).

5.6. Conclusion

The results of this study suggest moderate construct validity of the Canadian Food Intake Screener, supporting its use to rapidly assess overall alignment of adults' intake over the past month with the CFG-2019 healthy food choices recommendations. Use of the screener in research and surveillance, as appropriate, can contribute to a robust evidence base that can be synthesized to inform policies and programs to improve dietary quality. Administration of the screener to large, diverse samples in the future may enable refinement of the screener's scoring system to improve its alignment with the underlying recommendations.

Competing interests: This project was funded by and conducted in collaboration with Health Canada, through a contract to SIK. SIK has received funding from Agriculture and Agri-Food Canada, AI for Good, the Canadian Institutes of Health Research (CIHR), the Canadian Foundation for Dietetic Research, Health Canada, the National Institutes of Health, the Ontario Ministry of Research and Innovation, and the Social Sciences and Humanities Research Council of Canada. SIK is a member of the Health Canada Nutrition Science Advisory Committee and the CIHR Institute of Nutrition, Metabolism, and Diabetes Institute Advisory board. SL has received funding from CIHR. BL has received funding from CIHR (ongoing), the Fonds de recherche du Québec – Santé (FRQS) (ongoing), Fonds de recherche du Québec – Nature et technologies (NT) (ongoing), the Ministère de la santé et des services sociaux (MSSS) du Québec (ongoing), Health Canada (completed in 2021) and Atrium Innovations (completed in 2019). BL is a member of the Canadian Nutrition Society Advisory Board. JH has received funding from the Canadian Foundation for Dietetic Research, CIHR, Danone Institute International, Danone Institute North America, Health Canada, and the National Institutes of Health. The remaining authors have no competing interests to disclose.

Funding: This project was funded by Health Canada, as well as an Ontario Ministry of Research and Innovation Early Researcher Award held by SIK. JMH received funding from a Social Sciences and Humanities Research Council (SSHRC) Vanier Canada Graduate Scholarship and TEW received funding from a SSHRC Canada Graduate Scholarship – Master's.

Data availability: Data are available from the authors upon request.

5.7. References

Alkerwi, A., Baydarlioglu, B., Sauvageot, N., Stranges, S., Lemmens, P., Shivappa, N., and Hébert, J.R. 2017. Smoking status is inversely associated with overall diet quality: Findings from the ORISCAV-LUX study. Clinical Nutrition **36**(5): 1275–1282. Doi:10.1016/J.CLNU.2016.08.013.

Bland JM, Altman DG. 2015. Statistics notes: Bootstrap resampling methods. BMJ. **350**(h2622). Doi: 10.1136/bmj.h2622.

Bowling, A. 2005. Mode of questionnaire administration can have serious effects on data quality. J Public Health (Oxf) **27**(3): 281–291. Doi:10.1093/pubmed/fdi031.

Brassard, D., Elvidge Munene, L.-A., St Pierre, S., Gonzalez, A., Guenther, P.M., Jessri, M., Vena, J.E., Olstad, D.L., Vatanparast, H., Prowse, R., Lemieux, S., L'Abbé, M.R., Garriguet, D., Kirkpatrick, S.I., and Lamarche, B. 2022b. Evaluation of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Applied Physiology, Nutrition, and Metabolism **47**(5). Doi:10.1139/APNM-2021-0416.

Brassard, D., Elvidge Munene, L.-A., St Pierre, S., Guenther, P.M., Kirkpatrick, S.I., Slater, J., Lemieux, S., Jessri, M., Haines, J., Prowse, R., Olstad, D.L., Garriguet, D., Vena, J.E., Vatanparast, H., L'Abbé, M.R., and Lamarche, B. 2022a. Development of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Applied Physiology, Nutrition, and Metabolism **47**(5): 595–610. Doi:10.1139/APNM-2021-0415.

Cade, J., Thompson, R., Burley, V. and Warm, D., 2002. Development, validation and utilisation of food-frequency questionnaires–a review. Public Health Nutrition, **5**(4): 567-587. Doi: 10.1079/PHN2001318.

Centers for Disease Control and Prevention. (n.d.). Surveillance of fruit and vegetable intake using the Behavioral Risk Factor Surveillance System. Centers for Disease Control and Prevention, Atlanta, GA. Available from https://www.cdc.gov/brfss/pdf/fruits_vegetables.pdf [Accessed 3 August 2022].

Colby, S., Zhou, W., Allison, C., Mathews, A.E., Olfert, M.D., Morrell, J.S., Byrd-Bredbenner, C., Greene, G., Brown, O., Kattelmann, K., and Shelnutt, K. 2020. Development and validation of the short healthy eating index survey with a college population to assess dietary quality and intake. Nutrients **12**(9): 1–24. Doi:10.3390/nu12092611.

Dodd, K.W., Guenther, P.M., Freedman, L.S., Subar, A.F., Kipnis, V., Midthune, D., Tooze, J.A. and Krebs-Smith, S.M., 2006. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. Journal of the American Dietetic Association, **106**(10):1640-1650. Doi:10.1016/j.jada.2006.07.011.

Freedman, L.S., Commins, J.M., Moler, J.E., Arab, L., Baer, D.J., Kipnis, V., Midthune, D., Moshfegh, A.J., Neuhouser, M.L., Prentice, R.L., Schatzkin, A., Spiegelman, D., Subar, A.F., Tinker, L.F., and Willett, W. 2014. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. American Journal of Epidemiology **180**(2): 172–188. Doi:10.1093/aje/kwu116.

Freedman, L.S., Commins, J.M., Moler, J.E., Willett, W., Tinker, L.F., Subar, A.F., Spiegelman, D., Rhodes, D., Potischman, N., Neuhouser, M.L., Moshfegh, A.J., Kipnis, V., Arab, L., and Prentice, R.L. 2015. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for potassium and sodium intake. American Journal of Epidemiology **181**(7): 473–487. Doi:10.1093/aje/kwu325.

Frongillo, E.A., Baranowski, T., Subar, A.F., Tooze, J.A., and Kirkpatrick, S.I. 2019. Establishing Validity and Cross-Context Equivalence of Measures and Indicators. J Acad Nutr Diet **119**(11): 1817–1830. Doi:10.1016/J.JAND.2018.09.005.

Giavarina D. 2015. Understanding Bland Altman analysis. Biochemia medica **25**(2):141-51. Doi:10.11613/BM.2015.015.

Gicevic, S., Mou, Y., Bromage, S., Fung, T., and Willett, W. 2021. Development of a diet quality screener for global use: Evaluation in a sample of US women. J Acad Nutr Diet **121**(5): 854–871. Doi:10.1016/j.jand.2020.12.024.

Gilsing, A., Mayhew, A.J., Payette, H., Shatenstein, B., Kirkpatrick, S.I., Amog, K., Wolfson, C., Kirkland, S., Griffith, L.E., and Raina, P. 2018. 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 **10**(10): 1522. Doi:10.3390/NU10101522.

Government of Canada. 2021. Canadian Nutrient File. Available from http://www.hc-sc.gc.ca/fn-an/nutrition/fiche-nutri-data/cnf_aboutus-aproposdenous_fcen-eng.php [accessed 5 August 2022].

Government of Canada. 2022. While English and French are still the main languages spoken in Canada, the country's linguistic diversity continues to grow. Available from https://www150.statcan.gc.ca/n1/daily-quotidien/220817/dq220817a-eng.htm [accessed 7 September 2022].

Guenther, P.M., Kirkpatrick, S.I., Reedy, J., Krebs-Smith, S.M., Buckman, D.W., Dodd, K.W., Casavale, K.O. and Carroll, R.J. 2014. The Healthy Eating Index-2010 is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. The Journal of Nutrition **144**(3): 399-407. Doi:10.3945/jn.113.183079.

Haines J, Martin A, Perreault M. 2022. Questionnaire (screener) to measure eating practices based on Canada's Food Guide eating habits recommendations. Canadian Nutrition Society Webinar: How to use newly developed research tools to measure food choices and eating behaviours based on the 2019 Canada's Food Guide.

Hammond, D., Vanderlee, L., White, C.M., Acton, R.B., White, M., Roberto, C.A., Cameron, A., Sacks, G., Kirkpatrick, S.I., Dubin, J., Adams, J., Jauregui, A., and Thrasher, J.F. 2022. The conceptual framework for the International Food Policy Study: Evaluating the population-level impact of food policy. The Journal of Nutrition **152**(1): 1S–12S. doi:10.1093/jn/nxac042.

Health Canada. 2022. Canada's Food Guide. Health Canada, Ottawa, ON. Available from https://food-guide.canada.ca/en/ [Accessed 3 August 2022].

Health Canada. 2020. Healthy food choices - Canada's Food Guide. Available from https://food-guide.canada.ca/en/healthy-food-choices/ [accessed 4 August 2022].

Health Canada. 2017. Reference guide to understanding the data- 2015 Canadian Community Health Survey-Nutrition. Available from https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance/health-nutrition-surveys/canadian-community-health-survey-cchs/reference-guide-understanding-using-data-2015.html [Accessed 30 November 2022].

Hiza, H.A.B., Casavale, K.O., Guenther, P.M., and Davis, C.A. 2013. Diet Quality of Americans Differs by Age, Sex, Race/Ethnicity, Income, and Education Level. J Acad Nutr Diet **113**(2): 297–306. Doi:10.1016/J.JAND.2012.08.011.

Ho, D.K.N., Tseng, S.H., Wu, M.C., Shih, C.K., Atika, A.P., Chen, Y.C., and Chang, J.S. 2020. Validity of image-based dietary assessment methods: A systematic review and meta-analysis. Clinical Nutrition **39**(10): 2945–2959. Doi:10.1016/J.CLNU.2020.08.002.

Hutchinson J, Williams TE, Westaway AM, Bédard A, Pitre C, Lemieux S, Dodd KW, Lamarche B, Guenther PM, Haines J, Wallace A, Martin A, Hernandez KJ, St-Pierre S, Elvidge Munene L-A, da Costa Louzada ML, Jessri M, Olstad DL, Prowse R, Randall Simpson J, Vena JE, and Kirkpatrick SI. 2022. Development of the Canadian Food Intake Screener to assess alignment of adults' eating patterns with healthy food choices recommendations. MedXriv. Preprint. Doi: 10.1101/2022.12.19.22283569.

Institute of Medicine. 2006. Dietary Reference Intakes: The Essential Guide to Nutrient Requirements. National Academies Press, Washington, D.C. doi:10.17226/11537.

Jacques, S., Lemieux, S., Lamarche, B., Laramée, C., Corneau, L., Lapointe, A., Tessier-grenier, M., and Robitaille, J. 2016. Development of a Web-Based 24-h Dietary Recall for a French-Canadian Population. Nutrients **8**(724): 1–12. Doi:10.3390/nu8110724.

Kasuya, E. 2019. On the use of r and r squared in correlation and regression. Ecological Research **34(**1): 235-236. Doi:10.1111/1440-1703.1011.

Kirkpatrick, S.I., Baranowski, T., Subar, A.F., Tooze, J.A., and Frongillo, E.A. 2019. Best practices for conducting and interpreting studies to validate self-report dietary assessment methods. J Acad Nutr Diet **119**(11): 1801–1816. Doi:10.1016/j.jand.2019.06.010.

Kirkpatrick, S.I., Guenther, P.M., Subar, A.F., Krebs-Smith, S.M., Herrick, K.A., Freedman, L.S., and Dodd, K.W. 2022a. Using short-term dietary intake data to address research questions related to usual dietary intake among populations and subpopulations: Assumptions, statistical techniques, and considerations. J Acad Nutr Diet **122**(7): 1246-1262. Doi:10.1016/J.JAND.2022.03.010.

Kirkpatrick, S.I., Reedy, J., Ebonee, B.N., Dodd, K.W., Subar, A.F., Thompson, F.E., and McKinnon, R.A. 2014. Dietary assessment in food environment research: A systematic review. American Journal of Preventative Medicine **46**(1): 94–102. Doi:10.1016/j.amepre.2013.08.015.

Kirkpatrick, S.I., Reedy, J., Krebs-Smith, S.M., Pannucci, T.E., Subar, A.F., Wilson, M.M., Lerman, J.L., and Tooze, J.A. 2018. Applications of the Healthy Eating Index for surveillance, epidemiology, and intervention research: Considerations and caveats. J Acad Nutr Diet **118**(9): 1603–1621. Doi:10.1016/j.jand.2018.05.020.

Kirkpatrick, S.I., Vanderlee, L., Raffoul, A., Stapleton, J., Csizmadi, I., Boucher, B., Massarelli, I., Rondeau, I., and Robson, P. 2017. Self-report dietary assessment tools used in Canadian research: A scoping review. Advances in Nutrition **8**(2): 276–289. Doi:10.3945/AN.116.014027.

Kirkpatrick, S.I., Troiano, R.P., Barrett, B., Cunningham, C., Subar, A.F., Park, Y., Bowles, H.R., Freedman, L.S., Kipnis, V., Rimm, E.B., Willett, W.C., Potischman, N., Spiegelman, D., Baer, D.J., Schoeller, D.A., and Dodd, K.W. 2022b. Measurement error affecting web- and paper-based dietary assessment instruments: Insights from the Multi-Cohort Eating and Activity Study for Understanding Reporting Error. American Journal of Epidemiology **191**(6): 1125-1139. Doi:10.1093/AJE/KWAC026.

Kuczmarski, M., Adams, E., Cotugna, N., Pohlig, R., Beydoun, M., Zonderman, A., and Evans, M. 2016. Health literacy and education predict nutrient quality of diet of socioeconomically diverse, urban adults. Journal of Epidemiology and Preventive Medicine **2**(1). Doi:10.19104/JEPM.2016.115.

Lachat, C., Hawwash, D., Ocké, M.C., Berg, C., Forsum, E., Hörnell, A., Larsson, C., Sonestedt, E., Wirfält, E., Åkesson, A., Kolsteren, P., Byrnes, G., de Keyzer, W., van Camp, J., Cade, J.E., Slimani, N., Cevallos, M., Egger, M., and Huybrechts, I. 2016. Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): An extension of the STROBE Statement. PloS Med **13**(6). Doi:10.1371/JOURNAL.PMED.1002036.

Lafrenière, J., Lamarche, B., Laramée, C., Robitaille, J., and Lemieux, S. 2017. Validation of a newly automated web- based 24-hour dietary recall using fully controlled feeding studies. BMC Nutrition **3**(34): 1–10. Doi:10.1186/s40795-017-0153-3.

Litwin, H. and Sapir, E.V., 2009. Perceived income adequacy among older adults in 12 countries: findings from the survey of health, ageing, and retirement in Europe. The Gerontologist **49**(3): 397-406. Doi:10.1093/geront/gnp036.

Mansfield, E.D., Wahba, R., Gillis, D.E., Weiss, B.D., and L'Abbé, M. 2018. Canadian adaptation of the Newest Vital Sign©, a health literacy assessment tool. Public Health Nutrition **21**(11): 2038–2045. Doi:10.1017/S1368980018000253.

Mela, D.J. and Woolner, E.M., 2018. Perspective: total, added, or free? What kind of sugars should we be talking about? Advances in Nutrition **9**(2): 63-69. Doi:10.1093/advances/nmx020.

Millen, A.E., Midthune, D. Thompson, F.E., Kipnis, V., and Subar, A.F. 2006. The National Cancer Institute Diet History Questionnaire: Validation of pyramid food servings. American Journal of Epidemiology **163**(3): 279-288. Doi:10.1093/aje/kwj031.

Moubarac, J.C., Batal, M., Louzada, M.L., Martinez Steele, E., and Monteiro, C.A. 2017. Consumption of ultra-processed foods predicts diet quality in Canada. Appetite **108**: 512–520. Appetite. Doi:10.1016/J.APPET.2016.11.006.

Muff, S., Nilsen, E.B., O'Hara, R.B., and Nater, C.R. 2022. Rewriting results sections in the language of evidence. Trends in Ecology & Evolution **37**(3): 203–210. Doi:10.1016/J.TREE.2021.10.009.

Natarajan, L., Pu, M., Fan, J., Levine, R.A., Patterson, R.E., Thomson, C.A., Rock, C.L., and Pierce, J.P. 2010. Measurement error of dietary self-report in intervention trials. Am J Epidemiol **172**(7): 819–827. Doi:10.1093/AJE/KWQ216.

National Cancer Institute. 2022b. ASA24-Canada-2018. Available from https://epi.grants.cancer.gov/asa24/respondent/asa24-canada-2018.html [accessed 5 August 2022].

National Cancer Institute. 2022a. ASA24 Dietary Assessment Tool. Available from https://epi.grants.cancer.gov/asa24/ [accessed 5 August 2022].

National Cancer Institute. 2015. Dietary Assessment Primer. Available from https://dietassessmentprimer.cancer.gov/ [accessed 3 August 2022].

National Cancer Institute. 2022c. Diet History Questionnaire III (DHQ III). Available from https://epi.grants.cancer.gov/dhq3/ [accessed 5 August 2022].

National Cancer Institute. 2021. Dietary Screener Questionnaire in the NHANES 2009-10: Background. National Cancer Institute, Bethesda, MD. Available from https://epi.grants.cancer.gov/nhanes/dietscreen/ [accessed 3 August 2022].

National Cancer Institute. 2022d. Reviewing & Cleaning ASA24 Data. Available from https://epi.grants.cancer.gov/asa24/resources/cleaning.html [accessed 5 August 2022].

Neuhouser, M.L., Tinker, L., Shaw, P.A., Schoeller, D., Bingham, S.A., Horn, L. van, Beresford, S.A.A., Caan, B., Thomson, C., Satterfield, S., Kuller, L., Heiss, G., Smit, E., Sarto, G., Ockene, J., Stefanick, M.L., Assaf, A., Runswick, S., and Prentice, R.L. 2008. Use of recovery biomarkers to

calibrate nutrient consumption self-reports in the Women's Health Initiative. Am J Epidemiol **167**(10): 1247–1259. Doi:10.1093/AJE/KWN026.

Nusser, S.M., Carriquiry, A.L., Dodd, K.W., and Fuller, W.A. 1996. A semiparametric transformation approach to estimating usual daily intake distributions. J Am Stat Assoc **91**: 1440–1449. Doi:10.1080/01621459.1996.10476712.

Pei, W., Mayer, A., Tu, K., and Yue, C. 2020. Attention please: Your attention check questions in survey studies can be automatically answered. Proceedings of The Web Conference 2020. New York, USA. Doi:10.1145/3366423.

Reedy, J., Lerman, J.L., Krebs-Smith, S.M., Kirkpatrick, S.I., Pannucci, T.E., Wilson, M.M., Subar, A.F., Kahle, L.L. and Tooze, J.A., 2018. Evaluation of the Healthy Eating Index-2015. Journal of the Academy of Nutrition and Dietetics, **118**(9): 1622-1633. Doi:10.1016/j.jand.2018.05.019.

Ricart, G., Atoloye, A.T., Durward, C.M., and Guenther, P.M. 2022. New exponential scoring functions for diet quality indexes solve problems caused by truncation. J Nutr **152**(4): 1168–1173. Doi:10.1093/JN/NXAB431.

De Rijk, M.G., Slotegraaf, A.I., Brouwer-Brolsma, E.M., Perenboom, C.W.M., Feskens, E.J.M., and de Vries, J.H.M. 2021. Development and evaluation of a diet quality screener to assess adherence to the Dutch food-based dietary guidelines. Br J Nutr **128**(8): 1615-1625. Doi:10.1017/S0007114521004499.

Schröder, H., Benitez Arciniega, A., Soler, C., Covas, M.I., Baena-Díez, J.M., and Marrugat, J. 2012. Validity of two short screeners for diet quality in time-limited settings. Public Health Nutr **15**(4): 618–626. Doi:10.1017/S1368980011001923.

Schröder, H., Fitó, M., Estruch, R., Martínez-González, M.A., Corella, D., Salas-Salvadó, J., Lamuela-Raventós, R., Ros, E., Salaverría, I., Fiol, M., Lapetra, J., Vinyoles, E., Gómez-Gracia, E., Lahoz, C., Serra-Majem, L., Pintó, X., Ruiz-Gutierrez, V., and Covas, M.I. 2011. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. Journal of Nutrition **141**(6): 1140–1145. Doi:10.3945/jn.110.135566.

Shamon, H., and Berning, C. 2019. Attention check items and instructions in online surveys with incentivized and non-incentivized samples: Boon or bane for data quality? Survey Research Methods **14**(1): 55-77. Doi:10.2139/SSRN.3549789.

Statistics Canada. 2022. Canadian Community Health Survey – Annual component (CCHS). Available from: https://www.statcan.gc.ca/en/survey/household/3226 [accessed 3 August 2022].

Statistics Canada. 2021. Canadian Community Health Survey – Annual component (CCHS)-2021. Available from:

https://www23.statcan.gc.ca/imdb/p3Instr.pl?Function=assembleInstr&a=1&&lang=en&Item_Id=1293153#qb1293306 [accessed 3 August 2022].

Storozuk, A., Ashley, M., Delage, V., and Maloney, E.A. 2020. Got bots? Practical recommendations to protect online survey data from bot attacks. The Quantitative Methods for Psychology **16**(5): 472–481. Doi:10.20982/TQMP.16.5.P472.

Subar, A.F., Kirkpatrick, S.I., Mittl, B., Zimmerman, T.P., Thompson, F.E., Bingley, C., Willis, G., Islam, N.G., Baranowski, T., McNutt, S., and Potischman, N. 2012. 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 **112**(8): 1134–1137. Doi:10.1016/j.jand.2012.04.016.

Tarasuk, V., Fitzpatrick, S., and Ward, H. 2010. Nutrition inequities in Canada. Applied Physiology, Nutrition, and Metabolism **35**(2): 172–179. Doi:10.1139/H10-002.

Thompson, F.E., Kirkpatrick, S.I., Subar, A.F., Reedy, J., Schap, T.E., Wilson, M.M., and Krebs-Smith, S.M. 2015. The National Cancer Institute's Dietary Assessment Primer: A resource for diet research. J Acad Nutr Diet **115**(12): 1986–95. Doi:10.1016/j.jand.2015.08.016.

Thompson, F.E., Midthune, D., Kahle, L., and Dodd, K.W. 2017. Development and evaluation of the National Cancer Institute's Dietary Screener Questionnaire scoring algorithms. The Journal of Nutrition **147**(6): 1226–1233. Doi:10.3945/JN.116.246058.

Tooze, J.A., Kipnis, V., Buckman, D.W., Carroll, R.J., Freedman, L.S., Guenther, P.M., Krebs-Smith, S.M., Subar, A.F. and Dodd, K.W., 2010. A mixed-effects model approach for estimating the distribution of usual intake of nutrients: the NCI method. Statistics in Medicine, **29**(27): 2857-2868. Doi:10.1002/sim.4063.

Wang, Y., Wei, Y., Zhang, M., Liu, Y., and Wang, B. 2021. Make complex CAPTCHAs simple: A fast text captcha solver based on a small number of samples. Information Sciences **578**: 181–194. Doi:10.1016/J.INS.2021.07.040.

Weiss, B.D., Mays, M.Z., Martz, W., Castro, K.M., DeWalt, D.A., Pignone, M.P., Mockbee, J., and Hale, F.A. 2005. Quick assessment of literacy in primary care: the newest vital sign. Annals of Family Medicine. **3**(6): 514-522. Doi: 10.1370/afm.405. PMID: 16338915. Erratum in: 2006. Ann Fam Med. 4(1): 83.

Willett, W. Nutritional Epidemiology, 3rd ed.; Oxford University Press: New York, NY, USA, 2012.

Zhang, S., Midthune, D., Guenther, P., Krebs-Smith, S., Kipnis, V., Dodd, K., Buckman, D., Tooze, J., Freedman, L., and Carroll, R. 2011. A new multivariate measurement error model with zero-inflated dietary data, and its application to dietary assessment. The Annals of Applied Statistics **5**(2B): 1456–1487. Doi:10.1214/10-AOAS446.

Chapter 6: Advances in methods for characterizing dietary patterns: A scoping review

Status: Under review by the *British Journal of Nutrition*

Authors: Joy M. Hutchinson[1], Amanda Raffoul[2], Alexandra Pepetone[1], Lesley Andrade[1], Tabitha E. Williams[1], Sarah A. McNaughton[3], Rebecca M. Leech[4], Jill Reedy[5], Marissa M. Shams-White[6,7], Jennifer E. Vena[8], Kevin W. Dodd[9], Lisa M. Bodnar[10], Benoît Lamarche[11], Michael P. Wallace[12], Megan Deitchler[13], Sanaa Hussain[1], Sharon I. Kirkpatrick[1]

1. School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada
2. Department of Nutritional Sciences, University of Toronto, Toronto, ON, Canada
3. Health and Well-Being Centre for Research Innovation, School of Human Movement and Nutrition Sciences, University of Queensland, St. Lucia, QLD, Australia
4. Institute for Physical Activity and Nutrition, School of Exercise and Nutrition Sciences, Deakin University, Victoria, Geelong, Australia
5. National Cancer Institute, National Institutes of Health, Bethesda, MD, USA
6. Population Science Department, American Cancer Society, Washington DC, USA
7. Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA
8. Alberta's Tomorrow Project, Alberta Health Services, Edmonton, AB, Canada
9. Division of Cancer Prevention, National Cancer Institute, Bethesda, MD, USA
10. School of Public Health, University of Pittsburgh, Pittsburgh, PA, USA
11. Centre Nutrition, santé et société (NUTRISS), Institut sur la nutrition et les aliments fonctionnels (INAF), Université Laval, Québec City, QC, Canada
12. Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada

13. Intake – Center for Dietary Assessment, FHI Solutions, Washington, DC, USA

Acknowledgements: We thank research librarian Jackie Stapleton (JS) of the University of Waterloo for support with the search strategy.

6.1. Overview

There is a growing focus on better understanding the complexity of dietary patterns and how they relate to health and other factors. Approaches that have not traditionally been applied to characterize dietary patterns, such as machine learning algorithms and latent class analysis methods, may offer opportunities to measure and characterize dietary patterns in greater depth than previously considered. However, there has not been a formal examination of how this wide range of approaches has been applied to characterize dietary patterns. This scoping review synthesized literature from 2005-2022 applying methods not traditionally used to characterize dietary patterns, referred to as novel methods. MEDLINE, CINAHL, and Scopus were searched using keywords including machine learning, latent class analysis, and least absolute shrinkage and selection operator (LASSO). Of 5274 records identified, 24 met the inclusion criteria. Twelve of 24 articles were published since 2020. Studies were conducted across 17 countries. Nine studies used approaches that have applications in machine learning to identify dietary patterns. Fourteen studies assessed associations between dietary patterns that were characterized using novel methods and health outcomes, including cancer, cardiovascular disease, and asthma. There was wide variation in the methods applied to characterize dietary patterns and in how these methods were described. The extension of reporting guidelines and quality appraisal tools relevant to nutrition research to consider specific features of novel methods may facilitate complete and consistent reporting and enable evidence synthesis to inform policies and programs aimed at supporting healthy dietary patterns.

6.2. Introduction

Dietary intake is among the top risk factors for chronic diseases.[1,2] Research examining dietary intake has historically focused on single foods, nutrients, or other dietary constituents.[3] As the focus of public health nutrition shifted from the prevention of deficiency to include the prevention of chronic diseases, research likewise shifted towards the examination of dietary patterns, aiming to capture how foods and beverages are consumed in real life.[3–5] Humans typically do not consume foods or nutrients on their own, but in the context of a broader dietary pattern.[3,4] Accordingly, food-based dietary guidelines are now typically focused on patterns of intake rather than single dietary components.[6] It is likely the synergistic and antagonistic relationships among the multiple foods, beverages, and other dietary components that humans consume that influence health rather than individual components.[4] In addition to this multidimensionality, dietary patterns are dynamic, changing from meal to meal, day to day and across the life course.[4,7] Further, dietary patterns are shaped by culture, social position, and other contextual factors.[8,9] However, incorporating the domains of multidimensionality, dynamism, and contextual factors into dietary patterns analysis is a difficult task.

Traditional approaches to identify dietary patterns, including 'a priori' and 'a posteriori' approaches, are useful for understanding overall dietary patterns or the diet quality of populations and population subgroups.[10] For example, 'a priori' methods like the Healthy Eating Index-2020 and the Healthy Eating Food Index-2019 are generally investigator driven[11,12] and consider multiple components such as fruits and vegetables and whole grains as inputs, , but typically compress the multidimensional construct of dietary patterns to a single unidimensional score reflecting overall diet quality.[13,14] 'A posteriori' approaches are data-driven and have also been widely used to identify dietary patterns. Commonly applied data-driven approaches include clustering methods (e.g., k-means, Ward's method), principal component analysis, and factor analysis, providing opportunities to identify dietary patterns through statistical modeling

or clustering algorithms rather than relying on researcher hypotheses.[15] These approaches compress dietary components to key food groupings typically expressed as single scores.[10,16] By reducing the dimensionality of dietary patterns, these methods are limited in their ability to explain the wide variation in dietary intakes.[4] Methods employed to traditionally characterize dietary patterns using 'a priori' and 'a posteriori' approaches thus address multidimensionality to some extent, but do not allow for explorations of dietary patterns in their totality because they miss potential synergistic or antagonistic associations among dietary components.[4,14,17]

Novel methods that have not traditionally been used to identify dietary patterns, such as probabilistic graphical modeling, latent class analysis, and machine learning algorithms (e.g., random forest, neural networks), may capture complexities like dietary synergy. There is no clear delineation between traditional and novel methods, and specifically defining what is novel is challenging given it naturally implies an evolution of methods. Nonetheless, there is a growing interest among nutrition researchers in the application of methods that have not typically been used to capture dietary complexity, with these methods often centered in machine learning.[18] To date, there have been perspectives and narrative reviews on the application of machine learning in nutrition,[19–21] and a recent systematic review of studies that applied machine learning approaches to assess food consumption.[22] However, there has not been an assessment of studies applying novel methods to characterize dietary patterns. Given the rapid adoption of these methods within the field of health,[23–26] it is increasingly important for researchers to have a basic understanding of available methods and how they are being applied in the field. This will facilitate the synthesis of evidence from a range of methodological inputs to inform food-based dietary guidelines and other policies and programs that promote health. The objective of this scoping review was therefore to describe the use of novel methods not traditionally used to characterize dietary patterns in the published literature.

6.3. Methods

The review was conducted in accordance with the JBI Manual for Evidence Synthesis,[27] which was developed using the Arksey and O'Malley framework.[28] Reporting follows the Preferred Reporting Items for Systematic Reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR).[29]

Defining novel methods

The novel methods considered were based on a preliminary search of the literature and the expertise of the research team and included systems methods (e.g., agent-based modeling, system dynamics), least absolute shrinkage and selection operator (LASSO), machine learning algorithms, copulas, and data-driven statistical modeling approaches (e.g., treelet transformations, principal balances and coordinates). Novel methods could also include those that have been used previously in nutrition research if applied in new ways to characterize dietary patterns (e.g., linear programming used to model a modified dietary pattern rather than to test scenarios). Methods that were not considered to be novel were those that have been applied to assess dietary patterns in numerous studies and have been considered by prior reviews and commentaries,[2,10,30] including regression, 'a priori' approaches such as investigator-driven indices, and routinely used data-driven approaches, including factor analysis and cluster analysis.[10,31]

Identifying relevant studies

Articles were eligible for inclusion if they were: a primary research article; focused on dietary intake as an exposure or outcome, including examination of dietary patterns (i.e., multiple dietary components in combination rather than single nutrients, foods, or other dietary components); used at least one or more novel methods, as described above, to characterize dietary patterns; were published in English; and focused on humans. Ineligible studies included

those focused on individual foods or human milk rather than dietary patterns, and commentaries and reviews.

Searches of three research databases, MEDLINE (via PubMed), the Cumulative Index to Nursing and Allied Health Literature (CINAHL), and Scopus, were conducted in March 2022. These health-focused, specialized, and multidisciplinary databases were selected based on consultation with a research librarian (JS) to ensure a range of possibly relevant study types were included. The search strategies were developed in consultation with the research librarian using keywords and subject headings to capture diet-related constructs (e.g., dietary intake, patterns, recommendations, feeding behavior, food habits) and novel methods to characterize dietary patterns (e.g., machine learning, network science, system dynamics model). No date limits were applied to the searches, and articles were included until the end of the search in March 2022. The search strategies for MEDLINE, CINAHL, and Scopus are available in **Appendix J**.

Study selection

Two independent reviewers (two of AP, AR, SH, SIK) screened each record at the title and abstract and full-text screening stages using Covidence,[32] with one consistent reviewer (AP) participating throughout the entire process. At the title and abstract screening stage, an initial pilot screening (25 records) generated 100% agreement (AR and AP) and 92% agreement (AP and SH). A second pilot screening (100 records) generated 91% agreement (AR and AP) and 93% agreement (AP and SH). When applicable, discrepancies were discussed by reviewers and if needed deferred to a third reviewer (SIK) for decision. Following pilot screening, the reviewers independently reviewed the remaining articles (96% agreement, Kappa = 0.83).

The reviewers were intentionally liberal during the title and abstract screening stage because of the breadth of possible novel methods. This required iteratively revisiting the inclusion criteria. For example, reduced rank regression was initially considered to be novel but

was found to be prevalent in the literature based on title and abstract screening and was excluded during full-text review. Further, articles that used 'a posteriori' methods to identify dietary intake but did not specify the exact method in the title or abstract were included for full-text review.

Pilot screening of full-text reviews (50 records) generated 82% agreement (AR and AP) and 96% agreement (AP and SIK); after discrepancies were discussed, two reviewers independently screened the remaining full-text articles (93% agreement, Kappa = 0.60). The high agreement between reviewers but relatively low Cohen's Kappa is described as Cohen's paradox, with a larger number of studies excluded than included.[33–35]

Data extraction

Data extraction was completed by TEW and JMH using a pre-specified Excel template, with all extracted data subsequently verified by LA. Data extraction fields (**Appendix K**) included information pertaining to authorship, study title, journal, year of publication, funding source, contextual details (e.g., study location), sample size, and participant characteristics (e.g., age). Details relating to study methods (e.g., analysis input variables, measurement of dietary intake, analytic approaches) and results (e.g., findings related to dietary patterns and if applicable, health risk and outcomes) were also extracted.

6.4. Results

Summary of search

A total of 5274 unique articles were identified after removing duplicates. Of these, 436 were identified as potentially relevant based on the title and abstract review and underwent full-text screening (**Figure 5**). Studies excluded during full-text screening included those that did not include methods defined as novel, those that did not focus on dietary patterns,

commentaries, narrative reviews, systematic reviews, studies that were not published in English, studies that were not conducted with humans, and theses/dissertations. A final pool of 24 articles describing 24 unique studies met the inclusion criteria.

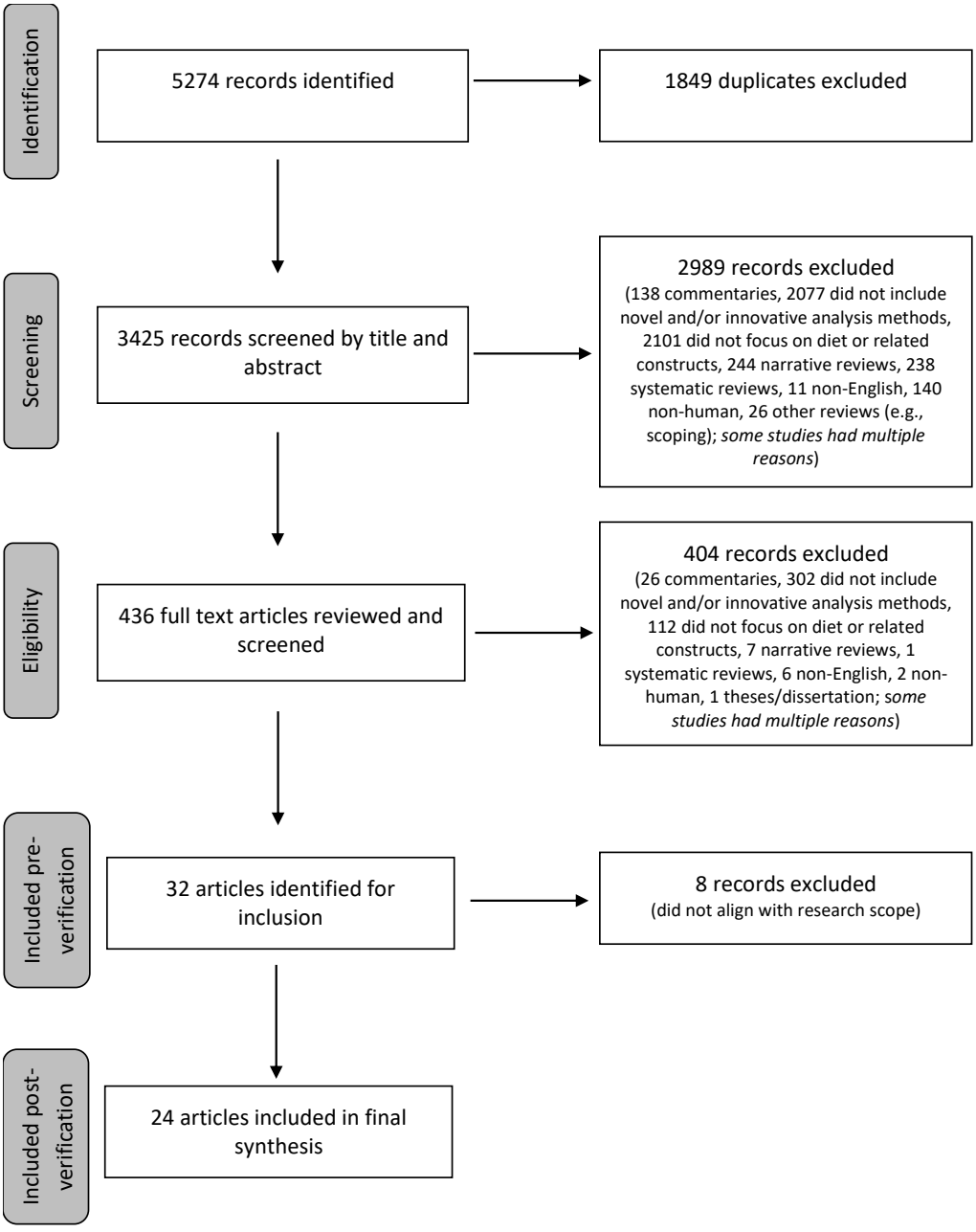


Figure 5. PRISMA diagram illustrating the screening process for a scoping review exploring innovative methods for the analysis of dietary intake data and characterization of dietary patterns

Characteristics of included studies

Across the 24 included studies, data from 17 countries were represented (**Table 8**). Half of the studies were published between 2005 and 2019,[36–47] and the remaining 12 were published between 2020 and March 2022.[48–59] Three studies used data from subsets of the European Prospective Investigation into Cancer and Nutrition,[36,45,46] two studies used waves of data from the National Health and Nutrition Examination Survey,[54,58] and two studies used data from the ELSA-Brasil cohort study (**Table 9**).[42,57] Sample sizes ranged from 250 to over 73,000 participants. Nineteen studies were conducted using data from cohort or cross-sectional studies and five studies applied a case-control design.

The majority (n=15) of studies used food frequency questionnaires to assess dietary intake.[36,37,39–43,45,47–49,52,55–57] Six studies used 24-hour recalls,[46,51,53,54,58,59] two studies used food records/diaries,[44,50] and one study used a food frequency questionnaire and a 24-hour recall.[38] Among the studies using 24-hour recalls and records/diaries, one used data from a single recall that was combined with data from a food frequency questionnaire.[38] The remaining studies including records or recalls averaged or combined data from two or more days of intake. Apart from averaging recalls or records, none of the included studies applied substantial efforts to mitigate measurement error present in dietary intake data. Several studies noted potential misreporting as a limitation, and only five studies specifically noted that findings may have been influenced by measurement error present in self-reported dietary assessment instruments.[40,43,51,52,59]

Table 8: Study characteristics across included studies applying novel methods to characterize dietary patterns

Characteristic	Number of articles
Country of study population[§]	
Austria	1
Brazil	2
China	2
Finland	1
France	3
Germany	3
Italy	1
Mexico	1
Netherlands	1
Northern Ireland	1
Poland	1
Portugal	3
Republic of Ireland	2
South Korea	1
Spain	1
Switzerland	1
USA	3
Year published	
2005-2009	2
2010-2014	3
2015-2019	7
2020-2022	12
Included health risk or outcomes	
Health risk/outcome (cross-sectional)	8
Health outcome (longitudinal)	6
No	10

[§]Some studies included more than one country

Novel methods applied to identify dietary patterns

The type of methods used and how they were implemented to identify dietary patterns varied widely (**Table 10**). Nine studies applied approaches that have applications in machine learning, including classification models, neural networks, and probabilistic graphical models (**Table 11**).[36,38,44–46,49,53,56,59] The earliest study included in this review was published in 2005 and applied neural networks to characterize dietary patterns.[38] Fifteen studies applied other novel methods, including latent class analysis, mutual information, and treelet transform.[37,39–43,47,48,50–52,54,55,57,58] Two studies identified dietary patterns using more than one novel method.[44,53] Five studies included comparisons of different novel methods, though these were

typically versions of the same model.[44,45,47,49,53] For example, Solans et al. compared three models for compositional data analysis and reported that the best-performing model incorporated both investigator- and data-driven methods.[47]

In twelve studies, two to eight distinct dietary patterns, such as the 'prudent' pattern or 'Western' pattern, were identified using methods such as latent class analysis, treelet transform, random forest with classification tree analysis, and multivariate finite mixture models.[36–38,41–43,48,50,51,54,57,58] Six studies applied network methods, including probabilistic graphical models and mutual information, to identify networks of dietary patterns among populations.[45,46,49,52,55,59]

Dynamism, or how dietary patterns vary across time, was incorporated into four studies' characterization or analysis of dietary patterns. Three studies incorporated stratification by meals to consider dynamism.[44,46,59] In two studies using graphical models, separate networks were created for each meal to provide insights into how patterns of intake vary throughout the day.[46,59] Hearty and Gibney used decision trees and neural networks and ran models by meals based on 62 food groups to predict diet quality.[44] Additionally, one study considered dynamism by using ANOVA and chi-square tests to descriptively show how a variety of characteristics were associated with stable or changing dietary patterns characterized using latent class analysis.[43]

Table 9: Characteristics of studies (n=24) identified in a scoping review of novel analytic methods to characterize dietary patterns

Authors, date, country(s)	Study aims	*n*	Study design and data source	Participant characteristics	Diet assessment method(s)
Affret et al., 2017, France[41]	To characterize the relationship of dietary patterns with socioeconomic factors	73 031	Cohort; E3N cohort[a]	Adults in France, females only, mean age 52.9 y	FFQ[b]
Bezerra et al., 2018, Brazil[42]	To examine how dietary patterns differ among generations in a Brazilian cohort	15 069	Cross-sectional analysis of cohort; ELSA-Brasil cohort[c]	Adults in Brazil, aged 35-74 y	FFQ
Biesbroek et al., 2015, Netherlands[36]	To compare if two methods (traditional vs. novel) derived differential dietary patterns and associations with health risk	34644	Cohort; EPIC[d] Prospect and EPIC MORGEN	Adults in the Netherlands, male and female, aged 20-70 y	FFQ
Dalmartello et al., 2020, Italy[48]	To identify dietary patterns associated with oropharyngeal cancer	946 cases, 2492 controls	Case-control; Unnamed study	Adults in Italy, admitted to major hospitals, 19-82 y	FFQ
De Almeida et al., 2022, Brazil[57]	To investigate how a variety of approaches identify dietary patterns in the same dataset	12 816	Cohort; ELSA-Brasil cohort	Adults in Brazil, aged 35-74 y	FFQ
Farmer et al., 2020, USA[58]	To explore how latent class analysis groups individuals into dietary patterns	11 481	Repeat cross-sectional; NHANES[e] 2007-2010	Adults in the USA, male and female, ≥19 y	Up to two 24HDRs[f] (per day)

Fonseca et al., 2012, Portugal[37]	To investigate the association between dietary patterns and metabolic syndrome	2167	Cross-sectional analysis of cohort study; EPIPorto cohort	Adults in Portugal, male and female, age not specified	FFQ
Harrington et al., 2014, Republic of Ireland[43]	To investigate changes in dietary patterns across 10 years among Irish adults	925 (baseline) 320 (follow-up)	Cohort; Cork and Kerry Diabetes and Heart Disease Study	Adults in the Republic of Ireland, 50-69 y	FFQ
Hearty & Gibney, 2008, Republic of Ireland & Northern Ireland[44]	To compare the ability of two supervised machine learning techniques to predict diet quality based on dietary patterns	1379	Cross-sectional; North-South Ireland Food Consumption Survey	Adults in Northern Ireland and the Republic of Ireland, aged 18-64 y	7-day food diary
Hoang et al., 2021, South Korea[49]	To describe interactions among dietary components, demographic characteristics, and chronic disease risk factors	7423	Cross-sectional analysis of cohort study; Cancer Screening Examination Cohort	Adults in South Korea, 30-70 y	FFQ
Hose et al., 2021, Austria, Finland, France, Germany, Switzerland[50]	To explore how feeding patterns across the first year of life relate to asthma in school-aged children	1246	Cohort; PASTURE[g] Birth Cohort and LUKAS2[h] Cohort	Children in 5 European countries, birth cohort	Food diaries (weekly & monthly), dietary screener (monthly)
Iqbal et al., 2016, Germany[45]	To identify dietary patterns among a German adult population	27 120	Cross-sectional analysis of cohort; EPIC Potsdam	Adults in Germany, male and female, aged 35-65 y	FFQ
Jamiolkowski et al., 2005, Poland[38]	To identify dietary patterns and their associations with	556	Cohort; Unnamed study	Adult males in Poland, age not specified	FFQ + 24HDR

	markers of health status				"24-hour consumption questionnaire"
Madeira et al., 2021, Portugal[51]	To investigate associations between dietary patterns, malnutrition, and body mass index among older adults	849	Cross-sectional; PEN-3S Study[i]	Adults in Portugal, ≥65 y	Two 24HDR (per day)
Oliveira et al., 2011, Portugal[39]	To assess associations between a posteriori dietary patterns and health outcomes	3016 *820 cases, 2196 controls	Case-control; Unnamed study	Adults in Portugal, male and female, aged ≥ 18 years	FFQ
Samieri et al., 2020, France[52]	To model dietary patterns prior to the onset of dementia	1522	Case-control; 3-City Study	Adults in France, ≥65 y	FFQ
Schwedhelm et al., 2018, Germany[46]	To describe dietary networks using machine learning methods across meals and usual patterns of eating	814	Cross-sectional analysis of cohort; EPIC Potsdam	Adults in Germany, male and female, average age 65.5 y	Up to three 24HDR (per day)
Schwedhelm et al., 2021, USA[59]	To compare dietary patterns by meals among pregnant women with high and low diet quality	365	Cohort; Pregnancy Eating Attributes Study	Women receiving prenatal care in the USA, 18-44 y	Up to three 24HDRs, pooled
Shang et al., 2020, China[53]	To characterize top dietary components that are associated with changes in cardiometabolic risk factors among children	5676	Cohort; Unnamed study	Children in China, 6-13 y	Three 24HDRs, assessed at baseline and one year. Average of recalls

Solans et al., 2019, Spain[47]	To compare three compositional data analysis approaches to identify dietary patterns	3471	Case-control; Multi-Caso Control Spain	Adults in Spain, male and female, aged 20-85 y (controls only)	FFQ
Wright et al., 2020, USA[54]	To examine how dietary patterns are associated with periodontitis	10 010	Repeat cross-sectional; NHANES 2009-2014	Adults in the USA, male and female, ≥30 y	Up to two 24HDRs Average of recalls
Wu et al., 2019, Mexico[40]	To explore how maternal and adolescent diet are associated with pubertal tempo	250 (baseline) 222 (follow-up)	Cohort; ELEMENT cohort[j]	Children in Mexico enrolled in a birth cohort	FFQ
Xia et al., 2020, China[55]	To examine how dietary patterns differ between individuals with non-alcoholic fatty liver disease and controls	4086, split 50/50 between cases and controls	Case-control; Tianjin Chronic Low-grade Systemic Inflammation and Health Cohort Study	Individuals in China	FFQ
Zhao et al., 2021, USA[56]	To explore the relationship between dietary components, dietary patterns, and cardiovascular disease	1928	Cross-sectional analysis of cohort; CARDIA cohort[k]	Adults in the USA, 18-30 y	FFQ

[a]Etude Epidémiologique auprès de femmes de la Mutuelle Générale de l'Education Nationale
[b]Food frequency questionnaire
[c]Brazilian Longitudinal Study of Adult Health
[d]European Prospective Investigation into Cancer and Nutrition
[e]National Health and Nutrition Examination Survey
[f]24 hour dietary recall
[g]Protection Against Allergy: Study in Rural Environments
[h]Finnish, rural-suburban birth cohort
[i]Portuguese Elderly and Nutritional Status Surveillance System
[j]Early Life Exposure in Mexico to Environmental Toxicants
[k]Coronary Artery Risk Development in Young Adults

Table 10: Description of dietary patterns (n=24) identified in a scoping review of novel analytic methods to characterize dietary patterns

Studies that characterized dietary patterns to assess health outcomes/risk					
Authors, date, country(s)	**Novel method(s) to analyze diet**	**Dietary input variables**	**Characterization of dietary patterns**	**Consideration of health outcomes/risk**	**Consideration of sociodemographic characteristics[§]**
Biesbroek et al., 2015, Netherlands[36]	Random forest with classification tree analysis	Food items were grouped into 35 food groups based on Dutch standard food groups, represented as % of total energy.	7 patterns were derived to explain cardiovascular disease risk factors: Prudent-like 1, Prudent-like 2, Traditional-like 1, Traditional-like 2, Traditional-like 3, Western-like 1, Western-like 2.	Investigated the association between dietary patterns and first occurrence of coronary artery disease, stroke, death from any cause through hazard ratios	All hazard ratios were adjusted for age and sex and one model additionally considered smoking status and education.
Dalmartello et al., 2020, Italy[48]	Latent class analysis	Items from the FFQ were grouped into 25 food groups and daily intake (g/d) was estimated using standard portion sizes.	Latent class analysis was used to identify distinct dietary patterns: Western, Prudent, Lower consumers- combination, Higher consumers-combination.	Investigated associations of dietary patterns with risk of oropharyngeal cancer using odds ratios.	Not stated how cases and controls were matched. Odds ratios were adjusted for age, sex, education, smoking status.
Fonseca et al., 2012, Portugal[37]	Multivariate finite mixture models	Items from the FFQ were grouped into 14 food groups based on similar nutritional characteristics.	Identified four dietary patterns among males and females. Males: healthy, fish, red meat and alcohol, intermediate. Females: healthy, low fruit and vegetable, red meat and alcohol, in transition to fast food.	Associations between metabolic syndrome and dietary patterns were assessed using crude and adjusted odds ratios	Identified dietary patterns by sex. Odds ratios were adjusted for age, education, smoking status.
Hoang et al., 2021, South Korea[49]	Gaussian graphical models and	16 food groups, assessed as g/d, were included in	A network of intake was identified among the 16 food groups, with edges	Graphical models also included comorbidity factors,	Sociodemographic characteristics including age, sex,

	mixed graphical models	graphical models. These food groups were also included in a graphical model as a single dietary score.	representing conditional dependencies between food groups. Food groups that were not connected by an edge were conditionally independent.	such as blood pressure or total cholesterol, to identify associations with dietary components.	marital status, education, and income were included in graphical models.
Hose et al., 2021, Austria, Finland, France, Germany, Switzerland[50]	Latent class analysis	Included 17 food items from 19-item screener that assessed time of introduction (screeners repeated at 2, 12, 18, 24, 36, 48, 60, 72 months of age) and frequency of consumption.	Four food introduction patterns were identified. Latent class 1 was characterized by daily fruit, vegetable, and milk consumption; Latent class 2 was characterized by daily fruit, vegetable, and meat and cereal consumption, but less frequent milk consumption; Latent class 3 was characterized by slower introduction to foods, a diverse range of food items, but less than daily fruit and vegetable consumption; Latent class 4 was similar to Latent class 3 but with quicker introduction to foods.	Linear or logistic/multinomial regression was used to estimate associations between latent class analysis food introduction patterns, possible determinants, and asthma outcomes.	Regression analyses adjusted for sex.
Jamiolkowski et al., 2005, Poland[38]	Kohonen neural networks	41-item FFQ was included as input.	Three dietary patterns were identified: 1- high frequency of margarine, fried vegetable oil, low consumption of butter, lard, bacon, most frequent consumption of poultry, fresh fish, fruit;	Impact of dietary habits on blood serum lipids assessed using Kruskall-Wallis procedure.	No.

			2- high frequency butter, lard, giblets, smoked 'second-rate' meats, sweets, sugary beverages; 3- low frequency butter, margarine, milk, cheese.		
Madeira et al., 2021, Portugal[51]	Latent class transition model	Items were grouped from 71 food items to 27 food groups.	Latent class transition models identified two dietary patterns, labelled as a protein-foods pattern and a Mediterranean pattern.	Linear regression was used to examine relationships between identified dietary patterns and nutritional status and BMI	Regression models adjusted for sex, age, education.
Oliveira et al., 2011, Portugal[39]	Multivariate and univariate finite fixture models	Items from an 82-item FFQ were grouped into 14 food groups (g/d) based on nutritional similarities.	Distinct dietary patterns were identified among males and females using multivariate and univariate fixture models. Males: Healthy, Fish, Red meat and alcohol, intermediate intake. Females: Healthy, Low fruit and vegetables, red meat and alcohol, in transition to fast food.	Logistic regression was used to examine the association between dietary patterns and acute myocardial infarction (with and without confounders)	Identified dietary patterns by sex. Multivariate models included age, education, smoking status as confounders.
Samieri et al., 2020, France[52]	Mutual information	148-item FFQ grouped into 62 dietary variables.	A mutual information matrix was computed to estimate the network relationships between dietary components. Networks of intake were estimated among cases and controls. The case network showed one mainly connected network of foods,	Two mutual information networks were developed and compared – one for dementia cases and one for controls. A third network representing the differences between the cases' and	Nested case-control. Matched cases and controls by age, sex, and education.

			whereas the controls network was represented by several disconnected subnetworks.	controls' dietary patterns was developed.	
Shang et al., 2020, China[53]	Random forest, gradient boost machine	Foods were categorized into 26 food groups (% energy/d)	The components from baseline dietary intake served as inputs into machine learning algorithms to determine the most important components for changes in cardiometabolic disease risk. The most important components were then summed to create a Healthy Diet Score.	Changes in cardiometabolic risk were evaluated by subtracting the results at baseline from those at follow-up. The Healthy Diet Score was then tested in an external dataset.	Regression models adjusted for age, sex, household income, parental education.
Wright et al., 2020, USA[54]	Treelet transform	Food and Nutrient Database for Dietary Studies food groups were used as inputs.	Eight treelet components (dietary patterns) were identified that explained substantial variation. 1: diets high in salads, moderate vegetables, poultry seafood, fruit, water, tea. 2: loaded on milk, ready-to-eat cereal, bananas. 3: loaded on bread, luncheon meat, table fat, coffee, sugar. 4: loaded on fruit juice, fruit-flavored drinks. 5: low intake bread, meat, table fats, high intake coffee and sugar. 6: loaded on beef, potatoes, beans, corn.	Quantile regression was used to estimate associations between dietary patterns and periodontitis.	Adjusted models for sex, ethnicity, age, education, smoking status.

			7: loaded on candy, crackers, pretzels, low-calorie carbonated soft drinks. 8: loaded on pancakes, eggs, French toast.		
Wu et al., 2019, Mexico[40]	LASSO	From a 116-item FFQ among adolescents and 104-item FFQ among pregnant women, 21 dietary components were identified as potential methyl-donor rich foods for inclusion in LASSO models.	Seven food items were retained from LASSO models for the maternal diet and eight from the adolescent diet. An Epigenetics Associated Diet Score was created for methyl donor-rich foods, with weights for dietary components obtained from LASSO.	Time-to-event analysis with interval-censored regression models was used to investigate the relationship between Epigenetics Associated Diet Scores and pubertal timing.	Final models adjusted for household socioeconomic status (included material wellbeing, highest education level in household).
Xia et al., 2020, China[55]	Mutual information	100-item FFQ condensed to 25 food groups (g/d) based on similar characteristics.	Dietary patterns were identified using mutual information, identifying five primary dietary components in the cases and six in the controls.	Separate networks were developed for non-alcoholic fatty liver disease cases and controls and were compared.	Case-control. Used propensity score matching to match cases and controls. This process considered sex, age, education, income, smoking status, employment status. Cases that could not be matched closely enough were not included in analyses.
Zhao et al., 2021, USA[56]	Bayesian kernel machine regression	FFQ items were assigned to 166 food groups and were further grouped to 12	Bayesian kernel machine regression simultaneously modeled the relationship between the 12 dietary	Bayesian kernel machine regression was used to jointly consider the 12 dietary components	Balanced recruitment by sociodemographic characteristics. 10-year CVD calculation

		dietary factors based on evidence of associations with cardiovascular disease.	components of interest and cardiovascular disease risk, including the change in each dietary component and disease risk and the dose-response relationship of each dietary component.	and cardiovascular disease risk.	considers age, sex, race/ethnicity, smoking status. Additionally adjusted regression models for income and education.
Studies that characterized dietary patterns (no health outcomes/risk)					
Authors, date, country(s)	**Analytic method(s)**	**Dietary input variables**	**Characterization of dietary patterns**	**Consideration of sociodemographic characteristics**	
Affret et al., 2017, France[41]	Latent Class Analysis	Created 57 food groups from the 208 food items available in the FFQ.	Used Latent Class Analysis to obtain distinct dietary patterns from these groups, deriving a 'Healthy' and 'Less Healthy' pattern.	Used logistic regression to assess the individual and socioeconomic factors associated with 'healthy diet'. These factors included education, employment status, income, number of children, marital status, place of residence, size of city, deprivation index, geographical location of birth.	
Bezerra et al., 2018, Brazil[42]	Latent Class Analysis	A 114-item FFQ was condensed to create 13 mutually exclusive groups based on nutritional content and level of processing (cereals and tubers and roots; skimmed milk and dairy products; white meats and fish; fruit; vegetables; legumes and nuts; whole milk and dairy products; red	Latent class models identified dietary patterns for the whole sample and each generation. Traditionalists: mixed pattern, restricted pattern, prudent pattern. Baby Boomers: mixed pattern, processed pattern, prudent pattern. Generation X: mixed pattern, processed pattern, prudent pattern.	Assessed how dietary patterns differed by three birth cohorts (age).	

		meats; processed meats; ultra-processed products; soft drinks and industrialized fruit drinks; sweets; coffee).		
De Almeida et al., 2022, Brazil[57]	Treelet transform	114-item FFQ condensed to 25 food groups based on nutritional similarity.	Used three analytic approaches to derive dietary patterns (treelet transform, factor analysis, reduced rank regression). Treelet transform identified three patterns: convenience, prudent, and rice and beans. Factor analysis identified three patterns: Convenience, Brazilian traditional, and prudent. Reduced rank regression identified three patterns but only retained one, the convenience pattern.	Assessed how dietary patterns differed by sociodemographic characteristics: age, sex, smoking status, education, race
Farmer et al., 2020, USA[58]	Latent class profile analysis	11 out of 12 components from the Healthy Eating Index-2010 served as inputs (1. total vegetable, 2. greens/beans, 3. total fruit, 4. whole grains, 5. dairy, 6. total protein foods, 7. seafood and	Distinct classes for total day intake and dinner intake were identified. Classes included (total day intake): Standard American Diet, Standard American Diet with low sodium, Healthy US, Healthy US with high vegetable, Healthy US low sodium.	Multinomial logistic regression used to assess how latent classes differed by sociodemographic characteristics, including age, sex/gender, race/ethnicity, marital status, education, employment status, poverty/income ratio, number of people in household.

		plant proteins, 8. fatty acid ratio, 9. sodium, 10. refined grains, 11. solid fats, alcohol and added sugars).		
Harrington et al., 2014, Republic of Ireland[43]	Latent class analysis	Food items were aggregated from a 159-item and 167-item FFQ into 23 mutually exclusive food groups (g/d, ml/d)	Latent class analysis identified three dietary patterns: Western, Healthy, and low energy.	Assessed relationships between latent classes and sociodemographic characteristics, including smoking status, education, marital status.
Hearty & Gibney, 2008, Republic of Ireland & Northern Ireland[44]	Artificial neural networks and decision trees	Food codes were aggregated to 62 food groups.	Food groups were provided to neural networks and decision trees to identify dietary patterns predictive of diet quality quintile by meal.	No.
Iqbal et al., 2016, Germany[45]	Gaussian graphical models (GGM); confirmed with Semiparametric Gaussian Copula graphical models	Items from the 148-item FFQ were grouped into 49 groupings and were log-transformed.	GGMs were applied to derive dietary patterns among males and females. Edges represented conditional dependencies between food groups. Food groups that were not connected by an edge were conditionally independent.	Identified sex-specific graphical networks.
Schwedhelm et al., 2018, Germany[46]	Semiparametric Gaussian Copula graphical models	Food codes were condensed to 39 food groups.	Semiparametric Gaussian Copula graphical models derived meal-specific networks. For example, the breakfast network was characterized by one primary network with	No.

			five communities and included foods such as breakfast cereals, margarine, and eggs. Edges represented conditional dependencies between food groups. Food groups that were not connected by an edge were conditionally independent.	
Schwedhelm et al., 2021, USA[59]	Gaussian graphical model	Foods were grouped into 40 food groups based on the Food and Nutrient Database for Dietary Studies.	Diet quality was assessed using the Healthy Eating Index 2015, and participants were split into tertiles of HEI 2015 scores, to compare dietary patterns by meal of women in the low and high HEI tertile generated using Gaussian graphical models. Edges represented conditional dependencies between food groups. Food groups that were not connected by an edge were conditionally independent.	No.
Solans et al., 2019, Spain[47]	3 Compositional data analysis approaches: compositional PC analysis, balances,	9 items (vegetables, fruit, legumes, seafood, cereals, meat, dairy, monounsaturated fats, and saturated fats) from the 140-	Three variations of compositional data analysis were used to explain the variance among the dietary components and identify dietary patterns.	No.

	principal balances	item FFQ were included as inputs.

§We considered sociodemographic characteristics that are related to social position or are indicators of equity including age, sex, gender, race/ethnicity, marital status, education, employment status, smoking status as examples. We did not include physical activity, BMI, or alcohol consumption.

Fourteen studies examined relationships between dietary patterns characterized using novel methods and variables indicative of health risk or outcomes, such as periodontitis, cardiovascular disease, and metabolic syndrome (Table 10).[36–40,48–56] Six studies included longitudinal analysis of the relationship between dietary patterns and health outcomes.[36,38,40,50,52,53] Most studies that examined health risk or outcomes first identified dietary patterns using a novel method and then investigated relationships with health outcomes using regression models.[36–40,48,50,51,54] In contrast, some studies incorporated variables indicative of health outcomes or risk directly into the machine learning models.[49,56] For example, Zhao et al.[56] applied Bayesian kernel machine regression, a machine learning model designed to incorporate high dimensional data, to jointly model the relationship between several dietary components and cardiovascular disease risk. Similarly, a paper by Hoang et al.[49] included health variables within mixed graphical models, though directionality of diet-health relationships could not be ascertained given the cross-sectional nature of the data. In two case-control studies, dietary patterns were identified using mutual information to estimate dietary pattern networks, with stratification by health outcomes.[52,55]

Nineteen studies considered sociodemographic characteristics, such as sex, age, race/ethnicity, education, and income.[36,37,39–43,45,48–58] In one case, sociodemographic characteristics were included in models used to characterize dietary patterns.[49] Two studies stratified by sociodemographic characteristics, examining dietary patterns by sex[45] or age groups.[42] Studies that used case-control designs typically considered sociodemographic characteristics through matching.[52,55] In the remaining studies that considered sociodemographic characteristics, these were incorporated in regression models to explore how dietary patterns characterized using novel methods were associated with health and other characteristics.

Two studies included comparisons of novel methods and traditional statistical approaches.[36,57] For instance, Biesbroek et al.[36] found that dietary patterns identified through reduced rank regression were more strongly associated with coronary artery disease compared to those identified through random forest with classification tree analysis.

Table 11: Novel methods applied to identify dietary patterns across included studies

Method of analysis	Description	References
Methods with applications in machine learning		
Classification models	Algorithms that can be used for prediction and classification.[60,61]	
Random forest	The random forest is generated from a wide array of trees that are randomly sampled and averaged. Random forest models can additionally be used on regressors.[60]	(Biesbroek et al. 2015; Shang et al. 2020)
Decision tree	Trees include nodes (variables) and branches, representing possible outcomes from a node. This forms an 'inverted tree', beginning with a root node and branching to many possible end nodes.[61]	(Hearty and Gibney 2008)
Neural network	The structure of this algorithm was based on the human brain and is well-suited to complex, non-linear problems.[62]	
Artificial neural network	The algorithm starts with input nodes, which communicate, forming hidden layers of connections within the network, and provide learned output.[62]	(Hearty and Gibney 2008)
Kohonen neural network	A neural network comprised of a single layer. Data enter through input nodes and are organized to a simple structure of output nodes.[63]	(Jamiołkowski, Szpak, and Pawłowska 2005)
Probabilistic graphical models	Algorithm that can be used to determine joint relationships between nodes, forming a graphical network, where connections between nodes are represented by edges. These connections may be full- or partial-order correlations.[64–66]	
Gaussian graphical models	Probabilistic graphical model where data follows a Gaussian distribution.[65]	(Hoang, Lee, and Kim 2021; Iqbal et al. 2016; Schwedhelm et al. 2021)
Mixed graphical models	Probabilistic graphical model where data are mixed (i.e., Gaussian and categorical).[65]	(Hoang, Lee, and Kim 2021)
Semiparametric Gaussian Copula graphical models	Probabilistic graphical model where the data are not required to be normally distributed.[67]	(Iqbal et al. 2016;

		Schwedhelm et al. 2018)
Other machine learning approaches		
Gradient boost machine	Machine learning algorithm that applies a boosting ensemble method to decision trees. This allows a weak learning algorithm to provide predictions through weighted averages of weak predictions.[53,68]	(Shang et al. 2020)
Bayesian kernel machine regression	Machine learning algorithm that uses a kernel function to consider the effects of multiple exposures, covariates, and confounders on an outcome.[69]	(Zhao et al. 2021)
Compositional data analysis	Method that identifies the relative magnitude of differences between variables to determine relative importance.[70,71]	(Solans et al. 2019)
Least Absolute Shrinkage and Selection Operator (LASSO)	A method that selects the most important variables and reduces the number of variables in the model.[72]	(Wu et al. 2019)
Multivariate finite mixture models		
Latent Class Analysis	A method that creates mutually exclusive groups (classes) of participants based on similar latent characteristics.[41,42]	(Affret et al. 2017; Bezerra et al. 2018; Dalmartello et al. 2020; Harrington et al. 2014; Hose et al. 2021)
Latent class transition model	Allows for transition between identified groups.[51,73]	(Madeira et al. 2021)
Latent class profile analysis	Latent class analysis when continuous variables are used.[58]	(Farmer et al. 2020)
Multivariate finite mixture models	Identifies mutually exclusive groupings where cluster shape is not limited.[74]	(Fonseca et al. 2012; Oliveira et al. 2011)
Mutual information	This method creates networks, estimating the amount that uncertainty is reduced in one variable by knowing another.[75]	(Samieri et al. 2020; Xia et al. 2020)
Treelet transform	Data-driven approach that reduces dimensionality and is a hybrid between hierarchical clustering and principal component analysis.[76,77]	(de Almeida Alves et al. 2022; Wright et al. 2020)

*Studies may have used more than one novel approach to characterize dietary patterns.

6.5. Discussion

The application of novel methods to dietary pattern research is rapidly expanding, with the aim of better understanding their complexity and how they are related to health and other factors. Many studies used methods that characterize distinct dietary patterns based on the population being studied, such as the 'prudent' pattern or the 'Western' pattern. Most studies used cross-sectional data, limiting opportunities to examine the effect of dietary patterns on health.

Methods newly being applied in this field offer promising capacity to better understand the totality of dietary patterns and synergistic relationships among dietary components when compared with traditional approaches that do not assume synergy.[4,14] Given the large variation in how dietary patterns were characterized using novel methods, multidimensionality and potential synergistic relationships between dietary components were considered and presented in a range of ways, from latent classes to networks. Several studies incorporated dynamism into their consideration of dietary patterns, though in most cases this was through stratification, for example, by meal, rather than through direct use of novel methods.[43,44,46,59] In these cases, it was a combination of input variables, stratification by time, and the novel method that enabled explorations of dynamism.

The methods highlighted have a range of strengths and limitations for the characterization of dietary patterns. Methods that focused on the classification of distinct patterns allowed for the assessment of relationships between these patterns and health outcomes or other indicators of interest but explored the interrelationships between dietary components to a lesser degree.[36–38,41–43,48,50,51,54,57,58] Other methods were designed to consider synergistic relationships among dietary components, such as compositional data analysis, mutual information, or probabilistic graphical models, but required further analyses, such as the development of a score, to assess relationships with health outcomes.[45–47,49,52,55,59]

There are trade-offs between novel and traditional methods that should be considered when contemplating the most appropriate methods for a given study. Though potential benefits such as a greater ability to discern multidimensionality may be desirable, these must be weighed against the implications for interpretability and computational costs. The application of novel methods may not always yield insights beyond those gained from traditional approaches. For example, Biesbroek et al.[36] found that random forest models did not outperform reduced rank regression when examining associations of dietary patterns with coronary artery disease. Conversely, a study that was not included in this review because it first identified dietary patterns using a traditional method—principal component analysis—found that machine learning algorithms were better able to classify the identified dietary patterns according to cardiometabolic risk compared to traditional approaches.[78]

Several sociodemographic characteristics are indicators of systemic health inequity and have been shown to be associated with dietary patterns among populations.[79–81] The degree to which studies incorporated sociodemographic characteristics into their consideration of dietary patterns or relationships between dietary patterns and health varied, with adjusted regression models applied after dietary patterns were characterized as the most common approach. Consistent with nutrition research more broadly,[81–83] there was little consideration of possible interactions among sociodemographic characteristics in relation to dietary patterns. Methods particularly suited to pattern recognition and complexity could be leveraged to simultaneously explore potential joint relationships among facets of social identity and dietary patterns[84] and advance our understanding of how broader systems of oppression and intersecting characteristics contribute to dietary patterns.[80]

Beyond the inclusion of sociodemographic characteristics in models, considering equity from the beginning of study design is a critical consideration given potential bias in data and algorithms that can have immense implications for those who already experience inequities

because of factors such as structural racism.[85–87] The included studies did not explicitly discuss the incorporation of equity into study design, and many conducted secondary analyses of existing datasets. The use of directed acyclic graphs has been identified as a potential solution to mitigate some possible issues with bias through careful model design[85] and has been applied in other domains of nutrition research using novel methods.[14] Engaging individuals with lived experience and the integration of interdisciplinary teams with broad expertise that can combine content knowledge with data-driven approaches can help to mitigate potential bias in algorithms.[84]

The level of description of methods varied and it was sometimes challenging to decipher the specifics of how novel methods were applied. Although the Strengthening the Reporting of Observational Studies in Epidemiology—Nutritional Epidemiology (STROBE-nut) reporting guidelines provide guidance for transparently reporting nutritional epidemiology and dietary assessment research,[88] it was not designed specifically for the methods used in the studies considered in this review and the ways in which they are being applied in dietary patterns research. Other reporting guidelines, such as the Consolidated Standards of Reporting Trials (CONSORT), have been extended to consider the application of artificial intelligence (AI).[89] Motivations related to the extension of CONSORT included inadequate reporting of studies using AI and the lack of full consideration of potential sources of bias specific to AI within existing reporting guidelines.[89] Relevant items added to CONSORT-AI pertain to the role of AI in the study, the nature of the data used in AI systems, and how humans interacted with AI systems, for example.[89] The extension of reporting guidelines such as STROBE-nut to consider applications of AI, including machine learning, and other methods that are becoming more commonly used, may facilitate consistent and complete reporting and improved comparability of studies. Reporting guidelines should continue to emphasize strategies applied to mitigate measurement error in dietary intake data,[88] as studies using novel methods are not immune to the effects of error on findings.[90] Along with reporting guidelines, the development of tailored

quality appraisal tools may facilitate synthesis of high-quality evidence to inform recommendations about dietary patterns and health.

This review provides a snapshot of a rapidly evolving field,[91,92] with the involvement of an interdisciplinary team of researchers lending to a robust consideration of emerging methods in dietary patterns research. While prior reviews have provided perspectives on the potential applications of machine learning within the field of nutrition,[19–21] this review considered dietary patterns in particular, as well as considering approaches beyond machine learning that have not traditionally been used in this area, broadening the scope compared to prior reviews.[22,93] The search terms were informed by preliminary searching, though it is unlikely that all relevant articles applying novel methods to characterize dietary patterns were captured. This is partially driven by the wide range of descriptors used for these methods and the lack of reporting standards. As well, determining whether a method is novel is somewhat subjective. Methods such as factor analysis and principal component analysis once revolutionized dietary pattern analysis, providing data-driven approaches to identify patterns.[15] Now, they are widely applied and recognized as limited in their capabilities to capture complexity compared to some newer approaches. Further, the search terms skewed toward multidimensionality versus dynamism, potentially overlooking some studies focusing on variation of dietary patterns over time or across eating occasions. Nonetheless, this review documents an acceleration of the application of a range of novel methods to dietary patterns research and captures a broad scope of methods being used to characterize these patterns, highlighting the need for researchers to develop the lexicon and knowledge needed to interpret the emerging literature.

6.6. Conclusion

The findings of this review indicate a strong motivation to apply novel methods, including but not limited to machine learning, to improve understanding of dietary patterns and

how they relate to health and other factors. The application of these methods may help us to learn about complex relationships that may not be possible to discern through traditional approaches. However, these methods may not be suitable for every question and do not necessarily overcome the limitations of more traditional approaches.

Given the proliferation of these methods, it is becoming increasingly worthwhile for nutrition researchers to have at least a basic understanding of novel methods such as machine learning and latent class analysis, so they can interpret the results of emerging studies. The development and implementation of reporting guidelines and quality appraisal mechanisms for studies that apply novel methods may improve the capacity for synthesis of evidence generated to inform strategies that promote improved population health and well-being.

Conflict of Interest: RML is a statistical editor for the British Journal of Nutrition. Other authors have none to declare.

Funding: This review was funded by the Canadian Institutes of Health Research, a University of Waterloo Research Incentive Fund award, an Ontario Ministry of Research and Innovation Early Researcher Award held by SIK, and Microsoft AI for Good. JMH was funded by a Vanier Canada Graduate Scholarship (Social Sciences and Humanities Research Council). RML is funded by a National Health and Medical Research Council Emerging Leadership Fellowship (APP1175250). LMB was funded by the National Institutes of Health (R01 HD102313, MPI Bodnar LM, Naimi AI).

6.7. References

1. Afshin A, Sur PJ, Fay KA, et al. Health effects of dietary risks in 195 countries, 1990–2017: A systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*. 2019;393(10184):1958-1972. doi:10.1016/S0140-6736(19)30041-8

2. English LK, Ard JD, Bailey RL, et al. Evaluation of dietary patterns and all-cause mortality: A systematic review. *JAMA network open*. 2021;4(8). doi:10.1001/JAMANETWORKOPEN.2021.22277

3. Mozaffarian D, Rosenberg I, Uauy R. History of modern nutrition science—implications for current research, dietary guidelines, and food policy. *BMJ*. 2018;361. doi:10.1136/BMJ.K2392

4. Reedy J, Subar AF, George SM, Krebs-Smith SM. Extending methods in dietary patterns research. *Nutrients*. 2018;10(5). doi:10.3390/NU10050571

5. Schulz CA, Oluwagbemigun K, Nöthlings U. Advances in dietary pattern analysis in nutritional epidemiology. *Eur J Nutr*. 2021;60(8):4115-4130. doi:10.1007/s00394-021-02545-9

6. Herforth A, Arimond M, Álvarez-Sánchez C, Coates J, Christianson K, Muehlhoff E. A global review of food-based dietary guidelines. *Advances in Nutrition*. 2019;10(4):590. doi:10.1093/ADVANCES/NMY130

7. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Food and Nutrition Board; Committee to Review the Process to Update the Dietary Guidelines for Americans. *Redesigning the Process for Establishing the Dietary Guidelines for Americans*. National Academies Press; 2017. doi:10.17226/24883

8. Imamura F, Micha R, Khatibzadeh S, et al. Dietary quality among men and women in 187 countries in 1990 and 2010: a systematic assessment. *The Lancet Global Health*. 2015;3(3):e132-e142. doi:10.1016/S2214-109X(14)70381-X

9. Delormier T, Frohlich K, Potvin L. Food and eating as social practice – understanding eating patterns as social phenomena and implications for public health. *Sociolgy of Health & Illness*. 2009;31(2):215-228. doi:10.1111/j.1467-9566.2008.01128.x

10. Ocké MC. Evaluation of methodologies for assessing the overall diet: dietary quality scores and dietary pattern analysis. *Proceedings of the Nutrition Society*. 2013;72(2):191-199. doi:10.1017/S0029665113000013

11. Shams-White MM, Pannucci TE, Lerman JL, et al. Healthy Eating Index-2020: Review and update process to reflect the Dietary Guidelines for Americans, 2020-2025. *J Acad Nutr Diet*. 2023;123(9):1280-1288. doi:10.1016/j.jand.2023.05.015

12. Brassard D, Munene LAE, Pierre SS, et al. Development of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. *Applied Physiology, Nutrition, and Metabolism*. 2022;47(5):595-610. doi:10.1139/APNM-2021-0415

13. Kirkpatrick SI, Reedy J, Krebs-Smith SM, et al. Applications of the Healthy Eating Index for surveillance, epidemiology, and intervention research: Considerations and caveats. *Journal of the Academy of Nutrition and Dietetics*. 2018;118(9):1603. doi:10.1016/J.JAND.2018.05.020

14. Bodnar LM, Cartus AR, Kirkpatrick SI, et al. Machine learning as a strategy to account for dietary synergy: an illustration based on dietary intake and adverse pregnancy outcomes. *The American Journal of Clinical Nutrition*. 2020;111(6):1235-1243. doi:10.1093/AJCN/NQAA027

15. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Current Opinion in Lipidology*. 2002;13(1):3-9. doi:10.1097/00041433-200202000-00002

16. Michels KB, Schulze MB. Can dietary patterns help us detect diet-disease associations? *Nutr Res Rev*. 2005;18(2):241-248. doi:10.1079/NRR2005107

17. Reedy J, Krebs-Smith SM, Hammond RA, Hennessy E. Advancing the science of dietary patterns research to leverage a complex systems approach. *Journal of the Academy of Nutrition and Dietetics*. 2017;117(7):1019-1022. doi:10.1016/J.JAND.2017.03.008

18. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Food and Nutrition Board; Alice Vorosmarti and Joe Alper, Rapporteurs. *The Role of Advanced Computation, Predictive Technologies, and Big Data Analytics in Food and Nutrition Research*. National Academies Press; 2024. Accessed May 10, 2024. https://www.nationalacademies.org/our-work/the-role-of-advanced-computation-predictive-technologies-and-big-data-analytics-in-research-related-to-food-and-nutrition-a-workshop-series

19. Kirk D, Kok E, Tufano M, Tekinerdogan B, Feskens EJM, Camps G. Machine learning in nutrition research. *Advances in Nutrition*. 2022;13(6):2573-2589. doi:10.1093/ADVANCES/NMAC103

20. Morgenstern JD, Rosella LC, Costa AP, Souza RJ de, Anderson LN. Perspective: Big data and machine learning could help advance nutritional epidemiology. *Advances in Nutrition*. 2021;12(3):621-631. doi:10.1093/ADVANCES/NMAA183

21. Côté M, Lamarche B. Artificial intelligence in nutrition research: perspectives on current and future applications. *Appl Physiol Nutr Metab*. Published online 2021:1-8. doi:10.1139/apnm-2021-0448

22. Oliveira Chaves L, Gomes Domingos AL, Louzada Fernandes D, Ribeiro Cerqueira F, Siqueira-Batista R, Bressan J. Applicability of machine learning techniques in food intake assessment: A systematic review. *Critical Reviews in Food Science and Nutrition*. 2023;63(7):902-919. doi:10.1080/10408398.2021.1956425

23. Le Glaz A, Haralambous Y, Kim-Dufor DH, et al. Machine learning and natural language processing in mental health: Systematic review. *J Med Internet Res*. 2021;23(5):e15708. doi:10.2196/15708

24. Payedimarri AB, Concina D, Portinale L, et al. Prediction models for public health containment measures on COVID-19 using artificial intelligence and machine learning: A

systematic review. *Int J Environ Res Public Health*. 2021;18(9):4499. doi:10.3390/ijerph18094499

25. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology*. 2019;19(1):64. doi:10.1186/s12874-019-0681-4

26. Morgenstern JD, Buajitti E, O'Neill M, et al. Predicting population health with machine learning: a scoping review. *BMJ Open*. 2020;10(10):e037860. doi:10.1136/BMJOPEN-2020-037860

27. Chapter 11: Scoping reviews. In: *JBI Manual for Evidence Synthesis*. JBI; 2020. doi:10.46658/JBIMES-20-12

28. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *International Journal of Social Research Methodology*. 2005;8(1):19-32. doi:10.1080/1364557032000119616

29. Tricco AC, Lillie E, Zarin W, et al. PRISMA Extension for Scoping Reviews (PRISMA-ScR): Checklist and Explanation. *Ann Intern Med*. 2018;169(7):467-473. doi:10.7326/M18-0850

30. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: A review. *Nutrition Reviews*. 2004;62(5):177-203. doi:10.1111/j.1753-4887.2004.tb00040.x

31. 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;132(9):790-793. doi:10.1161/CIRCULATIONAHA.115.018010

32. Covidence - Better systematic review management. Covidence. Accessed September 12, 2023. https://www.covidence.org/

33. Gwet KL. Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*. 2008;61(1):29-48. doi:10.1348/000711006X126600

34. Cicchetti DV, Feinstein AR. High agreement but low kappa: II. Resolving the paradoxes. *Journal of Clinical Epidemiology*. 1990;43(6):551-558. doi:10.1016/0895-4356(90)90159-M

35. Belur J, Tompson L, Thornton A, Simon M. Interrater reliability in systematic review methodology: Exploring variation in coder decision-making. *Sociological Methods & Research*. 2021;50(2):837-865. doi:10.1177/0049124118799372

36. Biesbroek S, van der A DL, Brosens MCC, et al. Identifying cardiovascular risk factor-related dietary patterns with reduced rank regression and random forest in the EPIC-NL cohort. *Am J Clin Nutr*. 2015;102(1):146-154. doi:10.3945/ajcn.114.092288

37. Fonseca MJ, Gaio R, Lopes C, Santos AC. Association between dietary patterns and metabolic syndrome in a sample of Portuguese adults. *Nutr J*. 2012;11:64. doi:10.1186/1475-2891-11-64

38. Jamiołkowski J, Szpak A, Pawłowska D. Dietary habits of men from Podlasie region of Poland in the years 1987-1998 analysed with self-organizing neural networks. *Rocz Akad Med Bialymst*. 2005;50 Suppl 1:220-224.

39. Oliveira A, Rodríguez-Artalejo F, Gaio R, Santos AC, Ramos E, Lopes C. Major habitual dietary patterns are associated with acute myocardial infarction and cardiovascular risk markers in a southern European population. *J Am Diet Assoc*. 2011;111(2):241-250. doi:10.1016/j.jada.2010.10.042

40. Wu Y, Sánchez BN, Goodrich JM, et al. Dietary exposures, epigenetics and pubertal tempo. *Environ Epigenet*. 2019;5(1):dvz002. doi:10.1093/eep/dvz002

41. Affret A, Severi G, Dow C, et al. Socio-economic factors associated with a healthy diet: results from the E3N study. *Public Health Nutr*. 2017;20(9):1574-1583. doi:10.1017/S1368980017000222

42. Bezerra IN, Bahamonde NMSG, Marchioni DML, et al. Generational differences in dietary pattern among Brazilian adults born between 1934 and 1975: a latent class analysis. *Public Health Nutr*. 2018;21(16):2929-2940. doi:10.1017/S136898001800191X

43. Harrington JM, Dahly DL, Fitzgerald AP, Gilthorpe MS, Perry IJ. Capturing changes in dietary patterns among older adults: a latent class analysis of an ageing Irish cohort. *Public Health Nutr*. 2014;17(12):2674-2686. doi:10.1017/S1368980014000111

44. Hearty AP, Gibney MJ. Analysis of meal patterns with the use of supervised data mining techniques--artificial neural networks and decision trees. *Am J Clin Nutr*. 2008;88(6):1632-1642. doi:10.3945/ajcn.2008.26619

45. Iqbal K, Buijsse B, Wirth J, Schulze MB, Floegel A, Boeing H. Gaussian graphical models identify networks of dietary intake in a German adult population. *The Journal of Nutrition*. 2016;146(3):646-652. doi:10.3945/JN.115.221135

46. Schwedhelm C, Knüppel S, Schwingshackl L, Boeing H, Iqbal K. Meal and habitual dietary networks identified through Semiparametric Gaussian Copula Graphical Models in a German adult population. *PLoS One*. 2018;13(8):e0202936. doi:10.1371/journal.pone.0202936

47. Solans M, Coenders G, Marcos-Gragera R, et al. Compositional analysis of dietary patterns. *Stat Methods Med Res*. 2019;28(9):2834-2847. doi:10.1177/0962280218790110

48. Dalmartello M, Decarli A, Ferraroni M, et al. Dietary patterns and oral and pharyngeal cancer using latent class analysis. *Int J Cancer*. 2020;147(3):719-727. doi:10.1002/ijc.32769

49. Hoang T, Lee J, Kim J. Network analysis of demographics, dietary intake, and comorbidity interactions. *Nutrients*. 2021;13(10):3563. doi:10.3390/NU13103563/S1

50. Hose AJ, Pagani G, Karvonen AM, et al. Excessive unbalanced meat consumption in the first year of life increases asthma risk in the PASTURE and LUKAS2 birth cohorts. *Front Immunol*. 2021;12:651709. doi:10.3389/fimmu.2021.651709

51. Madeira T, Severo M, Oliveira A, Gorjão Clara J, Lopes C. The association between dietary patterns and nutritional status in community-dwelling older adults-the PEN-3S study. *Eur J Clin Nutr*. 2021;75(3):521-530. doi:10.1038/s41430-020-00745-w

52. Samieri C, Sonawane AR, Lefèvre-Arbogast S, Helmer C, Grodstein F, Glass K. Using network science tools to identify novel diet patterns in prodromal dementia. *Neurology*. 2020;94(19):e2014-e2025. doi:10.1212/WNL.0000000000009399

53. Shang X, Li Y, Xu H, et al. Leading dietary determinants identified using machine learning techniques and a healthy diet score for changes in cardiometabolic risk factors in children: a longitudinal analysis. *Nutr J*. 2020;19(1):105. doi:10.1186/s12937-020-00611-2

54. Wright DM, McKenna G, Nugent A, Winning L, Linden GJ, Woodside JV. Association between diet and periodontitis: a cross-sectional study of 10,000 NHANES participants. *Am J Clin Nutr*. 2020;112(6):1485-1491. doi:10.1093/ajcn/nqaa266

55. Xia Y, Zhao Z, Zhang S, et al. Complex dietary topologies in non-alcoholic fatty liver disease: A network science analysis. *Front Nutr*. 2020;7:579086. doi:10.3389/fnut.2020.579086

56. Zhao Y, Naumova EN, Bobb JF, Claus Henn B, Singh GM. Joint associations of multiple dietary components with cardiovascular disease risk: A machine-learning approach. *Am J Epidemiol*. 2021;190(7):1353-1365. doi:10.1093/aje/kwab004

57. de Almeida Alves M, Molina MDCB, da Fonseca M de JM, Lotufo PA, Benseñor IM, Marchioni DML. Different statistical methods identify similar population-specific dietary patterns: an analysis of Longitudinal Study of Adult Health (ELSA-Brasil). *Br J Nutr*. 2022;128(11):2249-2257. doi:10.1017/S0007114522000253

58. Farmer N, Lee LJ, Powell-Wiley TM, Wallen GR. Cooking frequency and perception of diet among US adults are associated with US healthy and healthy Mediterranean-style dietary related classes: A latent class profile analysis. *Nutrients*. 2020;12(11):3268. doi:10.3390/nu12113268

59. Schwedhelm C, Lipsky LM, Shearrer GE, et al. Using food network analysis to understand meal patterns in pregnant women with high and low diet quality. *International Journal of Behavioral Nutrition and Physical Activity*. 2021;18(1):101. doi:10.1186/s12966-021-01172-1

60. Breiman L. Random Forests. *Machine Learning*. 2001;45(1):5-32. doi:10.1023/A:1010933404324

61. Song Y yan, Lu Y. Decision tree methods: applications for classification and prediction. *Shanghai Arch Psychiatry*. 2015;27(2):130-135. doi:10.11919/j.issn.1002-0829.215044

62. Grossi E, Buscema M. Introduction to artificial neural networks. *European Journal of Gastroenterology & Hepatology*. 2007;19(12):1046. doi:10.1097/MEG.0b013e3282f198a0

63. Bianchi D, Calogero R, Tirozzi B. Kohonen neural networks and genetic classification. *Mathematical and Computer Modelling*. 2007;45(1):34-60. doi:10.1016/j.mcm.2006.04.004

64. Drton M, Maathuis MH. Structure learning in graphical modeling. *https://doi.org/101146/annurev-statistics-060116-053803*. 2017;4:365-393. doi:10.1146/ANNUREV-STATISTICS-060116-053803

65. Altenbuchinger M, Weihs A, Quackenbush J, Grabe HJ, Zacharias HU. Gaussian and mixed graphical models as (multi-)omics data analysis tools. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2020;1863(6):194418. doi:10.1016/J.BBAGRM.2019.194418

66. Koller D, Friedman N. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press; 2009.

67. Liu H, Han F, Yuan M, Lafferty J, Wasserman L. High-dimensional semiparametric Gaussian copula graphical models. *The Annals of Statistics*. 2012;40(4):2293-2326.

68. Freund Y, Schapire RE. A decision-theoretic generalization of on-line learning and an application to boosting. *Journal of Computer and System Sciences*. 1997;55(1):119-139. doi:10.1006/jcss.1997.1504

69. Bobb JF, Valeri L, Claus Henn B, et al. Bayesian kernel machine regression for estimating the health effects of multi-pollutant mixtures. *Biostatistics*. 2015;16(3):493-508. doi:10.1093/biostatistics/kxu058

70. Aitchison J. The statistical analysis of compositional data. *Journal of the Royal Statistical Society Series B (Methodological)*. 1982;44(2):139-177.

71. Rodrigues LA, Daunis-i-Estadella P, Figueras G, Henestrosa S. Flying in compositional morphospaces: evolution of limb proportions in flying vertibrates. In: *Compositional Data Analysis: Theory and Applications*. ; 2011:235-254. doi:10.1002/9781119976462.ch17

72. Zhang CH, Huang J. The sparsity and bias of the Lasso selection in high-dimensional linear regression. *The Annals of Statistics*. 2008;36(4):1567-1594.

73. Oliveira A, Lopes C, Torres D, Ramos E, Severo M. Application of a latent transition model to estimate the usual prevalence of dietary patterns. *Nutrients*. 2021;13(1):133. doi:10.3390/nu13010133

74. Fahey MT, Thane CW, Bramwell GD, Coward WA. Conditional Gaussian mixture modelling for dietary pattern analysis. *Journal of the Royal Statistical Society Series A (Statistics in Society)*. 2007;170(1):149-166.

75. Cover T, Thomas J. *Elements of Information Theory*. Wiley; 2005. Accessed August 23, 2023. https://books-scholarsportal-info.proxy.lib.uwaterloo.ca/en/read?id=/ebooks/ebooks2/pda/2011-12-01/1/13568.9780471748816#page=1

76. Gorst-Rasmussen A, Dahm C, Dethlefsen C, Scheike T, Overvad K. Exploring dietary patterns by using the treelet transform. *Am J Epidemiol*. 2011;173:1097-1104. doi:10.1093/aje/kwr060

77. Lee AB, Nadler B, Wasserman L. Treelets—An adaptive multi-scale basis for sparse unordered data. *The Annals of Applied Statistics*. 2008;2(2):435-471. doi:10.1214/07-AOAS137

78. Panaretos D, Koloverou E, Dimopoulos AC, et al. A comparison of statistical and machine-learning techniques in evaluating the association between dietary patterns and 10-year cardiometabolic risk (2002-2012): the ATTICA study. *The British Journal of Nutrition*. 2018;120(3):326-334. doi:10.1017/S0007114518001150

79. Hanson KL, Connor LM. Food insecurity and dietary quality in US adults and children: A systematic review. *Am J Clin Nutr*. 2014;100(2):684-692. doi:10.3945/ajcn.114.084525

80. Doan N, Olstad DL, Vanderlee L, Hammond D, Wallace M, Kirkpatrick SI. Investigating the intersections of racial identity and perceived income adequacy in relation to dietary quality among adults in Canada. *The Journal of Nutrition*. 2022;13(152):67S-75S. doi:10.1093/JN/NXAC076

81. Hiza HAB, Casavale KO, Guenther PM, Davis CA. Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. *Journal of the Academy of Nutrition and Dietetics*. 2013;113(2):297-306. doi:10.1016/J.JAND.2012.08.011

82. Olstad DL, Nejatinamini S, Victorino C, Kirkpatrick SI, Minaker LM, McLaren L. Trends in socioeconomic inequities in diet quality between 2004 and 2015 among a nationally representative sample of children in Canada. *The Journal of Nutrition*. 2021;151(12):3781-3794. doi:10.1093/JN/NXAB297

83. Brassard D, Munene LAE, Pierre SS, et al. Evaluation of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. *Applied Physiology, Nutrition, and Metabolism*. 2022;47(5). doi:10.1139/APNM-2021-0416

84. Wang A, Ramaswamy VV, Russakovsky O. Towards intersectionality in machine learning: Including more identities, handling underrepresentation, and performing evaluation. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*. FAccT '22. Association for Computing Machinery; 2022:336-349. doi:10.1145/3531146.3533101

85. Robinson WR, Renson A, Naimi AI. Teaching yourself about structural racism will improve your machine learning. *Biostatistics*. 2020;21(2):339-344. doi:10.1093/biostatistics/kxz040

86. Rojas JC, Fahrenbach J, Makhni S, et al. Framework for integrating equity into machine learning models: A case study. *Chest*. 2022;161(6):1621-1627. doi:10.1016/j.chest.2022.02.001

87. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. *Annals of Internal Medicine*. 2018;169(12):866-872. doi:10.7326/M18-1990

88. Lachat C, Hawwash D, Ocké MC, et al. Strengthening the Reporting of Observational Studies in Epidemiology-Nutritional Epidemiology (STROBE-nut): An extension of the STROBE statement. *PLOS Medicine*. 2016;13(6). doi:10.1371/JOURNAL.PMED.1002036

89. Liu X, Cruz Rivera S, Moher D, et al. Reporting guidelines for clinical trial reports for interventions involving artificial intelligence: the CONSORT-AI extension. *Nature Medicine*. 2020;26(9):1364-1374. doi:10.1038/s41591-020-1034-x

90. Spicker D, Nazemi A, Hutchinson J, et al. Challenges for predictive modeling with neural network techniques using error-prone dietary intake data. Published online November 15, 2023. doi:10.48550/arXiv.2311.09338

91. Lampignano L, Tatoli R, Donghia R, et al. Nutritional patterns as machine learning predictors of liver health in a population of elderly subjects. *Nutr Metab Cardiovasc Dis*. 2023;33(11):2233-2241. doi:10.1016/j.numecd.2023.07.009

92. Slurink IA, Corpeleijn E, Bakker SJ, et al. Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines Study. *Am J Clin Nutr*. 2023;118(6):1077-1090. doi:10.1016/j.ajcnut.2023.10.002

93. Kirk D, Catal C, Tekinerdogan B. Precision nutrition: A systematic literature review. *Computers in Biology and Medicine*. 2021;133:104365. doi:10.1016/j.compbiomed.2021.104365

Chapter 7: Multidimensional dietary patterns and their joint associations with intersecting sociodemographic characteristics among adults in Canada

Status: Prepared for submission to *Advances in Nutrition*.

Authors: Joy M. Hutchinson[1], Dylan Spicker[2], Benoît Lamarche[3,4], Michael Wallace[5], Mélina Coté[3,4], Abel Torres-Espín[1], Sharon I. Kirkpatrick[1]

1. School of Public Health Sciences, University of Waterloo, Waterloo, ON, Canada
2. Department of Mathematics and Statistics, University of New Brunswick, NB, Canada
3. Centre Nutrition, santé et société (NUTRISS), Institut sur la nutrition et les aliments fonctionnels (INAF), Université Laval, QC, Canada
4. École de nutrition, Faculté des sciences de l'agriculture et de l'alimentation, Université Laval, Québec, QC, Canada
5. Department of Statistics and Actuarial Science, University of Waterloo, Waterloo, ON, Canada

Acknowledgements: The authors are grateful to the members of a working group that focused on the applicability of AI and related methods to expand understanding of dietary patterns, including Kevin W. Dodd, Amir Nazemi, Paul Fieguth, Jennifer Vena, Megan Deitchler, and Navreet Singh.

7.1. Overview

Dietary patterns, comprised of multidimensional components that have synergistic and antagonistic relationships with each other, are a complex construct to assess. Individuals have numerous characteristics that form their social identities, and several sociodemographic characteristics are associated with dietary intake. However, associations between sociodemographic characteristics and dietary intake typically do not consider potential joint relationships. This study applied mixed graphical models to examine joint relationships between dietary patterns and sociodemographic characteristics among adults in Canada. We used a single 24-hour dietary recall from the 2015 Canadian Community Health Survey Nutrition among adults ≥18 years with nonzero energy intakes who were not missing data on key sociodemographic characteristics (n=14 097). Mixed graphical models included 30 log-transformed food groups (grams), sex, age, household food security status, income, employment status, education, geographic region, and smoking status. Three models were identified: a sociodemographic characteristic model, a diet-only model, and a full model that included all variables. Sociodemographic characteristics shared several joint relationships. Likewise, several dietary components were jointly related with each other. In the full model, age was associated with grains (other) (-0.12; 95% CI [-0.16, -0.09], coffee/tea (0.21; 95% CI [0.17, 0.24], and whole grains (0.12; 95% CI [0.08, 0.15]). Sex was associated with sweet beverages (0.11; 95% CI [0.06, 0.17]), alcohol (0.18; 95% CI [0.13, 0.24]), cured meat (0.20; 95% CI [0.15, 0.26]), and red meat (0.16; 95% CI [0.11, 0.21]). Smoking status was associated with whole grains (0.11; 95% CI [0.06, 0.17]) and education was associated with alcohol (0.10; 95% CI [0.04, 0.17]). Other sociodemographic characteristics were weakly associated with dietary components. This research identified interrelationships between dietary patterns and sociodemographic characteristics among adults in Canada, considering multidimensional relationships among dietary components and allowing for intersections between sociodemographic characteristics.

7.2. Introduction

While nutrition research has historically been rooted in the examination of single dietary components, interest has shifted from researching single nutrients to broader dietary patterns.[1,2] Humans typically do not eat nutrients or foods in isolation, but in combination as part of the total diet.[3] In general, the overall dietary pattern, with potentially multidimensional synergistic or antagonistic interactions among dietary components, contributes more strongly to the development of diet-related chronic disease than single foods or nutrients.[4–6] For example, evidence suggests that eating a variety of berries in combination may be more beneficial for the prevention of cancer than when consumed individually.[7] More broadly, understanding the combinations in which foods are consumed as part of population dietary patterns could help to illuminate relationships with health and lead to nutrition policies and dietary guidance that focus on improving overall dietary patterns.

Much dietary patterns research has relied on *a priori* approaches, using expert knowledge and researcher hypotheses to create indices and scores indicative of dietary patterns associated with health.[8–10] However, *a priori*-identified indices may not be fully representative of the range of potentially high quality dietary patterns among diverse populations.[11–13] Further, using *a priori* approaches, multidimensional inputs are typically compressed to a unidimensional composite score, which is readily interpretable but does not capture complex interactions among different dietary components within the pattern. *A posteriori* approaches have also provided opportunities to investigate dietary patterns and do not rely on researcher hypotheses.[3,14] For instance, multidimensional analysis methods such as principal component analysis and factor analysis have been employed to characterize combinations of key components that constitute dietary patterns.[14] These approaches are useful for gaining insights into the dietary patterns of different populations and population subgroups, but also typically

provide condensed scores (e.g., factor loadings) based on multidimensional inputs.[3,15] To date, the methods predominantly used in the literature limit examinations of dietary synergy.[2,16,17]

Not only are dietary patterns inherently complex in terms of the combinations of foods consumed, but they are also shaped by sociodemographic characteristics and other factors such as culture.[18–22] Differences in dietary intake have been observed in relation to several characteristics, including age, sex, income, education, household food security status, and smoking status.[18,20–24] However, most research has not considered the potential combined effects of intersecting sociodemographic characteristics and their relationships to diet. Sociodemographic characteristics may intersect in ways that are compounded by forces of oppression to shape the dietary patterns that people are able to consume, as postulated by intersectionality theory.[25–27] Research considering these intersections suggests that sociodemographic characteristics jointly contribute to dietary intake.[28] For example, a study of adults in Canada found that when racial identity and perceived income adequacy were considered separately, only perceived income adequacy meaningfully associated with diet quality.[28] However, when the intersection of these characteristics was included, an association with diet quality was observed, suggesting the joint consideration of these characteristics is relevant.[28] In the realm of precision nutrition, considering intersecting sociodemographic characteristics is important to inform strategies to improve dietary patterns.[29,30] Considering intersections among characteristics can also have important implications for population-level dietary recommendations, such as the development of appropriate, tailored dietary guidance.

Undirected probabilistic graphical models offer the opportunity to examine how individual dietary components are interrelated with one another and with other factors.[31,32] Undirected graphical models identify joint probability distributions in which the absence of a link between variables means that those variables are independent conditional to other variables, and a link represents associations that remain unexplained after accounting for all

other variables. [33–35] Probabilistic graphical models are well-suited to investigating multidimensionality because they can identify conditional relationships between the foods and food groupings that comprise high-dimensional dietary patterns.[31,32,35–40] In other words, probabilistic graphical models allow for the examination of joint relationships conditioned on, or accounting for, all other variables in the model, whereas traditional approaches may not control for the effect of all other variables.[34,36,39] These joint relationships form patterns of components by food groups and allow for an exploration of the internal structure of dietary patterns.[39] An added benefit of this modeling approach is that results are presented using interpretable graphical representations, easing knowledge mobilization with diverse stakeholders, including policymakers.

Probabilistic graphical models can be extended to incorporate categorical variables, enabling explorations of joint relationships among mixed variable types.[34,41] A growing number of studies have used mixed graphical models to jointly model dietary components, sociodemographic characteristics, and health characteristics.[31,32,42] For example, Hoang et al.[31] applied mixed graphical models to explore joint relationships among 16 dietary components, demographic characteristics, and health parameters. When dietary components were condensed to a single dietary score, no relationships with demographic or health characteristics were observed. However, when components were considered separately, joint relationships were observed between dietary components and other characteristics.[31] Given the growing recognition of the need to account for dietary heterogeneity among populations,[43,44] it is particularly relevant to explore the joint relationships between intersecting sociodemographic characteristics and high-dimensional dietary patterns. The objective of this study therefore was to explore dietary patterns and their joint associations with sociodemographic characteristics using mixed graphical models.

7.3. Methods

Data source and analytic sample:

Data were drawn from the Public Use Microdata Files for the 2015 Canadian Community Health Survey (CCHS) Nutrition, which contains information on a range of sociodemographic characteristics and up to two 24-hour dietary recalls (24HDRs) per participant (first recall n=20 487, second recall n=7623). The 2015 CCHS Nutrition was conducted by Statistics Canada and Health Canada from January to December 2015 and is representative of individuals living in the 10 provinces in Canada.[45] The survey included individuals aged one year and older who lived in private dwellings, with sampling based on age, sex, geographic area, and socioeconomic status.[45] Individuals who lived in the territories, who were full-time members of the Canadian Forces, those living on reserves and in some remote areas, and those living in institutions were not included.[45] The analytic sample (n=14 097) consisted of adults aged 18 years and older who completed the first 24HDR, excluding those ≤17 years because the dietary patterns of children and youth were considered to be distinct from adults.[46] Those with zero energy intakes and/or missing data on characteristics of interest including sex, age, household food security status, income, employment status, education, and smoking status were also excluded from analysis.

Because publicly available data were used, Ethical Standards Disclosure was not applicable.

Sociodemographic characteristics:

Sociodemographic characteristics of interest were selected because they have been shown to be related to dietary intake, individually or in some combination, in analyses focused on individual dietary components or using diet quality indices.[21,22,24,47–51] A directed acyclic graph was created to explore which variables to include and their hypothesized directions with dietary patterns (**Appendix L**), while also considering what was available for analysis in the

public-use version of the 2015 CCHS Nutrition.[52,53] Because information on gender identity was not available, biological sex was included as a binary variable, with responses including male or female. Age in years was included as a continuous variable. Total household income was expressed using a categorical variable based on reported income, with categories of $0-19,999, $20,000-$39,999, $40,000-59,999, 60,000-79,999, $80,000-99,999, $100,000-119,000, $120,000-139,999, and $140,000 and higher. Household food security status was expressed using a categorical variable, determined based on affirmative responses to the Household Food Security Survey Module[54] and coded using Health Canada's scoring thresholds. Categorizations of food security status included living in households that were food secure, moderately food insecure, or severely food insecure.[55] Respondent education was included as a four-level categorical variable, with categorizations of less than a high school diploma or its equivalent, a high school diploma or its equivalent, a certificate or diploma at the trade, college, or below bachelor's level, and a bachelor's degree or higher. Employment status over the past week was condensed from a four-level variable to a binary variable reflecting those who did or did not work in the previous week. Geographic regions included the Maritimes, Québec, Ontario, the Prairies, and British Columbia. Smoking status was assessed as whether participants currently smoked cigarettes daily, occasionally, or not at all.

24-hour dietary recalls:

Dietary intake data were collected by trained interviewers using the Automated Multiple-Pass Method, originally designed by the United States Department of Agriculture[56–58] and adapted to the Canadian context by Health Canada.[45] The Automated Multiple-Pass Method probes for commonly forgotten foods and beverages to minimize misreporting. Data from the first 24HDR for each participant were used. Foods and beverages reported were assigned food codes using the 2015 Canadian Nutrient File and a recipe file[45,59] and categorized into 156 food groups—referred to as the Bureau of Nutritional Sciences food groups—by Health Canada.[60]

Consistent with similar research,[31,37] the number of food groups was reduced to 30 for analysis by condensing groups for which there was a high frequency of zero intakes and that included similar foods (e.g., potato chips, popcorn, and other salty snacks were grouped into a salty snack group) (**Appendix M**). Food group variables were expressed in grams. To account for the skewed distribution typically observed in nutrition data and better approximate normality, a log-transformation was applied to dietary variables.[61,62]

Analysis:

Network analysis:

Analyses were conducted using R and RStudio (RStudio, Boston, MA, 2020).[63] Mixed graphical models can be used to examine interrelationships between continuous and categorical variables, simultaneously modeling variables that follow a variety of distributions, including Gaussian and multinomial.[41,64] The models used linear regression for Gaussian variables (dietary intake variables and age) and penalized multinomial logistic regression for categorical variables (the remaining sociodemographic variables).[34,41] Variables, or nodes, are connected by edges, which represent dependencies between nodes, conditioned on all other variables in the network.[35,36] The absence of an edge represents conditional independence between nodes, conditioned on all other variables in the network.[35,36] The width of edges is proportionate to the strength of the edge parameter.[34] Directionality can be assessed between continuous nodes, where positive or negative linear relationships are inferred, but the direction of relationships is undefined for relationships involving a categorical node.[34] The 'mgm' package was used to model dependencies.[34,65]

To improve interpretability, a 'least absolute shrinkage and selection operator' (LASSO) regularization parameter was used to perform variable selection and generate a sparse network by allowing some coefficients to be exactly equal to zero.[65–67] To further improve interpretation of networks by not displaying very weak non-zero relationships, edges with weights below 0.05

were not shown.[68] Ten-fold cross-validation was applied to determine the regularization parameter. To assess the stability of the network, bootstrapping with 2500 replication weights was used with the 'bootnet' package.[69] The standard (root) estimates and bootstrapped confidence intervals are presented. Models were visualized using the 'qgraph' package based on the standard estimates.[68]

Variables in the models included the 30 food groups and the sociodemographic characteristics of interest. Three networks were modeled: network one included sociodemographic variables only, network two included dietary components only, and network three combined the dietary components and sociodemographic characteristics to explore joint relationships. To facilitate interpretation of the networks, variables were assigned color-coded group-level categories that included sociodemographic characteristics, fruit and vegetables, grains, protein foods, foods to limit, and other foods/beverages. Network stability was measured using the centrality stability coefficient, which indicates whether the ranking of nodes remain consistent when bootstrapping.[70] A threshold of 0.7 or higher was used to assess whether the network was suitably stable.[70]

Survey weights from the 2015 CCHS Nutrition were incorporated to improve population representativeness.[45]

7.4. Results

Sample characteristics:

Characteristics of the sample are presented in **Table 12**. The mean age was 48.5 (± 17.5) years and half of participants identified as female. About a quarter (28.1%) of participants had a bachelor's degree or higher and 60.1% worked in the previous week. Mean one-day intakes of food groups are presented in **Table 13**.

Table 12: Characteristics of adults in the 2015 Canadian Community Health Survey Nutrition

Characteristic	**Total (n=14,097)**
	Weighted %
Sex	
Male	49.2
Female	50.8
Age (mean, years) ± SD	48.5 (17.5)
Household food security status	
Food secure	92.2
Moderately food insecure	5.3
Severely food insecure	2.5
Income (household)	
$0-19,999	8.0
$20,000-39,999	16.7
$40,000-59,999	16.4
$60,000-79,999	14.1
$80,000-99,999	11.1
$100,000-119,999	10.3
$120,000-139,999	7.0
$>140,000	16.5
Education (respondent)	
Less than high school	11.7
High school diploma	26.7
Certificate/diploma below bachelor's	33.6
Bachelor's degree or higher	28.1
Worked in the past week	
Yes	60.1
No	39.9
Geographic region	
Maritimes	6.8
Québec	23.5
Ontario	38.9
Prairies	17.7
British Columbia	13.1
Smoking status	
Daily	13.7
Occasionally	4.4
Not at all	81.9

Table 13: Mean one-day intakes of 30 food groupings used to identify dietary pattern networks among adults in the 2015 Canadian Community Health Survey Nutrition

Food grouping*	Mean (grams) ± SD
Alcohol	141 ± 380.8
Coffee & tea	449.7 ± 454.1
Cured meats	22.9 ± 51.7
Higher fat dairy products	60.2 ± 121
Lower fat dairy products	157.8 ± 215.6
Eggs	25.4 ± 47.9
Fish and shellfish	19 ± 55.3
Other grains (refined or whole grain)	91.5 ± 119.3
Green vegetables	41.5 ± 70.8
Onions, garlic, leeks, green onion	14.2 ± 24.4
Orange vegetables	16.2 ± 49.5
Other foods/beverages	76 ± 188.2
Fruits (other)	97 ±155.7
Vegetables (other)	47.9 ± 82.4
Plant-based proteins/alternatives	35.2 ± 80.8
Potatoes	42 ± 89.1
Poultry	45.5 ± 86.9
Red meat	45.2 ± 77.9
Refined grains	63.5 ± 78.9
Salty snacks	8.9 ± 29.7
Saturated fats[a]	7.5 ± 13.2
Soups	44.3 ± 115.4
Sugars[b]	17.6 ± 28.2
Sweet beverages	178.5 ± 310.3
Sweet treats/desserts	32.7 ± 69.6
Tomatoes	44.5 ± 77.8
Tropical fruits	53.6 ± 87
Unsaturated fats	10.3 ± 13.3
Water	987.1 ± 902.6
Whole grains	40.9 ± 74.5

*Further details on food groupings are provided in Appendix M
[a]Represents food sources of saturated fat (i.e., butter)
[b]Represents food sources of sugars (i.e., jams/jellies/marmalade)

Network descriptions:

Appendix N contains the weighted adjacency matrix detailing weights of relationships between nodes for the three networks.

Sociodemographic characteristic network:

The first network considered only the sociodemographic characteristics (**Figure 6**). Sex shared an edge with household food security status (0.15; 95% CI [0.05, 0.25]), working in the

past week (0.20; 95% CI [0.12, 0.29]) and smoking status (0.14; 95% CI [0.07, 0.22]) and was weakly related with other variables (<0.1). Geographic region shared edges with household food insecurity (0.14; 95% CI [0.04, 0.24]), income (0.12; 95% CI [0.04, 0.20]), and education (0.16; 95% CI [0.06, 0.25]), and was weakly associated with other variables (<0.1). The strongest pairwise relationships were observed between household food security status and income (1.10; 95% CI [0.86, 1.32]), between income and working in the past week (0.67; 95% CI [0.54, 0.80]), and age and working in the past week (0.53; 95% CI [0.47, 0.59]). Income shared edges ≥0.1 with age (0.17; 95% CI [0.10, 0.23]) and education (0.40; 95% CI [0.28, 0.50]). Household food security status also shared edges ≥0.1 with age (0.34; 95% CI [0.26, 0.43]), smoking status (0.37; 95% CI [0.23, 0.51]), education (0.14; 95% CI [0.02, 0.27]), and working in the past week (0.11; 95% CI [0.00, 0.22]). Further, working in the past week shared an edge with education (0.38; 95% CI [0.27, 0.50]) and smoking status (0.14; 95% CI [0.05, 0.24]). Smoking status additionally shared an edge ≥0.1 education (0.29; 95% CI [0.17, 0.42]) and age (0.19; 95% CI [0.12, 0.26]). Age shared an additional edge with education (0.28; 95% CI [0.28, 0.35]).

Dietary pattern network:

The second network explored the joint relationships between dietary components (**Figure 7**). Several fruit and vegetables dietary components shared edges, with a positive linear relationship identified between tomatoes with onions and garlic (0.19; 95% CI [0.16, 0.23]). In turn, a positive linear relationship was observed between onions and garlic with soup (0.24; 95% CI [0.21, 0.28]) and with other vegetables (0.23; 95% CI [0.20, 0.27]). Both soup and other vegetables were positively associated with orange vegetables (0.17; 95% CI [0.14, 0.21], 0.13; 95% CI [0.09, 0.16] respectively). Orange vegetables and other vegetables were positively linearly related with green vegetables (0.12; 95% CI [0.08, 0.16], 0.10; 95% CI [0.07, 0.14] respectively). Tropical fruits and other fruits were also positively related (0.14; 95% CI [0.10, 0.18]).

Negative linear relationships were identified among other components. A network of negative linear relationships was identified between cured meat and fish (-0.13; 95% CI [-0.16, -0.10]) cured meat and poultry (-0.14; 95% CI [-0.17, -0.11]), poultry and fish (-0.18; 95% CI [-0.21, -0.14]), poultry and red meat (-0.29; 95% CI [-0.33, -0.26]), and fish and red meat (-0.16; 95% CI [-0.20, -0.12]). Refined grains and whole grains were also negatively related (-0.30; 95% CI [-0.33, -0.26]). Negative linear relationships were observed between coffee/tea and water (-0.12; 95% CI [-0.16, -0.09]), coffee/tea and sweet beverages (-0.14; 95% CI [-0.18, -0.11]), and water and sweet beverages (-0.12; 95% CI [-0.16, -0.09]).

Patterns were also observed outside of groupings of similar foods. Eggs were positively associated with unsaturated fats (0.26; 95% CI [0.23, 0.30]), and unsaturated fats were positively associated with potatoes (0.25; 95% CI [0.21, 0.28]) and grains (other) (0.21; 95% CI [0.18, 0.25]). Grains (other) and potatoes were negatively associated (-0.27; 95% CI [-0.30, -0.23]) and grains (other) and saturated fats were positively associated (0.22; 95% CI [0.19, 0.25]). A positive linear relationship was observed between saturated fats and sugars (0.19; 95% CI [0.15, 0.22]). Weak relationships <0.1 were observed among other sweet food groups (sugars, sweet beverages, sweet treats).

Dietary pattern and sociodemographic characteristic network:

The third network included both the food group and sociodemographic variables (**Figure 8**). Accounting for the dietary components, sociodemographic characteristics largely patterned together, with similar pairwise relationships compared to the network that explored only sociodemographic characteristics (Figure 6). Dietary components also maintained similar joint relationships as observed in the second model (Figure 7). For example, refined grains and whole grains were negatively linearly related (-0.29; 95% CI [-0.32, -0.25]) and a pattern among vegetable components was observed. Sociodemographic characteristics were associated with several dietary components, though edge weights were mostly <0.1. Age was negatively linearly

related with grains (other) (-0.12; 95% CI [-0.16, -0.09]) and positively related with coffee/tea (0.21; 95% CI [0.17, 0.24]) and whole grains (0.12; 95% CI [0.08, 0.15]). Sex shared edges with sweet beverages (0.11; 95% CI [0.06, 0.17]), alcohol (0.18; 95% CI [0.13, 0.24]), cured meat (0.20; 95% CI [0.15, 0.26]), and red meat (0.16; 95% CI [0.11, 0.21]). Smoking status shared an edge ≥0.1 with whole grains (0.11; 95% CI [0.06, 0.17]) and education shared an edge with alcohol (0.10; 95% CI [0.04, 0.17]). Weak pairwise relationships were observed between other sociodemographic characteristics and dietary components.

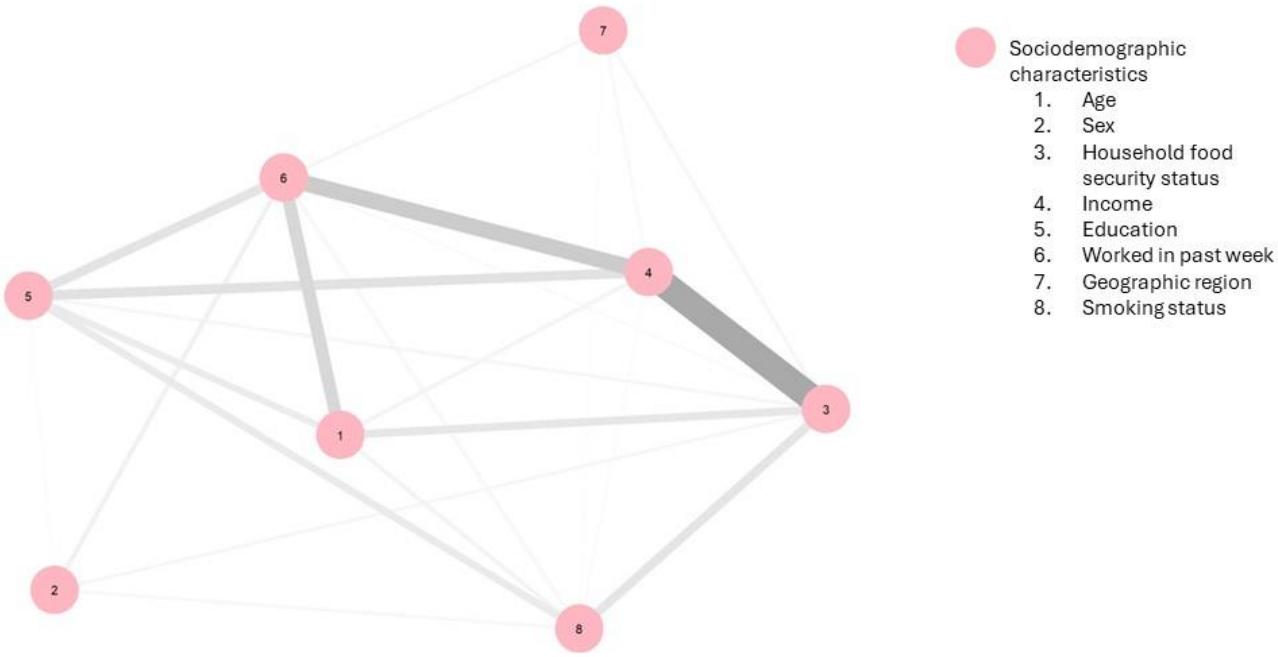


Figure 6. Network of sociodemographic characteristics among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition

Network illustrates the distribution of nodes conditional on all other nodes in the network. Green edges within the network represent positive interactions between continuous variables, red edges represent negative interactions between continuous variables, and grey edges indicate an undefined relationship and are shown when categorical variables are involved in the relationship.[34] Numbers assigned to variables do not indicate a ranking.

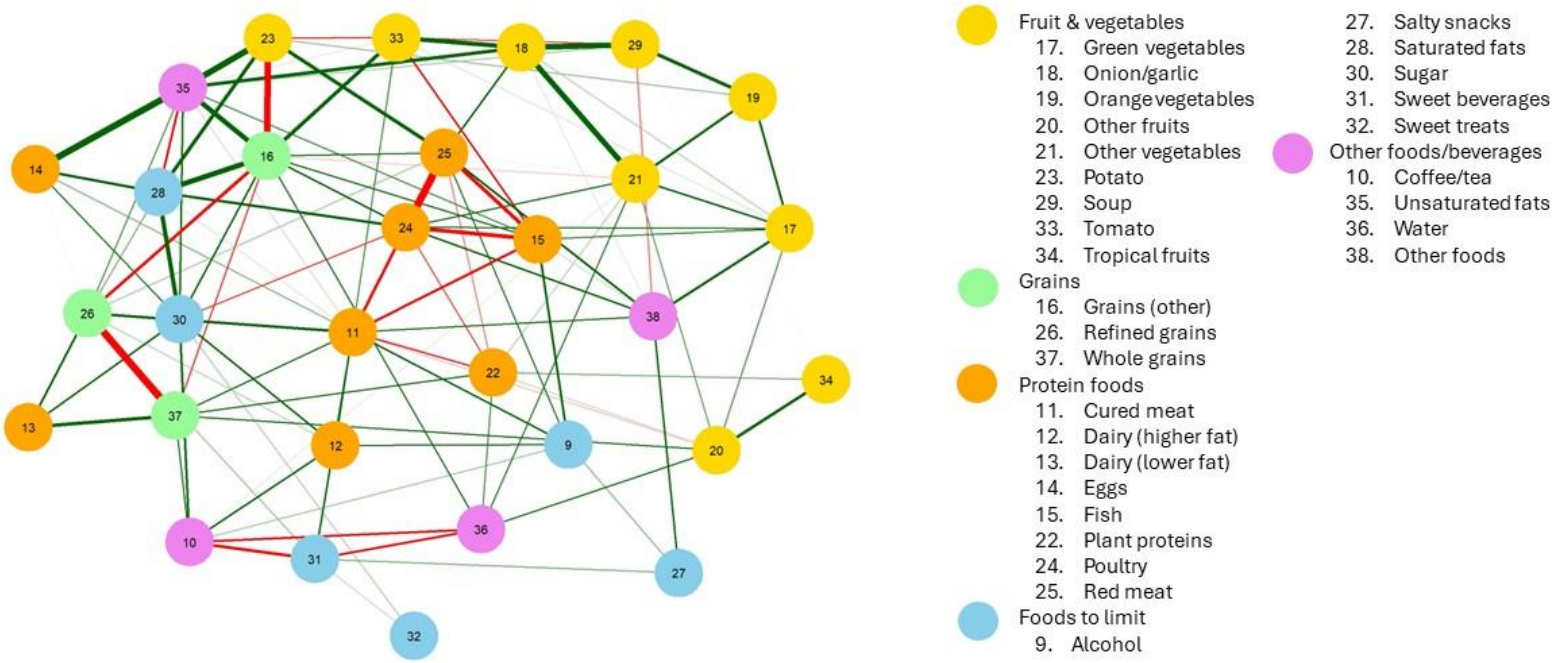


Figure 7. Network of dietary components among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition

Network illustrates the distribution of nodes conditional on all other nodes in the network. Green edges within the network represent positive interactions between continuous variables, red edges represent negative interactions between continuous variables, and grey edges indicate an undefined relationship and are shown when categorical variables are involved in the relationship.[34] Numbers assigned to variables do not indicate a ranking.

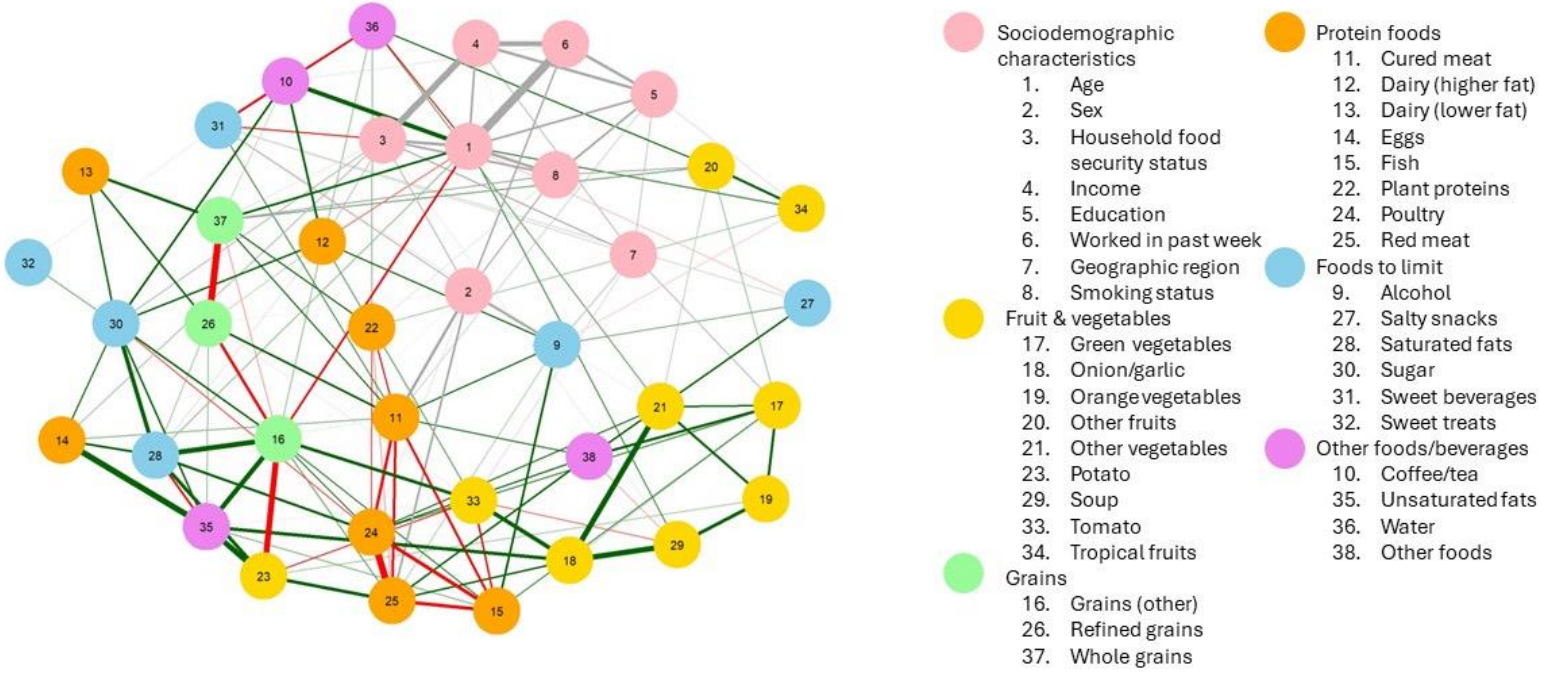


Figure 8. Full network of dietary components and sociodemographic characteristics among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition

Network illustrates the distribution of nodes conditional on all other nodes in the network. Green edges within the network represent positive interactions between continuous variables, red edges represent negative interactions between continuous variables, and grey edges indicate an undefined relationship and are shown when categorical variables are involved in the relationship.[34] Numbers assigned to variables do not indicate a ranking.

Network inference:

Network centrality indices for the full network that included sociodemographic characteristics and dietary components are presented in **Figure 9**. Household food insecurity and age were the strongest nodes in the model, followed by income and working in the past week. Nodes with less influence were salty snacks, sweet treats, and dairy (lower fat).

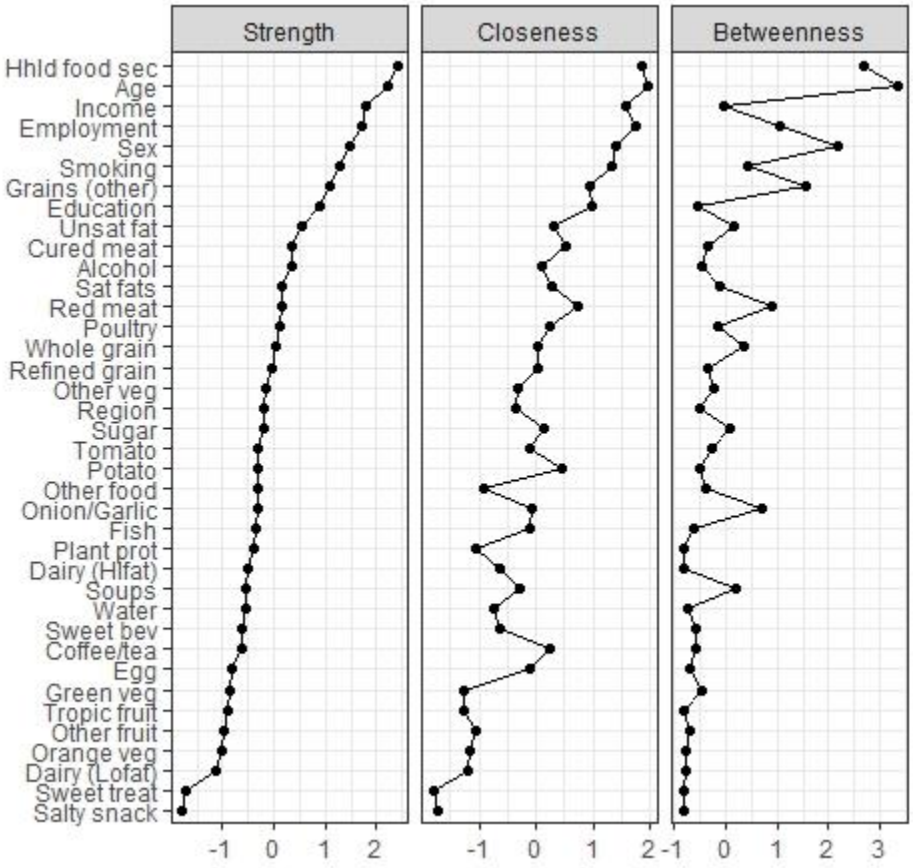


Figure 9. Centrality indices of variables included in the full network with sociodemographic and dietary components among adults ≥18 in the 2015 Canadian Community Health Survey Nutrition

Z scores are presented on the x axis, nodes are ordered by strength.

Network stability:

The central stability coefficient was above 0.7 for strength in the network that included sociodemographic characteristics and dietary components. Likewise, central stability coefficients were above 0.7 for the sociodemographic only network and the diet only network

(results not shown). Bootstrapped confidence intervals (CI) of estimated edge-weights for the full network are presented in **Appendix O**.

7.5. Discussion

This study explored two complex interrelated constructs, dietary patterns and intersecting sociodemographic characteristics, among adults in Canada. Mixed graphical models considered multidimensional relationships among dietary components and sociodemographic characteristics by conditioning each relationship between nodes on all other nodes in the network.[34] Models revealed patterns of dietary components and intersecting sociodemographic characteristics. The strongest pairwise relationships were observed among dietary components and among sociodemographic characteristics, rather than between dietary components and sociodemographic characteristics. For example, a strong edge was observed between income and household food security status, in line with associations observed through traditional approaches.[71] The key sociodemographic characteristics that had pairwise relationships with dietary components were age and sex.

Similar to Iqbal et al.,[39] who identified separate Gaussian graphical models for men and women in a German adult population, we identified positive linear relationships among vegetable groupings. We also identified a comparable negative linear relationship between whole grains and refined grains, which was observed in both the male and female networks in Iqbal et al.'s.[39] work. This negative linear relationship suggests that whole grains or refined grains displace each other within the dietary pattern after accounting for all other dietary components in the network, indicating a possible target for future nutrition policy or intervention. In contrast with prior research, we observed negative linear relationships among several protein foods, whereas Iqbal et al.[39] observed positive partial correlations between processed meat and red meat, and red meat and poultry. The negative linear relationships

observed among protein foods in this study suggests that individuals may rely on a primary protein food type, regardless of whether the protein food is recommended (i.e., fish, poultry) or suggested to limit (i.e., cured meat) in dietary guidance.

Negative linear relationships were also observed among each of the beverage groups examined in this study, suggesting potential displacement. The 2019 Canada's Food Guide recommends making 'Water your drink of choice',[72] aligning with the observed pattern of beverage consumption in this study, with water consumption negatively related to the consumption of other beverages, including sweet beverages and coffee. The observed negative linear relationship among beverage types suggests that the guidance to emphasize water is applicable to the Canadian population, given trade-offs with other beverage types that tend not to be recommended. The positive linear relationship observed between sugars and saturated fats is relevant in light of the guidance included in the 2019 Canada's Food Guide to limit foods that add excess sugars, sodium or saturated fats and suggests these dietary components are consumed in combination.[72] Of note, the foods captured in these groups are not highly processed, but are considered to be processed culinary ingredients (i.e., sugar, lard) according to NOVA.[48] Other studies have applied graphical models to understand meal patterning, exploring how multidimensional dietary patterns vary throughout the day.[37,38] The patterns observed by meal were overall not comparable to this analysis, likely in part due possible differential meal-specific patterns compared to our total day pattern and our inclusion of sociodemographic characteristics.[37,38]

Overall, our findings align with those of studies that have included sociodemographic characteristics and dietary components in mixed graphical models.[31,32] In agreement with the current findings, the strongest relationships identified by Hoang et al.[31] were within dietary component groups and within sociodemographic characteristic groups, rather than between sociodemographic characteristics and dietary components. Slurink et al.[32] included a range of

sociodemographic characteristics, dietary components, and health indicators in mixed graphical models and found several pairwise relationships among sociodemographic characteristics and several relationships among dietary components but relatively few relationships between sociodemographic characteristics and dietary components. Sex similarly had the strongest pairwise relationships with dietary components.[32] A possible explanation may be that the simultaneous inclusion of the selected sociodemographic characteristics may have muted possible observed individual relationships because the model conditioned each relationship on all other variables.

Some findings on relationships between dietary intake and sociodemographic characteristics are comparable with those from prior Canadian research that used traditional statistical approaches.[20–22,73] For example, Brassard et al.[21] observed differences in scores on the Healthy Eating Food Index-2019, which assesses alignment of dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations, by sex, age, and smoking status. Likewise, we observed pairwise relationships between sex and age and several dietary components, which were in turn related to several other dietary components. Evidence suggests dietary inequities in Canada in relation to household food security status, income, and education.[20,22,24,47,49] For example, using data from CCHS Nutrition, Olstad et al.[20] found that males and females with higher income and education had higher diet quality compared to those with lower income and education using multivariable linear regression. In contrast, the current study found few relationships between income or education and dietary components. We observed weak relationships between household food insecurity and several dietary components, with the strongest relationship with sweet beverages. However, strong relationships were observed between income and household food security and between income and working in the past week, consistent with the literature.[71,74]

As researchers look to novel methods, including probabilistic graphical models and machine learning approaches, to answer questions of increasing complexity, they must consider trade-offs with interpretability.[75] For example, while neural networks have excellent predictive capabilities, they have been criticized for being 'black box' models.[75,76] Within dietary patterns research, there is a tension between representing complexity and generating interpretable findings to inform public health nutrition policy. The adoption of novel methods within the field of nutrition is continuing to evolve,[30,76–79] and the balance of interpretability and complexity will continue to be an important consideration. Other types of network methods, such as mutual information, exponential random graphs, and neural networks, have been applied to assess intersecting relationships between sociodemographic characteristics and to characterize dietary patterns[25,80–83] and may be worth considering for future explorations of dietary patterns and sociodemographic characteristics. As the field progresses, extensions to existing quality appraisal guidelines will be essential to enable the synthesis of research using diverse analytic approaches that vary in complexity and interpretability.

There are instances in which an indication of overall diet quality, such as through the application of a diet quality index, may be sufficient. However, expanding our understanding of the internal structures of dietary patterns, including in groups that traditionally have been understudied, can potentially lead to more tailored policy and guidance. Future work using methods such as probabilistic graphical models could incorporate stratification by sociodemographic characteristics to explore further how multidimensional dietary patterns may be shaped by indicators of equity and oppression.

Although probabilistic graphical modeling is an *a posteriori* or data driven method rather than an *a priori* method and the algorithm itself is undirected/unsupervised, several stages in the process involved researcher insight and decisions. The inclusion of an interdisciplinary research team, with expertise in dietary assessment, statistics, and network

methods, helped to guide this decision-making process. Decisions included the selection of food groups and parameters to determine model sparsity. For example, food groupings could have been condensed further or differently, which may have changed the resulting dietary patterns and their associations with sociodemographic characteristics. We also could have applied Extended Bayesian Information Criterion (EBIC) to select the regularization parameter instead of cross-validation.[34] Extended Bayesian Information Criterion is typically more conservative than cross-validation,[34] and given our interest in learning potential relationships, a potentially less conservative regularization parameter was chosen to prioritize discovery and hypothesis generation. While networks hold appeal to address complexity in understanding dietary patterns, these models may not always match the question and may not consistently outperform traditional approaches.[78] However, as noted by Iqbal et al.,[39] graphical models have the potential to complement current methods to expand our understanding of multidimensional relationships, but findings require confirmation using traditional approaches.

Strength of this work were the use of a large, population-representative dataset with detailed dietary intake data and a range of sociodemographic variables, along with the involvement of an interdisciplinary team to guide the analysis and interpretation. Several caveats are also relevant to the interpretation of our findings. Although second recalls were available for a subset of the sample, we relied on data from the first recall and did not account for random within-person variation in intake or systematic error.[84] These sources of error may have distorted the observed relationships.[85] We considered a range of approaches to possibly mitigate the effects of the error, including incorporating energy intake in the model and including both available recalls. However, preliminary analysis (not shown) showed that energy was jointly correlated with nearly all dietary components, making it challenging to decipher other relationships with dietary components. This was also observed in another study applying mixed graphical models.[32] We were not able to include available recalls as separate observations in our estimation, as this would violate independence assumptions.[64] Future work could seek to

tailor probabilistic graphical models to the nuances of dietary intake data, including error, potentially expanding the insights to be gleaned from these approaches.[86] Such models will be applicable to many other exposures and outcomes within health that are measured with error.

Another possible limitation was the reliance on linear models, which may have missed possible complex non-linear relationships between variables of interest.[34] Further, several key population groups were not included in the 2015 CCHS Nutrition, including the territories and individuals living on reserves, limiting the representativeness of the findings.[45] Although other characteristics, such as racial and ethnic identity are relevant, we were limited by available characteristics in the Public Use Microdata Files. This study considered food groups, comprised of foods as they were eaten. However, future work could integrate key nutrients of interest, such as saturated fat, sugars, and sodium. In addition to sociodemographic characteristics, future analyses could be extended to include health outcomes using directed graphical models.

7.6. Conclusion

This research provides a lens into the internal structure of dietary patterns and how they are jointly related with sociodemographic characteristics among adults in Canada. By jointly modeling dietary components with sociodemographic characteristics, we were able to visualize these pairwise relationships in a network. The findings overall align with those of studies using traditional approaches, supporting the potential of these methods to explore dietary patterns while capturing greater complexity, such as demonstrating multidimensional relationships between dietary components.[39,87] Findings could be used to inform confirmatory research that models directions between relationships. In sum, mixed graphical models appear to be a useful tool to complement existing methods of dietary pattern analysis, offering the potential to learn internal structures of dietary patterns and other relevant variables of interest.

Conflicts of interest: None to declare.

Funding: This project was funded by an Ontario Ministry of Research and Innovation Early Researcher Award held by SIK. JMH was funded by a Vanier Canada Graduate Scholarship (Social Sciences and Humanities Research Council).

7.7. References

1. Mozaffarian D, Rosenberg I, Uauy R. History of modern nutrition science—implications for current research, dietary guidelines, and food policy. *BMJ*. 2018;361. doi:10.1136/BMJ.K2392

2. Reedy J, Subar AF, George SM, Krebs-Smith SM. Extending methods in dietary patterns research. *Nutrients*. 2018;10(5). doi:10.3390/NU10050571

3. Ocké MC. Evaluation of methodologies for assessing the overall diet: dietary quality scores and dietary pattern analysis. *Proceedings of the Nutrition Society*. 2013;72(2):191-199. doi:10.1017/S0029665113000013

4. World Cancer Research Fund/ American Institute for Cancer Research. *Diet, Nutrition, Physical Activity and Cancer: A Global Perspective*. Continuous Update Project Expert Report; 2018. http://gco.iarc.fr/today

5. 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(5):860-869. doi:10.1079/bjn20061731

6. Jacobs DR, Steffen LM. Nutrients, foods, and dietary patterns as exposures in research: a framework for food synergy. *The American Journal of Clinical Nutrition*. 2003;78(3):508S-513S. doi:10.1093/AJCN/78.3.508S

7. Zafra-Stone S, Yasmin T, Bagchi M, Chatterjee A, Vinson JA, Bagchi D. Berry anthocyanins as novel antioxidants in human health and disease prevention. *Mol Nutr Food Res*. 2007;51(6):675-683. doi:10.1002/mnfr.200700002

8. Rosato V, Temple NJ, La Vecchia C, Castellan G, Tavani A, Guercio V. Mediterranean diet and cardiovascular disease: a systematic review and meta-analysis of observational studies. *Eur J Nutr*. 2019;58(1):173-191. doi:10.1007/s00394-017-1582-0

9. Brassard D, Munene LAE, Pierre SS, et al. Development of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. *Applied Physiology, Nutrition, and Metabolism*. 2022;47(5):595-610. doi:10.1139/APNM-2021-0415

10. Shams-White MM, Pannucci TE, Lerman JL, et al. Healthy Eating Index-2020: Review and update process to reflect the Dietary Guidelines for Americans, 2020-2025. *J Acad Nutr Diet*. 2023;123(9):1280-1288. doi:10.1016/j.jand.2023.05.015

11. Burt K. The whiteness of the Mediterranean Diet: A historical, sociopolitical, and dietary analysis using Critical Race Theory. *Journal of Critical Dietetics*. 2021;5(2):41-52. doi:10.32920/cd.v5i2.1329

12. Olstad DL, Kirkpatrick SI. Planting seeds of change: reconceptualizing what people eat as eating practices and patterns. *Int J Behav Nutr Phys Act*. 2021;18(1):32. doi:10.1186/s12966-021-01102-1

13. Delormier T, Frohlich K, Potvin L. Food and eating as social practice – understanding eating patterns as social phenomena and implications for public health. *Sociolgy of Health & Illness*. 2009;31(2):215-228. doi:10.1111/j.1467-9566.2008.01128.x

14. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. *Current Opinion in Lipidology*. 2002;13(1):3-9. doi:10.1097/00041433-200202000-00002

15. Delisle H. Findings on dietary patterns in different groups of African origin undergoing nutrition transition. *Applied Physiology, Nutrition and Metabolism*. 2010;35(2):224-228. doi:10.1139/H10-008

16. Reedy J, Krebs-Smith SM, Hammond RA, Hennessy E. Advancing the science of dietary patterns research to leverage a complex systems approach. *Journal of the Academy of Nutrition and Dietetics*. 2017;117(7):1019-1022. doi:10.1016/J.JAND.2017.03.008

17. Bodnar LM, Cartus AR, Kirkpatrick SI, et al. Machine learning as a strategy to account for dietary synergy: an illustration based on dietary intake and adverse pregnancy outcomes. *The American Journal of Clinical Nutrition*. 2020;111(6):1235-1243. doi:10.1093/AJCN/NQAA027

18. Hiza HAB, Casavale KO, Guenther PM, Davis CA. Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. *Journal of the Academy of Nutrition and Dietetics*. 2013;113(2):297-306. doi:10.1016/J.JAND.2012.08.011

19. Hanson KL, Connor LM. Food insecurity and dietary quality in US adults and children: A systematic review. *Am J Clin Nutr*. 2014;100(2):684-692. doi:10.3945/ajcn.114.084525

20. Olstad DL, Nejatinamini S, Victorino C, Kirkpatrick SI, Minaker LM, McLaren L. Socioeconomic inequities in diet quality among a nationally representative sample of adults living in Canada: an analysis of trends between 2004 and 2015. *The American Journal of Clinical Nutrition*. 2021;114(5):1814-1829. doi:10.1093/AJCN/NQAB249

21. Brassard D, Munene LAE, Pierre SS, et al. Evaluation of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. *Applied Physiology, Nutrition, and Metabolism*. 2022;47(5). doi:10.1139/APNM-2021-0416

22. Hutchinson J, Tarasuk V. The relationship between diet quality and the severity of household food insecurity in Canada. *Public Health Nutrition*. 2022;25(4):1013-1026. doi:10.1017/S1368980021004031

23. Reedy J, Lerman JL, Krebs-Smith SM, et al. Evaluation of the Healthy Eating Index-2015. *Journal of the Academy of Nutrition and Dietetics*. 2018;118(9):1622-1633. doi:10.1016/J.JAND.2018.05.019

24. Tarasuk V, Fitzpatrick S, Ward H. Nutrition inequities in Canada. *Applied Physiology, Nutrition, and Metabolism*. 2010;35(2):172-179. doi:10.1139/H10-002

25. Wang A, Ramaswamy VV, Russakovsky O. Towards intersectionality in machine learning: Including more identities, handling underrepresentation, and performing evaluation. In: *Proceedings of the 2022 ACM Conference on Fairness, Accountability, and Transparency*.

FAccT '22. Association for Computing Machinery; 2022:336-349. doi:10.1145/3531146.3533101

26. Cho S, Williams Crenshaw K, McCall L. Toward a field of intersectionality studies: Theory, applications, and praxis. *Signs*. 2013;38(4):785-810. doi:https://doi.org/10.1086/669608

27. Crenshaw K. Demarginalizing the Intersection of Race and Sex: A Black Feminist Critique of Antidiscrimination Doctrine, Feminist Theory and Antiracist Politics. In: *Feminist Legal Theories*. Routledge; 1997.

28. Doan N, Olstad DL, Vanderlee L, Hammond D, Wallace M, Kirkpatrick SI. Investigating the intersections of racial identity and perceived income adequacy in relation to dietary quality among adults in Canada. *The Journal of Nutrition*. 2022;13(152):67S-75S. doi:10.1093/JN/NXAC076

29. Greenhill C. Towards precision nutrition. *Nat Rev Endocrinol*. 2020;16(9):473-473. doi:10.1038/s41574-020-0385-1

30. Kirk D, Catal C, Tekinerdogan B. Precision nutrition: A systematic literature review. *Computers in Biology and Medicine*. 2021;133:104365. doi:10.1016/j.compbiomed.2021.104365

31. Hoang T, Lee J, Kim J. Network analysis of demographics, dietary intake, and comorbidity interactions. *Nutrients*. 2021;13(10):3563. doi:10.3390/NU13103563/S1

32. Slurink IA, Corpeleijn E, Bakker SJ, et al. Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines Study. *Am J Clin Nutr*. 2023;118(6):1077-1090. doi:10.1016/j.ajcnut.2023.10.002

33. Keiding N. Book reviews: Graphical models. Steffen L. Lauritzen, Oxford University Press, 1996. No. of pages: 298. ISBN 0-19-852219-3. *Statistics in Medicine*. 1999;18(21):2983-2984. doi:10.1002/(SICI)1097-0258(19991115)18:21<2983::AID-SIM198>3.0.CO;2-A

34. Haslbeck JMB, Waldorp LJ. mgm: Estimating time-varying mixed graphical models in high-dimensional data. *Journal of Statistical Software*. 2020;93:1-46. doi:10.18637/jss.v093.i08

35. Koller D, Friedman N. *Probabilistic Graphical Models: Principles and Techniques*. MIT Press; 2009.

36. Drton M, Maathuis MH. Structure learning in graphical modeling. *https://doi.org/101146/annurev-statistics-060116-053803*. 2017;4:365-393. doi:10.1146/ANNUREV-STATISTICS-060116-053803

37. Schwedhelm C, Lipsky LM, Shearrer GE, et al. Using food network analysis to understand meal patterns in pregnant women with high and low diet quality. *International Journal of Behavioral Nutrition and Physical Activity*. 2021;18(1):101. doi:10.1186/s12966-021-01172-1

38. Schwedhelm C, Knüppel S, Schwingshackl L, Boeing H, Iqbal K. Meal and habitual dietary networks identified through Semiparametric Gaussian Copula Graphical Models in a German adult population. *PLoS One*. 2018;13(8):e0202936. doi:10.1371/journal.pone.0202936

39. Iqbal K, Buijsse B, Wirth J, Schulze MB, Floegel A, Boeing H. Gaussian graphical models identify networks of dietary intake in a German adult population. *The Journal of Nutrition*. 2016;146(3):646-652. doi:10.3945/JN.115.221135

40. Iqbal K, Schwingshackl L, Floegel A, et al. Gaussian graphical models identified food intake networks and risk of type 2 diabetes, CVD, and cancer in the EPIC-Potsdam study. *European Journal of Nutrition*. 2019;58(4):1673-1686. doi:10.1007/S00394-018-1714-1/TABLES/4

41. Altenbuchinger M, Weihs A, Quackenbush J, Grabe HJ, Zacharias HU. Gaussian and mixed graphical models as (multi-)omics data analysis tools. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*. 2020;1863(6):194418. doi:10.1016/J.BBAGRM.2019.194418

42. Solmi M, Koyanagi A, Thompson T, Fornaro M, Correll CU, Veronese N. Network analysis of the relationship between depressive symptoms, demographics, nutrition, quality of life and medical condition factors in the Osteoarthritis Initiative database cohort of elderly North-American adults with or at risk for osteoarthritis. *Epidemiol Psychiatr Sci*. 2019;29:e14. doi:10.1017/S204579601800077X

43. Rasmussen KM, Yaktine AL, Delaney KM, Callahan EA, eds. *Evaluating the Process to Develop the Dietary Guidelines for Americans, 2020-2025: Final Report*. National Academies Press; 2023. doi:10.17226/26653

44. Government of Canada, Canadian Institutes of Health Research. Heterogeneity in Nutrition and Metabolic Health: Virtual Workshop Report Summary - CIHR. Published January 24, 2023. Accessed May 16, 2024. https://cihr-irsc.gc.ca/e/53359.html

45. Health Canada. *Reference Guide to Understanding and Using the Data: 2015 Canadian Community Health Survey-Nutrition*. Health Canada; 2017. https://www.canada.ca/content/dam/hc-sc/documents/services/food-nutrition/food-nutrition-surveillance/ReferenceGuide2015CCHS-Nutr_Eng_Final_06192017.pdf

46. Winpenny EM, van Sluijs EMF, White M, Klepp KI, Wold B, Lien N. Changes in diet through adolescence and early adulthood: longitudinal trajectories and association with key life transitions. *Int J Behav Nutr Phys Act*. 2018;15(1):86. doi:10.1186/s12966-018-0719-8

47. Hutchinson JM, Tarasuk V. Prevalence of micronutrient inadequacy differs by severity of food insecurity among adults living in Canada in 2015. *Appl Physiol Nutr Metab*. Published online 2024. doi:10.1139/apnm-2023-0614

48. Moubarac JC, Batal M, Louzada ML, Steele EM, Monteiro CA. Consumption of ultra-processed foods predicts diet quality in Canada. *Appetite*. 2017;108:512-520. doi:10.1016/J.APPET.2016.11.006

49. La Mantia C. *Diet Quality in Relation to Income, Education, and Food Insecurity among Canadian Adults*. University of Toronto; 2020.

50. Valdes M, Conklin A, Veenstra G, Black JL. Plant-based dietary practices in Canada: examining definitions, prevalence and correlates of animal source food exclusions using nationally representative data from the 2015 Canadian Community Health Survey–Nutrition. *Public Health Nutrition*. 2021;24(5):777-786. doi:10.1017/S1368980020003444

51. Milicic S, DeCicca P. The impact of economic conditions on healthy dietary intake: Evidence from fluctuations in Canadian unemployment rates. *Journal of Nutrition Education and Behavior*. 2017;49(8):632-638.e1. doi:10.1016/j.jneb.2017.06.010

52. Shrier I, Platt RW. Reducing bias through directed acyclic graphs. *BMC Medical Research Methodology*. 2008;8(1):70. doi:10.1186/1471-2288-8-70

53. Robinson WR, Renson A, Naimi AI. Teaching yourself about structural racism will improve your machine learning. *Biostatistics*. 2020;21(2):339-344. doi:10.1093/biostatistics/kxz040

54. United States Department of Agriculture. U.S. Household Food Security Survey Module. Accessed April 18, 2023. https://www.ers.usda.gov/topics/food-nutrition-assistance/food-security-in-the-u-s/survey-tools/#household

55. Health Canada. Determining food security status. Published 2020. Accessed April 27, 2023. https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance/health-nutrition-surveys/canadian-community-health-survey-cchs/household-food-insecurity-canada-overview/determining-food-security-status-food-nutrition-surveillance-health-canada.html

56. Blanton CA, Moshfegh AJ, Baer DJ, Kretsch MJ. The USDA Automated Multiple-Pass Method accurately estimates group total energy and nutrient intake. *The Journal of Nutrition*. 2006;136(10):2594-2599. doi:10.1093/JN/136.10.2594

57. Moshfegh AJ, Rhodes DG, Baer DJ, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. *American Journal of Clinical Nutrition*. 2008;88(2):324-332. doi:10.1093/AJCN/88.2.324

58. Rhodes DG, Murayi T, Clemens JC, Baer DJ, Sebastian RS, Moshfegh AJ. The USDA Automated Multiple-Pass Method accurately assesses population sodium intakes. *American Journal of Clinical Nutrition*. 2013;97(5):958-964. doi:10.3945/AJCN.112.044982

59. Health Canada. Canadian Nutrient File (CNF). Published 2023. Accessed March 9, 2023. https://food-nutrition.canada.ca/cnf-fce/?lang=eng

60. Statistics Canada. *2015 Canadian Community Health Survey (CCHS) - Nutrition User Guide*. Statistics Canada; 2018.

61. Kirkpatrick SI, Guenther PM, Subar AF, et al. Using short-term dietary intake data to address research questions related to usual dietary intake among populations and subpopulations: Assumptions, statistical techniques, and considerations. *Journal of the Academy of Nutrition and Dietetics*. 2022;122(7):1246-1262. doi:10.1016/J.JAND.2022.03.010

62. Dodd KW, Guenther PM, Freedman LS, et al. Statistical methods for estimating usual intake of nutrients and foods: a review of the theory. *Journal of the American Dietetic Association*. 2006;106(10):1640-1650. doi:10.1016/J.JADA.2006.07.011

63. R Core Team. R: A language and environment for statistical computing. Published online 2021. https://www.R-project.org/

64. Chen S, Witten DM, Shojaie A. Selection and estimation for mixed graphical models. *Biometrika*. 2015;102(1):47-64. doi:10.1093/biomet/asu051

65. Haslbeck J. mgm: Estimating Time-Varying k-Order Mixed Graphical Models. Published online September 5, 2023. Accessed February 1, 2024. https://cran.r-project.org/web/packages/mgm/index.html

66. Friedman J, Tibshirani TH and R. glasso: Graphical Lasso: Estimation of Gaussian Graphical Models. Published online October 1, 2019. Accessed February 12, 2024. https://cran.r-project.org/web/packages/glasso/index.html

67. Tibshirani R. Regression Shrinkage and Selection Via the Lasso. *Journal of the Royal Statistical Society: Series B (Methodological)*. 1996;58(1):267-288. doi:10.1111/j.2517-6161.1996.tb02080.x

68. Epskamp S, Costantini G, Haslbeck J, et al. qgraph: Graph Plotting Methods, Psychometric Data Visualization and Graphical Model Estimation. Published online November 3, 2023. Accessed February 1, 2024. https://cran.r-project.org/web/packages/qgraph/index.html

69. Epskamp S, Fried EI. bootnet: Bootstrap Methods for Various Network Estimation Routines. Published online October 4, 2023. Accessed February 12, 2024. https://cran.r-project.org/web/packages/bootnet/index.html

70. Epskamp S, Borsboom D, Fried EI. Estimating psychological networks and their accuracy: A tutorial paper. *Behav Res*. 2018;50(1):195-212. doi:10.3758/s13428-017-0862-1

71. PROOF. New data on household food insecurity in 2022. PROOF. Published May 2, 2023. Accessed October 11, 2023. https://proof.utoronto.ca/2023/new-data-on-household-food-insecurity-in-2022/

72. Health Canada. Healthy Food Choices - Canada's Food Guide. Published 2020. Accessed May 1, 2022. https://food-guide.canada.ca/en/healthy-food-choices/

73. Hutchinson JM, Dodd KW, Guenther PM, et al. The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: scoring system and construct validity. *Appl Physiol Nutr Metab*. 2023;48(8):620-633. doi:10.1139/apnm-2023-0018

74. Stronks K, van de Mheen H, van den Bos J, Mackenbach JP. The interrelationship between income, health and employment status. *Int J Epidemiol*. 1997;26(3):592-600. doi:10.1093/ije/26.3.592

75. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. *BMC Medical Research Methodology*. 2019;19(1):64. doi:10.1186/s12874-019-0681-4

76. Kirk D, Kok E, Tufano M, Tekinerdogan B, Feskens EJM, Camps G. Machine learning in nutrition research. *Advances in Nutrition*. 2022;13(6):2573-2589. doi:10.1093/ADVANCES/NMAC103

77. Oliveira Chaves L, Gomes Domingos AL, Louzada Fernandes D, Ribeiro Cerqueira F, Siqueira-Batista R, Bressan J. Applicability of machine learning techniques in food intake

assessment: A systematic review. *Critical Reviews in Food Science and Nutrition*. 2023;63(7):902-919. doi:10.1080/10408398.2021.1956425

78. Côté M, Lamarche B. Artificial intelligence in nutrition research: perspectives on current and future applications. *Appl Physiol Nutr Metab*. Published online 2021:1-8. doi:10.1139/apnm-2021-0448

79. Hutchinson JM, Raffoul A, Pepetone A, et al. Advances in methods for characterizing dietary patterns: A scoping review.

80. Birds of a feather, or friend of a friend? using exponential random graph models to investigate adolescent social networks* | Demography. Accessed November 3, 2023. https://link.springer.com/article/10.1353/dem.0.0045

81. Hearty AP, Gibney MJ. Analysis of meal patterns with the use of supervised data mining techniques--artificial neural networks and decision trees. *Am J Clin Nutr*. 2008;88(6):1632-1642. doi:10.3945/ajcn.2008.26619

82. Xia Y, Zhao Z, Zhang S, et al. Complex dietary topologies in non-alcoholic fatty liver disease: A network science analysis. *Front Nutr*. 2020;7:579086. doi:10.3389/fnut.2020.579086

83. Samieri C, Sonawane AR, Lefèvre-Arbogast S, Helmer C, Grodstein F, Glass K. Using network science tools to identify novel diet patterns in prodromal dementia. *Neurology*. 2020;94(19):e2014-e2025. doi:10.1212/WNL.0000000000009399

84. National Cancer Institute. Dietary Assessment Primer. Accessed April 2, 2023. https://dietassessmentprimer.cancer.gov/

85. Kipnis V, Subar AF, Midthune D, et al. Structure of dietary measurement error: results of the OPEN biomarker study. *American Journal of Epidemiology*. 2003;158(1):14-21. doi:10.1093/AJE/KWG091

86. Spicker D, Nazemi A, Hutchinson J, et al. Challenges for predictive modeling with neural network techniques using error-prone dietary intake data. Published online November 15, 2023. doi:10.48550/arXiv.2311.09338

87. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning, Second Edition. *Springer New York, NY*. 2009;27(2):337-387.

Chapter 8: General Discussion

8.1. Overview

Given the strong links between diet and health,(4,6,7) using dietary assessment methods that are well matched to research questions and regularly monitoring dietary patterns are important aspects of nutrition epidemiology and public health surveillance. Dietary assessment methods that are designed to capture and analyze dietary patterns and consider heterogeneity are essential for the advancement of public health nutrition, with policies and guidance increasingly focused on the overall dietary pattern as opposed to single components. As methods rapidly progress, it is imperative to understand the scope of available tools to capture and analyze dietary patterns and to access these tools as needed to address relevant questions, depending on the study context and population.

This dissertation considered a range of dietary assessment methods to advance our understanding of dietary patterns, considering the links between dietary patterns and sociodemographic characteristics throughout, with insights for the Canadian context and beyond. The key objectives of this dissertation were to: (1) develop a brief screener to assess overall alignment of dietary patterns of adults with the 2019 Canada's Food Guide healthy food choices recommendations; (2) assess the construct validity of the screener and its scoring system among a sample of adults; (3) describe the use of novel methods to characterize dietary patterns; and (4) identify dietary patterns among adults in Canada in relation to sociodemographic characteristics, including age, sex, income, education, household food security status, geographic region, employment status, and smoking status, using probabilistic graphical models.

8.2. Summary of key findings

Chapters 4 and **5** describe the development and evaluation of a brief dietary screener to assess alignment of adults' dietary patterns with the 2019 Canada's Food Guide healthy food choices recommendations. This study was motivated by the updated 2019 Canada's Food Guide, which represented a large shift in the style of dietary guidance compared to prior iterations of the Guide.(51,55) Health Canada expressed an interest in assessing alignment with the dietary guidance represented in the Food Guide.(117) To address situations in which comprehensive measures of intake are not feasible, Health Canada contracted the development of a brief dietary screener. This research was undertaken in collaboration with Health Canada and represents the first known Canada-specific dietary screener that assesses alignment with dietary guidance. We followed key steps throughout the development and evaluation process, outlined *a priori*. Through the development, which included cognitive testing and face and content validity testing, we iteratively revised the screener to ensure it would best meet the needs of the target population and researchers and policymakers.

When evaluating the screener, we compared scores to the HEFI-2019,(56,57) which, in the absence of a reference biomarker for alignment with guidance or total dietary intake, was the most appropriate available comparator. We also investigated differences in screener scores by several sociodemographic characteristics previously shown to be associated with differential diet quality.(56,104,105,118) Based on these assessments, the Canadian Food Intake Screener demonstrated reasonable construct validity among a sample of adults in Canada. We observed a correlation of 0.53 between screener and HEFI-2019 scores, which is similar to what has been observed for other dietary screeners.(119–121) Screener scores ranged from 25.1 to 45.2 (1st-99th percentile). Although the screener was scored out of a maximum possible 65 points, it would be unlikely for anyone to achieve the maximum points because of the frequency-based scoring approach. Similarly, individuals reported consuming at least some of the food groups some of

the time, making scores close to zero also unlikely. This attenuated range in scores aligns with other screeners and diet quality indices.(56,122)

Overall, this process led to the development of a simple tool that meets the needs outlined by Health Canada. Already, the screener has been implemented in large cohorts and is freely available on the Government of Canada's Open Government Portal.(123) The Canadian Food Intake Screener is appropriate for instances in which time or financial resources are constrained and fills a gap in the dietary assessment landscape in Canada.

In addition to needing tools to capture dietary patterns, methods to analyze dietary patterns are central to dietary assessment. Methods to analyze dietary patterns using traditional *a priori* and *a posteriori* methods have been well-described.(4,63,124) With advances in methods, the field of health has rapidly adopted novel approaches to address complex, often non-linear research questions.(19,125–127) Within the field of nutrition, these methods have been adopted over the past several years, with reviews indicating their potential uses for precision nutrition, image-based dietary assessment, and more.(13,17,18,65,66) It is important to understand the range of novel methods that are being used to identify dietary patterns, addressed through a scoping review presented in **Chapter 6**. We considered how researchers integrated health outcomes, individual level characteristics, and the types of dietary assessment instruments used in research using novel methods to characterize dietary patterns.

A range of methods were used to identify dietary patterns in this review, with nine of the 24 included studies using approaches that have applications in machine learning. Half of the included studies used methods that characterized distinct dietary patterns, while the remaining studies characterized dietary patterns as networks or through a range of other approaches. We found that these methods are being applied at an increasing pace. This wide range of approaches and their rapid adoption suggests the need for guidance in the application of novel methods to dietary data. This will help to ensure comparability within and across approaches.(63) Efforts to

support improved interpretability and comparability across methods will be essential for enabling the uptake of evidence from novel methods into practice to promote improved population dietary patterns and health. Without guidance, it is challenging for researchers to determine when they have selected the ideal approach for their analysis, and it will be challenging for those synthesizing and using evidence in practice to develop policy informed by evidence generated using novel approaches.

From our review in Chapter 6, we noted an opportunity to explore dietary patterns using a novel method in the Canadian population. A possible advantage of using a novel method to characterize dietary patterns is the prospect of better capturing the inherent complexity of dietary patterns.(15,16) Further, while it is well-established that sociodemographic characteristics are associated with differential dietary intake,(9,56,104,128) analyses rarely consider possible interactions between these characteristics. However, people have multidimensional identities that interact with each other to shape their experiences and outcomes in life.(129) We sought to explore the capacity of probabilistic graphical models to capture the complexity of multidimensional dietary patterns and intersecting sociodemographic characteristics.

Chapter 7 describes the application of mixed graphical models, applied to data on adults from the 2015 Canadian Community Health Survey Nutrition, the most recent and robust population-representative nutrition dataset of individuals in Canada that is available.(40) We applied mixed graphical models, which can incorporate both categorical and continuous variables, to explore multidimensionality in dietary patterns and intersecting relationships among sociodemographic characteristics. We identified a sociodemographic network, dietary components network, and a network that included dietary components and sociodemographic characteristics. This layered approach allowed us to explore how dietary components and sociodemographic characteristics were jointly related with each other prior to applying the full

model. Within the full model, several strong joint relationships were observed between dietary components, with several patterning by food group type. Likewise, sociodemographic characteristics were most strongly jointly related with each other. Among the sociodemographic characteristics included in the model, age and sex were the most strongly connected with dietary components. Overall, this research provides insights into a method that allows for insights into the internal structures of dietary patterns and could complement traditional approaches.

Throughout the development and analysis of these models, efforts to balance interpretability and complexity were weighed. There is a substantial amount of expertise and judgment required to develop and interpret models. While there may be a perception that these methods are more 'objective' because they are data-driven, researchers make numerous decisions to implement such analyses. As noted by Ocké,(63) it is essential to evaluate the methods being applied to assess overall dietary patterns to enable reproducibility, evaluate validity, and foster comparisons between methods. The development of quality appraisal guidance and tailored reporting guidelines will improve the ability of researchers to compare different types of analyses more easily. It is also important to develop interdisciplinary research teams, as we did in this study, to promote the optimal application of research and practice. Novel methods like probabilistic graphical models hold promise in the field of public health nutrition to better understand the complexity of dietary patterns, but continued exploration and evaluation is needed to determine how findings can be translated and synthesized to inform public health.

8.3. Contributions of the dissertation

Within the public health nutrition landscape, this dissertation makes three key contributions to advance our understanding of dietary patterns:

1) **A tool to quickly assess alignment with dietary guidance among adults in Canada.**

Healthy eating patterns have been identified as a top priority by Health Canada through the Healthy Eating Strategy and the 2019 Canada's Food Guide.(50,51) Health Canada identified the need for tailored tools to assess alignment with the 2019 dietary guidance.(117) As discussed in **Chapters 4** and **5**, the development and evaluation of a new brief screener that has demonstrated moderate construct validity presents many opportunities to quickly and consistently assess alignment with the 2019 Canada's Food Guide healthy food choices recommendations. Open-access, low burden tools like the Canadian Food Intake Screener provide opportunities to promote reproducibility within the field.(130) The screener is freely available on the Government of Canada's Open Government portal and has been integrated into studies including a cohort.(123)

While a screener should not replace more detailed methods of assessment, such as 24HDRs, when it is possible to use such methods, its simplicity and low burden offers many possibilities for wide implementation, with potential applications in surveillance and epidemiology. In Canada, comprehensive assessments of dietary intake at the population level are relatively rare, with the two most recent surveys occurring in 2004 and 2015 and no date for an upcoming survey.(40,41) In the absence of frequent comprehensive assessments, a brief tool can be easily incorporated into existing national surveillance systems, such as the annual CCHS. Such an effort can provide consistent, ongoing evidence about alignment with dietary guidance with relatively low cost and effort, including in the Territories, which are not part of the nutrition-focused cycles of CCHS.

2) **Identifying novel methods to characterize dietary patterns.**

While traditional approaches to characterize dietary patterns have provided insights into overall patterns and how they relate to health and sociodemographic characteristics, these methods do not capture the 'totality' of dietary patterns, overlooking potential synergistic or antagonistic relationships among dietary components.(15,16) Novel methods, such as those with applications in machine learning, may be more insightful when attempting to learn complex, non-linear relationships that often make up dietary patterns.(13,17,18,131) With the rapid adoption of novel methods throughout the field of nutrition and health,(65,66,125–127) **Chapter 6** provides a review of novel methods that have been applied to identify dietary patterns.

This review highlighted the diverse range in methods that are being applied to characterize dietary patterns and examination of relationships between these patterns and sociodemographic characteristics or health outcomes. There are a range of appropriate methods that researchers can use to answer questions that involve the characterization of dietary patterns. Depending on the research objective, a traditional approach, such as a diet quality index or principal component analysis, may be well-suited to answer the question at hand. If researchers wish to consider the complex, synergistic relationships among dietary components or to capture non-linear relationships, novel methods may be more appropriate.(13,15,17)

In a rapidly shifting public health and nutrition landscape, the scoping review sheds light on current approaches that are being adopted to characterize dietary patterns and how these can be used to further public health. In the future, the creation of quality assessment frameworks and guidelines to assist those conducting research using these methods may enhance the quality and comparability of these studies, lending to synthesis of the evidence they generate to inform policy and programs.

3) **An exploration of the dietary patterns of adults using a novel method of assessment, considering potential intersections with sociodemographic characteristics.**

While there have been several explorations of the dietary patterns of individuals in Canada through *a priori* or *a posteriori* approaches, these have not considered potential synergistic relationships between dietary components.(15,16) **Chapter 7** provides an illustration of how novel methods can be applied to consider complexity on multiple axes. In this case, we modeled multidimensional dietary patterns with intersecting sociodemographic characteristics to characterize patterns among adults in Canada, expanding on prior work characterizing patterns in Canada (9,56,98,106,112,113,132–138).

Diet assessment methods that consider the diverse dietary patterns of the Canadian population will provide much-needed evidence to inform policies that can meaningfully reduce inequities. The chosen method of assessment can have major implications for a study's findings and interpretation, and the development and application of appropriate tools and analytic approaches can provide evidence to better represent the population and promote healthy eating for all Canadians. Our investigation using networks to model dietary patterns provides an illustration of how we can consider a novel method that provides insight into the internal structure of dietary patterns as one evidence input when informing future public health directions. The outputs from this research complement traditional methods to analyze dietary patterns and can be used as a starting point for confirmatory research.

This analysis provides an initial step towards better understanding the dietary patterns of adults in Canada, with all dietary components conditioned on each other, as well as considering sociodemographic characteristics shown to be associated with dietary intake. Further work layering clustering algorithms such as the Louvain algorithm onto identified networks may

increase interpretability.(139) Another possible future direction is a two-stage analysis, first clustering participants by their intersecting characteristics and then exploring dietary patterns using undirected probabilistic graphical models among clusters. Such work can help to inform dietary guidance and other policies and programs that address the complexity of dietary patterns in the population and population subgroups.

8.4. Overall strengths and limitations of the dissertation

This dissertation provides an examination of methods to capture and analyze dietary patterns, largely focusing on the Canadian context. Dietary assessment methods are broad and encompass a wide range of potential approaches. This dissertation explored a few approaches to advance how dietary assessment methods can be used to characterize dietary patterns from simple tools to complex methods. These approaches included a brief dietary screener that can be used in surveillance and epidemiology when comprehensive assessments of dietary intake are not available, a review that provides insights into new ways in which dietary patterns are being characterized, and an exploration of the dietary patterns of adults in Canada and their joint associations with sociodemographic characteristics. While a process of reflection was adopted throughout this research process, it is possible that I missed relevant perspectives to inform this work.

Strengths of included studies

Overall strengths of this dissertation include the range of dietary assessment methods applied or considered, including a dietary screener in **Chapters 4** and **5**, food frequency questionnaires and food records in **Chapter 6**, and 24HDRs in **Chapters 5, 6,** and **7**. Within this dissertation, a range of methods were applied to analyze dietary intake data, from simple techniques to more complex statistical approaches. These methods were carefully selected to

match research questions, with consideration of measurement error as possible. Within **Chapter 5**, a careful cleaning approach was implemented to reduce mismatches between reported intakes on the 24HDRs and assigned food codes, although this required substantial time to manually clean. This dissertation focused overall on dietary patterns as opposed to single dietary components, reflecting shifts to consider the whole diet as it is eaten rather than prior reductionist perspectives.(16,70) Another strength was the inclusion of interdisciplinary research teams throughout. For example, advisors for the development and evaluation of the Canadian Food Intake Screeners in **Chapter 4** and **5** represented experts in academia, government, and practitioners. In **Chapters 6** and **7**, an interdisciplinary team of experts from nutrition, statistics, and machine learning provided insights throughout study development and analysis.

Further strengths of **Chapters 4** and **5** include that they were led by guiding principles and a core team of advisors to ensure the development of a screener that would fill a needed gap in the dietary assessment landscape in Canada. **Chapter 6**, meanwhile, recognized the evolving and ranging nature of novel methods and thus adopted an intentionally liberal approach to capture the broadest range of studies characterizing dietary patterns. Finally, **Chapter 7** applied a complex statistical technique to jointly model dietary components and sociodemographic characteristics, enabling the consideration of synergistic relationships.

Limitations of included studies

The studies included in this dissertation were also subject to several limitations. Across these studies, we sought to consider a range of sociodemographic characteristics and how they may be associated with dietary patterns or the reporting of dietary intake. While several characteristics' relationships to dietary patterns were considered; certain characteristics, including disability, Indigeneity, and race or ethnicity, are relevant indicators of equity and were under- or unexplored. Depending on the study and characteristic, we were limited in our ability

for examination due to small cell sizes or because questions about these characteristics were not asked. Future research should prioritize the inclusion of questions that enable exploration of a broad range of characteristics and sampling strategies to ensure analyses of subgroups and of intersections among characteristics are possible.

The individual studies in this dissertation each had unique limitations that are also worth noting. In **Chapter 4**, we aimed to recruit a diverse sample for cognitive testing; however, we were not able to recruit a fully balanced sample, for example, with respect to gender identity. We sought to develop a screener that considered equity through guiding principles that prioritized a screener that was simple to use and score, considered the numeracy and literacy levels of the target population, and considered equivalence in capturing the construct across subgroups of the population. Though we attempted to engage a wide variety of experts with expertise in nutrition, dietary guidance, and dietary assessment and validation, perspectives that may have influenced the design of the screener and its evaluation may not have been considered. It is also possible that the informal coding approach missed details from the cognitive interviews but a comparison between this approach and a more in-depth method found they were comparable overall.(140) While tailored screeners could potentially be developed for different subgroups, this would make it very challenging to harmonize and compare results of studies conducted with different subpopulations.

Chapter 5 was conducted using data from 24HDRs, which are subject to recall bias and are not a perfect reference instrument for evaluation of the screener.(141,142) While we applied quotas, we did not reach our desired sample size, nor did we achieve all of our quotas. We did achieve quotas for language, race/ethnicity, and educational attainment and were close to reaching the sample size goal. The screener demonstrates moderate construct validity,(143) suggesting it can be used to provide insights into overall alignment with the 2019 Canada's Food

Guide healthy food choices recommendations, but that the HEFI-2019 should be used along with a comprehensive data collection method whenever possible.

Another limitation of Chapter 5 was the age of the Canadian Nutrient File.(144) The Canadian Nutrient File was last updated in 2015, and the Canadian food supply has shifted substantially since that time, with new products available that are not represented. For example, the availability of processed plant-based protein options has rapidly increased since 2015,(145) but these additions are not represented in the current Canadian Nutrient File. While this limitation was mitigated as possible by assigning food codes that were the closest match when a participant's entry was not available (e.g., replacing entered oat milk with rice milk), an update to the Canadian Nutrient File by Health Canada is essential to ensure that the dietary intakes of individuals in Canada are represented as accurately as possible. Given that there is no estimated timeline for an updated CCHS Nutrition, an updated Canadian Nutrient File is crucial to support robust, ongoing dietary assessment in Canada using other data sources that link to the Canadian Nutrient File, such as studies that apply the Automated Self-Administered 24-Hour Dietary Assessment Tool (ASA24).(30,146)

Given the rapidly evolving nature of novel methods being applied to identify dietary patterns and the wide range of terms used to describe these methods, it is unlikely that the scoping review in **Chapter 6** captured all articles that have used novel methods to characterize dietary patterns. However, the review provides a snapshot of the range of novel methods that are being applied to characterize dietary patterns.

Chapter 7 was subject to several limitations. It drew upon data from 24HDRs, which are subject to bias.(141,142) A single 24HDR is not representative of usual intake and it is likely that the random error in a single day of intake may have impacted findings.(26,29) We considered pre-processing the intake data used in the graphical models by estimating usual

intake distributions,(147) but determined that this would impose a structure that would restrict the graphical models in freely learning from the inputs.

Data were drawn from the 2015 CCHS Nutrition, the most recent population-representative survey of the dietary intakes of individuals living in Canada. This survey is already nearly a decade old and does not include those living in the territories, those experiencing homelessness, those living in institutions, full-time members of the Canadian military, individuals living on reserves, or those in some remote areas.(40) There is no known strategy to mitigate the age of these data or the absence of key demographic groups in this survey, as these strategies would need to be implemented at the time of data collection. At the same time, this survey remains the most current and representative data source to examine the dietary patterns of individuals in Canada, pointing to the dearth of comprehensive population-representative dietary data in Canada. Further, the 2015 CCHS Nutrition does not have information on all sociodemographic characteristics of interest, including gender identity and disability and the Public Use Microdata Files do not include information on characteristics such as race/ethnicity. It is also possible that participants may have differentially reported dietary patterns by sociodemographic characteristics, such as food security status.(148) This was noted in the interpretation of models, though it is not possible to determine precisely how this error may have influenced final models, as differential reporting may cause bias in either direction from truth.(149)

8.5. Implications for public health policy and research

This dissertation has notable implications for public health nutrition policy and research. The field of diet assessment has shifted from its roots in the study of single dietary components towards dietary patterns (70) and this dissertation sought to advance methods to better understand dietary patterns. Dietary assessment methods may not seem as immediately

relevant to public health as other subfields within nutrition. However, appropriate tools to capture and analyze dietary patterns are fundamental to nutrition research and are essential for the development, implementation, and evaluation of nutrition policy, and for monitoring the dietary patterns of populations. As noted in the National Cancer Institute's Diet Assessment Primer, using the appropriate instrument and acknowledging possible limitations is critical for appropriate analyses.(14) Diet assessment methods that are designed to answer questions of public health importance, including the methods in this dissertation, are essential for effectively conducting surveillance and evaluation to promote health and well-being. The key policy and research implications of this dissertation include:

1) Flexible approaches to evaluate nutrition policy

A key theme of this dissertation lies in the necessity of having a range of approaches to capture and analyze dietary patterns, which in turn can be used to monitor, develop, and evaluate nutrition policy. In some cases, comprehensive measures that enable complex analytic techniques will match research objectives and resources, whereas in other instances, it is necessary to use a simple, quick tool to rapidly assess the dietary patterns of a population or population subgroups. Through the development and evaluation of a dietary screener to capture alignment of the dietary patterns of adults quickly and easily in Canada with the 2019 Canada's Food Guide healthy food choices recommendations, we provide a low-barrier method to evaluate nutrition policy that can easily be applied in a range of studies.

Beyond the screener, this dissertation points to the potential for novel analytic approaches to complement current analytic techniques in the evaluation of dietary patterns and targets for nutrition policy. Our scoping review of the literature identifying studies that have applied novel methods to characterize dietary patterns found that a range of novel methods are being rapidly adopted. This review points to the need for continued monitoring of the adoption

of these methods and suggests a need for quality appraisal frameworks and reporting guidelines that will enable the synthesis of evidence from novel and traditional approaches. As the application of a range of novel and traditional methods for dietary pattern analysis becomes easier,(150) engaging interdisciplinary experts in projects helps to ensure the appropriate application of methods.

Novel methods show some promise for capturing complexity within dietary patterns that may not be uncovered using traditional approaches.(18) As researchers seek to understand overall multidimensional dietary patterns, methods that account for complexity, including potential synergistic or antagonistic relationships, can clarify diet-health relationships and inform improved nutrition policies and guidance.(16) All approaches to capture and analyze dietary patterns have associated strengths and limitations, and assessing patterns from a range perspectives, incorporating both simple and complex approaches, can lead to an overall better understanding of dietary patterns.

2) Consideration of heterogeneity in dietary patterns

As noted in the Canadian Institute for Health Research's Institute of Nutrition, Metabolism and Diabetes' Strategic Plan for 2021-2026, it is important to better understand heterogeneity across the field of nutrition to allow for improved, targeted public health nutrition initiatives among Canadians.(151) This dissertation provides a brief dietary screener that can easily be integrated into a range of study types to understand the heterogeneity of dietary patterns across sociodemographic characteristics and at the population-level.

Mixed graphical models, applied in Chapter 7, also offer the opportunity to explore heterogeneity in dietary patterns and help to advance this strategic priority. The application of probabilistic graphical models, a novel approach to identify dietary patterns and their joint associations with sociodemographic characteristics, provided a practical exploration of how a

novel method could be applied in the Canadian context to investigate dietary patterns and their intersecting relationships with sociodemographic characteristics. This illustration largely supported prior studies of relationships between diet and sociodemographic characteristics while integrating heterogeneity by jointly modeling all components rather than condensing diet to a unidimensional score and looking at its relationships with individual characteristics. These methods can be used in the future as a complement to the traditional confirmatory approaches that form the basis of nutrition evidence. This dissertation thus provides advancements for diet assessment methods to integrate heterogeneity to enable improved development, implementation, and evaluation of public health nutrition policy and research that represents the Canadian population.

8.6. Conclusion

Through methods to capture and analyze dietary patterns, this dissertation advances our understanding of dietary patterns, focusing overall on the Canadian context. As interest in understanding multidimensional dietary patterns that reflect the heterogeneity within populations continues to grow,(114,151) it is important to have a range of methods to capture and analyze dietary patterns to allow an evolving understanding of the complex construct of dietary patterns. Having flexibility in the choice of approaches to explore and understand dietary patterns is particularly useful considering there is no 'gold-standard' method available to assess total dietary patterns at the population level. In the face of rapidly shifting methods to capture and analyze dietary patterns, it is important to be aware of both traditional and novel approaches for data capture and analysis, to engage researchers with expertise in dietary assessment methods in related research, and above all, to appropriately match approaches with the research question and study context.

While this research extended insights into dietary patterns, there is room for future studies to optimize methods and interpretation. The Canadian Food Intake Screener, presented in Chapter 4 and 5, is now freely available on the Government of Canada's Open Government portal and could be tested and adapted for other population subgroups.(123) The screener can be used to evaluate alignment of dietary patterns with dietary guidance and future research can explore relationships with health outcomes. With the proliferation of new, varied methods to characterize dietary patterns, an essential step will be to develop quality appraisal frameworks and reporting guidelines that specifically promote improved comparability across the wide range of methods that are being applied to identify dietary patterns. Building from the work in Chapter 7, explorations of the multidimensional dietary patterns of individuals in Canada can continue to be extended to inform nutrition policy and guidance that is as applicable as possible to the diverse population. Finally, while Chapter 7 describes the interrelationships between dietary components and sociodemographic characteristics through undirected models, future studies can explore causal relationships between dietary patterns and health outcomes through directed models.(86)

Taken together, this dissertation provides an in-depth exploration into a range of dietary assessment approaches that capture and analyze dietary patterns. There is a great need for continued exploration into dietary patterns, a highly complex construct, to continue to improve our understanding of how diet relates to health and other factors. By enhancing our understanding of these methods, this dissertation provides opportunities to better understand dietary patterns, whether through simple or complex approaches, with potential for findings to help inform more targeted nutrition policies in Canada and beyond.

References

1. Rajkomar A, Hardt M, Howell MD, Corrado G, Chin MH. Ensuring fairness in machine learning to advance health equity. Ann Intern Med. 2018 Dec;169(12):866–72.

2. Robinson WR, Renson A, Naimi AI. Teaching yourself about structural racism will improve your machine learning. Biostatistics. 2020 Apr 1;21(2):339–44.

3. Rojas JC, Fahrenbach J, Makhni S, Cook SC, Williams JS, Umscheid CA, et al. Framework for integrating equity into machine learning models: A case study. Chest. 2022 Jun;161(6):1621–7.

4. English LK, Ard JD, Bailey RL, Bates M, Bazzano LA, Boushey CJ, et al. Evaluation of dietary patterns and all-cause mortality: A systematic review. JAMA Netw Open [Internet]. 2021;4(8). Available from: https://pubmed.ncbi.nlm.nih.gov/34463743/

5. Schwingshackl L, Bogensberger B, Hoffmann G. Diet quality as assessed by the Healthy Eating Index, Alternate Healthy Eating Index, Dietary Approaches to Stop Hypertension Score, and health outcomes: An updated systematic review and meta-analysis of cohort studies. J Acad Nutr Diet. 2018;118(1):74-100.e11.

6. Alam S, Lang JJ, Drucker AM, Gotay C, Kozloff N, Mate K, et al. Assessment of the burden of diseases and injuries attributable to risk factors in Canada from 1990 to 2016: An analysis of the Global Burden of Disease Study. Can Med Assoc Open Access J. 2019;7(1):E140–8.

7. 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. The Lancet. 2019;393(10184):1958–72.

8. Tarasuk V, Fitzpatrick S, Ward H. Nutrition inequities in Canada. Appl Physiol Nutr Metab. 2010;35(2):172–9.

9. Hutchinson J, Tarasuk V. The relationship between diet quality and the severity of household food insecurity in Canada. Public Health Nutr. 2022;25(4):1013–26.

10. Olstad DL, Nejatinamini S, Victorino C, Kirkpatrick SI, Minaker LM, McLaren L. Socioeconomic inequities in diet quality among a nationally representative sample of adults living in Canada: an analysis of trends between 2004 and 2015. Am J Clin Nutr. 2021 Nov;114(5):1814–29.

11. La Mantia C. Diet quality in relation to income, education, and food insecurity among Canadian adults. University of Toronto; 2020.

12. Thompson FE, Kirkpatrick SI, Subar AF, Reedy J, Schap TRE, Wilson MM, et al. The National Cancer Institute's Dietary Assessment Primer: A resource for diet research. J Acad Nutr Diet. 2015;115(12):1986–95.

13. Kirk D, Kok E, Tufano M, Tekinerdogan B, Feskens EJM, Camps G. Machine learning in nutrition research. Adv Nutr. 2022;13(6):2573–89.

14. National Cancer Institute. Dietary Assessment Primer [Internet]. [cited 2023 Apr 2]. Available from: https://dietassessmentprimer.cancer.gov/

15. Bodnar LM, Cartus AR, Kirkpatrick SI, Himes KP, Kennedy EH, Simhan HN, et al. Machine learning as a strategy to account for dietary synergy: an illustration based on dietary intake and adverse pregnancy outcomes. Am J Clin Nutr. 2020;111(6):1235–43.

16. Reedy J, Subar AF, George SM, Krebs-Smith SM. Extending methods in dietary patterns research. Nutrients [Internet]. 2018 May;10(5). Available from: https://pubmed.ncbi.nlm.nih.gov/29735885/

17. Morgenstern JD, Rosella LC, Costa AP, Souza RJ de, Anderson LN. Perspective: Big data and machine learning could help advance nutritional epidemiology. Adv Nutr. 2021 Jun;12(3):621–31.

18. Côté M, Lamarche B. Artificial intelligence in nutrition research: perspectives on current and future applications. Appl Physiol Nutr Metab. 2021;1–8.

19. Sidey-Gibbons JAM, Sidey-Gibbons CJ. Machine learning in medicine: a practical introduction. BMC Med Res Methodol. 2019 Mar 19;19(1):64.

20. Frongillo EA, Baranowski T, Subar AF, Tooze JA, Kirkpatrick SI. Establishing validity and cross-context equivalence of measures and indicators. J Acad Nutr Diet. 2019 Nov;119(11):1817–30.

21. Kirkpatrick SI, Baranowski T, Subar AF, Tooze JA, Frongillo EA. Best practices for conducting and interpreting studies to validate self-report dietary assessment methods. J Acad Nutr Diet. 2019 Nov;119(11):1801–16.

22. Clifton PM, Keogh JB. A systematic review of the effect of dietary saturated and polyunsaturated fat on heart disease. Nutr Metab Cardiovasc Dis NMCD. 2017 Dec;27(12):1060–80.

23. Schoeller DA, Westerterp M, editors. Advances in the Assessment of Dietary Intake. Boca Raton: CRC Press; 2017. 384 p.

24. Freedman LS, Schatzkin A, Midthune D, Kipnis V. Dealing with dietary measurement error in nutritional cohort studies. J Natl Cancer Inst. 2011 Jul;103(14):1086–92.

25. Keogh RH, Shaw PA, Gustafson P, Carroll RJ, Deffner V, Dodd KW, et al. STRATOS guidance document on measurement error and misclassification of variables in observational epidemiology: Part 1 – basic theory and simple methods of adjustment. Stat Med. 2020 Jul;39(16):2197.

26. Kirkpatrick SI, Guenther PM, Subar AF, Krebs-Smith SM, Herrick KA, Freedman LS, et al. Using short-term dietary intake data to address research questions related to

usual dietary intake among populations and subpopulations: Assumptions, statistical techniques, and considerations. J Acad Nutr Diet. 2022 Mar;122(7):1246–62.

27. Krebs-Smith SM, Guenther PM, Subar AF, Kirkpatrick SI, Dodd KW. Americans do not meet federal dietary recommendations. J Nutr. 2010 Oct;140(10):1832.

28. National Research Council (US) Subcommittee on Criteria for Dietary Evaluation. Nutrient Adequacy: Assessment Using Food Consumption Surveys [Internet]. Washington (DC): National Academies Press (US); 1986 [cited 2023 Apr 23]. Available from: http://www.ncbi.nlm.nih.gov/books/NBK217533/

29. 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.

30. ASA24® Dietary Assessment Tool | EGRP/DCCPS/NCI/NIH [Internet]. [cited 2023 Mar 15]. Available from: https://epi.grants.cancer.gov/asa24/

31. 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 Apr 5;3(1):34.

32. 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.

33. Moshfegh AJ, Rhodes DG, Baer DJ, Murayi T, Clemens JC, Rumpler WV, et al. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. Am J Clin Nutr. 2008 Aug;88(2):324–32.

34. Tangney CC, Rasmussen HE, Richards C, Li M, Appelhans BM. Evaluation of a brief sodium screener in two samples. Nutrients [Internet]. 2019 Jan;11(1). Available from: https://pubmed-ncbi-nlm-nih-gov.proxy.lib.uwaterloo.ca/30646541/

35. Using the New BRFSS Module | Nutrition | CDC [Internet]. 2020 [cited 2023 Apr 2]. Available from: https://www.cdc.gov/nutrition/data-statistics/using-the-new-BRFSS-modules.html

36. Colby S, Zhou W, Allison C, Mathews AE, Olfert MD, Morrell JS, et al. Development and validation of the short healthy eating index survey with a college population to assess dietary quality and intake. Nutrients. 2020 Sep;12(9):1–24.

37. Institute for Health Metrics and Evaluation. Canada profile [Internet]. IHME, University of Washington; 2021. Available from: https://www.healthdata.org/research-analysis/health-by-location/profiles/canada

38. Lieffers JRL, Ekwaru JP, Ohinmaa A, Veugelers PJ. The economic burden of not meeting food recommendations in Canada: The cost of doing nothing. PloS One [Internet]. 2018 Apr;13(4). Available from: https://pubmed.ncbi.nlm.nih.gov/29702680/

39. Overview of Canadians' Eating Habits, no. 2 - ARCHIVED [Internet]. [cited 2023 Apr 2]. Available from: https://www150.statcan.gc.ca/n1/en/catalogue/82-620-M2006002

40. Health Canada. Reference guide to understanding and using the data: 2015 Canadian Community Health Survey-Nutrition [Internet]. Health Canada; 2017. Available from: https://www.canada.ca/content/dam/hc-sc/documents/services/food-nutrition/food-nutrition-surveillance/ReferenceGuide2015CCHS-Nutr_Eng_Final_06192017.pdf

41. Health Canada. Canadian Community Health Survey Cycle 2.2, Nutrition (2004) - A guide to accessing and interpreting the data. Office of Nutrition Policy and Promotion, Health Canada; 2006.

42. Kirkpatrick SI, Vanderlee L, Raffoul A, Stapleton J, Csizmadi I, Boucher BA, et al. Self-report dietary assessment tools used in Canadian research: A scoping review. Adv Nutr. 2017 Mar;8(2):276–89.

43. Government of Canada, Statistics Canada. Canadian Health Measures Survey (CHMS) [Internet]. 2022 [cited 2023 Apr 2]. Available from: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=5071

44. Government of Canada, Statistics Canada. Canadian Community Health Survey - Annual Component (CCHS) [Internet]. 2022 [cited 2023 Apr 2]. Available from: https://www23.statcan.gc.ca/imdb/p2SV.pl?Function=getSurvey&SDDS=3226

45. Ye M, Robson PJ, Eurich DT, Vena JE, Xu JY, Johnson JA. Cohort profile: Alberta's Tomorrow Project. Int J Epidemiol. 2017 Aug;46(4):1097–1098l.

46. Lapointe A, Laramée C, Belanger-Gravel A, Buckeridge DL, Desroches S, Garriguet D, et al. NutriQuébec: a unique web-based prospective cohort study to monitor the population's eating and other lifestyle behaviours in the province of Québec. BMJ Open. 2020 Oct 1;10(10):e039889.

47. Raina P, Wolfson C, Kirkland S, Griffith LE, Balion C, Cossette B, et al. Cohort Profile: The Canadian Longitudinal Study on Aging (CLSA). Int J Epidemiol. 2019 Dec 1;48(6):1752–1753j.

48. Haines J, Douglas S, Mirotta JA, O'Kane C, Breau R, Walton K, et al. Guelph Family Health Study: pilot study of a home-based obesity prevention intervention. Can J Public Health. 2018 Aug;109(4):549–60.

49. Wiggers D, Vanderlee L, White CM, Reid JL, Minaker L, Hammond D. Food sources among young people in five major Canadian cities. Can J Public Health. 2018 Aug 1;109(4):506–15.

50. Government of Canada. Healthy Eating Strategy [Internet]. Health Canada; 2016 Oct. Available from: https://www.canada.ca/content/dam/canada/health-canada/migration/publications/eating-nutrition/healthy-eating-strategy-canada-strategie-saine-alimentation/alt/pub-eng.pdf

51. Health Canada. Canada's Food Guide [Internet]. 2019 [cited 2022 Oct 6]. Available from: https://food-guide.canada.ca/en/

52. Health Canada. Voluntary sodium reduction targets for processed foods 2020-2025 [Internet]. 2020 [cited 2023 Apr 17]. Available from: https://www.canada.ca/en/health-canada/services/publications/food-nutrition/sodium-reduced-targets-2020-2025.html

53. Health Canada. Changes to the nutrition facts table [Internet]. 2022 [cited 2023 Apr 17]. Available from: https://www.canada.ca/en/health-canada/services/food-labelling-changes/nutrition-facts-table.html

54. Health Canada. Healthy Food Choices - Canada's Food Guide [Internet]. 2020 [cited 2022 May 1]. Available from: https://food-guide.canada.ca/en/healthy-food-choices/

55. Health Canada. Canada's Dietary Guidelines for Health Professionals and Policy Makers. 2019.

56. Brassard D, Munene LAE, Pierre SS, Gonzalez A, Guenther PM, Jessri M, et al. Evaluation of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Appl Physiol Nutr Metab [Internet]. 2022 Jan;47(5). Available from: https://pubmed.ncbi.nlm.nih.gov/35030069/

57. Brassard D, Munene LAE, Pierre SS, Guenther PM, Kirkpatrick SI, Slater J, et al. Development of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Appl Physiol Nutr Metab. 2022 Jan;47(5):595–610.

58. Breiman L. Statistical modeling: The two cultures (with comments and a rejoinder by the author). https://doi.org/101214/ss/1009213726. 2001 Aug;16(3):199–231.

59. Efron B. Prediction, estimation, and attribution. Int Stat Rev. 2020 Dec;88(S1):S28–59.

60. Newby PK, Tucker KL. Empirically derived eating patterns using factor or cluster analysis: A review. Nutr Rev. 2004 May 1;62(5):177–203.

61. 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 Nov;29(27):2857–68.

62. 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 Jul;158(1):14–21.

63. Ocké MC. Evaluation of methodologies for assessing the overall diet: dietary quality scores and dietary pattern analysis. Proc Nutr Soc. 2013 May;72(2):191–9.

64. Jacobs DR, Steffen LM. Nutrients, foods, and dietary patterns as exposures in research: a framework for food synergy. Am J Clin Nutr. 2003 Sep;78(3):508S-513S.

65. Kirk D, Catal C, Tekinerdogan B. Precision nutrition: A systematic literature review. Comput Biol Med. 2021 Jun 1;133:104365.

66. Oliveira Chaves L, Gomes Domingos AL, Louzada Fernandes D, Ribeiro Cerqueira F, Siqueira-Batista R, Bressan J. Applicability of machine learning techniques in food intake assessment: A systematic review. Crit Rev Food Sci Nutr. 2023 Mar 11;63(7):902–19.

67. Bzdok D, Altman N, Krzywinski M. Statistics versus machine learning. Nat Methods. 2018 Apr 1;15(4):233–4.

68. Bi Q, Goodman KE, Kaminsky J, Lessler J. What is machine learning? A primer for the epidemiologist. Am J Epidemiol. 2019;188(12):2222–39.

69. Markus AF, Kors JA, Rijnbeek PR. The role of explainability in creating trustworthy artificial intelligence for health care: A comprehensive survey of the terminology, design choices, and evaluation strategies. J Biomed Inform. 2021 Jan;113:103655.

70. Mozaffarian D, Rosenberg I, Uauy R. History of modern nutrition science—implications for current research, dietary guidelines, and food policy. BMJ [Internet]. 2018 Jun;361. Available from: https://www.bmj.com/content/361/bmj.k2392

71. Guenther PM, Reedy J, Krebs-Smith SM, Reeve BB, Basiotis PP. Development and evaluation of the Healthy Eating Index. 2007; Available from: http://www.cnpp.usda.gov/HealthyEatingIndex.htm.

72. Guenther PM, Casavale KO, Reedy J, Kirkpatrick SI, Hiza HA, Kuczynski KJ, et al. Update of the Healthy Eating Index: HEI-2010. J Acad Nutr Diet. 2013;(12):2049–56.

73. Krebs-Smith SM, Pannucci TRE, Subar AF, Kirkpatrick SI, Lerman JL, Tooze JA, et al. Update of the Healthy Eating Index: HEI-2015. J Acad Nutr Diet. 2018 Sep;118(9):1591–602.

74. Tucker KL. Dietary patterns, approaches, and multicultural perspective. Appl Physiol Nutr Metab. 2010 Apr;35(2):211–8.

75. Hu FB. Dietary pattern analysis: a new direction in nutritional epidemiology. Curr Opin Lipidol. 2002;13(1):3–9.

76. Behrouzi P, Grootswagers P, Keizer PLC, Smeets ETHC, Feskens EJM, de Groot LCPGM, et al. Dietary intakes of vegetable protein, folate,and vitamins B-6 and B-12 are partially correlated with physical functioning of Dutch older adults using copula graphical models. J Nutr. 2020 Mar 1;150(3):634–43.

77. Dietary Guidelines Advisory Committee. Scientific Report of the 2015 Dietary Guidelines Advisory Committee: Advisory Report to the Secretary of Health and Human Services and the Secretary of Agriculture [Internet]. Washington (DC): U.S. Department of Agriculture, Agricultural Research Service; 2015. Available from: https://health.gov/our-work/nutrition-physical-activity/dietary-guidelines/previous-dietary-guidelines/2015/advisory-report

78. Hoang T, Lee J, Kim J. Network analysis of demographics, dietary intake, and comorbidity interactions. Nutrients. 2021 Oct;13(10):3563.

79. Schwedhelm C, Lipsky LM, Shearrer GE, Betts GM, Liu A, Iqbal K, et al. Using food network analysis to understand meal patterns in pregnant women with high and low diet quality. Int J Behav Nutr Phys Act. 2021 Jul 23;18(1):101.

80. Iqbal K, Buijsse B, Wirth J, Schulze MB, Floegel A, Boeing H. Gaussian graphical models identify networks of dietary intake in a German adult population. J Nutr. 2016;146(3):646–52.

81. Iqbal K, Schwingshackl L, Floegel A, Schwedhelm C, Stelmach-Mardas M, Wittenbecher C, et al. Gaussian graphical models identified food intake networks and risk of type 2 diabetes, CVD, and cancer in the EPIC-Potsdam study. Eur J Nutr. 2019 Jun;58(4):1673–86.

82. Schwedhelm C, Knüppel S, Schwingshackl L, Boeing H, Iqbal K. Meal and habitual dietary networks identified through Semiparametric Gaussian Copula Graphical Models in a German adult population. PloS One. 2018;13(8):e0202936.

83. Hastie T, Tibshirani R, Friedman J. The Elements of Statistical Learning, Second Edition. Springer N Y NY. 2009;27(2):337–87.

84. Drton M, Maathuis MH. Structure learning in graphical modeling. Httpsdoiorg101146annurev-Stat-060116-053803. 2017 Mar;4:365–93.

85. Altenbuchinger M, Weihs A, Quackenbush J, Grabe HJ, Zacharias HU. Gaussian and mixed graphical models as (multi-)omics data analysis tools. Biochim Biophys Acta BBA - Gene Regul Mech. 2020 Jun;1863(6):194418.

86. Koller D, Friedman N. Probabilistic Graphical Models: Principles and Techniques. MIT Press; 2009.

87. Haslbeck JMB, Waldorp LJ. mgm: Estimating time-varying mixed graphical models in high-dimensional data. J Stat Softw. 2020 Apr 27;93:1–46.

88. Fereidani SS, Sedaghat F, Eini-Zinab H, Heidari Z, Jalali S, Mohammadi E, et al. Gaussian graphical models identified food intake networks among Iranian women with and without breast cancer: A case-control study. Nutr Cancer. 2021 Nov 26;73(10):1890–7.

89. Gunathilake M, Lee J, Choi IJ, Kim YI, Kim J. Identification of dietary pattern networks associated with gastric cancer using Gaussian graphical models: A case-control study. Cancers. 2020 Apr;12(4):1044.

90. Gunathilake M, Hoang T, Lee J, Kim J. Association between dietary intake networks identified through a Gaussian graphical model and the risk of cancer: a prospective cohort study. Eur J Nutr. 2022 Dec 1;61(8):3943–60.

91. Hoang T, Lee J, Kim J. Differences in dietary patterns identified by the Gaussian graphical model in Korean adults with and without a self-reported cancer diagnosis. J Acad Nutr Diet. 2021 Aug 1;121(8):1484-1496.e3.

92. Jayedi A, Janbozorgi N, Djafarian K, Yekaninejad MS, Shab-Bidar S. Dietary networks identified by Gaussian graphical model and general and abdominal obesity in adults. Nutr J. 2021 Oct 27;20(1):86.

93. Jahanmiri R, Djafarian K, Janbozorgi N, Dehghani-Firouzabadi F, Shab-Bidar S. Saturated fats network identified using Gaussian graphical models is associated with metabolic syndrome in a sample of Iranian adults. Diabetol Metab Syndr. 2022 Aug 26;14(1):123.

94. Landaeta-Díaz L, Durán-Agüero S, González-Medina G. Exploring food intake networks and anhedonia symptoms in a Chilean Adults sample. Appetite. 2023 Nov 1;190:107042.

95. Slurink IA, Corpeleijn E, Bakker SJ, Jongerling J, Kupper N, Smeets T, et al. Dairy consumption and incident prediabetes: prospective associations and network models in the large population-based Lifelines Study. Am J Clin Nutr. 2023 Dec;118(6):1077–90.

96. Côté M, Osseni MA, Brassard D, Carbonneau É, Robitaille J, Vohl MC, et al. Are machine learning algorithms more accurate in predicting vegetable and fruit consumption than traditional statistical models? An exploratory analysis. Front Nutr. 2022 Feb;9:740898.

97. Morgenstern JD, Rosella LC, Costa AP, Anderson LN. Development of machine learning prediction models to explore nutrients predictive of cardiovascular disease using Canadian linked population-based data. Appl Physiol Nutr Metab. 2022 May;47(5):529–46.

98. Hosseini SH, Farag M, Hosseini SZ, Vatanparast H. Behavioral factors are perhaps more important than income in determining diet quality in Canada. SSM - Popul Health [Internet]. 2021 Mar;17. Available from: https://pubmed.ncbi.nlm.nih.gov/35005185/

99. Hu G, Flexner N, Tiscornia MV, L'Abbé MR. Accelerating the classification of NOVA food processing levels using a fine-tuned language model: A multi-country study. Nutrients. 2023 Jan;15(19):4167.

100. Hu G, Ahmed M, L'Abbé MR. Natural language processing and machine learning approaches for food categorization and nutrition quality prediction compared with traditional methods. Am J Clin Nutr. 2023 Mar 1;117(3):553–63.

101. Shah N, Srivastava G, Savage DW, Mago V. Assessing Canadians health activity and nutritional habits through social media. Front Public Health [Internet]. 2020 Jan 14 [cited 2024 Apr 23];7. Available from: https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2019.00400/full

102. Hutchinson JM, Tarasuk V. Prevalence of micronutrient inadequacy differs by severity of food insecurity among adults living in Canada in 2015. Appl Physiol Nutr Metab [Internet]. 2024 [cited 2024 May 2]; Available from: https://cdnsciencepub.com/doi/abs/10.1139/apnm-2023-0614

103. Hanson KL, Connor LM. Food insecurity and dietary quality in US adults and children: A systematic review. Am J Clin Nutr. 2014 Aug;100(2):684–92.

104. Hiza HAB, Casavale KO, Guenther PM, Davis CA. Diet quality of Americans differs by age, sex, race/ethnicity, income, and education level. J Acad Nutr Diet. 2013 Feb;113(2):297–306.

105. Guenther PM, Kirkpatrick SI, Reedy J, Krebs-Smith SM, Buckman DW, Dodd KW, et al. The Healthy Eating Index-2010 Is a valid and reliable measure of diet quality according to the 2010 Dietary Guidelines for Americans. J Nutr. 2014 Mar;144(3):399.

106. Moubarac JC, Batal M, Louzada ML, Steele EM, Monteiro CA. Consumption of ultra-processed foods predicts diet quality in Canada. Appetite. 2017 Jan;108:512–20.

107. Reedy J, Lerman JL, Krebs-Smith SM, Kirkpatrick SI, Pannucci TRE, Wilson MM, et al. Evaluation of the Healthy Eating Index-2015. J Acad Nutr Diet. 2018 Sep;118(9):1622–33.

108. Tarasuk VS. Household food insecurity with hunger is associated with women's food intakes, health and household circumstances. J Nutr. 2001 Oct;131(10):2670–6.

109. McIntyre L, Glanville NT, Raine KD, Dayle JB, Anderson B, Battaglia N. Do low-income lone mothers compromise their nutrition to feed their children? CMAJ. 2003 Mar 18;168(6):686–91.

110. Tarasuk V, McIntyre L, Li J. Low-income women's dietary intakes are sensitive to the depletion of household resources in one month. J Nutr. 2007 Aug;137(8):1980–7.

111. Kirkpatrick SI, Tarasuk V. Food insecurity is associated with nutrient inadequacies among Canadian adults and adolescents. J Nutr. 2008;138(3):604–12.

112. Garriguet D. Diet quality in Canada. Health Rep. 2009;20(3):41–52.

113. Doan N, Olstad DL, Vanderlee L, Hammond D, Wallace M, Kirkpatrick SI. Investigating the intersections of racial identity and perceived income adequacy in relation to dietary quality among adults in Canada. J Nutr. 2022;13(152):67S-75S.

114. Rasmussen KM, Yaktine AL, Delaney KM, Callahan EA, editors. Evaluating the Process to Develop the Dietary Guidelines for Americans, 2020-2025: Final Report [Internet]. Washington, D.C.: National Academies Press; 2023 [cited 2024 May 15]. Available from: https://www.nap.edu/catalog/26653

115. Shams-White MM, Pannucci TE, Lerman JL, Herrick KA, Zimmer M, Meyers Mathieu K, et al. Healthy Eating Index-2020: Review and update process to reflect the Dietary Guidelines for Americans, 2020-2025. J Acad Nutr Diet. 2023 Sep;123(9):1280–8.

116. Edefonti V, De Vito R, Dalmartello M, Patel L, Salvatori A, Ferraroni M. Reproducibility and validity of a posteriori dietary patterns: A systematic review. Adv Nutr. 2020 Mar 1;11(2):293–326.

117. Health Canada. Food and Nutrition Highlights - Helping you maintain and improve your health: 2020 [Internet]. 2021 [cited 2024 Feb 21]. Available from: https://www.canada.ca/content/dam/hc-sc/documents/services/publications/food-nutrition/highlights-faits-saillants-2020/highlights-2020.pdf

118. Alkerwi A, Baydarlioglu B, Sauvageot N, Stranges S, Lemmens P, Shivappa N, et al. Smoking status is inversely associated with overall diet quality: Findings from the ORISCAV-LUX study. Clin Nutr. 2017 Oct;36(5):1275–82.

119. Centers for Disease Control and Prevention. Surveillance of Fruit and Vegetable Intake Using the Behavioral Risk Factor Surveillance System [Internet]. Available from: https://www.cdc.gov/brfss/pdf/fruits_vegetables.pdf

120. Rijk MGD, Slotegraaf AI, Brouwer-Brolsma EM, Perenboom CWM, Feskens EJM, Vries JHMD. Development and evaluation of a diet quality screener to assess adherence to the Dutch food-based dietary guidelines. Br J Nutr [Internet]. 2021; Available from: https://pubmed.ncbi.nlm.nih.gov/34776025/

121. Schröder H, Fitó M, Estruch R, Martínez-González MA, Corella D, Salas-Salvadó J, et al. A short screener is valid for assessing Mediterranean diet adherence among older Spanish men and women. J Nutr. 2011 Jun;141(6):1140–5.

122. Gicevic S, Mou Y, Bromage S, Fung T, Willett W. Development of a diet quality screener for global use: Evaluation in a sample of US women. J Acad Nutr Diet. 2021;121(5):854–71.

123. Canada H. Canada's Food Guide Research Tools - Data and Resources - Open Government Portal [Internet]. [cited 2024 Feb 23]. Available from: https://open.canada.ca/data/en/dataset/4f1c44a6-1ecf-4da7-a1da-058b1ff9ce06

124. Kirkpatrick SI, Reedy J, Krebs-Smith SM, Pannucci TRE, Subar AF, Wilson MM, et al. Applications of the Healthy Eating Index for surveillance, epidemiology, and intervention research: Considerations and caveats. J Acad Nutr Diet. 2018 Sep;118(9):1603.

125. Le Glaz A, Haralambous Y, Kim-Dufor DH, Lenca P, Billot R, Ryan TC, et al. Machine learning and natural language processing in mental health: Systematic review. J Med Internet Res. 2021 May 4;23(5):e15708.

126. Payedimarri AB, Concina D, Portinale L, Canonico M, Seys D, Vanhaecht K, et al. Prediction models for public health containment measures on COVID-19 using artificial intelligence and machine learning: A systematic review. Int J Environ Res Public Health. 2021 Apr 23;18(9):4499.

127. Morgenstern JD, Buajitti E, O'Neill M, Piggott T, Goel V, Fridman D, et al. Predicting population health with machine learning: a scoping review. BMJ Open. 2020 Oct;10(10):e037860.

128. Olstad DL, Nejatinamini S, Victorino C, Kirkpatrick SI, Minaker LM, McLaren L. Trends in socioeconomic inequities in diet quality between 2004 and 2015 among a nationally representative sample of children in Canada. J Nutr. 2021 Dec;151(12):3781–94.

129. Green MA, Evans CR, Subramanian SV. Can intersectionality theory enrich population health research? Soc Sci Med. 2017 Apr 1;178:214–6.

130. Holscher HD. Diving into dietary pattern and dietary diversity analyses. Am J Clin Nutr. 2024 May 1;119(5):1095–6.

131. Russo S, Bonassi S. Prospects and pitfalls of machine learning in nutritional epidemiology. Nutrients [Internet]. 2022 May;14(9). Available from: /pmc/articles/PMC9105182/ /pmc/articles/PMC9105182/?report=abstract https://www.ncbi.nlm.nih.gov/pmc/articles/PMC9105182/

132. Nshimyumukiza L, Lieffers JRL, Ekwaru JP, Ohinmaa A, Veugelers PJ. Temporal changes in diet quality and the associated economic burden in Canada. PLOS ONE. 2018 Nov;13(11):e0206877.

133. Bouchard-Mercier A, Rudkowska I, Lemieux S, Couture P, Vohl MC. The metabolic signature associated with the Western dietary pattern: A cross-sectional study. Nutr J. 2013 Dec 11;12(1):158.

134. Chen Z, Liu L, Roebothan B, Ryan A, Colbourne J, Baker N, et al. Four major dietary patterns identified for a target-population of adults residing in Newfoundland and Labrador, Canada. BMC Public Health. 2015 Jan 31;15(1):69.

135. Chen Z, Wang PP, Woodrow J, Zhu Y, Roebothan B, Mclaughlin JR, et al. Dietary patterns and colorectal cancer: results from a Canadian population-based study. Nutr J. 2015 Jan 15;14(1):8.

136. Langsetmo L, Poliquin S, Hanley DA, Prior JC, Barr S, Anastassiades T, et al. Dietary patterns in Canadian men and women ages 25 and older: relationship to demographics, body mass index, and bone mineral density. BMC Musculoskelet Disord. 2010 Jan 28;11(1):20.

137. Trudeau K, Rousseau MC, Barul C, Csizmadi I, Parent MÉ. Dietary Patterns Are Associated with Risk of Prostate Cancer in a Population-Based Case-Control Study in Montreal, Canada. Nutrients. 2020 Jul;12(7):1907.

138. Walker M, Aronson KJ, King W, Wilson JWL, Fan W, Heaton JPW, et al. Dietary patterns and risk of prostate cancer in Ontario, Canada. Int J Cancer. 2005;116(4):592–8.

139. Zhang J, Fei J, Song X, Feng J. An improved Louvain Algorithm for community detection. Math Probl Eng [Internet]. 2021 [cited 2024 Jun 18]; Available from: https://onlinelibrary.wiley.com/doi/full/10.1155/2021/1485592

140. Williams TE. Cognitive Testing of a Brief Dietary Screener: A Comparison of Analysis Methods. University of Waterloo; 2022.

141. 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 Apr;181(7):473–87.

142. Freedman LS, Commins JM, Moler JE, Arab L, Baer DJ, Kipnis V, et al. Pooled results from 5 validation studies of dietary self-report instruments using recovery biomarkers for energy and protein intake. Am J Epidemiol. 2014 Jul;180(2):172–88.

143. Hutchinson JM, Dodd KW, Guenther PM, Lamarche B, Haines J, Wallace A, et al. The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019 Canada's Food Guide healthy food choices recommendations: scoring system and construct validity. Appl Physiol Nutr Metab. 2023 Aug;48(8):620–33.

144. Health Canada. Canadian Nutrient File (CNF) [Internet]. 2023 [cited 2023 Mar 9]. Available from: https://food-nutrition.canada.ca/cnf-fce/?lang=eng

145. National Research Council Canada. Plant-based protein market: global and Canadian market analysis [Internet]. 2019 [cited 2023 Apr 17]. Available from: https://nrc.canada.ca/en/research-development/research-collaboration/programs/plant-based-protein-market-global-canadian-market-analysis

146. National Cancer Institute. ASA24-Canada-2018 [Internet]. 2022 [cited 2022 Aug 5]. Available from: https://epi.grants.cancer.gov/asa24/respondent/asa24-canada-2018.html

147. Zhang S, Midthune D, Guenther PM, Krebs-Smith SM, Kipnis V, Dodd KW, et al. A new multivariate measurement error model with zero-inflated dietary data, and its application to dietary assessment. Ann Appl Stat. 2011 Jun;5(2B):1456–87.

148. Measurement Error in Epidemiology [Internet]. [cited 2023 Apr 23]. Available from: https://prevention.cancer.gov/research-groups/biometry/measurement-error-impact/measurement-error-epidemiology

149. Keogh RH, White IR. A toolkit for measurement error correction, with a focus on nutritional epidemiology. Stat Med. 2014 May 30;33(12):2137–55.

150. Dietary pattern and diversity analysis using DietDiveR in R: a cross-sectional evaluation in the National Health and Nutrition Examination Survey - ScienceDirect [Internet]. [cited 2024 May 3]. Available from: https://www.sciencedirect.com/science/article/pii/S000291652400159X

151. Government of Canada, Canadian Institutes of Health Research. Heterogeneity in Nutrition and Metabolic Health: Virtual Workshop Report Summary - CIHR [Internet]. 2023 [cited 2024 May 16]. Available from: https://cihr-irsc.gc.ca/e/53359.html

Appendices

Appendix A: Advisors who provided input on the development of the Canadian Food Intake Screener

- Meghan Day, British Columbia Ministry of Health
- Kevin Dodd, U.S. National Cancer Institute
- Patricia Guenther, University of Utah
- Jess Haines, University of Guelph
- Mahsa Jessri, University of British Columbia
- Mary L'Abbé, University of Toronto
- Benoît Lamarche, Université Laval
- Simone Lemieux, Université Laval
- Maria Laura Louzada, University of São Paulo
- Dana Lee Olstad, University of Calgary
- Rachel Prowse, Memorial University of Newfoundland
- Janis Randall Simpson, University of Guelph
- Jill Reedy, U.S. National Cancer Institute
- Hassan Vatanparast, University of Saskatchewan
- Jennifer Vena, Alberta Health Services

Appendix B: Initial questions in the Canadian Food Intake Screener/ Questionnaire court canadien sur les apports alimentaires to assess alignment of adults' intake with the Canada Food Guide-2019 healthy food choices recommendations

NOTE: THIS IS THE VERSION OF THE SCREENER THAT WAS TESTED IN COGNITIVE INTERVIEWS AND FACE AND CONTENT VALIDITY TESTING.

SEE APPENDIX C FOR THE FINAL SCREENER.

La version Française suit

These questions are about foods and beverages you ate or drank in the past month. When answering, please include meals and snacks consumed at home, at work or school, in restaurants, and anyplace else.

1. Over the past month, how often did you consume fresh, frozen, and canned fruit? *Do **not** include fruit juices or drinks*.
   - ☐ Never
   - ☐ 1 time in the past month
   - ☐ 2-3 times in the past month
   - ☐ 1-2 times per week
   - ☐ 3-4 times per week
   - ☐ 5-6 times per week
   - ☐ 1 time per day
   - ☐ 2-3 times per day
   - ☐ 4-5 times per day
   - ☐ 6 or more times per day

2. Over the past month, how often did you consume cooked, raw, frozen, and canned vegetables? *Do **not** include deep-fried vegetables or vegetable juices or drinks*.
   - ☐ Never
   - ☐ 1 time in the past month
   - ☐ 2-3 times in the past month
   - ☐ 1-2 times per week
   - ☐ 3-4 times per week
   - ☐ 5-6 times per week
   - ☐ 1 time per day
   - ☐ 2-3 times per day
   - ☐ 4-5 times per day

- 6 or more times per day

3. Over the past month, how often did you consume deep-fried foods and ready-to-heat or ready-to-eat dishes? Include frozen, canned, and packaged meat-based and vegetarian/vegan dishes, such as soups, pre-made pasta and rice dishes, and plant-based meats. *Do **not** include hot dogs, sausages, ham, corned beef, beef jerky or other deli or luncheon meats*.
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day
   - 6 or more times per day

4. Over the past month, how often did you consume hot dogs, sausages, ham, corned beef, beef jerky, and other deli or luncheon meats? *Do **not** include canned fish or canned poultry or packaged plant-based meats.*
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day
   - 6 or more times per day

5. Over the past month, how often did you consume lean red meat or pork, wild game, poultry, fish, shellfish, and eggs? Include canned fish and canned poultry. *Do **not** include hot dogs, sausages, ham, corned beef, beef jerky, or other deli or luncheon meats*.
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day

- 2-3 times per day
- 4-5 times per day
- 6 or more times per day

6. Over the past month, how often did you consume nuts (including nut butters), seeds, tofu, beans (e.g., chickpeas, hummus, lentils, black beans), and other plant-based sources of protein? *Do **not** include green beans or ready-to-heat foods, such as pre-made veggie burgers and packaged plant-based meats*.
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day
   - 6 or more times per day

7. Over the past month, how often did you drink milk and unsweetened plant-based beverages (e.g., soy, almond, and oat milk)? *Do **not** include small amounts in coffee or tea, or chocolate and other flavoured milk*.
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day
   - 6 or more times per day

8. Over the past month, how often did you consume yogurt, kefir, and cheese?
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day

- 6 or more times per day

9. Over the past month, how often did you drink flavoured milk, sweetened plant-based beverages (e.g., soy, almond, and oat milk), and sweetened coffee or tea (including bottled)? *Do **not** include diet or sugar-free beverages*.
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day
   - 6 or more times per day

10. Over the past month, how often did you drink soda or pop containing sugar, fruit juice and fruit-flavoured drinks, sweetened sports drinks, and sweetened waters? *Do **not** include diet or sugar-free beverages, such as diet soda and plain water*.
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day
   - 6 or more times per day

11. Over the past month, how often did you consume sugary breakfast cereals, cookies, cakes, pastries, granola bars, ice cream, candy, and chocolate?
   - Never
   - 1 time in the past month
   - 2-3 times in the past month
   - 1-2 times per week
   - 3-4 times per week
   - 5-6 times per week
   - 1 time per day
   - 2-3 times per day
   - 4-5 times per day
   - 6 or more times per day

12. Over the past month, how often did you consume chips, pretzels, popcorn, crackers, and other salty snacks?

- Never
- 1 time in the past month
- 2-3 times in the past month
- 1-2 times per week
- 3-4 times per week
- 5-6 times per week
- 1 time per day
- 2-3 times per day
- 4-5 times per day
- 6 or more times per day

13. Over the past month, how often did you consume white breads, bagels, rice, pasta, and noodles? *Do **not** include whole wheat or whole grain foods*.

- Never
- 1 time in the past month
- 2-3 times in the past month
- 1-2 times per week
- 3-4 times per week
- 5-6 times per week
- 1 time per day
- 2-3 times per day
- 4-5 times per day
- 6 or more times per day

14. Over the past month, how often did you consume whole wheat or whole grain breads, rice, pasta, noodles, and cereals? *Do **not** include white breads, rice, pasta, noodles, or cereals*.

- Never
- 1 time in the past month
- 2-3 times in the past month
- 1-2 times per week
- 3-4 times per week
- 5-6 times per week
- 1 time per day
- 2-3 times per day
- 4-5 times per day
- 6 or more times per day

15. Over the past month, how often did you cook with or add vegetable oils (e.g., canola, olive, sunflower) or soft margarines to your foods? *Do **not** include coconut oil, palm oil, or butter*.

- Never
- 1 time in the past month
- 2-3 times in the past month
- 1-2 times per week
- 3-4 times per week
- 5-6 times per week
- 1 time per day
- 2-3 times per day
- 4-5 times per day
- 6 or more times per day

Ces questions portent sur les aliments et les boissons que vous avez consommés au cours du **dernier mois**, c'est-à-dire au cours des 30 derniers jours. Lorsque vous répondez, veuillez inclure dans vos réponses les repas et les collations consommés à la maison, au travail, à l'école, au restaurant et à tout autre endroit.

1. Au cours du dernier mois, à quelle fréquence avez-vous consommé des fruits frais, congelés ou en conserve? **N**'incluez **pas** les jus et les boissons de fruits.

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

2. Au cours du dernier mois, à quelle fréquence avez-vous consommé des légumes cuits, crus, congelés ou en conserve? **N'**incluez **pas** les légumes frits, les jus de légumes et les boissons de légumes.

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

3. Au cours du dernier mois, à quelle fréquence avez-vous consommé des aliments frits et des plats prêts à réchauffer ou prêts à manger?

Incluez les plats congelés, en conserve et emballés à base de viande et les plats végétariens/végétaliens, tels que les soupes, les plats de pâtes et de riz préparés, et les viandes *à base de produits d'origine végétale*. ***N'****incluez* ***pas*** *les hot-dogs, les saucisses, le jambon, le bœuf salé, le bœuf séché et d'autres viandes de charcuterie ou froides*.

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

4. Au cours du dernier mois, à quelle fréquence avez-vous consommé des hot-dogs, des saucisses, du jambon, du bœuf salé, du bœuf séché et d'autres viandes de charcuterie ou froides? ***N'****incluez* ***pas*** *le poisson en conserve, la volaille en conserve et les viandes à base de produits d'origine végétale emballées.*

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

5. Au cours du dernier mois, à quelle fréquence avez-vous consommé de la viande rouge maigre, du porc maigre, du gibier, de la volaille, du poisson, des crustacés et des œufs? Incluez le poisson en conserve et la volaille en conserve. ***N'****incluez* ***pas*** *les hot-dogs, les saucisses, le jambon, bœuf salé, le bœuf séché ou d'autres viandes de charcuterie ou froides.*

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour

- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

6. Au cours du dernier mois, à quelle fréquence avez-vous consommé des noix (y compris des beurres de noix), des graines, du tofu, des légumineuses (par exemple, des pois chiches, du houmous, des lentilles, des haricots noirs) et d'autres sources de protéines d'origine végétale? ***N'****incluez* ***pas*** *les haricots verts ou les aliments prêts à réchauffer, comme les hamburgers végétariens prépréparés et les viandes à base de produits d'origine végétale emballées.*

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

7. Au cours du dernier mois, à quelle fréquence avez-vous bu du lait et des boissons d'origine végétale non sucrées (par exemple, du lait de soja, aux amandes et d'avoine)? ***N'****incluez* ***pas*** *les petites quantités dans le café ou le thé, ni le lait au chocolat et autres laits aromatisés*.

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

8. Au cours du dernier mois, à quelle fréquence avez-vous consommé du yogourt, du kéfir et du fromage?

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour

- 4 - 5 fois par jour
- Au moins 6 fois par jour

9. Au cours du dernier mois, à quelle fréquence avez-vous bu du lait aromatisé, des boissons d'origine végétale sucrées (par exemple, du lait de soja, aux amandes et d'avoine) et du café ou thé sucré (y compris en bouteille)? ***N'****incluez* ***pas*** *les boissons diètes ou sans sucre.*

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

10. Au cours du dernier mois, à quelle fréquence avez-vous bu des boissons gazeuses contenant du sucre, des jus de fruits et des boissons aromatisées aux fruits, des boissons sportives sucrées et des eaux sucrées? ***N'****incluez* ***pas*** *les boissons diètes ou sans sucre, comme les boissons gazeuses diètes et l'eau plate.*

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

11. Au cours du dernier mois, à quelle fréquence avez-vous consommé des céréales sucrées pour le déjeuner, des biscuits, des gâteaux, des pâtisseries, des barres tendres, de la crème glacée, des bonbons et du chocolat?

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

12. Au cours du dernier mois, à quelle fréquence avez-vous consommé des croustilles, des bretzels, du maïs soufflé, des craquelins et autres collations salées?

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

13. Au cours du dernier mois, à quelle fréquence avez-vous consommé des pâtes, des nouilles, du pain, des bagels et du riz **blancs**? *N'incluez **pas** les aliments à base de blé entier ou de grains entiers*.

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

14. Au cours du dernier mois, à quelle fréquence avez-vous consommé du pain, du riz, des pâtes, des nouilles et des céréales **de blé entier ou de grains entiers**? ***N'****incluez **pas** le pain blanc, le riz blanc, les pâtes et les nouilles blanches et les céréales raffinées.*

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

15. Au cours du dernier mois, combien de fois avez-vous cuisiné avec des huiles végétales (par exemple, de canola, d'olive, de tournesol) ou des margarines molles ou en avez-vous ajoutées à vos aliments? ***N'****incluez **pas** l'huile de coco, l'huile de palme ou le beurre.*

- Jamais
- 1 fois au cours du dernier mois
- 2 - 3 fois au cours du dernier mois
- 1 - 2 fois par semaine
- 3 - 4 fois par semaine
- 5 - 6 fois par semaine
- 1 fois par jour
- 2 - 3 fois par jour
- 4 - 5 fois par jour
- Au moins 6 fois par jour

Appendix C: Canadian Food Intake Screener in English and French

Canadian Food Intake Screener

These questions are about foods and beverages you ate or drank in the past month, that is, the past 30 days. When answering, please include meals and snacks consumed at home, at work or school, in restaurants, and anyplace else.

1. Over the past month, how often did you eat fresh, frozen, canned, or dried fruit?

   Do not include fruit juices and drinks.

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

2. Over the past month, how often did you eat potatoes, including baked, boiled, or mashed potatoes, or sweet potatoes?

   Do not include french fries, poutine, home fries, or hash browns.

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

3. Over the past month, how often did you eat fresh, cooked, frozen, or canned vegetables?

   Do not include potatoes, french fries, poutine, or other deep-fried vegetables, or vegetable juices and drinks.

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

4. Over the past month, how often did you eat food from fast food restaurants, such as burgers, french fries, poutine, pizza, submarine sandwiches, fried chicken, burritos, or tacos?

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

5. Over the past month, how often did you eat hot dogs, sausages, beef jerky, bacon, ham or other deli or luncheon meats?

   Do not include fast food, canned fish, canned poultry, or packaged veggie burgers and plant-based meats.

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

6. Over the past month, how often did you eat eggs, beef, pork, wild meat, chicken or other poultry, fish, shellfish, or other animal-based sources of protein? Include canned fish and canned poultry.

   Do not include fast food, hot dogs, sausages, beef jerky, bacon, ham, or other deli or luncheon meats.

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

7. Over the past month, how often did you eat nuts, seeds, tofu, beans and lentils, peanut butter or other nut butters, or other plant-based sources of protein?

   Do not include green beans or packaged veggie burgers and plant-based meats.

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

8. Over the past month, how often did you eat yogurt, kefir, or cheese?

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

9. Over the past month, how often did you have **white** cows' milk or **unsweetened** plant-based beverages (e.g., soy, almond, or oat milk)?

   Do not include small amounts in coffee or tea, or chocolate and other sweetened milk.

   - ❑ Never
   - ❑ 1 time in the past month
   - ❑ 2-3 times in the past month
   - ❑ 1-2 times per week
   - ❑ 3-4 times per week
   - ❑ 5-6 times per week
   - ❑ 1 time per day
   - ❑ 2-3 times per day
   - ❑ 4-5 times per day
   - ❑ 6 or more times per day

10. Over the past month, how often did you have chocolate milk or other **flavoured** milk or **sweetened** plant-based beverages (e.g., soy, almond, or oat milk)?

    Do not include small amounts in coffee or tea, or diet/artificially sweetened or sugar-free beverages.

    - ❑ Never
    - ❑ 1 time in the past month
    - ❑ 2-3 times in the past month
    - ❑ 1-2 times per week
    - ❑ 3-4 times per week
    - ❑ 5-6 times per week
    - ❑ 1 time per day
    - ❑ 2-3 times per day
    - ❑ 4-5 times per day
    - ❑ 6 or more times per day

11. Over the past month, how often did you drink fruit juice, fruit-flavoured drinks, soda or pop, **sweetened** sports or energy drinks, **sweetened** hot or iced coffee or tea, or **sweetened** waters?

    Do not include diet/artificially sweetened or sugar-free beverages, such as diet soda.

    - ❑ Never
    - ❑ 1 time in the past month
    - ❑ 2-3 times in the past month
    - ❑ 1-2 times per week
    - ❑ 3-4 times per week
    - ❑ 5-6 times per week
    - ❑ 1 time per day
    - ❑ 2-3 times per day
    - ❑ 4-5 times per day
    - ❑ 6 or more times per day

12. Over the past month, how often did you eat cookies, cakes, muffins, pastries, granola bars, protein bars, ice cream, candy, chocolate, sugary breakfast cereals, or other sugary foods?

- ❑ Never
- ❑ 1 time in the past month
- ❑ 2-3 times in the past month
- ❑ 1-2 times per week
- ❑ 3-4 times per week
- ❑ 5-6 times per week
- ❑ 1 time per day
- ❑ 2-3 times per day
- ❑ 4-5 times per day
- ❑ 6 or more times per day

13. Over the past month, how often did you eat crackers, chips, pretzels, popcorn, or other salty snacks?

- ❑ Never
- ❑ 1 time in the past month
- ❑ 2-3 times in the past month
- ❑ 1-2 times per week
- ❑ 3-4 times per week
- ❑ 5-6 times per week
- ❑ 1 time per day
- ❑ 2-3 times per day
- ❑ 4-5 times per day
- ❑ 6 or more times per day

14. Over the past month, how often did you eat **white** breads, bagels, rice, pasta, noodles, or other refined grains, such as breakfast cereals?

Do not include whole wheat or whole grain foods.

- ❑ Never
- ❑ 1 time in the past month
- ❑ 2-3 times in the past month
- ❑ 1-2 times per week
- ❑ 3-4 times per week
- ❑ 5-6 times per week
- ❑ 1 time per day
- ❑ 2-3 times per day
- ❑ 4-5 times per day
- ❑ 6 or more times per day

15. Over the past month, how often did you eat **whole wheat or whole grain** breads, bagels, pasta, noodles, quinoa, oats, brown or wild rice, breakfast cereals, or other whole wheat or whole grain foods?

    Do not include white breads, bagels, pasta, noodles, rice, or refined breakfast cereals.

    - ❑ Never
    - ❑ 1 time in the past month
    - ❑ 2-3 times in the past month
    - ❑ 1-2 times per week
    - ❑ 3-4 times per week
    - ❑ 5-6 times per week
    - ❑ 1 time per day
    - ❑ 2-3 times per day
    - ❑ 4-5 times per day
    - ❑ 6 or more times per day

16. Over the past month, how often did you have margarine or vegetable oils (e.g., olive, canola, or sunflower oil)?

    Do **not** include lard, coconut oil, palm oil, or butter.

    - ❑ Never
    - ❑ 1 time in the past month
    - ❑ 2-3 times in the past month
    - ❑ 1-2 times per week
    - ❑ 3-4 times per week
    - ❑ 5-6 times per week
    - ❑ 1 time per day
    - ❑ 2-3 times per day
    - ❑ 4-5 times per day
    - ❑ 6 or more times per day

Questionnaire court canadien sur les apports alimentaires

Ces questions portent sur les aliments et les boissons que vous avez consommés **au cours du dernier mois**, c'est-à-dire au cours des 30 derniers jours. Lorsque vous répondez, veuillez inclure dans vos réponses **tous** les repas et **toutes** les collations consommés à la maison, au travail, à l'école, au restaurant et à tout autre endroit.

1. Au cours du dernier mois, à quelle fréquence avez-vous consommé des fruits frais, congelés, en conserve ou séchés?

   N'incluez pas les jus de fruits et les boissons aux fruits.

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

2. Au cours du dernier mois, à quelle fréquence avez-vous consommé des pommes de terre, y compris des pommes de terre au four, bouillies ou en purée, ou des patates douces?

   N'incluez pas les frites, les frites maison, la poutine, les pommes de terre rissolées et les galettes de pommes de terre (hash browns).

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

3. Au cours du dernier mois, à quelle fréquence avez-vous consommé des légumes crus, cuits, congelés ou en conserve?

   N'incluez pas les pommes de terre, les frites ou autres légumes frits, la poutine, les jus de légumes et les boissons de légumes.

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

4. Au cours du dernier mois, à quelle fréquence avez-vous consommé de la nourriture provenant d'un restaurant de type fast-food, comme des hamburgers, des frites, de la poutine, de la pizza, des sous-marins, du poulet frit, des burritos ou des tacos?

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

5. Au cours du dernier mois, à quelle fréquence avez-vous consommé des hot-dogs, des saucisses, du bœuf séché (*beef jerky*), du bacon, du jambon ou autres viandes de charcuterie ou froides?

   N'incluez pas le fast-food, le poisson en conserve, le poulet ou autre volaille en conserve, les hamburgers végétariens prépréparés du commerce et les simili-viandes d'origine végétale prépréparées du commerce.

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

6. Au cours du dernier mois, à quelle fréquence avez-vous consommé des œufs, du bœuf, du porc, du gibier, du poulet ou autre volaille, du poisson, des fruits de mer ou d'autres sources de protéines d'origine animale? Incluez le poisson en conserve et le poulet ou autre volaille en conserve.

   N'incluez pas le fast-food, les hot-dogs, les saucisses, le bœuf séché (*beef jerky*), le bacon, le jambon et autres viandes de charcuterie ou froides.

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

7. Au cours du dernier mois, à quelle fréquence avez-vous consommé des arachides, des noix, des graines, du tofu, des lentilles ou autres légumineuses, du beurre d'arachides ou autres beurres de noix, ou d'autres sources de protéines d'origine végétale?

   N'incluez pas les haricots verts et les hamburgers végétariens prépréparés du commerce et les simili-viandes d'origine végétale prépréparées du commerce.

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

8. Au cours du dernier mois, à quelle fréquence avez-vous consommé du yogourt, du kéfir ou du fromage?

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

9. Au cours du dernier mois, à quelle fréquence avez-vous consommé du lait de vache ou des boissons d'origine végétale **non sucrées** (par exemple, du lait de soya, du lait aux amandes, du lait d'avoine)?

   N'incluez pas les petites quantités dans le café ou le thé, le lait au chocolat et autres laits aromatisés sucrés.

   - ❒ Jamais
   - ❒ 1 fois au cours du dernier mois
   - ❒ 2 - 3 fois au cours du dernier mois
   - ❒ 1 - 2 fois par semaine
   - ❒ 3 - 4 fois par semaine
   - ❒ 5 - 6 fois par semaine
   - ❒ 1 fois par jour
   - ❒ 2 - 3 fois par jour
   - ❒ 4 - 5 fois par jour
   - ❒ 6 fois ou plus par jour

10. Au cours du dernier mois, à quelle fréquence avez-vous consommé du lait au chocolat ou autres laits **aromatisés sucrés** ou des boissons d'origine végétale **sucrées** (par exemple, du lait de soya, du lait aux amandes, du lait d'avoine).

    N'incluez pas les petites quantités dans le café ou le thé, les boissons diètes, les boissons sucrées avec des édulcorants et les boissons sans sucre.

    - ❒ Jamais
    - ❒ 1 fois au cours du dernier mois
    - ❒ 2 - 3 fois au cours du dernier mois
    - ❒ 1 - 2 fois par semaine
    - ❒ 3 - 4 fois par semaine
    - ❒ 5 - 6 fois par semaine
    - ❒ 1 fois par jour
    - ❒ 2 - 3 fois par jour
    - ❒ 4 - 5 fois par jour
    - ❒ 6 fois ou plus par jour

11. Au cours du dernier mois, à quelle fréquence avez-vous consommé des jus de fruits, des boissons aromatisées aux fruits, des boissons gazeuses **contenant du sucre**, des boissons sportives **sucrées**, des boissons énergisantes **sucrées**, des cafés **sucrés** ou thés **sucrés,** chauds ou glacés, ou des eaux **sucrées**?

    N'incluez pas les boissons diètes, les boissons sucrées avec des édulcorants et les boissons sans sucre, comme les boissons gazeuses diètes.

    - ❒ Jamais
    - ❒ 1 fois au cours du dernier mois
    - ❒ 2 - 3 fois au cours du dernier mois
    - ❒ 1 - 2 fois par semaine
    - ❒ 3 - 4 fois par semaine
    - ❒ 5 - 6 fois par semaine
    - ❒ 1 fois par jour
    - ❒ 2 - 3 fois par jour
    - ❒ 4 - 5 fois par jour
    - ❒ 6 fois ou plus par jour

12. Au cours du dernier mois, à quelle fréquence avez-vous consommé des biscuits, des gâteaux, des muffins, des pâtisseries, des barres tendres, des barres protéinées, de la crème glacée, des bonbons, du chocolat, des céréales à déjeuner sucrées ou d'autres aliments sucrés?

    - ❒ Jamais
    - ❒ 1 fois au cours du dernier mois
    - ❒ 2 - 3 fois au cours du dernier mois
    - ❒ 1 - 2 fois par semaine
    - ❒ 3 - 4 fois par semaine
    - ❒ 5 - 6 fois par semaine
    - ❒ 1 fois par jour
    - ❒ 2 - 3 fois par jour
    - ❒ 4 - 5 fois par jour
    - ❒ 6 fois ou plus par jour

13. Au cours du dernier mois, à quelle fréquence avez-vous consommé des craquelins, des croustilles, des bretzels, du maïs soufflé ou autres collations salées?

- ❒ Jamais
- ❒ 1 fois au cours du dernier mois
- ❒ 2 - 3 fois au cours du dernier mois
- ❒ 1 - 2 fois par semaine
- ❒ 3 - 4 fois par semaine
- ❒ 5 - 6 fois par semaine
- ❒ 1 fois par jour
- ❒ 2 - 3 fois par jour
- ❒ 4 - 5 fois par jour
- ❒ 6 fois ou plus par jour

14. Au cours du dernier mois, à quelle fréquence avez-vous consommé du pain **blanc**, des bagels **blancs**, du riz **blanc**, des pâtes et des nouilles **blanches** ou d'autres grains **raffinés**, comme les céréales à déjeuner raffinées?

    N'incluez pas les aliments à base de blé entier ou de grains entiers.

- ❒ Jamais
- ❒ 1 fois au cours du dernier mois
- ❒ 2 - 3 fois au cours du dernier mois
- ❒ 1 - 2 fois par semaine
- ❒ 3 - 4 fois par semaine
- ❒ 5 - 6 fois par semaine
- ❒ 1 fois par jour
- ❒ 2 - 3 fois par jour
- ❒ 4 - 5 fois par jour
- ❒ 6 fois ou plus par jour

15. Au cours du dernier mois, à quelle fréquence avez-vous consommé du pain, des bagels, des pâtes, des nouilles, du quinoa, du gruau, du riz brun ou sauvage, des céréales à déjeuner ou tout autre **aliment fait de blé entier ou de grains entiers**?

    N'incluez pas le pain blanc, les bagels blancs, les pâtes ou les nouilles blanches, le riz blanc et les céréales à déjeuner raffinées.

    - ❒ Jamais
    - ❒ 1 fois au cours du dernier mois
    - ❒ 2 - 3 fois au cours du dernier mois
    - ❒ 1 - 2 fois par semaine
    - ❒ 3 - 4 fois par semaine
    - ❒ 5 - 6 fois par semaine
    - ❒ 1 fois par jour
    - ❒ 2 - 3 fois par jour
    - ❒ 4 - 5 fois par jour
    - ❒ 6 fois ou plus par jour

16. Au cours du dernier mois, à quelle fréquence avez-vous consommé de la margarine ou des huiles végétales (par exemple, l'huile d'olive, l'huile de canola, l'huile de tournesol)?

    N'incluez pas le saindoux, l'huile de coco, l'huile de palme et le beurre.

    - ❒ Jamais
    - ❒ 1 fois au cours du dernier mois
    - ❒ 2 - 3 fois au cours du dernier mois
    - ❒ 1 - 2 fois par semaine
    - ❒ 3 - 4 fois par semaine
    - ❒ 5 - 6 fois par semaine
    - ❒ 1 fois par jour
    - ❒ 2 - 3 fois par jour
    - ❒ 4 - 5 fois par jour
    - ❒ 6 fois ou plus par jour

Appendix D: Details about the Automated Self-Administered Dietary Assessment Tool (ASA24-Canada-2018) and cleaning of 24-hour dietary recall data

About ASA24

ASA24 was developed by the National Cancer Institute (NCI) (Subar et al. 2012) and adapted to the Canadian food supply by Health Canada (National Cancer Institute 2022a). ASA24-Canada-2018 prompts participants to report everything they ate and drank from midnight to midnight on the previous day using multiple passes adapted from the interviewer-administered Automated Multiple-Pass Method (Blanton et al. 2006; Moshfegh et al. 2008), used in the nutrition-focused cycles of the CCHS (Health Canada 2006, 2017). ASA24 recall data have been shown to be comparable to those based on interviewer-administered 24HDR and to observed intake in feeding studies (Kirkpatrick et al. 2014, 2019).

The two 24HDRs were spread over approximately one month to align with the period captured by the screener, as well as across days of the week at the group level to permit accounting for potential weekend day vs. weekday effects. Although the screener captured frequency of intake during the month prior to the survey, the 24HDRs were administered after the survey to avoid the possibility that engaging in a detailed accounting of dietary intake might influence responses to the screener. As well, it was hypothesized that overall diet quality would not vary substantially over a few months (Bernstein et al. 2016). Invitations to complete the first recall were sent via email eight to ten days after completion of the survey, with up to three reminders, every two days, in cases in which the recall was not completed in response to the prior email.

Cleaning ASA24-Canada-2018 data

Recalls were reviewed by a registered dietitian (JMH), with input from a second registered dietitian (SIK) and a Health Canada nutritionist. Based on this review, four first 24HDRs and three second 24HDRs were removed because they were completed in less than five minutes (n=2), were incomplete (n=1), or had low or high estimated energy intakes (≤600 kcal or ≥4400 kcal for women and ≤650 kcal or ≥5700 kcal for men) combined with an indication by the participant that intake on the recalled day was usual (n=4). This resulted in 128 first recalls (83% of the analytic sample) and 102 second recalls (66%) included in the analyses.

Known ASA24 database issues identified by NCI (National Cancer Institute 2022b) were manually corrected, and portion size outliers were investigated and adjusted in two instances. For example, a respondent reported consuming 6.25 cups of potato salad. This was decreased to three cups, reflecting a more plausible portion size but also that the participant selected greater than the largest option, more than two cups. The default food codes applied by ASA24-Canada-2018 to free-text entries, entered when a participant indicated they could not find the food they were looking for, were investigated and corrected in cases in which they did not match (Zimmerman et al. 2015). For example, a code for chicken breast was assigned to a free-text entry of burrito bowl, based on the food category indicated by the respondent, and manually adjusted to burrito. Corrections to free-text entries caused changes to 116 food codes (1.8% of all food codes in the dataset).

Appendix D References

Bernstein, S., Zambell, K., Amar, M.J., Arango, C., Kelley, R.C., Miszewski, S.G., Tryon, S., and Courville, A.B. 2016. Dietary Intake Patterns Are Consistent Across Seasons in a Cohort of Healthy Adults in a Metropolitan Population. J Acad Nutr Diet 116(1): 38–45. Doi:10.1016/J.JAND.2015.08.008.

Blanton, C.A., Moshfegh, A.J., Baer, D.J., and Kretsch, M.J. 2006. The USDA Automated Multiple-Pass Method accurately estimates group total energy and nutrient intake. J Nutr 136(10): 2594–2599. Doi:10.1093/jn/136.10.2594.

Health Canada. 2006. Canadian Community Health Survey Cycle 2.2, Nutrition (2004) – A guide to accessing and interpreting the data. Available from https://www.canada.ca/content/dam/hc-sc/migration/hc-sc/fn-an/alt_formats/hpfb-dgpsa/pdf/surveill/cchs-guide-escc-eng.pdf [Accessed 30 November 2022].

Health Canada. 2017. Reference guide to understanding the data- 2015 Canadian Community Health Survey-Nutrition. Available from https://www.canada.ca/en/health-canada/services/food-nutrition/food-nutrition-surveillance/health-nutrition-surveys/canadian-community-health-survey-cchs/reference-guide-understanding-using-data-2015.html [Accessed 30 November 2022].

Kirkpatrick, S.I., Subar, A.F., Douglass, D., Zimmerman, T.P., Thompson, F.E., Kahle, L.L., George, S.M., Dodd, K.W., and Potischman, N. 2014. Performance of the Automated Self-Administered 24-hour recall relative to a measure of true intakes and to an interviewer-administered 24-h recall. Am J Clin Nutr 100(1): 233–40. Doi:10.3945/ajcn.114.083238.

Kirkpatrick, S.I., Guenther, P.M., Douglass, D., Zimmerman, T., Kahle, L.L., Atoloye, A., Marcinow, M., Savoie-Roskos, M.R., Dodd, K.W., and Durward, C. 2019. The provision of assistance does not substantially impact the accuracy of 24-hour dietary recalls completed using the Automated Self-Administered 24-H dietary assessment tool among women with low incomes. The Journal of Nutrition 149(1): 114–122. Doi:10.1093/jn/nxy207.

Moshfegh, A.J., Rhodes, D.G., Baer, D.J., Murayi, T., Clemens, J.C., Rumpler, W. v, Paul, D.R., Sebastian, R.S., Kuczynski, K.J., Ingwersen, L.A., Staples, R.C., and Cleveland, L.E. 2008. The US Department of Agriculture Automated Multiple-Pass Method reduces bias in the collection of energy intakes. Am J Clin Nutr 88(2): 324–332. Doi:10.1093/ajcn/88.2.324.

National Cancer Institute. 2022a. ASA24-Canada-2018. Available from https://epi.grants.cancer.gov/asa24/respondent/asa24-canada-2018.html [accessed 5 August 2022].

National Cancer Institute. 2022b. ASA24® Known Issues & Workarounds. Available from https://epi.grants.cancer.gov/asa24/resources/issues.html [5 August 2022].

Subar, A.F., Kirkpatrick, S.I., Mittl, B., Zimmerman, T.P., Thompson, F.E., Bingley, C., Willis, G., Islam, N.G., Baranowski, T., McNutt, S., and Potischman, N. 2012. 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 112(8): 1134–1137. doi:10.1016/j.jand.2012.04.016.

Zimmerman, T.P., Potischman, N., Douglass, D., Dixit-Joshi, S., Kirkpatrick, S.I., Subar, A.F., McNutt, S., Coleman, L.A., Alexander, G.L., Kushi, L.H., and Thompson, F.E. 2015. 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 Science 4: 160–172. doi:10.1016/J.PROFOO.2015.06.021.

Appendix E: Canadian Food Intake Screener/Questionnaire court canadien sur les apports alimentaires: Scoring System READ ME File

READ ME

Background

The Canadian Food Intake Screener/Questionnaire court canadien sur les apports alimentaires was developed to rapidly assess alignment of adults' dietary intake over the past month with the 2019 Canada Food Guide's healthy food choices recommendations. The screener includes 16 questions. Arriving at screener scores involves creating components by summing and deriving ratios based on responses to screener questions. Scores on each component are summed to arrive at a total possible score with a maximum of **65 points**.

The screener is designed to provide one total score assessing alignment. The component scores are meant only for the purpose of arriving at total scores.

Details on the scoring system can be found in the accompanying manuscript (Hutchinson et al. 2022).

Suggested layout for input dataset

The sample SAS code was created with the assumption that observations are rows and screener responses are columns.

Sample code

Sample SAS code is provided in the accompanying Appendix (Appendix F), demonstrating how components used in arriving at screener scores are created and summed.

Output

From the sample SAS code, several variables are created:

- Intermediate variables used to create components for scoring:
    - CFIS_FVSUM
    - CFIS_PROT_TOT
    - CFIS_PB_RATIO
    - CFIS_GR_TOT
    - CFIS_GR_RATIO
    - CFIS_4S
    - CFIS_5S
    - CFIS_10S
    - CFIS_11S
    - CFIS_12S

    - CFIS_13S
- Component variables:
    - CFIS_FV = vegetables and fruit component
    - CFIS_PB_PRO = plant-based protein: total protein component
    - CFIS_PROT_SUM = total protein component
    - CFIS_WG = whole grain component
    - CFIS_GR_Score = whole grain: total grain component
    - CFIS_SW_Score = sugary foods component
    - CFIS_SALT_Score = salty foods component
    - CFIS_OIL = unsaturated oils component
- Total score:
    - CFIS_Total = total screener score

Appendix F: Sample Code to Score the Canadian Food Intake Screener


```
*************************************************************************************************************************;
*************************************************************************************************************************;
***The Canadian Food Intake Screener for assessing alignment of adults' dietary intake with the 2019**********************;
***Canada's Food Guide healthy food choices recommendations: Scoring system and construct validity**********************;

***Hutchinson JM, et al.*************************************************************************************************;

***Supplementary File S5. Sample Code to Score the Canadian Food Intake Screener*****************************************;
*************************************************************************************************************************;
*************************************************************************************************************************;

*NOTE: Components are intended only for the purpose of summing to achieve a total score;

*Screener variable names:;
*CFIS_1:  Fruits;
*CFIS_2:  Potatoes;
*CFIS_3:  Vegetables;
*CFIS_4:  Fast food;
*CFIS_5:  Processed meat;
*CFIS_6:  Animal-based protein foods;
*CFIS_7:  Plant-based protein foods;
*CFIS_8:  Yogurt, kefir, cheese;
*CFIS_9:  Unsweetened milk;
*CFIS_10: Sweetened milk;
*CFIS_11: Sweetened beverages;
*CFIS_12: Sugary snacks;
*CFIS_13: Salty snacks;
*CFIS_14: Refined grain foods;
*CFIS_15: Whole grain foods;
*CFIS_16: Unsaturated oils;

data samplescoring;
set inputdata;

***Vegetables and fruit***;
*Sum vegetable + potato + fruit items
*Weight to /20;
CFIS_FVSUM=(CFIS_1+CFIS_2+CFIS_3);
CFIS_FV=CFIS_FVSUM/1.35;

***Proteins***;
*Score comprised of two components: the ratio of;
*plant-based protein foods/total protein and total protein;


*Ratio of plant-based protein foods item/total proteins (sum of animal-based protein foods,
*plant-based protein foods, yogurt, kefir, cheese, and unsweetened milk);
*Weight to /5;
CFIS_PROT_TOT = (CFIS_6 + CFIS_7 + CFIS_8 + CFIS_9);
CFIS_PB_RATIO = CFIS_7/CFIS_PROT_TOT;
if CFIS_PB_RATIO =1 then CFIS_PB_PRO = 5;
else if CFIS_PB_RATIO = 0 then CFIS_PB_PRO = 0;
else if CFIS_PB_RATIO >0 & CFIS_PB_RATIO <1 then CFIS_PB_PRO=5*(CFIS_PB_RATIO);

*Total protein (sum of animal-based protein foods, plant-based protein foods,;
*yogurt, kefir, cheese, and unsweetened milk);
*Weight to /5;
CFIS_PROT_SUM = CFIS_PROT_TOT/7.2;

***Whole grains***;
*Score comprised of two components: total whole grain and the ratio;
*of whole-grain foods over total grains (whole-grain foods plus refined-grain foods);

*Total whole-grain foods;
*Weight to /5;
CFIS_WG = CFIS_15/1.8;

*Grain foods ratio (Whole-grain foods/(Whole-grain foods + Refined-grain foods));
*Weight to /5;
CFIS_GR_TOT = CFIS_14 + CFIS_15;
if CFIS_GR_TOT ne 0 then CFIS_GR_RATIO = CFIS_15/CFIS_GR_TOT;
    else if CFIS_GR_TOT = 0 then CFIS_GR_RATIO = 0;
if CFIS_GR_RATIO = 0 then CFIS_GR_Score = 0;
    else if CFIS_GR_RATIO =1 then CFIS_GR_Score = 5;
    else if CFIS_GR_RATIO >0 & CFIS_GR_RATIO <1 then CFIS_GR_Score=5*(CFIS_GR_RATIO);
```

```
***Foods and beverages high in sugars***;
*Weight to /10;
*Foods to limit are reverse-scored;

*scoring sweetened milk (reverse), CFIS_10;
    if CFIS_10 = 0 then CFIS_10S = 9;
    else if CFIS_10 = 1 then CFIS_10S = 8;
    else if CFIS_10 = 2 then CFIS_10S = 7;
    else if CFIS_10 = 3 then CFIS_10S = 6;
    else if CFIS_10 = 4 then CFIS_10S = 5;
    else if CFIS_10 = 5 then CFIS_10S = 4;
    else if CFIS_10 = 6 then CFIS_10S = 3;
    else if CFIS_10 = 7 then CFIS_10S = 2;
    else if CFIS_10 = 8 then CFIS_10S = 1;
    else if CFIS_10 = 9 then CFIS_10S = 0;
    else CFIS_10S=".";
*scoring sweetened beverages (reverse), CFIS_11;
    if CFIS_11 = 0 then CFIS_11S = 9;
    else if CFIS_11 = 1 then CFIS_11S = 8;
    else if CFIS_11 = 2 then CFIS_11S = 7;
    else if CFIS_11 = 3 then CFIS_11S = 6;
    else if CFIS_11 = 4 then CFIS_11S = 5;
    else if CFIS_11 = 5 then CFIS_11S = 4;
    else if CFIS_11 = 6 then CFIS_11S = 3;
    else if CFIS_11 = 7 then CFIS_11S = 2;
    else if CFIS_11 = 8 then CFIS_11S = 1;
    else if CFIS_11 = 9 then CFIS_11S = 0;
    else CFIS_11S=".";
*scoring sugary treats (reverse), CFIS_12;
    if CFIS_12 = 0 then CFIS_12S = 9;
    else if CFIS_12 = 1 then CFIS_12S = 8;
    else if CFIS_12 = 2 then CFIS_12S = 7;
    else if CFIS_12 = 3 then CFIS_12S = 6;
    else if CFIS_12 = 4 then CFIS_12S = 5;
    else if CFIS_12 = 5 then CFIS_12S = 4;
    else if CFIS_12 = 6 then CFIS_12S = 3;
    else if CFIS_12 = 7 then CFIS_12S = 2;
    else if CFIS_12 = 8 then CFIS_12S = 1;
    else if CFIS_12 = 9 then CFIS_12S = 0;
    else CFIS_12S=".";


CFIS_SW_SCORE=(CFIS_10S + CFIS_11S + CFIS_12S)/2.7;

***Foods high in sodium/saturated fat***;
*Weight to /10;
*Foods to limit are reverse-scored;

*scoring fast food (reverse), CFIS_4;
    if CFIS_4 = 0 then CFIS_4S = 9;
    else if CFIS_4 = 1 then CFIS_4S = 8;
    else if CFIS_4 = 2 then CFIS_4S = 7;
    else if CFIS_4 = 3 then CFIS_4S = 6;
    else if CFIS_4 = 4 then CFIS_4S = 5;
    else if CFIS_4 = 5 then CFIS_4S = 4;
    else if CFIS_4 = 6 then CFIS_4S = 3;
    else if CFIS_4 = 7 then CFIS_4S = 2;
    else if CFIS_4 = 8 then CFIS_4S = 1;
    else if CFIS_4 = 9 then CFIS_4S = 0;
    else CFIS_4S=".";
*scoring processed meat (reverse), CFIS_5;
    if CFIS_5 = 0 then CFIS_5S = 9;
    else if CFIS_5 = 1 then CFIS_5S = 8;
    else if CFIS_5 = 2 then CFIS_5S = 7;
    else if CFIS_5 = 3 then CFIS_5S = 6;
    else if CFIS_5 = 4 then CFIS_5S = 5;
    else if CFIS_5 = 5 then CFIS_5S = 4;
    else if CFIS_5 = 6 then CFIS_5S = 3;
    else if CFIS_5 = 7 then CFIS_5S = 2;
    else if CFIS_5 = 8 then CFIS_5S = 1;
    else if CFIS_5 = 9 then CFIS_5S = 0;
    else CFIS_5S=".";
*scoring salty snacks (reverse), CFIS_13;
    if CFIS_13 = 0 then CFIS_13S = 9;
    else if CFIS_13 = 1 then CFIS_13S = 8;
```

```
    else if CFIS_13 = 2 then CFIS_13S = 7;
    else if CFIS_13 = 3 then CFIS_13S = 6;
    else if CFIS_13 = 4 then CFIS_13S = 5;
    else if CFIS_13 = 5 then CFIS_13S = 4;
    else if CFIS_13 = 6 then CFIS_13S = 3;
    else if CFIS_13 = 7 then CFIS_13S = 2;
    else if CFIS_13 = 8 then CFIS_13S = 1;
    else if CFIS_13 = 9 then CFIS_13S = 0;
    else CFIS_13S=".";

CFIS_SALT_SCORE=(CFIS_4S + CFIS_5S + CFIS_13S)/2.7;

***Unsaturated oils***;
*Weight to /5;

CFIS_OIL = CFIS_16/1.8;

*delete unnecessary variables;
delete CFIS_FVSUM CFIS_PROT_TOT CFIS_PB_RATIO CFIS_GR_TOT CFIS_GR_RATIO;

***Sum components to get total score***;
CFIS_TOTAL= (CFIS_FV + CFIS_PB_PRO + CFIS_PROT_SUM + CFIS_WG + CFIS_GR_Score +  CFIS_SW_SCORE + CFIS_SALT_SCORE + CFIS_OIL);

run;
```

Appendix G: Deriving Healthy Eating Food Index-2019 variables

Deriving Healthy Eating Food Index-2019 (HEFI-2019) variables

For five of the ten HEFI-2019 components, the numerators and denominators are based on reference amounts, which represent the amount of food typically consumed at one time (Health Canada 2016). A database developed by Health Canada specifying the reference amounts for food codes applied to the 2015 CCHS-Nutrition data, which was used for evaluation of the HEFI-2019 (Brassard et al. 2022b), was drawn upon. Unmatched food codes were identified, and reference amounts were assigned by Health Canada staff. Reference amounts were then merged into the ASA24-Canada-2018 data by food code, the amounts reported were divided by the reference amount for each dietary constituent of interest, and the resulting quantities were summed to arrive at total reference amounts consumed of the categories of interest, such as Vegetables and fruits, per recall. In some cases, it was necessary to disaggregate recipes into ingredients, assign reference amounts at the level of ingredients based on their proportions within the recipe, and recompile the recipe and associated reference amounts. Recipe information underlying the ASA24-Canada-2018 system was used for this purpose. The denominator for Vegetables and fruits, Whole-grain foods, and Protein foods is total foods expressed as reference amounts, computed by summing the HEFI-2019 Vegetables and fruit component, Whole-grain foods component, non-whole grain foods (used in the Grain foods ratio), animal protein foods (used in the Protein foods component), Plant-based protein foods component, unsweetened milk and unsweetened plant-based protein foods (used in the Beverages component), plus an "other" foods component that captures foods not otherwise included in HEFI-2019 components, such as condiments, processed meats, and desserts. Total foods does not include culinary ingredients, beverages without protein, and oils and spreads (Brassard et al., 2022a).

The Beverages component of the HEFI-2019 is computed as the ratio of grams of water (including carbonated water), unsweetened milk, and unsweetened plant-based beverages containing protein to grams of all beverages, including alcohol. Except for the numerator for Free sugars, the remaining numerators and denominators–unsaturated fats, saturated fats, sodium, and energy–were available from the ASA24-Canada-2018 data. The process to calculate amounts of free sugars per recall was similar to that for reference amounts, with information on the free sugars content (per 100g) for each food code obtained from a Health Canada database (Rana et al. 2021) and estimates for codes for which free sugar values had not been determined provided by Health Canada staff.

Appendix G References

Brassard, D., Elvidge Munene, L.-A., St Pierre, S., Gonzalez, A., Guenther, P.M., Jessri, M., Vena, J.E., Olstad, D.L., Vatanparast, H., Prowse, R., Lemieux, S., L'Abbé, M.R., Garriguet, D., Kirkpatrick, S.I., and Lamarche, B. 2022b. Evaluation of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Applied Physiology, Nutrition, and Metabolism **47**(5). doi:10.1139/APNM-2021-0416.

Brassard, D., Elvidge Munene, L.-A., St Pierre, S., Guenther, P.M., Kirkpatrick, S.I., Slater, J., Lemieux, S., Jessri, M., Haines, J., Prowse, R., Olstad, D.L., Garriguet, D., Vena, J.E., Vatanparast, H., L'Abbé, M.R., and Lamarche, B. 2022a. Development of the Healthy Eating Food Index (HEFI)-2019 measuring adherence to Canada's Food Guide 2019 recommendations on healthy food choices. Applied Physiology, Nutrition, and Metabolism **47**(5): 595–610. doi:10.1139/APNM-2021-0415.

Health Canada. 2016. Table of Reference Amounts for Food. Available from https://www.canada.ca/en/health-canada/services/technical-documents-labelling-requirements/table-reference-amounts-food.html [Accessed 8 February 2022].

Rana, H., Mallet, M.C., Gonzalez, A., Verreault, M.F., and St-Pierre, S. 2021. Free sugars consumption in Canada. Nutrients 2021, Vol. 13, Page 1471 **13**(5): 1471. doi:10.3390/NU13051471.

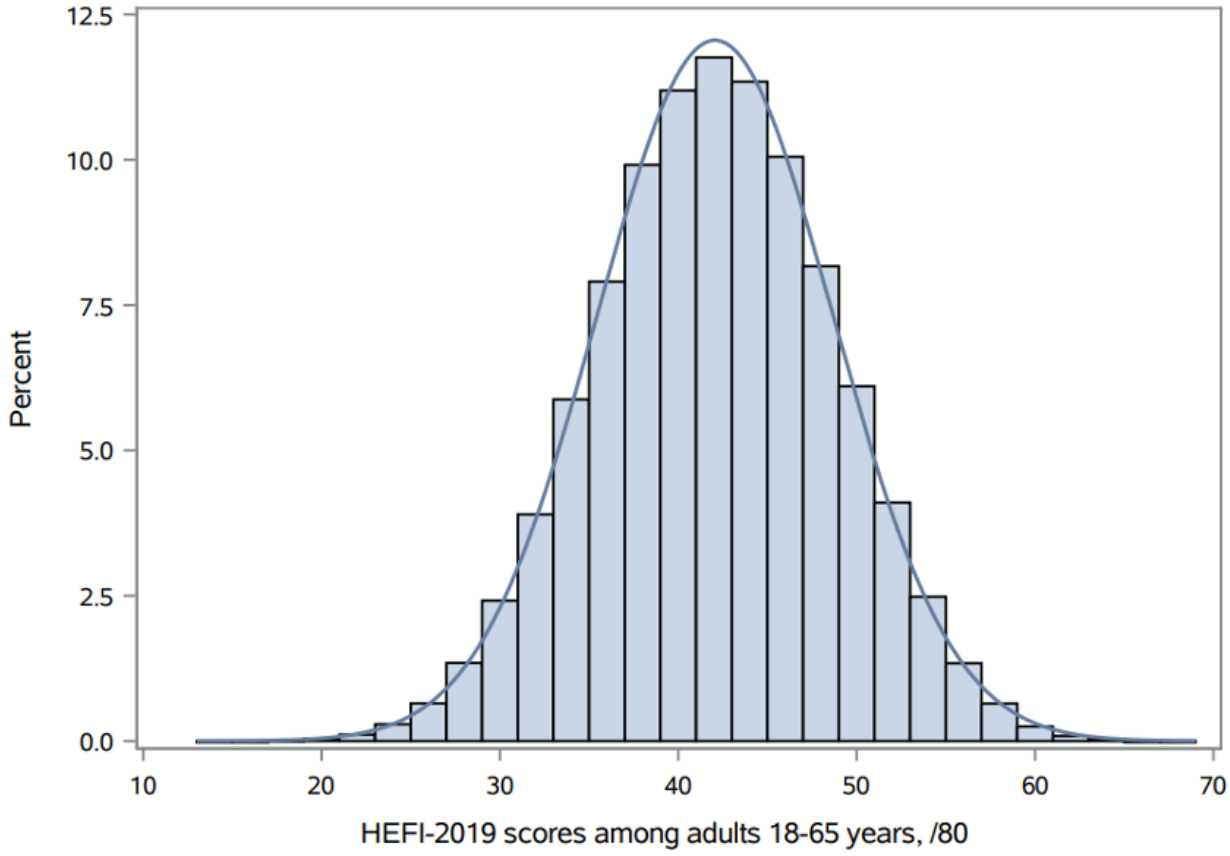
Appendix H: Hypothesized differences in Canadian Food Intake Screener scores across subgroups based on known differences in diet quality

Hypothesized differences in Canadian Food Intake Screener scores across subgroups based on known differences in diet quality

Characteristic	Hypothesized association	References
Age	Higher screener scores among older versus younger adults.	Brassard et al. 2022b; Guenther et al. 2014; Hiza et al. 2013; Moubarac et al. 2017; Reedy et al. 2018
Gender	Higher screener scores among women versus men.	Brassard et al. 2022b; Guenther et al. 2014; Hiza et al. 2013; Moubarac et al. 2017; Reedy et al. 2018
Income adequacy	Higher screener scores among individuals with higher versus lower income adequacy.	Tarasuk et al. 2010; Hiza et al., 2013
Educational attainment	Higher screener scores among individuals with higher versus lower educational attainment.	Tarasuk et al. 2010; Hiza et al. 2013
Health literacy	Higher screener scores among individuals with higher versus lower health literacy.	Kuczmarski et al. 2016
Smoking status	Higher screener scores among individuals who report not smoking versus those who report smoking.	Alkerwi et al. 2017; Brassard et al. 2022b; Guenther et al. 2014; Reedy et al. 2018

Appendix I: Distribution of total Healthy Eating Food Index-2019 scores[1] among adults aged 18-65 years in a study to examine the construct validity of the Canadian Food Intake Screener

Distribution of total Healthy Eating Food Index-2019 scores[1] among adults aged 18-65 years in a study to examine the construct validity of the Canadian Food Intake Screener



[1] Healthy Eating Food Index-2019 scores were based on simulated usual intakes, using data from up to two non-consecutive 24-hour dietary recalls (n=230 recalls completed by 128 individuals).

Appendix J: Search strategies for MEDLINE, CINAHL, and SCOPUS

Due to the emerging nature of this field, an updated search of the literature was conducted in March 2022. The search for relevant articles was applied in the same three academic databases: MEDLINE, CINAHL, and Scopus. No date limits were applied to the searches and the same search was applied with minor changes, such as the removal of body weight and reduced rank regression.

MEDLINE 2019 search
((systems modelling [tiab] OR systems science [tiab] OR systems thinking [tiab] OR systems approach*[tiab] OR systems theory*[tiab] OR systems analysis [tiab] OR agent-based model*[tiab] OR system dynamics model*[tiab] OR systems-based [tiab]) OR (LASSO [tiab] OR "least absolute shrinkage and selection operator" [tiab]) OR (reduced rank regression [tiab] OR RRR [tiab]) OR (machine learning [tiab] OR neural networks [tiab] OR decision tree analys* [tiab] OR deep learning [tiab] OR deep neural networks [tiab] OR network science [tiab] OR network analys* [tiab] OR topological data analys* [tiab] OR network-based [tiab]) OR (copula* [tiab] OR copular function* [tiab]) OR ("data driven" [tiab] OR posteriori [tiab] OR priori [tiab] OR "hypothesis driven" [tiab]))

AND

(eating pattern*[tiab] OR diet pattern*[tiab] OR dietary pattern*[tiab] OR food pattern*[tiab] OR food habit*[tiab] OR diet quality score*[tiab] OR diet ind*[tiab] OR dietary ind*[tiab] OR dietary guide*[tiab] OR dietary recommend* [tiab] OR diet scor* [tiab] OR "healthy eating index" [tiab] OR dietary intake*[tiab] OR diet intake*[tiab] OR obes*[tiab] OR body weight [tiab])

Filter: Humans; No date limiter.

MEDLINE 2022 search with no body weight terms

((systems modelling [tiab] OR systems science [tiab] OR systems thinking [tiab] OR systems approach*[tiab] OR systems theory*[tiab] OR systems analysis [tiab] OR agent-based model*[tiab] OR system dynamics model*[tiab] OR systems-based [tiab]) OR (LASSO [tiab] OR "least absolute shrinkage and selection operator" [tiab]) OR (reduced rank regression [tiab] OR RRR [tiab]) OR (machine learning [tiab] OR neural networks [tiab] OR decision tree analys* [tiab] OR deep learning [tiab] OR deep neural networks [tiab] OR network science [tiab] OR network analys* [tiab] OR topological data analys* [tiab] OR network-based [tiab]) OR (copula* [tiab] OR copular function* [tiab]) OR ("data driven" [tiab] OR posteriori [tiab] OR priori [tiab] OR "hypothesis driven" [tiab]))

AND

(feeding behavior[mesh:noexp] OR eating pattern*[tiab] OR diet pattern*[tiab] OR dietary pattern*[tiab] OR food pattern*[tiab] OR food habit*[tiab] OR diet quality score*[tiab] OR diet ind*[tiab] OR dietary ind*[tiab] OR dietary guide*[tiab] OR dietary recommend* [tiab] OR diet scor* [tiab] OR "healthy eating index" [tiab] OR dietary intake*[tiab] OR diet intake*[tiab])

Filter: Humans; Date limiter: AND ("2019/10/25"[Date - Entry] : "3000"[Date - Entry])

MEDLINE 2022 search with subject heading added

(((systems modelling [tiab] OR systems science [tiab] OR systems thinking [tiab] OR systems approach*[tiab] OR systems theory*[tiab] OR systems analysis [tiab] OR agent-based model*[tiab] OR system dynamics model*[tiab] OR systems-based [tiab]) OR (LASSO [tiab] OR "least absolute shrinkage and selection operator" [tiab]) OR (reduced rank regression [tiab] OR RRR [tiab]) OR (machine learning [tiab] OR neural networks [tiab] OR decision tree analys* [tiab] OR deep learning [tiab] OR deep neural networks [tiab] OR network science [tiab] OR network analys* [tiab] OR topological data analys* [tiab] OR network-based [tiab]) OR (copula* [tiab] OR copular function* [tiab]) OR ("data driven" [tiab] OR posteriori [tiab] OR priori [tiab] OR "hypothesis driven" [tiab]))

AND

(feeding behavior[mesh:noexp] OR eating pattern*[tiab] OR diet pattern*[tiab] OR dietary pattern*[tiab] OR food pattern*[tiab] OR food habit*[tiab] OR diet quality score*[tiab] OR diet ind*[tiab] OR dietary ind*[tiab] OR dietary guide*[tiab] OR dietary recommend* [tiab] OR diet scor* [tiab] OR "healthy eating index" [tiab] OR dietary intake*[tiab] OR diet intake*[tiab] OR obes*[tiab] OR body weight [tiab]))

NOT

(((systems modelling [tiab] OR systems science [tiab] OR systems thinking [tiab] OR systems approach*[tiab] OR systems theory*[tiab] OR systems analysis [tiab] OR agent-based model*[tiab] OR system dynamics model*[tiab] OR systems-based [tiab]) OR (LASSO [tiab] OR "least absolute shrinkage and selection operator" [tiab]) OR (reduced rank regression [tiab] OR RRR [tiab]) OR (machine learning [tiab] OR neural networks [tiab] OR decision tree analys* [tiab] OR deep learning [tiab] OR deep neural networks [tiab] OR network science [tiab] OR network analys* [tiab] OR topological data analys* [tiab] OR network-based [tiab]) OR (copula* [tiab] OR copular function* [tiab]) OR ("data driven" [tiab] OR posteriori [tiab] OR priori [tiab] OR "hypothesis driven" [tiab]))

AND

(eating pattern*[tiab] OR diet pattern*[tiab] OR dietary pattern*[tiab] OR food pattern*[tiab] OR food habit*[tiab] OR diet quality score*[tiab] OR diet ind*[tiab] OR dietary ind*[tiab] OR dietary guide*[tiab] OR dietary recommend* [tiab] OR diet scor* [tiab] OR "healthy eating index" [tiab] OR dietary intake*[tiab] OR diet intake*[tiab] OR obes*[tiab] OR body weight [tiab])

Filter: Humans; No date limiter.

CINAHL 2019 search
TI (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR "healthy eating index" OR dietary intake* OR diet intake* OR obes* OR body weight) OR AB (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR "healthy eating index" OR dietary intake* OR diet intake* OR obes* OR body weight)

AND

TI (systems modelling OR systems science OR systems thinking OR systems approach* OR systems theory* OR systems analysis OR agent-based model* OR system dynamics model* OR systems-based) OR AB (systems modelling OR systems science OR systems thinking OR systems approach* OR systems theory* OR systems analysis OR agent-based model* OR system dynamics model* OR systems-based)) OR (TI (LASSO OR "least absolute shrinkage and selection operator") OR AB (LASSO OR "least absolute shrinkage and selection operator")) OR (TI (reduced rank regression OR RRR) OR AB (reduced rank regression OR RRR)) OR (TI (machine learning OR neural networks OR decision tree analys* OR deep learning OR deep neural networks OR network science OR network analys* OR topological data analys* OR network-based) OR AB (machine learning OR neural networks OR decision tree analys* OR deep learning OR deep neural networks OR network science OR network analys* OR topological data analys* OR network-based)) OR (TI (copula* OR copular function*) OR AB (copula* OR copular function*)) OR (TI ("data driven" OR posteriori OR priori OR "hypothesis driven") OR AB ("data driven" OR posteriori OR priori OR "hypothesis driven")

Filter: Human; No date limiter.

CINAHL 2022 search with no body weight terms

TI (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR “healthy eating index” OR dietary intake* OR diet intake*) OR AB (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR “healthy eating index” OR dietary intake* OR diet intake*)

AND

TI (systems modelling OR systems science OR systems thinking OR systems approach* OR systems theory* OR systems analysis OR agent-based model* OR system dynamics model* OR systems-based) OR AB (systems modelling OR systems science OR systems thinking OR systems approach* OR systems theory* OR systems analysis OR agent-based model* OR system dynamics model* OR systems-based)) OR (TI (LASSO OR “least absolute shrinkage and selection operator”) OR AB (LASSO OR “least absolute shrinkage and selection operator”)) OR (TI (reduced rank regression OR RRR) OR AB (reduced rank regression OR RRR)) OR (TI (machine learning OR neural networks OR decision tree analys* OR deep learning OR deep neural networks OR network science OR network analys* OR topological data analys* OR network-based) OR AB (machine learning OR neural networks OR decision tree analys* OR deep learning OR deep neural networks OR network science OR network analys* OR topological data analys* OR network-based)) OR (TI (copula* OR copular function*) OR AB (copula* OR copular function*)) OR (TI (“data driven” OR posteriori OR priori OR “hypothesis driven”) OR AB (“data driven” OR posteriori OR priori OR “hypothesis driven”)

Filter: Human; Date limiter: AND ZD (201910* OR 202* OR "in process")

CINAHL 2022 search with subject heading added

TI (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR “healthy eating index” OR dietary intake* OR diet intake* OR obes* OR body weight) OR AB (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR “healthy eating index” OR dietary intake* OR diet intake* OR obes* OR body weight) OR MH (food habits)

NOT

TI (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR “healthy eating index” OR dietary intake* OR diet intake* OR obes* OR body weight) OR AB (eating pattern* OR diet pattern* OR dietary pattern* OR food pattern* OR food habit* OR diet quality score* OR diet ind* OR dietary ind* OR dietary guide* OR dietary recommend* OR diet scor* OR “healthy eating index” OR dietary intake* OR diet intake* OR obes* OR body weight)

AND

TI (systems modelling OR systems science OR systems thinking OR systems approach* OR systems theory* OR systems analysis OR agent-based model* OR system dynamics model* OR systems-based) OR AB (systems modelling OR systems science OR systems thinking OR systems approach* OR systems theory* OR systems analysis OR agent-based model* OR system dynamics model* OR systems-based)) OR (TI (LASSO OR “least absolute shrinkage and selection operator”) OR AB (LASSO OR “least absolute shrinkage and selection operator”)) OR (TI (reduced rank regression OR RRR) OR AB (reduced rank regression OR RRR)) OR (TI (machine learning OR neural networks OR decision tree analys* OR deep learning OR deep neural networks OR network science OR network analys* OR topological data analys* OR network-based) OR AB (machine learning OR neural networks OR decision tree analys* OR deep learning OR deep neural networks OR network science OR network analys* OR topological data analys* OR network-based)) OR (TI (copula* OR copular function*) OR AB (copula* OR copular function*)) OR (TI (“data driven” OR posteriori OR priori OR “hypothesis driven”) OR AB (“data driven” OR posteriori OR priori OR “hypothesis driven”)

Filter: Human; No date limiter.

Scopus 2019 search

TITLE-ABS-KEY("eating pattern" OR "diet pattern" OR "dietary pattern" OR "food pattern" OR "food habit" OR "diet quality score" OR "diet index" OR "dietary index" OR "dietary guide" OR "dietary recommend" OR "diet score" OR “healthy eating index” OR "dietary intake" OR "diet intake" OR obes OR “body weight”)

AND

TITLE-ABS-KEY("systems modelling" OR "systems science" OR "systems thinking" OR "systems approach" OR "systems theory" OR "systems analysis" OR "agent-based model" OR "system dynamics model" OR "systems-based" OR LASSO OR “least absolute shrinkage and selection operator” OR "reduced rank regression" OR RRR OR "machine learning" OR "neural networks" OR "decision tree analysis" OR "deep learning" OR "deep neural networks" OR "network science" OR "network analysis" OR "topological data analysis" OR "network-based" OR "copula" OR "copular function" OR “data driven” OR "posteriori" OR "priori" OR “hypothesis driven”)

No date limiter.

Scopus 2022 search with no body weight terms

TITLE-ABS-KEY("eating pattern" OR "diet pattern" OR "dietary pattern" OR "food pattern" OR "food habit" OR "diet quality score" OR "diet index" OR "dietary index" OR "dietary guide" OR "dietary recommend" OR "diet score" OR “healthy eating index” OR "dietary intake" OR "diet intake")

AND

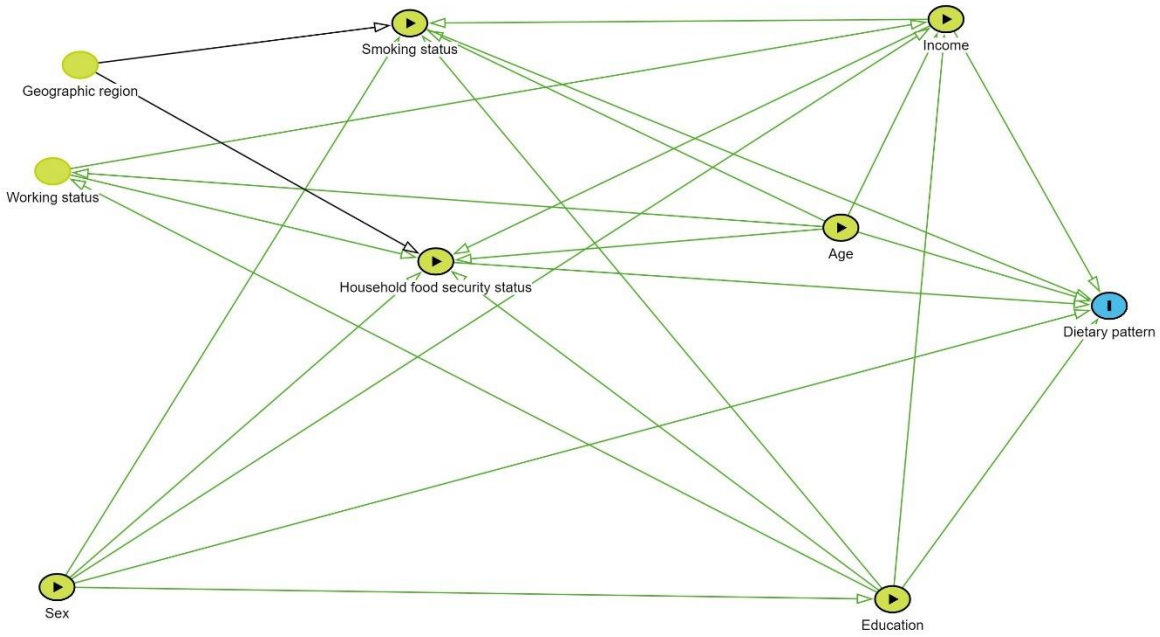
TITLE-ABS-KEY("systems modelling" OR "systems science" OR "systems thinking" OR "systems approach" OR "systems theory" OR "systems analysis" OR "agent-based model" OR "system dynamics model" OR "systems-based" OR LASSO OR “least absolute shrinkage and selection operator” OR "reduced rank regression" OR RRR OR "machine learning" OR "neural networks" OR "decision tree analysis" OR "deep learning" OR "deep neural networks" OR "network science" OR "network analysis" OR "topological data analysis" OR "network-based" OR "copula" OR "copular function" OR “data driven” OR "posteriori" OR "priori" OR “hypothesis driven”)

Date limiter: 2019 to present

Appendix K: Key extraction fields for review of novel analytic methods to characterize dietary patterns

Field	Description
Study identification	Authors
	Title
	Journal
	Year published
	Funding source for study
Study population	Sample size
	Sample characteristics
Methods	Research objectives/questions
	Study design
	Study name
	Measurement of diet
	Dietary components analyzed
	Novel methods of analysis
	Consideration of equity
Results	Findings re: diet or diet-related outcomes
	Author description of method utility and implications
Critical review	Author stated limitations
	Reviewer observed limitations

Appendix L: Directed acyclic graph of relationships between exposures of interest and dietary patterns



Appendix M: Bureau of Nutritional Science Condensed Variables for analysis of dietary patterns in the 2015 CCHS Nutrition

Suggested new variable	New variable description	% Zero intakes: condensed variable	Original Variable	Description (gram weight)	% Zero intakes: original variable
Alcohol	Alcohol	74.9	BNSD47A	Spirits	95.4
			BNSD47B	Liqueurs	99.4
			BNSD48A	Wine	88.2
			BNSD49A	Beer	88.7
			BNSD49B	Coolers	100.0
CoffeeTea	Coffee & tea	17.8	BNSD51A	Tea (Incl Iced Tea)	65.6
			BNSD51B	Coffee	34.6
Curedmeat	Cured meats	62.8	BNSD25C	Bacon	89.7
			BNSD25D	Ham: cured - lean only	97.0
			BNSD25E	Ham: cured - lean and fat	98.1
			BNSD30A	Sausage	88.5
			BNSD32A	Luncheon meat	80.7
DairyHifat	Higher fat dairy products	31.0	BNSD10A	Milk: whole	73.8
			BNSD10E	Milk: Evaporated whole	98.6
			BNSD10K	Milk: Goat/sheep	99.9
			BNSD13A	Whipping cream	95.9
			BNSD13B	Table cream	91.3
			BNSD13C	Half & Half cream	89.2
			BNSD13D	Sour cream	94.6
			BNSD14D	Cheese: More than 25% B.F.	59.8
			BNSD15B	Yoghurts: More than 2.1% B.F.	87.5
DairyLofat		15.9	BNSD10B	Milk: 2%	39.5

	Lower fat dairy products		BNSD10C	Milk: 1%	68.5
			BNSD10D	Milk: Skim	69.2
			BNSD10F	Milk: Evaporated 2%	99.7
			BNSD10G	Milk: Evaporated skim	100.0
			BNSD10H	Milk: Condensed	99.7
			BNSD10I	Milk: other (whey/buttermilk)	99.2
			BNSD14A	Cottage cheese	98.1
			BNSD14B	Cheese: less than 10% B.F.	97.3
			BNSD14C	Cheese: 10% B.F. to 25% B.F.	69.5
			BNSD15A	Yoghurts: Less than 2% B.F.	92.5
Egg	Eggs	42.4	BNSD16A	Egg	42.4
Fish	Fish and shellfish	81.4	BNSD34A	Fish: less than 6% total fat	89.3
			BNSD34B	Fish: superior or equal to 6% total fat	91.4
			BNSD35A	Shellfish	94.5
Grainsother	Other grains (refined or whole grain)	27.0	BNSD01A	Pasta	79.6
			BNSD01B	Rice	77.0
			BNSD01C	Cereal/grains/flours	45.3
Greenveg	Green vegetables	42.2	BNSD36A	Beans	92.8
			BNSD36B	Broccoli	88.7
			BNSD36C	Cabbage/Kale	86.5
			BNSD36H	Lettuce/Leafy greens (spinach/mustard greens)	64.7
			BNSD36K	Peas/snow peas	89.5
Oniongar	Onion, garlic, leeks, green onion	41.5	BNSD36J	Onion/green onion/leeks/garlic	41.5
Orangeveg	Orange vegetables	65.7	BNSD36E	Carrots	68.8
			BNSD36M	Squashes	94.4
Other		4.1	BNSD16B	Egg substitutes	99.9
			BNSD41D	Sugar substitutes	95.0

	Other foods/beverages		BNSD46B	Soft drink: diet	92.0
			BNSD52A	Babyfood products	99.7
			BNSD52B	Infant formula	100.0
			BNSD53B	Others (Baking soda/baking powder/yeast)	59.0
			BNSD54A	Energy bar	99.6
			BNSD54B	Protein bar/shake	99.4
			BNSD54C	Meal replacements	96.0
			BNSD99A	Mexican recipes	100.0
			BNSD36O	Juices: Tomato & vegetables	97.4
			BNSD50C	Gravies	96.2
			BNSD50D	Sauces (White/Bearnaise/Soya/Tartar/Ketchup)	67.3
			BNSD50E	Salad dressings (with or without oil)	62.0
			BNSD50F	Seasonings (salt/vinegar)	14.8
			BNSD53A	Spices	71.1
Otherfruit	Fruits (apples, cherries, grapes, melons, peaches, pears, plums, strawberries, etc.)	43.7	BNSD40B	Apple	77.5
			BNSD40D	Cherries	97.1
			BNSD40E	Grapes/raisins	86.4
			BNSD40F	Melons (canteloup/honeydew/watermelon)	95.8
			BNSD40G	Peaches/nectarines	96.8
			BNSD40H	Pears	96.5
			BNSD40J	Plums/prunes	98.2
			BNSD40K	Strawberries	90.7
			BNSD40L	Other fruits (blueberries/date/kiwi/fruit salad)	72.6
Otherveg	Vegetables (cauliflower, celery, corn, mushrooms, peppers, etc.)	36.4	BNSD36D	Cauliflower	95.4
			BNSD36F	Celery	79.0
			BNSD36G	Corn	92.7
			BNSD36I	Mushrooms	87.8
			BNSD36L	Peppers: red/green	77.7

			BNSD36P	Other veg (cucumber/beet/turnip)	55.2
Plantproteins	Plant-based proteins/alternatives (plant-based beverages, nuts, seeds, nut butters, legumes, tofu)	56.3	BNSD10J	Plant-based beverage (soy/almond/coconut)	95.7
			BNSD33A	Nuts	78.5
			BNSD33B	Seeds	95.2
			BNSD33C	Peanut butter/other nut spreads	85.7
			BNSD37A	Legume	89.3
			BNSD37B	Foods made with vegetable proteins (tofu)	97.1
Potato	Potatoes	60.6	BNSD38B	Fried/roasted potatoes	88.9
			BNSD39A	Potato	69.8
Poultry	Poultry	60.9	BNSD27A	Chicken: meat only	70.0
			BNSD27B	Chicken: meat and skin	92.2
			BNSD27C	Turkey: meat only	97.4
			BNSD27D	Turkey: meat + skin/ground	98.8
			BNSD27E	Other birds: Duck/pheasant/pigeon	99.7
			BNSD27F	Birds: skin only	100.0
Redmeat	Red meat	54.6	BNSD22A	Beef: Lean only	89.8
			BNSD22B	Beef: Lean + fat	96.0
			BNSD22C	Beef: ground	78.1
			BNSD23A	Veal: lean only	99.7
			BNSD23B	Veal: lean + fat/ground	99.7
			BNSD24A	Lamb: Lean only	99.5
			BNSD24B	Lamb: Lean + fat/ground	99.7
			BNSD25A	Pork: Fresh - lean only	92.5
			BNSD25B	Pork: Fresh - lean + fat/ground	94.1
			BNSD28A	Liver	99.7
			BNSD28B	Liver pate	99.9
			BNSD29A	Offal	99.8
			BNSD31A	Game meat	99.0

Refinedgrain	Refined grains (probable)	30.9	BNSD02A	White breads	70.2
			BNSD04A	Rolls/bagels/pita/croutons/dumplings/matzo/tortilla	63.6
			BNSD04B	Crackers/crispbreads	88.5
			BNSD04C	Muffins/English muffins	97.2
			BNSD04D	Pancakes/Waffles	98.6
			BNSD04E	Croissants/Piecrusts/Phyllo Dough	98.3
			BNSD04F	Dry Mixes (Cakes/Muffins/Pancakes)	95.9
			BNSD06A	Breakfast cereal (other)	91.7
Saltysnack	Salty snacks	79.3	BNSD38A	Potato chips	90.2
			BNSD42A	Plain popcorn/pretzels	97.1
			BNSD42B	Salty/high-fat snacks (incl tortilla chips)	90.9
Satfats	Saturated fats	44.4	BNSD17A	Butter	71.3
			BNSD20A	Block margarine	99.4
			BNSD21B	Animal fats	94.7
			BNSD21C	Shortening	63.5
Soup	Soups	74.4	BNSD50A	Soups with vegetables	93.0
			BNSD50B	Soups without vegetables	80.3
Sugar	Sugars (sugar, jams, syrup, honey, etc.)	28.1	BNSD41A	Sugars: white/brown	36.4
			BNSD41B	Jams/jellies/marmalade	88.7
			BNSD41C	Other sugars (syrups/molasses/honey)	77.7
Sweetbev	Sweet beverages	50.6	BNSD45A	Fruit juice	70.0
			BNSD46A	Soft drink: regular	82.8
			BNSD46C	Fruit drinks	92.4
			BNSD46D	Other beverages (malted milk/chocolate beverage)	97.9
			BNSD46E	Energy drink	99.4
			BNSD46F	Vitamin water	99.6
			BNSD46G	Sports drink	98.7
Sweettreat		50.7	BNSD07A	Cookies (commercial)	83.7

	Sweet treats/desserts		BNSD07B	Biscuits: commercial	99.4
			BNSD07C	Granola bar	93.8
			BNSD08A	Pies: commercial	99.7
			BNSD08B	Cakes: commercial (frozen cake)	99.4
			BNSD08C	Danishes/doughnuts/other pastries: commercial	99.8
			BNSD09A	Ice cream	90.1
			BNSD09B	Ice milk	99.5
			BNSD09C	Frozen yoghurt	99.1
			BNSD43A	Candy/gum	90.9
			BNSD43B	Ice pop/sherbert	98.7
			BNSD43C	Gelatin/dessert toppings/pudding mixes: commercial	97.0
			BNSD44A	Chocolate bar	82.5
Tomato	Tomatoes	54.0	BNSD36N	Tomatoes	54.0
Tropicalfruit	Tropical fruits (citrus, banana, pineapple)	59.8	BNSD40A	Citrus fruit (oranges/lemons/grapefruits)	81.9
			BNSD40C	Banana	73.1
			BNSD40I	Pineapple	95.6
Unsatfat	Unsaturated fats	22.2	BNSD18A	Regular margarine	47.2
			BNSD18B	Calorie-reduced margarine	96.5
			BNSD21A	Vegetable oils	49.3
Water	Water	7.7	BNSD51C	Water (Well/mineral)	7.7
Wholegrain	Whole grains	52.8	BNSD03A	Whole wheat breads	74.5
			BNSD03B	Other whole grain breads	75.9
			BNSD05A	Whole Grain/Oats/High Fibre Breakfast cereals	75.2

Day 1 intakes only

Appendix N: Weighted adjacency matrices

Table N1: Weighted adjacency matrix network only sociodemographic characteristics among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition

	1	**2**	**3**	**4**	**5**	**6**	7	**8**
1	0.00	0.00	0.34	0.17	0.28	0.53	0.05	0.19
2	0.00	0.00	0.15	0.07	0.08	0.20	0.00	0.14
3	0.34	0.15	0.00	1.10	0.14	0.11	0.14	0.37
4	0.17	0.07	1.10	0.00	0.40	0.67	0.12	0.09
5	0.28	0.08	0.14	0.40	0.00	0.38	0.16	0.29
6	0.53	0.20	0.11	0.67	0.38	0.00	0.06	0.14
7	0.05	0.00	0.14	0.12	0.16	0.06	0.00	0.08
8	0.19	0.14	0.37	0.09	0.29	0.14	0.08	0.00

Legend: 1-Age, 2-Sex, 3-Household food insecurity status, 4-Income, 5-Education, 6-Worked in the past week, 7-Geography, 8-Smoking status

Table N2: Weighted adjacency matrix network only dietary components among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition

	1	**2**	**3**	**4**	**5**	**6**	7	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
1	0.00	0.06	0.10	0.09	-0.02	0.02	0.12	-0.03	0.00	0.03	0.00	-0.02	0.05	0.00	0.04
2	0.06	0.00	0.01	0.10	0.00	0.05	0.02	0.00	0.02	0.00	0.00	0.00	0.05	0.03	0.00
3	0.10	0.01	0.00	0.03	0.03	0.06	-0.13	-0.03	-0.01	0.00	-0.04	-0.06	-0.02	-0.09	0.00
4	0.09	0.10	0.03	0.00	0.02	0.00	-0.04	0.00	0.02	0.00	0.00	0.02	0.04	0.03	0.00
5	-0.02	0.00	0.03	0.02	0.00	0.02	-0.01	0.03	-0.01	0.00	0.01	0.03	0.01	-0.05	0.03
6	0.02	0.05	0.06	0.00	0.02	0.00	0.00	0.00	0.02	-0.01	0.00	0.01	0.00	-0.03	-0.03
7	0.12	0.02	-0.13	-0.04	-0.01	0.00	0.00	0.08	0.07	-0.02	-0.01	0.02	0.02	-0.05	0.00
8	-0.03	0.00	-0.03	0.00	0.03	0.00	0.08	0.00	-0.02	0.04	0.04	0.00	-0.06	0.00	-0.27

9	0.00	0.02	-0.01	0.02	-0.01	0.02	0.07	-0.02	0.00	0.06	0.12	0.07	0.10	0.02	0.00
10	0.03	0.00	0.00	0.00	0.00	-0.01	-0.02	0.04	0.06	0.00	0.04	0.02	0.23	0.05	0.02
11	0.00	0.00	-0.04	0.00	0.01	0.00	-0.01	0.04	0.12	0.04	0.00	0.04	0.13	0.04	0.06
12	-0.02	0.00	-0.06	0.02	0.03	0.01	0.02	0.00	0.07	0.02	0.04	0.00	0.07	0.06	0.00
13	0.05	0.05	-0.02	0.04	0.01	0.00	0.02	-0.06	0.10	0.23	0.13	0.07	0.00	0.06	-0.03
14	0.00	0.03	-0.09	0.03	-0.05	-0.03	-0.05	0.00	0.02	0.05	0.04	0.06	0.06	0.00	-0.03
15	0.04	0.00	0.00	0.00	0.03	-0.03	0.00	-0.27	0.00	0.02	0.06	0.00	-0.03	-0.03	0.00
16	0.01	0.00	-0.14	-0.02	-0.02	0.00	-0.18	0.09	0.08	0.00	0.02	0.00	0.08	-0.07	0.02
17	0.07	-0.01	-0.12	-0.02	0.00	0.00	-0.16	0.08	0.01	0.10	0.00	-0.02	0.05	-0.06	0.16
18	-0.04	0.03	0.12	0.06	0.11	-0.05	0.00	-0.15	-0.02	0.01	-0.04	-0.01	-0.01	0.02	-0.01
19	0.06	0.02	0.04	0.03	0.00	0.00	-0.02	-0.02	-0.01	-0.02	0.00	-0.04	-0.02	0.00	0.00
20	0.03	0.00	0.05	0.05	0.04	0.12	0.00	0.22	-0.04	0.03	0.00	0.00	-0.01	-0.04	0.17
21	-0.02	0.02	-0.03	-0.04	-0.01	-0.03	-0.03	0.01	0.01	0.24	0.17	-0.03	0.00	-0.03	-0.03
22	0.00	0.12	-0.03	0.09	0.10	0.08	-0.02	0.10	-0.04	0.00	-0.01	0.04	-0.03	0.02	0.01
23	-0.05	-0.14	0.09	0.03	-0.01	0.04	0.01	0.03	0.00	0.02	-0.01	-0.03	0.01	0.00	0.03
24	-0.04	-0.02	0.00	0.04	0.03	-0.04	0.00	-0.01	0.00	-0.04	0.00	0.00	0.02	0.04	0.00
25	0.03	0.00	0.02	0.07	0.03	-0.03	-0.11	0.18	0.05	0.19	-0.01	0.02	0.05	0.00	-0.08
26	-0.04	0.02	0.01	0.02	0.03	0.02	0.03	0.00	0.04	-0.01	0.03	0.14	0.02	0.06	-0.01
27	-0.01	0.00	0.05	0.02	0.05	0.26	0.07	0.21	0.00	0.15	0.05	0.00	0.03	0.03	0.25
28	-0.02	-0.12	0.01	0.01	-0.02	0.02	0.01	0.08	0.03	0.01	0.01	0.08	0.07	0.07	-0.02
29	-0.02	0.08	0.09	0.00	0.16	-0.02	0.02	-0.07	0.00	0.00	0.00	0.08	0.03	0.08	-0.02
30	0.00	-0.02	0.08	0.05	0.03	0.03	0.05	-0.02	0.12	0.05	-0.03	0.00	0.04	0.02	0.03

	16	**17**	**18**	**19**	**20**	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
1	0.01	0.07	-0.04	0.06	0.03	-0.02	0.00	-0.05	-0.04	0.03	-0.04	-0.01	-0.02	-0.02	0.00
2	0.00	-0.01	0.03	0.02	0.00	0.02	0.12	-0.14	-0.02	0.00	0.02	0.00	-0.12	0.08	-0.02
3	-0.14	-0.12	0.12	0.04	0.05	-0.03	-0.03	0.09	0.00	0.02	0.01	0.05	0.01	0.09	0.08
4	-0.02	-0.02	0.06	0.03	0.05	-0.04	0.09	0.03	0.04	0.07	0.02	0.02	0.01	0.00	0.05
5	-0.02	0.00	0.11	0.00	0.04	-0.01	0.10	-0.01	0.03	0.03	0.03	0.05	-0.02	0.16	0.03

6	0.00	0.00	-0.05	0.00	0.12	-0.03	0.08	0.04	-0.04	-0.03	0.02	0.26	0.02	-0.02	0.03
7	-0.18	-0.16	0.00	-0.02	0.00	-0.03	-0.02	0.01	0.00	-0.11	0.03	0.07	0.01	0.02	0.05
8	0.09	0.08	-0.15	-0.02	0.22	0.01	0.10	0.03	-0.01	0.18	0.00	0.21	0.08	-0.07	-0.02
9	0.08	0.01	-0.02	-0.01	-0.04	0.01	-0.04	0.00	0.00	0.05	0.04	0.00	0.03	0.00	0.12
10	0.00	0.10	0.01	-0.02	0.03	0.24	0.00	0.02	-0.04	0.19	-0.01	0.15	0.01	0.00	0.05
11	0.02	0.00	-0.04	0.00	0.00	0.17	-0.01	-0.01	0.00	-0.01	0.03	0.05	0.01	0.00	-0.03
12	0.00	-0.02	-0.01	-0.04	0.00	-0.03	0.04	-0.03	0.00	0.02	0.14	0.00	0.08	0.08	0.00
13	0.08	0.05	-0.01	-0.02	-0.01	0.00	-0.03	0.01	0.02	0.05	0.02	0.03	0.07	0.03	0.04
14	-0.07	-0.06	0.02	0.00	-0.04	-0.03	0.02	0.00	0.04	0.00	0.06	0.03	0.07	0.08	0.02
15	0.02	0.16	-0.01	0.00	0.17	-0.03	0.01	0.03	0.00	-0.08	-0.01	0.25	-0.02	-0.02	0.03
16	0.00	-0.29	0.00	0.00	0.13	0.02	-0.07	0.02	0.00	-0.03	-0.01	0.03	0.03	-0.02	0.09
17	-0.29	0.00	0.06	0.00	-0.04	-0.02	0.00	0.03	0.00	0.00	-0.02	0.01	0.01	-0.03	0.11
18	0.00	0.06	0.00	0.01	0.06	0.01	0.04	0.06	0.02	0.06	-0.01	0.07	0.02	-0.30	0.05
19	0.00	0.00	0.01	0.00	0.01	0.00	0.01	0.06	0.03	0.00	0.03	-0.01	0.01	-0.02	0.10
20	0.13	-0.04	0.06	0.01	0.00	0.00	0.19	0.03	0.02	-0.02	0.01	-0.12	0.00	0.00	0.03
21	0.02	-0.02	0.01	0.00	0.00	0.00	0.00	-0.01	0.00	-0.07	0.05	0.06	-0.03	0.04	-0.07
22	-0.07	0.00	0.04	0.01	0.19	0.00	0.00	0.04	0.06	-0.02	0.00	0.10	0.00	-0.01	0.00
23	0.02	0.03	0.06	0.06	0.03	-0.01	0.04	0.00	0.06	0.01	-0.04	0.02	-0.12	-0.01	-0.04
24	0.00	0.00	0.02	0.03	0.02	0.00	0.06	0.06	0.00	0.00	0.00	0.02	0.02	0.03	0.04
25	-0.03	0.00	0.06	0.00	-0.02	-0.07	-0.02	0.01	0.00	0.00	0.02	-0.05	0.03	0.02	-0.01
26	-0.01	-0.02	-0.01	0.03	0.01	0.05	0.00	-0.04	0.00	0.02	0.00	0.00	0.02	0.04	-0.04
27	0.03	0.01	0.07	-0.01	-0.12	0.06	0.10	0.02	0.02	-0.05	0.00	0.00	0.00	-0.01	0.04
28	0.03	0.01	0.02	0.01	0.00	-0.03	0.00	-0.12	0.02	0.03	0.02	0.00	0.00	0.03	0.00
29	-0.02	-0.03	-0.30	-0.02	0.00	0.04	-0.01	-0.01	0.03	0.02	0.04	-0.01	0.03	0.00	0.02
30	0.09	0.11	0.05	0.10	0.03	-0.07	0.00	-0.04	0.04	-0.01	-0.04	0.04	0.00	0.02	0.00

Legend: 1-Alcohol, 2-Coffee/tea, 3-Cured meat, 4-Dairy (higher fat), 5-Dairy (lower fat), 6-Eggs, 7-Fish, 8-Grains (other), 9-Green vegetables, 10-Onion/garlic, 11-Orange vegetables, 12-Other fruits, 13-Other vegetables, 14-Plant proteins, 15-Potato, 16-Poultry, 17-

Red meat, 18-Refined grains, 19-Salty snacks, 20-Saturated fats, 21-Soup, 22-Sugar, 23-Sweet beverages, 24-Sweet treats, 25-Tomato, 26-Tropical fruits, 27-Unsaturated fats, 28-Water, 29-Whole grains, 30-Other foods

Table N3: Weighted adjacency matrix of full network including dietary components and sociodemographic characteristics among adults ≥18 years in the 2015 Canadian Community Health Survey Nutrition

	1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**
1	0.00	0.03	0.28	0.15	0.25	0.50	0.02	0.19	0.04	0.21	-0.02	-0.07	0.00	0.00	0.00	-0.12	0.00	0.00	0.01
2	0.03	0.00	0.13	0.05	0.08	0.20	0.03	0.15	0.18	0.03	0.20	0.05	0.04	0.00	0.03	0.06	0.05	0.00	0.00
3	0.28	0.13	0.00	0.69	0.05	0.12	0.12	0.27	0.07	0.01	0.00	0.09	0.00	0.09	0.00	0.07	0.04	0.00	0.03
4	0.15	0.05	0.69	0.00	0.20	0.58	0.09	0.06	0.08	0.03	0.02	0.02	0.01	0.02	0.01	0.02	0.02	0.01	0.02
5	0.25	0.08	0.05	0.20	0.00	0.35	0.08	0.20	0.10	0.04	0.07	0.03	0.01	0.03	0.01	0.00	0.02	0.01	0.01
6	0.50	0.20	0.12	0.58	0.35	0.00	0.03	0.12	0.03	0.09	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.02	0.00
7	0.02	0.03	0.12	0.09	0.08	0.03	0.00	0.04	0.08	0.03	0.01	0.02	0.05	0.02	0.02	0.02	0.07	0.00	0.03
8	0.19	0.15	0.27	0.06	0.20	0.12	0.04	0.00	0.04	0.09	0.04	0.05	0.01	0.03	0.08	0.05	0.05	0.00	0.04
9	0.04	0.18	0.07	0.08	0.10	0.03	0.08	0.04	0.00	0.05	0.08	0.08	-0.02	0.01	0.11	-0.04	0.00	0.02	0.00
10	0.21	0.03	0.01	0.03	0.04	0.09	0.03	0.09	0.05	0.00	0.02	0.11	0.00	0.04	0.02	0.02	0.02	0.00	0.00
11	-0.02	0.20	0.00	0.02	0.07	0.00	0.01	0.04	0.08	0.02	0.00	0.03	0.02	0.06	-0.13	-0.03	0.00	0.00	-0.04
12	-0.07	0.05	0.09	0.02	0.03	0.00	0.02	0.05	0.08	0.11	0.03	0.00	0.02	0.00	-0.04	0.00	0.02	0.00	0.00
13	0.00	0.04	0.00	0.01	0.01	0.00	0.05	0.01	-0.02	0.00	0.02	0.02	0.00	0.01	0.00	0.02	0.00	0.00	0.00
14	0.00	0.00	0.09	0.02	0.03	0.01	0.02	0.03	0.01	0.04	0.06	0.00	0.01	0.00	0.00	0.00	0.01	-0.01	0.00
15	0.00	0.03	0.00	0.01	0.01	0.00	0.02	0.08	0.11	0.02	-0.13	-0.04	0.00	0.00	0.00	0.07	0.07	-0.02	-0.01
16	-0.12	0.06	0.07	0.02	0.00	0.00	0.02	0.05	-0.04	0.02	-0.03	0.00	0.02	0.00	0.07	0.00	-0.01	0.03	0.04
17	0.00	0.05	0.04	0.02	0.02	0.00	0.07	0.05	0.00	0.02	0.00	0.02	0.00	0.01	0.07	-0.01	0.00	0.05	0.12
18	0.00	0.00	0.00	0.01	0.01	0.02	0.00	0.00	0.02	0.00	0.00	0.00	0.00	-0.01	-0.02	0.03	0.05	0.00	0.04
19	0.01	0.00	0.03	0.02	0.01	0.00	0.03	0.04	0.00	0.00	-0.04	0.00	0.00	0.00	-0.01	0.04	0.12	0.04	0.00
20	0.03	0.05	0.00	0.02	0.05	0.00	0.03	0.09	-0.02	-0.01	-0.04	0.02	0.02	0.00	0.02	0.00	0.06	0.01	0.03
21	0.00	0.08	0.00	0.02	0.05	0.01	0.03	0.02	0.05	0.05	0.00	0.04	0.00	0.00	0.01	-0.04	0.09	0.23	0.12
22	-0.03	0.03	0.06	0.01	0.06	0.00	0.05	0.08	0.00	0.04	-0.08	0.03	-0.05	-0.04	-0.05	-0.01	0.01	0.04	0.04
23	0.03	0.01	0.03	0.01	0.03	0.01	0.04	0.00	0.04	0.00	0.00	0.00	0.02	-0.02	0.00	-0.25	0.00	0.01	0.06

24	-0.05	0.08	0.00	0.01	0.01	0.03	0.03	0.00	0.01	0.00	-0.14	-0.02	-0.01	0.00	-0.17	0.07	0.07	0.00	0.02
25	0.00	0.16	0.00	0.02	0.02	0.00	0.03	0.00	0.06	0.00	-0.13	-0.02	0.00	0.00	-0.16	0.07	0.01	0.10	0.00
26	0.00	0.08	0.03	0.03	0.00	0.01	0.05	0.04	-0.04	0.03	0.11	0.06	0.10	-0.04	0.00	-0.14	-0.01	0.01	-0.04
27	-0.06	0.01	0.01	0.00	0.01	0.00	0.03	0.02	0.07	0.03	0.04	0.02	0.00	0.00	-0.02	-0.02	0.00	-0.02	0.00
28	0.07	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.03	-0.01	0.05	0.05	0.03	0.12	0.00	0.22	-0.04	0.03	0.00
29	0.07	0.01	0.07	0.01	0.02	0.00	0.04	0.01	-0.02	0.00	-0.02	-0.04	0.00	-0.03	-0.03	0.02	0.01	0.24	0.17
30	0.00	0.07	0.09	0.02	0.00	0.00	0.01	0.07	0.00	0.11	-0.03	0.10	0.10	0.08	-0.02	0.09	-0.03	0.00	-0.01
31	-0.08	0.11	0.02	0.01	0.02	0.00	0.07	0.08	-0.06	-0.11	0.07	0.03	0.00	0.04	0.00	0.01	0.00	0.01	-0.01
32	0.01	0.00	0.07	0.04	0.00	0.01	0.03	0.03	-0.04	-0.02	0.00	0.03	0.02	-0.04	0.00	0.00	0.00	-0.03	0.00
33	-0.03	0.03	0.01	0.02	0.01	0.04	0.06	0.03	0.03	0.00	0.02	0.07	0.02	-0.03	-0.10	0.16	0.05	0.19	-0.01
34	0.07	0.03	0.07	0.01	0.10	0.05	0.01	0.03	-0.06	0.00	0.01	0.02	0.03	0.01	0.03	0.00	0.04	-0.01	0.03
35	0.01	0.08	0.01	0.02	0.02	0.00	0.01	0.02	-0.02	0.00	0.04	0.02	0.04	0.26	0.06	0.21	0.00	0.15	0.04
36	-0.08	0.06	0.02	0.05	0.02	0.07	0.08	0.03	-0.03	-0.10	0.01	0.00	-0.02	0.02	0.01	0.06	0.02	0.01	0.02
37	0.12	0.03	0.09	0.01	0.03	0.04	0.01	0.11	-0.02	0.05	0.08	0.00	0.15	-0.02	0.01	-0.06	0.00	0.00	0.00
38	-0.01	0.04	0.02	0.04	0.01	0.06	0.04	0.01	-0.02	-0.01	0.07	0.04	0.03	0.03	0.05	-0.02	0.13	0.05	-0.03

	20	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**	**32**	**33**	**34**	**35**	**36**	**37**	**38**
1	0.03	0.00	-0.03	0.03	-0.05	0.00	0.00	-0.06	0.07	0.07	0.00	-0.08	0.01	-0.03	0.07	0.01	-0.08	0.12	-0.01
2	0.05	0.08	0.03	0.01	0.08	0.16	0.08	0.01	0.00	0.01	0.07	0.11	0.00	0.03	0.03	0.08	0.06	0.03	0.04
3	0.00	0.00	0.06	0.03	0.00	0.00	0.03	0.01	0.00	0.07	0.09	0.02	0.07	0.01	0.07	0.01	0.02	0.09	0.02
4	0.02	0.02	0.01	0.01	0.01	0.02	0.03	0.00	0.02	0.01	0.02	0.01	0.04	0.02	0.01	0.02	0.05	0.01	0.04
5	0.05	0.05	0.06	0.03	0.01	0.02	0.00	0.01	0.00	0.02	0.00	0.02	0.00	0.01	0.10	0.02	0.02	0.03	0.01
6	0.00	0.01	0.00	0.01	0.03	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.04	0.05	0.00	0.07	0.04	0.06
7	0.03	0.03	0.05	0.04	0.03	0.03	0.05	0.03	0.01	0.04	0.01	0.07	0.03	0.06	0.01	0.01	0.08	0.01	0.04
8	0.09	0.02	0.08	0.00	0.00	0.00	0.04	0.02	0.00	0.01	0.07	0.08	0.03	0.03	0.03	0.02	0.03	0.11	0.01
9	-0.02	0.05	0.00	0.04	0.01	0.06	-0.04	0.07	0.03	-0.02	0.00	-0.06	-0.04	0.03	-0.06	-0.02	-0.03	-0.02	-0.02
10	-0.01	0.05	0.04	0.00	0.00	0.00	0.03	0.03	-0.01	0.00	0.11	-0.11	-0.02	0.00	0.00	0.00	-0.10	0.05	-0.01
11	-0.04	0.00	-0.08	0.00	-0.14	-0.13	0.11	0.04	0.05	-0.02	-0.03	0.07	0.00	0.02	0.01	0.04	0.01	0.08	0.07
12	0.02	0.04	0.03	0.00	-0.02	-0.02	0.06	0.02	0.05	-0.04	0.10	0.03	0.03	0.07	0.02	0.02	0.00	0.00	0.04

13	0.02	0.00	-0.05	0.02	-0.01	0.00	0.10	0.00	0.03	0.00	0.10	0.00	0.02	0.02	0.03	0.04	-0.02	0.15	0.03
14	0.00	0.00	-0.04	-0.02	0.00	0.00	-0.04	0.00	0.12	-0.03	0.08	0.04	-0.04	-0.03	0.01	0.26	0.02	-0.02	0.03
15	0.02	0.01	-0.05	0.00	-0.17	-0.16	0.00	-0.02	0.00	-0.03	-0.02	0.00	0.00	-0.10	0.03	0.06	0.01	0.01	0.05
16	0.00	-0.04	-0.01	-0.25	0.07	0.07	-0.14	-0.02	0.22	0.02	0.09	0.01	0.00	0.16	0.00	0.21	0.06	-0.06	-0.02
17	0.06	0.09	0.01	0.00	0.07	0.01	-0.01	0.00	-0.04	0.01	-0.03	0.00	0.00	0.05	0.04	0.00	0.02	0.00	0.13
18	0.01	0.23	0.04	0.01	0.00	0.10	0.01	-0.02	0.03	0.24	0.00	0.01	-0.03	0.19	-0.01	0.15	0.01	0.00	0.05
19	0.03	0.12	0.04	0.06	0.02	0.00	-0.04	0.00	0.00	0.17	-0.01	-0.01	0.00	-0.01	0.03	0.04	0.02	0.00	-0.03
20	0.00	0.06	0.04	0.00	0.00	-0.01	0.00	-0.04	0.00	-0.03	0.05	-0.03	0.00	0.02	0.13	0.00	0.08	0.07	0.00
21	0.06	0.00	0.05	-0.02	0.07	0.05	0.00	-0.02	-0.01	0.00	-0.02	0.01	0.02	0.05	0.02	0.03	0.07	0.03	0.04
22	0.04	0.05	0.00	-0.03	-0.07	-0.06	0.01	0.00	-0.04	-0.03	0.02	0.00	0.03	0.00	0.06	0.02	0.06	0.08	0.02
23	0.00	-0.02	-0.03	0.00	0.03	0.16	0.00	0.00	0.16	-0.03	0.00	0.03	0.00	-0.07	-0.01	0.24	0.00	-0.02	0.03
24	0.00	0.07	-0.07	0.03	0.00	-0.29	0.00	0.00	0.13	0.02	-0.07	0.01	0.00	-0.02	0.00	0.02	0.02	0.00	0.08
25	-0.01	0.05	-0.06	0.16	-0.29	0.00	0.05	0.00	-0.03	-0.02	0.00	0.02	0.00	0.00	-0.02	0.00	0.01	-0.03	0.10
26	0.00	0.00	0.01	0.00	0.00	0.05	0.00	0.00	0.06	0.00	0.03	0.05	0.02	0.05	0.00	0.06	0.02	-0.29	0.04
27	-0.04	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.02	0.00	0.00	0.06	0.03	0.00	0.04	0.00	0.00	0.00	0.09
28	0.00	-0.01	-0.04	0.16	0.13	-0.03	0.06	0.02	0.00	0.00	0.19	0.03	0.02	-0.02	0.01	-0.11	0.01	0.00	0.03
29	-0.03	0.00	-0.03	-0.03	0.02	-0.02	0.00	0.00	0.00	0.00	0.00	0.00	0.00	-0.07	0.04	0.05	-0.02	0.02	-0.06
30	0.05	-0.02	0.02	0.00	-0.07	0.00	0.03	0.00	0.19	0.00	0.00	0.03	0.06	-0.02	0.00	0.09	0.00	0.00	0.00
31	-0.03	0.01	0.00	0.03	0.01	0.02	0.05	0.06	0.03	0.00	0.03	0.00	0.06	0.00	-0.03	0.02	-0.13	0.00	-0.04
32	0.00	0.02	0.03	0.00	0.00	0.00	0.02	0.03	0.02	0.00	0.06	0.06	0.00	0.00	0.00	0.02	0.02	0.02	0.04
33	0.02	0.05	0.00	-0.07	-0.02	0.00	0.05	0.00	-0.02	-0.07	-0.02	0.00	0.00	0.00	0.02	-0.05	0.03	0.02	0.00
34	0.13	0.02	0.06	-0.01	0.00	-0.02	0.00	0.04	0.01	0.04	0.00	-0.03	0.00	0.02	0.00	0.00	0.03	0.03	-0.04
35	0.00	0.03	0.02	0.24	0.02	0.00	0.06	0.00	-0.11	0.05	0.09	0.02	0.02	-0.05	0.00	0.00	0.00	-0.01	0.03
36	0.08	0.07	0.06	0.00	0.02	0.01	0.02	0.00	0.01	-0.02	0.00	-0.13	0.02	0.03	0.03	0.00	0.00	0.04	-0.01
37	0.07	0.03	0.08	-0.02	0.00	-0.03	-0.29	0.00	0.00	0.02	0.00	0.00	0.02	0.02	0.03	-0.01	0.04	0.00	0.02
38	0.00	0.04	0.02	0.03	0.08	0.10	0.04	0.09	0.03	-0.06	0.00	-0.04	0.04	0.00	-0.04	0.03	-0.01	0.02	0.00

1-Age, 2-Sex, 3-Household food insecurity status, 4-Income, 5-Education, 6-Worked in the past week, 7-Geography, 8-Smoking status, 9-Alcohol, 10-Coffee/tea, 11-Cured meat, 12-Dairy (higher fat), 13-Dairy (lower fat), 14-Eggs, 15-Fish, 16-Grains (other), 17-Green vegetables, 18-Onion/garlic, 19-Orange vegetables, 20-Other fruits, 21-Other vegetables, 22-Plant proteins, 23-Potato, 24-

Poultry, 25-Red meat, 26-Refined grains, 27-Salty snacks, 28-Saturated fats, 29-Soup, 30-Sugar, 31-Sweet beverages, 32-Sweet treats, 33-Tomato, 34-Tropical fruits, 35-Unsaturated fats, 36-Water, 37-Whole grains, 38-Other foods

Appendix O: Bootstrapped confidence intervals of edge weights for the full network including sociodemographic characteristics and dietary components among adults ≥18 y in the 2015 Canadian Community Health Survey Nutrition

