ADDRESSING LIMITATIONS IN FOODBORNE OUTBREAK INVESTIGATION: RECALL BIAS AND THE FEASIBILITY OF NEW SURVEILLANCE STRATEGIES

A thesis submitted to the

College of Graduate and Postdoctoral Studies

In Partial Fulfillment of the Requirements

For the Degree of Master of Public Health – Thesis

In the School of Public Health

University of Saskatchewan

Saskatoon

By

PATRICK J. SEITZINGER

PERMISISON TO USE

In presenting this thesis in partial fulfilment of the requirements for a Postgraduate degree from the University of Saskatchewan, I agree that the Libraries of this University may make it freely available for inspection. I further agree that permission for copying of this thesis in any manner, in whole or in part, for scholarly purposes may be granted by the professor or professors who supervised my thesis work or, in their absence, by the Head of the Department or the Dean of the College in which my thesis work was done. It is understood that any copying or publication or use of this thesis or parts thereof for financial gain shall not be allowed without my written permission. It is also understood that due recognition shall be given to me and to the University of Saskatchewan in any scholarly use which may be made of any material in my thesis.

Requests for permission to copy or to make other use of material in this thesis in whole or part should be addressed to:

Executive Director, School of Public Health

University of Saskatchewan

Saskatoon, Saskatchewan S7N 2Z4

ABSTRACT

Accurate data on the incidence of foodborne illness and food histories for affected individuals represent two important barriers to enteric outbreak surveillance and response. Innovative tools to collect and analyze this type of public health intelligence will play an important role in research efforts to improve understanding of the extent, impact of and risk factors for foodborne disease in Canada and around the world. Ethica, a smartphone based application used to acquire, store, and analyze data on human behaviour, provided an opportunity to gather information on the occurrence of enteric illness and the food consumption behaviour of 96 university students over a 10-week period. Nausea or vomiting were reported by 34% of participants, and 29% reported diarrhea at least once during the study using at least one of the available reporting options, but only 7% reported they sought medical care. Real-time data collected through digital images, meal descriptions, and microsurveys were used as a reference to measure the sensitivity and specificity of traditional food history questionnaires administered through an email link after 7 or 18 days (2.5 weeks). The validity of food history data collected after 7 days was found to be consequentially low with sensitivities ranging from 14.3% for sprouts to 100% for leafy greens and specificities ranging from 30.4% for beef to 80.4% for peanuts. Similarly, the sensitivities of questions administered after 18 days ranged from 15.8% for sprouts to 77.8% for tomatoes, with specificities ranging from 21.2% for leafy greens to 92.1% for melons. The impact of recall bias on the accuracy of food history data was found to vary with food type. Bayesian latent class analysis was conducted to determine the sensitivities and specificities in the absence of a true gold standard – the results support those of frequentist approach. These findings serve as a first step in measuring the occurrence of self-reported foodborne illness and the implications of recall bias on outbreak investigations so that these biases can be accounted for research and public health practice.

ACKNOWLEDGMENTS

I would like to sincerely thank the Saskatchewan Health Research Foundation for funding this project. I would also like to express my gratitude to the Integrated Training Program in Infectious Diseases, Food Safety of the Collaborative Training and Research Program – Natural Sciences and Engineering Research Council of Canada, the University of Saskatchewan Western College of Veterinary Medicine and University of Saskatchewan School of Public Health for their continued funding and support.

CONTRIBUTIONS

The Ethica study that took place at the University of Saskatchewan was a collaboration between the School of Public Health, the College of Nursing and the Department of Computer Science.

Patrick Seitzinger's contributions to this project

- For phase 2 and 3 of data collection starting in January 2016, Patrick was part of the team that:
    - Oversaw recruitment of participants for the study together with the Social Sciences Research Laboratory at the University of Saskatchewan.
    - Introduced study participants to the software, the pause functionality, and how to answer questions posed by the on-phone survey tool.
    - Provided technical assistance to participants experiencing technical difficulties with the Ethica software.
    - Assisted in debriefing participants after their involvement in the study is concluded. This debriefing session will include a series of focus groups with study participants led by Dr. Wanda Martin from the College of Nursing.
- Compiled the information collected on foodborne illness, the PhotoFoodDiaries, and the microsurveys from Ethica, and linked it using a numeric identifier to the Qualtrics surveys containing participant demographics, food eating behaviors, and the traditional postoutbreak food consumption survey.
- Designed the analysis strategy to address both objectives for the study, analyzed the data, and reported the findings in a format suitable for peer-reviewed publication as well as for communicating with local and national public health authorities.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF ABBREVIATIONS

PHAC	Public Health Agency of Canada
BCCDC	British Columbia Centre for Disease Control
US CDC	United States Centers for Disease Control and Prevention
WHO	World Health Organization
PFGE	Pulse Field Electrophoresis
MLVA	Multi Loci VNTR Analysis
WGS	Whole Genome Sequencing
NSAGI	National Studies on Acute Gastrointestinal Illness
GPHIN	Global Public Health Intelligence Network

CHAPTER 1 – INTRODUCTION

1.1 Background

Foodborne illness is a preventable problem. Caused by the ingestion of contaminated foods, most foodborne illnesses are characterized by enteric symptoms including vomiting and diarrhea lasting anywhere from a few hours to several weeks. Contaminants that have been linked to foodborne illness include bacteria, parasites, viruses, metals, toxins and prions (Thomas et al., 2013). While mild cases often resolve on their own, severe cases may result in a variety of life threatening consequence including, but not limited to: electrolyte imbalances, dehydration, septicemia, renal failure, and permanent neurological damage (Thomas et al., 2013). The incubation period – the time between consumption of the contaminated food and the onset of illness – ranges from hours to days, depending on the pathogen involved (CIFOR, 2014), but can be weeks for infectious agents such as listeriosis. These delays often make it difficult to link an episode of illness with the consumption of a particular food.

The global burden of food borne illness is significant in terms of the occurrence of disease and mortality as well as the economic burden associated with costs of treatment and lost days from work. The World Health Organization estimated 600 million cases of foodborne illness in 2010 alone (WHO, 2015b). The Public Health Agency of Canada (PHAC) has estimated that each year there are 4.0 million occurrences of domestically acquired foodborne illness in Canada with 1.6 million (40%) of these related to 30 known pathogens (Thomas et al., 2013). Of the total cases of foodborne illness in Canada, 11,600 are estimated to result in hospitalizations and 238 in death with 4000 (34%) of the hospitalizations and 105 (44%) of the deaths associated with domestically acquired illness due to the 30 known pathogens (Thomas et al., 2015). In an earlier study, the frequency of gastroenteritis had been described for Hamilton,

Ontario where there were estimated to be 126,300 cases per year per 100,000 people (Majowicz et al., 2006). The mean annual cost associated with each of these episodes was estimated to be $1089 CAN resulting in an annual economic burden of $115 CAN per capita in the community (Majowicz et al., 2006).

It should be noted that the definition of foodborne illness varies across studies – an issue that has been widely acknowledged (de Wit et al., 2001; Majowicz et al., 2004; Wheeler et al., 1999). In a study on the prevalence of foodborne illness in Canada, Thomas et al. (2013) defined acute cases of foodborne illness as having >3 loose stools in 24h with duration lasting >1 day. In a similar study in the US, Scallan et al. (2011) had used a slightly expanded case definition of >3 loose stools in 24 hours lasting >1 day or resulting in restricted daily activities. Other studies, such as Majowicz et al. (2004), defined foodborne illness broadly as diarrhea (loose stool or stool that is unusually liquid) or vomiting within the past 28 days. The variability in the definition of foodborne illness creates challenges in making accurate inferences about the occurrence of foodborne illness and in comparing incidence over time and differences among regions and countries.

1.1.1 Foodborne Illness Surveillance

A number of surveillance systems exist to collect information on foodborne illness in Canada. The Canadian Notifiable Disease Surveillance System (CNDSS) aggregates and summarizes data on laboratory confirmed cases that have been reported to provincial public health authorities (Government of Canada, 2012). The National Enteric Surveillance Program collects data on subtype and species of select bacteria, parasites and viruses on a weekly basis (Government of Canada, 2012). FoodNet Canada collects data on foodborne illness occurrence

and tests for pathogens in retail foods, water sources and agricultural operations in three sentinel sites – Middlesex-London Health Unit (Ontario), Fraser Health Region (BC), and Alberta Health Services: Calgary and Central Zones (Alberta) (Huang et al., 2015). National Studies on Acute Gastrointestinal Illness (NSAGI) have conducted population surveys on the occurrence of diarrhea and vomiting (Thomas et al., 2008). Pathogen specific surveillance programs also exist such as the Enhanced National Listeriosis Surveillance system which collects data specific to the occurrence and spread of invasive listeriosis (Government of Canada, 2016). Communication between laboratories regarding molecular diagnostics is facilitated by PulseNet Canada, an electronic database of pulsed-field gel electrophoresis (PFGE) results for *E. coli* 0157:H7, Salmonella, *Listeria monoctogenes*, *Campylobacter*, *Shigella* and *Vibrio* (Government of Canada, 2015). This network allows public health practitioners to quickly match cases infected with the same outbreak strain, regardless of where the cases were reported. Data on foodborne illness is also collected by the Canadian Integrated Program for Antimicrobial Resistance Surveillance (CIPARS) which assesses trends in antimicrobial resistance in bacterial pathogens such as *Salmonella* (Nesbitt et al., 2012). Although there are many different programs to collect data related to foodborne illness in Canada, limitations to surveillance still exist. Certain subgroups of the Canadian population including rural (Herikstad et al., 2002), Indigenous (Clarke, 2016; O'Neil et al., 1998), and immigrant populations (Clarke, 2016; Sanmartin and Ross, 2006) have been found to be under-represented by most surveillance initiatives. Furthermore, there is a lack of mechanisms in place to capture information on cases who do not seek medical care (Tam et al., 2003).

Due to the nature of the illness, most mild or moderate cases do not visit a health care provider creating serious challenges for the surveillance of foodborne disease (Flint et al., 2004).

In order for cases to be captured in national surveillance systems, the patient must visit a healthcare provider and submit a biological specimen (usually a sample of stool or vomit). Furthermore, a pathogen must be identified and the laboratory results must be reported to the local health department and then the provincial health authority (Flint et al., 2004). Examples of important pathogens resulting in illness that are not typically captured by routine surveillance include agents such as: norovirus, rotavirus, *C. perfringens, B. cereus, Y. enterocolitica*, and *S. aureus*. Since data on these pathogens are incomplete for Canada, national estimates often rely on data from other countries such as the UK and USA and individual provinces with enhanced surveillance or research initiatives (Thomas et al., 2013).

Some types of foodborne illness are more successfully captured by existing surveillance systems. Illnesses caused by pathogens that are considered to be severe, such as *Vibrio culnificus, E. coli O157:H7 and L. monocytogenes*, as well as pathogens that are well understood, are more likely to be recognized and reported (Thomas et al., 2015). Pathogens that are included in existing surveillance systems, such as *Campylobacter* spp., and nontyphoidal *Salmonella* spp., are also more likely to be captured (Thomas et al., 2015). Differences exist between laboratories in regards to which pathogens are included in routine diagnostic panels. A study involving 87% of laboratories in Canada found that 67% routinely tested stool samples for enteric bacteria, 31% tested for parasites, and only 10% tested for viruses (Flint et al., 2004). Pathogens that were tested for by at least 95% of laboratories included *Salmonella* spp., *Shigella* spp., *Camplobacter* spp., *E. coli* and *Yersinia* spp. Differences in the severity of symptoms, laboratory testing practices and reporting policies create challenges for surveillance systems and make it difficult to obtain complete estimates of the occurrence of foodborne illness in Canada.

1.1.2 Surveillance of Food Choices

In 2014, the Public Health Agency of Canada conducted the national survey known as Foodbook to assess food, water and animal exposures in the Canadian population (MacDonald, 2016). The objective of the Foodbook study was to fill gaps in data on these exposures in a Canadian context. Interviews were administered over the phone to 11,016 participants with a strong focus on food exposures (MacDonald, 2016). This data was used to provide a reference for the frequency of food exposures in the Canadian population. If an outbreak investigation indicates that the proportion of cases exposed to a particular food item is considerably higher than the Canadian average, the food is flagged as a possible source of infection. The results provide practical information on frequency of exposure to high-risk foods across different provinces/territories and demographics in Canada and continues to inform outbreak investigation and response.

1.1.3 Foodborne Illness Outbreak Detection and Investigation in Canada

When surveillance systems detect unusual cases of illness, or a greater number of illnesses than is expected, a food-related outbreak may be suspected (Vik and Hexemer, 2014). A foodborne outbreak refers to an incident that involves 2 or more individuals becoming ill due to the consumption of a common food (Lukacsovics et al., 2014). When an outbreak is confined within the borders of health region or a province, local and provincial epidemiologists may conduct an investigation to identify the source of illness. Although epidemiologists aim to collect food history data by the end of the first week after the onset of illness, in practice this timeframe is often extended considerably. Findings by Fong et al. (2017) indicate that the median time to initiate an outbreak investigation at a provincial level is 36 days. When outbreaks span multiple

provinces, these timeframes may be further extended to several weeks. To help mitigate challenges in communication and coordination, the Public Health Agency of Canada may activate an Outbreak Investigation Coordination Committee (Vik and Hexemer, 2014). This coordination committee facilitates collaboration and information sharing among jurisdictions. Canada's Food-borne Illness Outbreak Response Protocol (FIORP) was developed to maximize the efficiency and effectiveness of national enteric outbreak investigation by outlining the roles and responsibilities of federal, provincial and territorial partners and outline protocols to guide public health action (Vik and Hexemer, 2014).

Public health decisions during outbreak investigations are informed by a combination of epidemiological, laboratory and food safety evidence (Vik et al., 2014). The type and quality of evidence that is available differs from one outbreak to the next. If samples of suspect foods are available, laboratory molecular-typing techniques can be used to match the genetic fingerprint of the outbreak strain isolated from cases to the pathogen in the food. However, if no food samples are available, investigators must rely on the accounts of affected individuals to identify the source of the outbreak (Vik et al., 2014). This data is collected through a food history questionnaire, which typically prompts cases to recall foods that were consumed before the onset of the illness.

Current methods of collecting food histories typically involve questionnaires being administered in-person or via telephone. These traditional methods require considerable time commitment by staff, cases, and controls. During case-control studies, controls are interviewed in a similar manner, increasing the resources and time required to collect information. Studies that administer similar questionnaires over the phone have shown relatively low response rates (MacDonald, 2016). Possible challenges in contacting participants for interviews may include

decreasing numbers of geographically referenced landlines, lack of time, suspicion of telemarketers, and the option to refuse calls based on call display.

The number of households that have mobile phones has surpassed the number that have landlines (Radio-television and Commission, 2015). Including mobile phones into food exposure studies has been used to increase coverage of younger Canadians (MacDonald, 2016). This strategy is appropriate considering that the Millennial Generation (aged 18-35) represents a large portion of smartphone users (Poushter, 2016). University students in particular have shown to be among the most frequent users of these new technologies; they have been found to spend an average of 5-6 hours per day interacting with their smartphones (Lepp et al., 2015).

1.1.4 Emerging Methods to Collect Enteric Illness and Food Histories

New technologies have provided opportunities for innovation in methods to collect food histories. As an alternative to in-person and telephone interview strategies, several studies have begun to assess the feasibility of web-based food surveys as well as digital image food diaries (Arab et al., 2011; Kikunaga et al., 2007; Six et al., 2010; Wang et al., 2010). Online data collection techniques were found to aid the recall of food consumption, and digital images have provided an efficient means by which to capture detailed information about food consumption (Arab et al., 2011). Web-based surveys have been implemented in a number of outbreaks (Beatty et al., 2009; Srikantiah, 2005) and have been shown to reduce investigation time without affecting response rates (Ghosh, 2008). However, retrospective dietary histories are only as accurate as an individual's memory of the meal in question, regardless of the method of administration.

The wide range of capabilities of smartphones has proven useful to the investigation of human behaviours and administration of online surveys which includes the potential for real-time collection of high resolution digital images, text and audio files tagged with time and geographic location (Hashemain et al., 2012). The Ethica smartphone app provides a tool for the practical collection of data regarding enteric illness and food history. The Ethica Survey Tool allows researchers to administer app-triggered and user-triggered surveys throughout the study period to collect illness and risk factors. Various other features of this app serve as a user-friendly interfaces through which participants can record data on food consumption behaviour (Features - Ethica Data, 2017). For example, the PhotoFoodDiary feature of the application allows participants to capture digital images of foods and meal descriptions in real-time.

1.1.5 Emerging Methods to Detect Foodborne Outbreaks

Innovative uses of big data from a variety of sources including news reports, internet forums and social media have been utilized to detect disease outbreaks. Systems of particular relevance to foodborne illness detection include Yelp, the HealthMap Food Dashboard, and nEmesis. For example, restaurant reviews such as those on Yelp, were used to identify unreported cases of foodborne illness in New York City in 2012-13. The potential of this strategy is illustrated in a study by which data from Yelp reviews was used to identify three previously unreported restaurant-linked outbreaks over a course of nine months (Harris et al., 2014).

HealthMap Food Dashboard identifies Tweets that pertain to foodborne illness by searching for references to food poisoning (Harris et al., 2017). This system can identify potential cases at a global level or can be focused on any specific location. Taking this concept a step further, an adaptive algorithm based system known as nEmesis applies a more complex

machine learning algorithm to the recognition of posts on Twitter describing foodborne illness and links them to food service establishments visited in the previous five days by the person making the post (Sadilek et al., 2016). By matching each flagged location with a control site, researchers conducted a double blind trial and identified a 64% improvement in inspection efficiency over existing methods.

While all of these methods have provided important insights into foodborne illness distribution, they share limitations including self-reporting and misinformation, low specificity of signals, and a disproportional sensitivity to external forces such as the media, selection biases, poor specificity of syndromic definitions, and problems associated with inconsistent participation (Brownstein et al., 2009; Wójcik et al., 2014).

1.2 The Problem

Collecting accurate food histories and mitigating recall bias have been recognized by PHAC (Thomas et al., 2013) and the US CDC (Scallan et al., 2011) as key challenges in the area of foodborne illness investigation. Few studies have attempted to measure the effect of recall bias on the validity of data collected during outbreak investigations or more specifically how recall bias is affected by time since exposure. Decker et al. (1986) investigated the accuracy of dietary recall 2 to 3 days after a buffet meal that was videotaped. Sensitivities were found to range from 81.2% to 95.2% and specificities ranged from 93.1% to 98.5% for 5 different food groups. Mann (1981) assessed the accuracy of dietary recall to potato salad and quiche five days after a work luncheon where nurses observed and tracked which participants consumed the two foods. The sensitivity of the food history questionnaire was found to be 88% for both foods; the specificity for potato salad was 75% and for quiche was 93%. In the context of past outbreak investigation in Canada, outbreak reports indicate that even cases infected with the same

outbreak strain that was isolated from the contaminated food source – as confirmed with laboratory serotyping methods such as PFGE, MLVA and WGS – often misclassify their exposure status on food history questionnaires (Angelo et al., 2015; Calder et al., 2003; Gupta et al., 2007; Slayton et al., 2013). These misclassifications limit the power to detect meaningful associations between certain food exposures and the onset of illness. Considering the findings of Decker et al. (1986) and Mann (1981) and the fact that the median time to start an outbreak investigation in British Columbia is 36 days (Fong et al., 2017), it is reasonable to predict that recall bias will have an important impact on the validity of data collected during outbreak investigations.

Addressing the challenges of recall bias in a feasible and effective way would increase the efficiency of foodborne illness surveillance systems as well as inform investigative strategies to stop outbreaks in a timely manner and prevent further illness. As the food distribution network in Canada and the US continues to grow in size and complexity, so do the potential consequences of incomplete reporting. The development of innovative methods and tools for self-reporting food exposures and illness is both timely and necessary.

1.3 Need for an Interdisciplinary Approach

Addressing the issues of incomplete reporting of foodborne illness and recall bias resulting from delays in collecting food histories will require a multidisciplinary approach. The issue of recall bias, for example, encompasses aspects of psychology and epidemiology. Insights gained into the mechanisms underlying psychological processes such as memory encoding, storage and retrieval have guided strategies used to collect data for epidemiological studies. Of particular relevance to this study is progress made in understanding how memory decay occurs (Jenkins et al., 2002) and which types of memory are most susceptible. Studies by Janssen et al.

(2006) suggest that the length of time that passes between an event and when the memory is retrieved may impact the validity of data that is obtained. More specifically, the more time that passes between exposure and recall, the more likely it is that the event is to be misattributed to an earlier or later date. In the context of food consumption, it has been shown that strategies used to retrieve details about meals eaten in the past differ depending on food type (Johnson-Kozlow et al., 2006). The methods that participants used to remember exposures to different food types were each characterized by a distinct pattern of strategies such as using time and location cues, knowledge of habitual dietary routines and external influences (associations with holidays, activities, friends and family) (Johnson-Kozlow et al., 2006).

1.4 The Research Opportunity

There is a need for research on foodborne consumption behaviour and the occurrence of foodborne illness. Gaps in data are particularly evident in the demographic of university students and for the province of Saskatchewan. Conducting a study to address this gap would require considerable collaboration and resources including access to university students in Saskatchewan, access to appropriate data collection tools, and funding support. The Social Science and Research Laboratories (SSRL) at the University of Saskatchewan is a research support unit that provides the infrastructure, research space and resources to support research projects. During this study, the SSRL provided an effective way to recruit and interact with student volunteers at the university and supported the data collection process. Data were collected using the Ethica app developed by Drs. Nathaniel Osgood, Kevin Stanley, Mohammad Hashemian and colleagues at the University of Saskatchewan. Collaboration with the developers of the app provided the access, support and expertise needed to apply this new technology to the

field of foodborne illness. Funding for the project was received from the Saskatchewan Health Research Foundation through the Collaborative Innovation Development Grant.

1.5 Objectives

1.5.1 Assessing the Feasibility of New Technologies for the Collection of Food Histories

By providing a convenient method of triggering foodborne illness surveys whenever an episode of enteric illness occurs, the prevalence and underreporting associated with such illness can be more completely measured. The objective of Chapter 2 was to assess the feasibility of new smartphone technology known as Ethica to report symptoms of enteric illness and to collect comprehensive accounts of what each participant consumed. Different features of the application were utilized including the PhotoFoodDiary and the Survey Tool to allow participants to capture digital images of foods, describe meals in written text or through audio voice recordings and to complete microsurveys.

1.5.2 Quantifying Recall Bias of Retrospective Food History Questionnaires Through Comparison to Data Collected using the Ethica App

The objective of Chapter 3 was to quantify recall bias by assessing the accuracy of food history data collected using traditional retrospective questionnaires. The questionnaire was administered at time intervals designed to resemble a range of plausible local, provincial and national enteric outbreak investigations conducted by public health officials in Canada. Data collected from digital images, meal descriptions and microsurveys collected using the Ethica app were used as a reference standard. The results of the food history questionnaire were compared to this smartphone-based reference standard using a traditional, frequentist approach. The effect

of factors, such as differences in time delays from exposure to recall, experiencing symptoms of enteric illness, and repeated assessments, were assessed on the accuracy of dietary recall. The prevalences of food exposures in the sample population were then compared to the general Canadian population. These results will allow for a better understanding of how recall bias affects the validity of data collected during foodborne outbreak investigation and whether this effect is consistent across food types.

1.5.3 Assessing the Sensitivity and Specificity of Different Data Collection Methods through Bayesian Latent Class Modelling

The purpose of the third research chapter was to apply an alternative approach – Bayesian latent class modelling – to determine the sensitivity and specificity of test 1 (digital images/meal descriptions), test 2 (microsurveys) and test 3 (food history questionnaire) for five food items including tomatoes, cucumbers, lettuce and leafy greens, nuts, and berries. This approach allowed for previous knowledge regarding the measures of recall uncertainty to be considered in the model and reduced the need to assume there was a gold standard test for comparison. The results obtained from the Bayesian approach for the retrospective food history questionnaire were compared to the results of the more traditional, frequentist approach applied in the second research chapter.

1.6 Implications

This study has important theoretical practical implications for foodborne illness detection and investigation. Since the time periods and questionnaires used were designed to resemble those currently being used, the findings of these studies will be directly applicable to current public health practice in Canada. Measures of the sensitivity and specificity of these

questionnaires will allow outbreak investigators to account for recall bias when evaluating data from outbreak investigations and to generate more informed hypothesis about the source of illness. Results on the feasibility of different data collection methods, including new and innovative technologies, will allow national surveillance systems to further develop strategies to collect in-depth data from target groups of interest. These investigations will contribute to a better understanding of the extent, risk factors and impact of foodborne illnesses in Canada.

From a methodological perspective, this study also explored new applications of Bayesian latent class modelling; this is the first identified example where Bayesian Latent Class modelling was used to investigate the accuracy of food history questionnaires. By utilizing both frequentist and Bayesian methods, direct comparisons can be drawn between the two methods. The innovative combination of data collection methods, analysis techniques and investigation strategies may foster further interdisciplinary collaborations to address the problem of foodborne illness.

1.7 References

Angelo, K. M., Chu, A., Anand, M., Nguyen, T. A., Bottichio, L., Wise, M., Williams, I., Seelman, S., Bell, R., Fatica, M., Lance, S., Baldwin, D., Shannon, K., Lee, H., Trees, E., Strain, E., Gieraltowski, L., Centers for Disease, C., & Prevention. (2015). Outbreak of Salmonella Newport infections linked to cucumbers--United States, 2014. *MMWR Morb Mortal Wkly Rep, 64*(6), 144-147.

Arab, L., Estrin, D., Kim, D. H., Burke, J., & Goldman, J. (2011). Feasibility testing of an automated image-capture method to aid dietary recall. *Eur J Clin Nutr, 65*(10), 1156-1162. doi:10.1038/ejcn.2011.75

Beatty, M., Shevick, G., Shupe-Ricksecker, K., Bannister, E., Tulu, A., Lancaster, K., Alexander, N., Zellner, D., Lyszkowicz, E., & Braden, C. (2009). Large Salmonella Enteritidis outbreak with prolonged transmission attributed to an infected food handler, Texas, 2002. *Epidemiology and Infection, 137*(03), 417-427.

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection—harnessing the Web for public health surveillance. *New England Journal of Medicine, 360*(21), 2153-2157.

Calder, L., Simmons, G., Thornley, C., Taylor, P., Pritchard, K., Greening, G., & Bishop, J. (2003). An outbreak of hepatitis A associated with consumption of raw blueberries. *Epidemiology & Infection, 131*(1), 745-751.

CIFOR. (2014). Guidelines for foodborne disease outbreak response [Internet]. Atlanta: Council of State and Territorial Epidemiologists.

Clarke, J. (2016). Difficulty accessing health care services in Canada.

de Wit, Koopmans, M., Kortbeek, L., Wannet, W., Vinje, J., Van Leusden, F., Bartelds, A., & Van Duynhoven, Y. (2001). Sensor, a population-based cohort study on gastroenteritis in the Netherlands: incidence and etiology. *American journal of epidemiology, 154*(7), 666-674.

Decker, M., Booth, A., Dewey, M., Fricker, R., Hutcheson, R., & Schaffner, W. (1986). Validity of food consumption histories in a foodborne outbreak investigation. *American journal of epidemiology, 124*(5), 859-863.

Features - Ethica Data. (2017). Retrieved from https://www.ethicadata.com/features

Flint, J. A., Doré, K., Majowicz, S. E., Edge, V. L., & Sockett, P. (2004). From stool to statistics: reporting of acute gastrointestinal illnesses in Canada. *Canadian Journal of Public Health/Revue Canadienne de Sante'e Publique*, 309-313.

Fong, D., Otterstatter, M., Taylor, M., & Galanis, E. (2017). Analysis of enteric disease outbreak metrics, British Columbia Centre for Disease Control, 2005-2014. *Canada Communicable Disease Report, 43*(1), 1.

Ghosh, T. S. (2008). Internet-versus Telephone-based Local Outbreak Investigations-Volume 14, Number 6—June 2008-Emerging Infectious Disease journal-CDC.

Government of Canada. (2012). *Canadian National Enteric Pathogen Surveillance National Report (2005-2008): Public Health Agency of Canada*. Retrieved from Ottawa:

Government of Canada. (2015). PulseNet Canada. Retrieved from https://www.nml-lnm.gc.ca/Pulsenet/index-eng.htm

Government of Canada. (2016). Surveillance of food-borne illness in Canada. Retrieved from https://www.canada.ca/en/public-health/services/food-borne-illness-canada/surveillance-food-borne-illness-canada.html

Gupta, S. K., Nalluswami, K., Snider, C., Perch, M., Balasegaram, M., Burmeister, D., Lockett, J., Sandt, C., Hoekstra, R. M., & Montgomery, S. (2007). Outbreak of Salmonella Braenderup infections associated with Roma tomatoes, northeastern United States, 2004: a useful method for subtyping exposures in field investigations. *Epidemiol Infect, 135*(7), 1165-1173. doi:10.1017/S0950268807007911

Harris, J. K., Hawkins, J. B., Nguyen, L., Nsoesie, E. O., Tuli, G., Mansour, R., & Brownstein, J. S. (2017). Using Twitter to Identify and Respond to Food Poisoning: The Food Safety STL Project. *Journal of public health management and practice: JPHMP*.

Harris, J. K., Mansour, R., Choucair, B., Olson, J., Nissen, C., Bhatt, J., Control, C. f. D., & Prevention. (2014). Health department use of social media to identify foodborne illness-Chicago, Illinois, 2013-2014. *MMWR Morb Mortal Wkly Rep, 63*(32), 681-685.

Hashemian, M., Knowles, D. C., Calver, J., Qian, W., Bullock, M. C., Bell, S., Mandryk, R. L., Osgood, N., & Stanley, K. G. (2012). iEpi: an end to end solution for collecting, conditioning and utilizing epidemiologically relevant data. *Proceedings of the 2nd ACM international workshop on Pervasive Wireless Healthcare*.

Herikstad, H., Yang, S., Van Gilder, T., Vugia, D., Hadler, J., Blake, P., Deneen, V., Shiferaw, B., Angulo, F., & Group, F. W. (2002). A population-based estimate of the burden of diarrhoeal illness in the United States: FoodNet, 1996–7. *Epidemiology & Infection, 129*(1), 9-17.

Huang, H., Brooks, B. W., Lowman, R., & Carrillo, C. D. (2015). Campylobacter species in animal, food, and environmental sources, and relevant testing programs in Canada. *Canadian journal of microbiology, 61*(10), 701-721.

Janssen, S. M., Chessa, A. G., & Murre, J. M. (2006). Memory for time: How people date events. *Memory & Cognition, 34*(1), 138-147.

Jenkins, P., Earle-Richardson, G., Slingerland, D. T., & May, J. (2002). Time dependent memory decay. *American journal of industrial medicine, 41*(2), 98-101.

Johnson-Kozlow, M., Matt, G. E., & Rock, C. L. (2006). Recall strategies used by respondents to complete a food frequency questionnaire: an exploratory study. *J Am Diet Assoc, 106*(3), 430-433. doi:10.1016/j.jada.2005.12.007

Kikunaga, S., Tomoe, T., Ishibashi, G., & Da-Hong, W. (2007). The application of a handheld personal digital assistant with camera and mobile phone card (Wellnavi) to the general population in a dietary survey. *Journal of nutritional science and vitaminology, 53*(2), 109-116.

Lepp, A., Barkley, J. E., & Karpinski, A. C. (2015). The relationship between cell phone use and academic performance in a sample of US college students. *Sage Open, 5*(1), 2158244015573169.

Lukacsovics, A., Nesbitt, A., Marshall, B., Asplin, R., Stone, J., Embree, G., Hurst, M., & Pollari, F. (2014). Using environmental health officers' opinions to inform the source attribution of enteric disease: further analysis of the "most likely source of infection". *BMC Public Health, 14*, 1258. doi:10.1186/1471-2458-14-1258

MacDonald, D. (2016). *Foodbook: The Canadian Food, Water and Animal Exposure Study.* Paper presented at the IAFP 2016 Annual Meeting.

Majowicz, S., Dore, K., Flint, J., Edge, V., Read, S., Buffett, M., McEwen, S., McNab, W., Stacey, D., & Sockett, P. (2004). Magnitude and distribution of acute, self-reported gastrointestinal illness in a Canadian community. *Epidemiology and Infection, 132*(04), 607-617.

Majowicz, S., McNab, W., Sockett, P., Henson, S., Dore, K., Edge, V., Buffett, M., Fazil, A., Read, S., & McEwen, S. (2006). Burden and cost of gastroenteritis in a Canadian community. *Journal of food protection, 69*(3), 651-659.

Mann, J. M. (1981). A prospective study of response error in food history questionnaires: implications for foodborne outbreak investigation. *American journal of public health, 71*(12), 1362-1366.

Nesbitt, A., Ravel, A., Murray, R., McCormick, R., Savelli, C., Finley, R., Parmley, J., Agunos, A., Majowicz, S., & Gilmour, M. (2012). Integrated surveillance and potential sources of Salmonella Enteritidis in human cases in Canada from 2003 to 2009. *Epidemiology and Infection, 140*(10), 1757.

O'Neil, J., Reading, J., & Leader, A. (1998). Changing the relations of surveillance: the development of a discourse of resistance in Aboriginal epidemiology. *Human organization, 57*(2), 230-237.

Poushter, J. (2016). Smartphone ownership and Internet usage continues to climb in emerging economies. *Pew Research Center*.

Radio-television, C., & Commission, T. (2015). Communications Monitoring Report 2015: Canada's Communications System: An Overview for Citizens, Consumers, and Creators. Ottawa: Government of Canada.

Sadilek, A., Kautz, H. A., DiPrete, L., Labus, B., Portman, E., Teitel, J., & Silenzio, V. (2016). *Deploying nEmesis: Preventing Foodborne Illness by Data Mining Social Media.* Paper presented at the AAAI.

Sanmartin, C., & Ross, N. (2006). Experiencing difficulties accessing first-contact health services in Canada. *Healthcare Policy, 1*(2).

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., Jones, J. L., & Griffin, P. M. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis, 17*(1).

Six, B. L., Schap, T. E., Zhu, F. M., Mariappan, A., Bosch, M., Delp, E. J., Ebert, D. S., Kerr, D. A., & Boushey, C. J. (2010). Evidence-based development of a mobile telephone food record. *Journal of the American Dietetic Association, 110*(1), 74-79.

Slayton, R. B., Turabelidze, G., Bennett, S. D., Schwensohn, C. A., Yaffee, A. Q., Khan, F., Butler, C., Trees, E., Ayers, T. L., Davis, M. L., Laufer, A. S., Gladbach, S., Williams, I., & Gieraltowski, L. B. (2013). Outbreak of Shiga toxin-producing Escherichia coli (STEC) O157:H7 associated with romaine lettuce consumption, 2011. *PLoS One, 8*(2), e55300. doi:10.1371/journal.pone.0055300

Srikantiah, P. (2005). Web-based Investigation of Multistate Salmonellosis Outbreak-Volume 11, Number 4—April 2005-Emerging Infectious Disease journal-CDC.

Tam, C. C., Rodrigues, L. C., & O'Brien, S. J. (2003). The study of infectious intestinal disease in England: what risk factors for presentation to general practice tell us about potential for selection bias in case-control studies of reported cases of diarrhoea. *International Journal of Epidemiology, 32*(1), 99-105.

Thomas, M. K., Majowicz, S. E., Sockett, P. N., & Pollari, F. (2008). Burden of acute gastrointestinal illness in Canada, 1999-2007: interim summary of NSAGI activities. *Canada communicable disease report= Releve des maladies transmissibles au Canada, 34*(5), 8-15.

Thomas, M. K., Murray, R., Flockhart, L., Pintar, K., Fazil, A., Nesbitt, A., Marshall, B., Tataryn, J., & Pollari, F. (2015). Estimates of Foodborne Illness–Related Hospitalizations

and Deaths in Canada for 30 Specified Pathogens and Unspecified Agents. *Foodborne pathogens and disease, 12*(10), 820-827.
Thomas, M. K., Murray, R., Flockhart, L., Pintar, K., Pollari, F., Fazil, A., Nesbitt, A., & Marshall, B. (2013). Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne pathogens and disease, 10*(7), 639-648.
Vik, J., & Hexemer, A. (2014). Summary: Canada's Food-borne Illness Outbreak Response Protocol. *Canada Communicable Disease Report, 40*(14), 306.
Vik, J., Hexemer, A., & Farber, J. (2014). Summary: Weight of evidence-Factors to consider when investigating a food-borne illness outbreak. *Canada Communicable Disease Report, 40*(14), 303.
Wang, L., Xu, Y., Wang, Y., Dong, S., Cao, Z., Zhou, W., Sun, H., Huo, D., Zhang, H., Sun, Y., Huang, L., Yuan, Z., Zeng, D., & Song, H. (2010). The Epidemiological Investigation and Intelligent Analytical System for foodborne disease. *Food Control, 21*(11), 1466-1471. doi:10.1016/j.foodcont.2010.04.015
Wheeler, J. G., Sethi, D., Cowden, J. M., Wall, P. G., Rodrigues, L. C., Tompkins, D. S., Hudson, M. J., & Roderick, P. J. (1999). Study of infectious intestinal disease in England: rates in the community, presenting to general practice, and reported to national surveillance. *Bmj, 318*(7190), 1046-1050.
WHO. (2015b). *WHO estimates of the global burden of foodborne diseases: foodborne disease burden epidemiology reference group 2007-2015*: WHO.
Wójcik, O. P., Brownstein, J. S., Chunara, R., & Johansson, M. A. (2014). Public health for the people: participatory infectious disease surveillance in the digital age. *Emerging themes in epidemiology, 11*(1), 7.

CHAPTER 2 – FEASIBILITY OF SMARTPHONE-BASED TECHNOLOGY TO SUPPORT FOODBORNE DISEASE SURVEILLANCE

2.1 Abstract

Accurate data on the incidence of enteric illness is crucial to understanding the extent, impact of and risk factors for foodborne disease in Canada. Traditional surveillance methods rely on data collected from health practitioners, yet only a fraction of cases seek medical care for enteric illnesses. Self-reported enteric illness data are required to better quantify and develop interventions to reduce the burden of illness in Canada. Previous studies have been limited by the availability and feasibility of data collection tools. Ethica, a smartphone based application used to acquire, store, and analyze data on human behaviour has been proposed as a supplement to current collection strategies. The purpose of this study was to assess the feasibility of new technologies and strategies within the Ethica app for gathering data on the occurrence of symptoms consistent with illness and risk factors for foodborne disease in Canada. By way of user-triggered and prompted microsurveys, meal descriptions and PhotoFoodDiaries, the occurrence of enteric symptoms and food consumption behavior in 96 university students was collected over a period of 10 weeks. Approximately 34% of participants reported at least one episode of vomiting or nausea during this period using at least one of the reporting features and 29% reported at least one episode of diarrhea; only 7% sought medical care. During the first 10 days of the study, food consumption history features on the app were used an average of at least 2 times per day by 95% of participants for the time-triggered microsurveys and by 51% of participants for digital images. Ethica served as an effective tool for collecting data on enteric symptoms, typically seen with many foodborne illnesses, as well as uniquely detailed food history data.

2.2 Introduction

Foodborne illness is a ubiquitous problem, affecting individuals of all demographic groups and across all regions of Canada and the world. Symptoms of foodborne illness range from vomiting and diarrhea to severe neurological conditions and can be life threatening (Thomas et al., 2013). Severe cases may last for several days to several weeks and often require medical attention. Such cases are documented primarily through physician and hospital records as well as laboratory test results. National surveillance programs utilize these data sources to monitor disease rates, detect outbreaks and plan interventions efforts (Keusch, 2013). Many mild cases do not require medical attention and therefore go unreported (Tam et al., 2003). Particular subgroups of the Canadian population are less likely to utilize healthcare services and are therefore further under-represented by traditional surveillance methods. These subgroups include those that identify as Indigenous (Clarke, 2016; O'Neil et al., 1998), immigrant (Clarke, 2016; Sanmartin and Ross, 2006) and those living in rural areas (Herikstad et al., 2002). Furthermore, one of the primary sources of information on foodborne disease in Canada is based on a sentinel site surveillance system with FoodNet sites located in Ontario (Middlesex-London Health Unit), British Columbia (Fraser Health Region), and Alberta (Calgary). Traditional surveillance strategies are limited in their ability to provide an accurate representation of the risks influencing foodborne illness in all regions and populations across Canada.

The nature of foodborne illness creates inherent barriers to data collection and reporting. Lack of specificity with respect to the signs and symptoms of foodborne illness make it difficult for health care practitioners to correctly classify the origin of disease without laboratory and epidemiologic support. The short duration of many foodborne illnesses further limits the timeframe in which a diagnosis can be made. Finally, the limited severity of symptoms may

influence an individual's choice as to whether to seek medical attention or just to wait out the illness.

Discrepancies exist among reports of the prevalence of foodborne illness in Canada – ranging from 4 million (Thomas et al., 2013) to 6.8 million per year (Munro et al., 2012). The Public Health Agency of Canada (PHAC) estimates that 4 million – 1 in 8 - Canadians suffer from domestically acquired foodborne illness every year (Thomas et al., 2013). This estimate is based on the assumption that 14% of acute cases (>3 loose stools in 24h with duration lasting >1 day) and 44% of severe cases (acute diarrheal illness and bloody diarrhea or diarrhea lasting >7 days) seek medical care. These approximations are based on results from National Studies on Acute Gastrointestinal Illness (NSAGI) population surveys completed in 2001–2002, 2002–2003, and 2005–2006 (Majowicz et al., 2004; Sargeant et al., 2008; Thomas et al., 2006). The US Centers for Disease Control (US CDC) and Prevention estimates that 18% of mild cases (>3 loose stools in 24 hours lasting >1 day) and 35% of severe cases (>3 loose stools in 24 hours lasting >1 day and bloody diarrhea) seek medical care (Scallan et al., 2011). These approximations are based on findings from the FoodNet Population Surveys in 2000–2001, 2002–2003 (Jones et al., 2007) and 2006–2007 (Centers for Disease Control and Prevention, 2006). Based on multiplying factors suggested by the US CDC, the Conference Board of Canada estimates that 6.8 million cases of foodborne illness occur in Canada each year (Munro et al., 2012). The need for estimates and extrapolation bring uncertainty to the extent of the burden of illness and make it difficult to measure the progress of interventions aimed to reduce the burden. More precise estimates of domestically acquired foodborne illness are essential for developing new practices and policies in industry, and for encouraging best food safety practices by consumers (Centres for Disease Control and Prevention, 2017).

Alternative methods of gathering large amounts of data on outbreaks of enteric illness include mining big data from sources such as news reports and social media. Examples of such systems include the Global Public Health Intelligence Network that queries relevant public health information from online news articles from across the globe (Mawudeku and Blench, 2006) and HealthMap, which monitors, summarizes, and maps information from Prom-Med-mail, WHO and various national and international agencies (Bahk et al., 2015). In a review of web/based disease surveillance strategies Brownstein et al. (2009) noted that although internet-based systems provide high quantities of data in a timely manner, these methods suffer from a wide range of limitations including misinformation, minimal specificity of signals, information overload and a disproportional sensitivity to external forces such as the media.

Other systems to monitor and analyze big data have also contributed to the detection of foodborne illness. The system nEmesis uses an adaptive algorithm to improve its own ability to recognize Twitter posts that are of relevance to foodborne illness (Sadilek et al., 2016). New York City restaurant reviews on Yelp were used to identify three new restaurant-linked outbreaks over a course of none months (Harris et al., 2014). It is unlikely that these outbreaks would have been detected by traditional surveillance systems because only 3% of reviews of foodborne illness were also reported to local public health authorities.

An alternative to analysis of data passively reported on other platforms is the creation of sites for active reporting of disease information by volunteers. Examples of participatory systems that request information on nausea/vomiting and diarrhea include Influezanet, Reporta, and Flu Near You (Wójcik et al., 2014). Participatory web-based surveillance systems for infectious disease have high degree of sensitivity and timeliness and are independent from health seeking behaviors, but may suffer from selection biases due to who choose to participate, difficulty in

adjusting for confounders, limited specificity of syndromic definitions, and issues with inconsistent participation (Wójcik et al., 2014). Participation in this type of system could be limited especially in rural and remote areas by regular access to the internet.

A practical tool for self-reporting illness is urgently needed to better quantify the burden of foodborne illness particularly in groups not captured by the existing systems. The capacity to detect illnesses where the case does not seek medical advice, and therefore are not reported to public health, would allow for more accurate estimates of the burden of disease and provide a basis to measure progress in food safety interventions. By providing volunteers from the general public or target groups at-risk with a convenient, user-friendly and efficient method by which to report illness, greater rates of participation and compliance might be obtained. The development and implementation of such a tool would be an important step in filling gaps in our understanding of the extent, impact and risk factors for emerging foodborne illnesses in Canada.

Limitations also exist in current options for collection of food history data. Current strategies involve in-person and telephone interviews or questionnaires administered several days to several weeks after the onset of symptoms. The delay period between the onset of symptoms and the collection of data will introduce recall bias. These strategies also involve a substantial time commitment by the public health practitioner administering the interview as well as the affected individuals and those selected as part of a reference or control group. Finally, the limited coverage and participation by the general public in traditional surveillance strategies results in incomplete estimates of domestically acquired foodborne illness and a limited picture of food consumption patterns.

Several studies have begun to assess the feasibility of innovative tools for data collection. As an alternative to in-person and telephone interview strategies, Arab et al. (2011) investigated

the feasibility of web-based food surveys. This online data collection technique was found to improve dietary recall. The collection of high-resolution images of foods has also been implemented to aid the dietary recall (Arab et al., 2011; Kikunaga et al., 2007; Six et al., 2010; Wang et al., 2010). These high-resolution images have provided detailed accounts of food intake in real-time and with minimal burden on the participants. Arab et al. (2011) described limitations of this strategy including difficulty interpreting images and the potential for reporting bias associated with under-reporting of less socially desired foods. These studies utilized a variety of devices and platforms to carry out the specific objectives of the study.

Smartphone technology has been proposed as one platform through which to collect data on self-reported illness. The Millennial Generation (aged 18-35) represent some of the most frequent users of smartphone technology. The wide range of capabilities of smartphones including the ability to capture high resolution digital images, record and securely store text and audio files, and administer online surveys has proven useful to the investigation of human behaviours (Hashemian et al., 2012). The wide range of capabilities of these devices may be harnessed to facilitate research of food consumption behaviours. The purpose of this study is to assess the feasibility of new technologies and strategies within the Ethica app for gathering data on the extent and risk factors for foodborne illness and food consumption patterns in a group of university students. Although Ethica had not previously been used to study foodborne illness, the application had proven useful in collecting related food consumption data (Hashemian et al., 2012).

2.3 Methods

2.3.1 Participant recruitment

Recruitment was coordinated through the Social Sciences Research Laboratory at the University of Saskatchewan, Saskatoon, Saskatchewan who advertised the study to their pool of research volunteers. Announcements regarding information on the study were also made in classrooms across the University of Saskatchewan as well as distributed on the campus-wide Personalized Access to Web Services (PAWS) platform. All interested participants with access to an Android phone version 4.0 or greater were included in the study. A briefing session was held to inform potential participants of the purpose of the study and data collection process the week before the study began. Participants were instructed on how to download the Ethica app from the Google Play store and how to use the various data collection features. The weekend following the orientation session was provided as a test period so participants could become familiar with the process of recording images, text and audio files through the application.

Ethica (www.ethicadata.com) is a smartphone application used to acquire, store, and analyze data. Developed by Drs. Nathaniel Osgood, Kevin Stanley, Mohammad Hashemian and colleagues at the University of Saskatchewan, this application provides a user friendly interface through which to collect real-time data on human behaviour (Hashemian et al., 2012). Participants were given the option to pause or discontinue data collection at any time during the study and were shown how to delete or discontinue responses they did not wish to submit.

Upon completion of the study, each participant received compensation of up to $100 CAD depending on the proportion of the study they completed. Informed consent was obtained from each participant at the time of enrollment. The protocol was approval by the University of Saskatchewan Behavioural Ethics Review Committee (BEH #15-187).

Three cohorts of participants were recruited to the study. Seventeen students were initially recruited to participate in a pilot study from September 14 to November 29, 2015. The initial pilot study was followed by a larger group which was divided into two staggered cohorts for ease of enrollment and management. The first cohort consisted of 40 participants that completed the study between January 15 and March 27, 2016. The second cohort consisted of 39 participants that completed data collection between January 22 and April 4, 2016.

2.3.2 Data Collection Protocol

For each cohort, the 10-week study began on a Monday with a demographic survey and a 10-day period of intensive data collection (Figure 2.1).

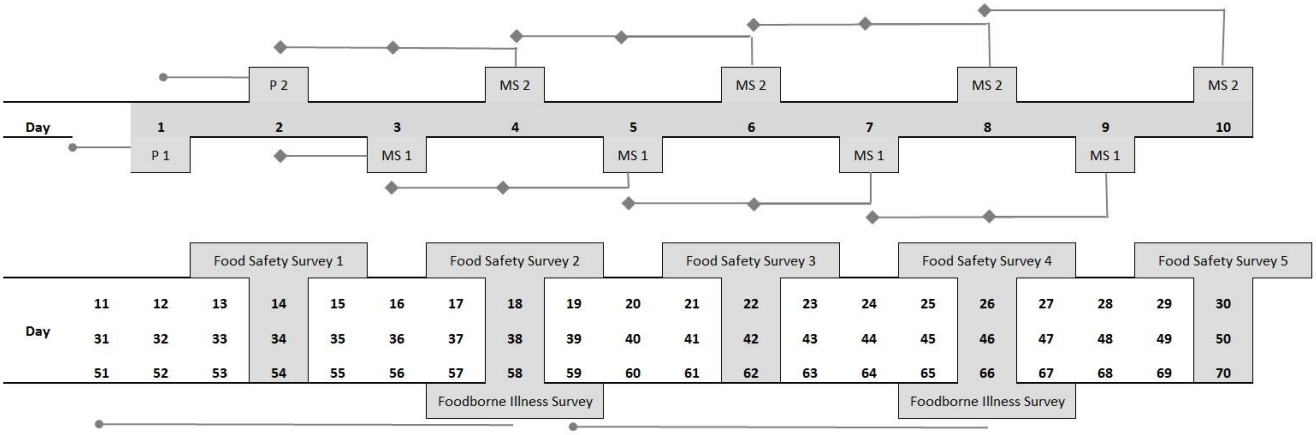

Figure 2.1: Timeline of data collection for the 70-day study period. P1 and P2 were practice microsurveys. M1 represents Microsurvey Type 1 and M2 represents Microsurvey Type 2. Microsurveys administered on days 1, 2, 3 and day 10 assessed food exposures for one day. The remainder of the microsurveys assessed specific food exposures of interest from the 2 previous days. Days 2 through 8 were a period of intensive focus for a concurrent survey on food exposure recall bias. Diamonds (♦) indicate the target for food intake data collection.

On the first day of the study a link to an online survey (Qualtrics Survey Software, Washington, DC) was emailed to participants to collect basic demographic information from participants. A link to a second follow up online survey requesting additional background

information on study participants was sent after day 10 of the study. Questions assessing enteric illnesses within the past week were administered together with Food Safety Surveys 2 and 4 (Figure 2.1).

2.3.2.1 User-triggered illness surveys

Throughout the study participants were asked to report anytime they felt ill by pressing a button on the app labeled "I'm feeling sick". This button was accessible any time that the app was open. These surveys allowed participants to select from a list of symptoms that might be associated with their illness such as diarrhea, vomiting, abdominal pain and cramps, nausea, fever and other. Note that participants were simply asked to self-report the occurrence of these symptoms – details on frequency, duration, and severity were not requested. This user-triggered interface also contained the questions "Did you consult a health-care professional regarding this illness?" and "Did you suspect your illness might be related to consumption of alcoholic beverages?"

2.3.2.2 User-triggered self-reported eating behavior

Also throughout the study participants were asked to report when they were eating and the source of the food using a button followed by a multiple choice question. Participants were asked if they were eating food prepared at home, eating at a restaurant, eating ready-to-eat food purchased off campus, and each food purchased on campus.

The self-report screen contained an option to take pictures of each meal or snack using the PhotoFoodDiary feature. Each image was digitally stamped with the time and geographical coordinates at which the picture was taken. All participants had the option to record a voice description of the meal or snack that was stored in the app database as an audio file. Based on

feedback from the pilot study, participants in the latter two cohorts were also given the option to accompany images with a written description in a text box directly under the image. This option enabled participants to provide relevant information on ingredients and other pertinent details that might not have been conveyed through pictures alone.

2.3.2.3 Initial 10-Day Period Focusing on Food Consumption

The Ethica Survey Tool allowed researchers to automatically administer surveys at predetermined times. Microsurveys were administered each day over the course of the initial 10 days of the study. Each microsurvey consisted of up to four multiple answer questions that assessed food exposures that may have taken place on the previous day or the day before the previous. To minimize the burden on participants, a limited number of foods were targeted with this data collection option. Foods identified as relevant to foodborne illness were given precedence. Participants were given a timeframe of four hours to complete each microsurvey to encourage prompt responses and minimize recall bias. Practice microsurveys were administered on day one (P1) and day two (P2) to give participants a chance to practice filling out surveys within the Ethica Survey feature of the application app before the start of the surveys focusing on food exposures for the 1-week period of greatest interest from day 2 to 8 (Figure 2.1).

Microsurvey 1 (MS 1) was administered on days 3, 5, 7 and 9 of the study (Figure 2.1). The survey consisted of three parts, each pertaining to a different group of foods released at 08:00, 12:45 and 18:30. This survey included questions such as 'On Friday or Saturday, Did you eat any servings of fresh fruit? Please check all that apply.' Other food groups assessed on Microsurvey 1 included nutrition bars, snacks, nuts, salads, sprouts and fish (see Appendix 1).

Microsurvey 2 (MS 2) was administered on days 4, 6, 8 and 10 of the study (Figure 2.1). Again, the survey consisted of three parts which were administered at 08:00, 12:45 and 18:30

each day. Microsurvey 2 assessed exposures to nutritional supplements, cereals, eggs and dairy products, cheese, nut butters, noodle mixes, tofu, poultry and vegetables (see Appendix 1).

2.3.2.4 Data Collection from Day 11 to Day 70

Participants were encouraged to continue tracking illnesses and eating behavior and food exposures through the PhotoFoodDiary feature of the app for the duration of the study. The second portion of the study differed from the first in that microsurveys were no longer administered to prompt recall of previous food exposures. Rather, five different food safety and eating behavior and preference surveys were introduced and were administered every 4 days on a rotating schedule. Food safety surveys were always administered at 18:30 and consisted of three to four multiple choice questions that assessed the participant's food preferences and knowledge about food safety and proper food handling processes.

Food safety survey 1 was administered on days 14, 34 and 54 of the study (Figure 2.1) and asked participants to self-report whether they had eaten in a restaurant or eaten take-out or ready to eat food. If ready to eat foods were purchased, participants were asked if were the food was stored at room temperature for more than 1 hour before eating. Participants were also asked if left overs taken home from a restaurant to eat later. See Appendix 3 for complete list of questions for food safety survey 1.

Food safety survey 2 was administered on days 18, 38, and 58 of the study (Figure 2.1) and assessed preferences such as organic, local, antibiotic-free and steroid/hormone free (Appendix 4). Food Safety Survey 3 was administered on days 22, 42 and 62 of the study and contained question such as "Are organic foods safer than foods from conventional production system?" and "Are locally grown foods safer…?" (Appendix 5).

Food safety survey 4 was administered on days 26, 46 and 66 of the study (Figure 2.1) and contained true or false questions such as `Washing the kitchen sponge with soap will get rid of all the bacteria" (Appendix 6). Food safety survey 5 was administered on days 30, 50 and 70 of the study (Figure 2.1) and contained additional true of false questions such as: "It is important to use a thermometer to check the temperature of meat before serving, but you don't need a thermometer for precooked leftovers" (Appendix 7).

To cross-validate a portion of the data that was collected from user-triggered illness surveys, additional illness surveys were administered at six predetermined times throughout the study period. Time-triggered enteric illness questions were administered at on days 18, 26, 38, 46, 58, and 66 of the study (Figure 2.1). These questionnaires asked questions such as "Did you feel nauseous or vomit in the last week?" and "Did you have diarrhea in the last week?" as well as assessing if medical care was sought for the illness and whether or not the user suspected that the illness might have been alcohol related. The enteric illness questions on the six microsurveys requested information for 42 of the 70 total study days (Figure 2.1).

2.3.2.5 Study Conclusion and Debriefing and Focus Groups

Debriefing sessions were held at the end of the study to answer any participant questions, assist in removal of the app and data from phones as necessary and to pay participant incentives.

Focus groups were also held at the end of the study to collect qualitative data on the user-friendliness and feasibility of the Ethica app with a subset of participants. Participants were recruited for focus groups through announcements directed to study participants by the Social Sciences Research Laboratory at the University of Saskatchewan. Focus groups were held on March 28, 2016 (n=15) and April 5, 2016 (n=13). Informed consent was given by all focus group

participants. Questions centered on challenges to compliance and the potential utility of the tool for monitoring enteric illness. Focus group discussions were recorded using audio recorder and manually transcribed to text by trained research assistants. Additional compensation was provided to participants in the form of $25 CAD for taking part in focus groups. Faculty investigators involved in developing the research project and app did not attend the focus groups to encourage open discussion. Rather the groups were facilitated by trained staff from the Social Sciences Research Laboratory and a faculty member experienced with qualitative research.

2.3.3 Data Extraction/Management

Data were encrypted and stored on the phones until a network connection became available. Once a Wi-Fi connection was available the encrypted data were transmitted to secure servers at the University of Saskatchewan (Figure 2.2).

At the end of the study, data were downloaded by researchers from the secure Ethica website as CSV files and transformed as necessary for further analysis in Microsoft Excel.

2.3.3.1 Enteric illness data – user triggered and time-triggered

All individual symptoms of enteric illness reported on the app and combinations of the symptoms consistent with enteric illness were summarized. However in this study, enteric illness was operationally defined as experiencing an episode of vomiting or nausea or diarrhea where symptoms were not associated with the consumption of alcohol to allow comparison of the self-reported data to that collected later in the study from prompted microsurveys. Any false positives in which the survey was triggered but not completed were removed. Any illnesses that were related to the consumption of alcohol were excluded from the analysis.

2.3.3.2 User-triggered self-reported food histories

Digital images collected between day two and day eight of the study were manually analyzed for relevance (i.e. the presence of clearly visible foods). Images in which foods were missing, partially eaten or unrecognizable were flagged. The remaining images were considered to have complete data and to be relevant to the analysis of food consumption. Audio files were manually transcribed to text. Term extraction was conducted for text descriptions and transcribed audio files (Figure 2).

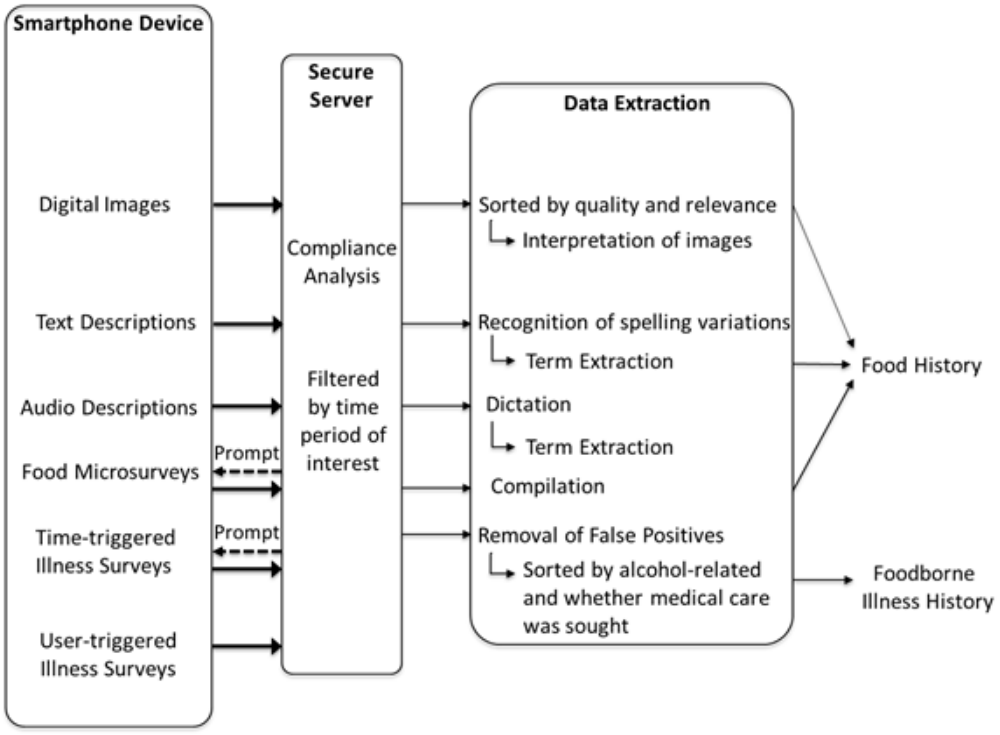

Figure 2.2: Overview of information flow and data extraction.

2.3.4 Data Analysis

2.3.4.1 Enteric illness data

The relative frequency of participants that reported each recorded symptom was summarized for the entire study period based on data provided by the user-triggered surveys. The data were then summarized based on important combinations of symptoms and whether or not cases were reported to health care practitioners.

Data on user-triggered illness surveys were then compared to the data collected from prompted time-triggered illness surveys. The relative frequency of participants that reported nausea or vomiting as well as those that reported diarrhea was assessed specifically for the 42-day period covered by the time-triggered questions. Similarly, the user-triggered enteric illness data was summarized specifically for each 1-week observation period where the time-triggered survey data were also available together with the number of individuals who reported seeking medical care.

Exact McNemar chi square tests for paired data were used to determine whether there was a significant difference in the odds of reporting any enteric illness, diarrhea, or vomiting or nausea between the user-triggered and timed-survey options. Kappa statistics were also used to summarize the agreement in reporting between the two options. All statistical analyses were completed with a commercial software program (Stata/SE 14.1 for Windows, StataCorp LP, College Station, TX).

To optimize the sensitivity of reported enteric symptoms, the data from the user-triggered surveys was combined with that from the time-triggered surveys and summarized for diarrhea, nausea or vomiting, and for enteric illness defined by reporting any of the three symptoms.

Because illness was not a rare outcome in this study, a Poisson model with a robust variance estimate rather than logistic regression was used to estimate the relative risk of enteric

illness for women as compared to men, those under 25 years of age compared to those greater than 25, and undergraduate students as compared to graduate students.

2.3.4.2 Food history data

The percentage of participants that took part in each data collection method – pictures, food descriptions and microsurveys – was summarized. In regards to capturing digital images of meals, compliance was defined as the number of images reported by the participant divided by the total number of images that were expected during that time period (i.e. an estimated three meals per day multiplied by the number of days). Excellent compliance was defined as capturing three images per day (breakfast, lunch and dinner) for each day of the study. The percentage of participants that collected an average of 1, 2 and 3 images per day was computed for the initial 10 days of the study as well as for the full 70-day study period.

Compliance for the reporting of food descriptions were calculated in a similar manner for participants in the second and third cohorts. Excellent compliance for the collection of food descriptions was defined as three descriptions per day (i.e. breakfast lunch and dinner). The number of descriptions collected were divided by the expected number of responses for the 10-day and 70-day study periods. The percentage of participants that reported an average of 1, 2, and 3 food meal description per day was determined for each respective time period.

Compliance on microsurveys was calculated for the initial ten days of the study. Again, excellent compliance was defined as having completed all microsurveys that were administered during this time period (i.e. one microsurvey per day, each consisting of three segments). The percentage of participants that completed an average of one segment per day (33.3% compliant)

was computed along with an average of 2 per day (66.7% compliant) and an average of three per day (100% compliant).

In the context of this study, good compliance was defined for each respective data collection method as having completed at least two thirds of required tasks – having submitted at least two images per day, two written meal descriptions per day, or as having completed at least two of the three microsurveys that were administered each day. This value was selected as not all students would be expected to eat three meals per day. Adequate compliance was defined as having completed one third of the required tasks. Linear regression was used to assess whether compliance, measured as a percentage, varied based on gender, age, student status, or history of illness during the study.

Finally, compliance was calculated for surveys administered between day 11 and day 70 of the study. These included five food safety questionnaires. Excellent compliance was defined as having completed all 15 surveys (five food safety questionnaires, each administered three times). The percentage of participants that filled out each survey at least once (33.3% compliant) was calculated along with the percentage that completed each survey at least twice (66.6% compliant) and the percentage that completed all 15 surveys (100% compliant).

Thematic analysis was conducted on data gathered from focus groups pertaining to the advantages, disadvantages and recommendations for improvement of the data collection strategies and the Ethica app. Databases were created and managed using Microsoft Access 2016 and Microsoft Excel 2016.

2.4 Results

2.4.1 Study Population

A total of 96 individuals participated in the study. There was a high proportion of both women as well as newcomers to Canada (Table 2.1). The initial online demographic survey was completed by 95.8% (92/96) of the study participants and a second follow up online survey after day 10 was completed by 91.7% (88/96) of participants.

Table 2.1: Demographic and other background characteristics of the study population

Participant attributes	Relative frequency
Gender identified as female (%)	67.8% (61/90)
Age less than 25 years (%)	50.0% (46/92)
Undergraduate students (%)	44.6% (41/92)
Recently moved from rural area to the city (%)	8.0% (7/88)
Newcomer to Canada (%)	39.8% (33/88)
Had used a food diary before (%)	26.1% (24/92)
Reported being very concerned about privacy (%)	6.5% (6/92)
Had access to facilities to prepare meals and lives independently (%)	59.1% (52/88)

2.4.2 User-triggered Enteric Illness Data

Episodes of vomiting, not related to alcohol consumption, were reported by 6.3% (6/96) of study participants and diarrhea was reported by 30.2% (29/96) of study participants with the user-triggered reporting feature on the app during the 70-day study period. Only 2.1% (2/96) of the individuals reported either vomiting or diarrhea to a health care professional (Table 2.2).

Table 2.2: The number of participants who reported illness through user-triggered surveys during the 70-day study period (n=96).

	List of symptoms provided to participants					
	Vomiting	Diarrhea	Abdominal pain and cramps	Nausea	Fever	Other
Reports of illness	6 (6.3%)	29 (30.2%)	32 (33.3%)	32 (33.3%)	15 (15.6%)	43 (44.8%)
Sought medical care	2 (2.1%)	2 (2.1%)	1 (1.0%)	1 (1.0%)	2 (2.1%)	5 (5.2%)

During the 70-day study period, 46 (47.9%) of participants reported diarrhea, vomiting or nausea; 33 (34.4%) reported vomiting or diarrhea; and 33 (34.4%) reported vomiting or nausea through user-triggered illness surveys. Abdominal pain and nausea were more commonly reported than diarrhea or vomiting. Only 3 (3.2%) of the study participants who reported diarrhea, vomiting or nausea indicated that they sought medical care for their condition during the 70-day study period using the user-reporting feature of the app. There was no obvious time trend in the reporting of symptoms with user-triggered option during the study period (data not shown).

2.4.3 Time-triggered Enteric Illness Data

Twenty participants (20.8%) reported experiencing vomiting or nausea through the time-based enteric illness questions with the number for each of the weeks examined being relatively consistent over the course of the study. Similarly, the responses from the 20 participants (20.8%) reporting diarrhea were fairly uniformly distributed across the 6 weeks where specific questions were sent to participants' phones (Table 2.3).

Table 2.3: Number of participants who reported illness during the 42 days of the 70-day study period where time-triggered surveys requested illness data – a comparison of responses using the time-triggered and user-triggered options.

	Time-triggered survey		User-triggered survey	
Time period	Vomiting or nausea	Diarrhea	Vomiting or nausea	Diarrhea
Days 11-18	4 (4.2%)	6 (6.3%)	4 (4.2%)	3 (3.1%)
Days 19-26	6 (6.3%)	6 (6.3%)	4 (4.2%)	0 (0.0%)
Days 31-38	3 (3.1%)	3 (3.1%)	6 (6.3%)	6 (6.3%)
Days 39-46	5 (5.2%)	9 (9.4%)	5 (5.2%)	2 (2.1%)
Days 51-58	3 (3.1%)	8 (8.3%)	4 (4.2%)	2 (2.1%)
Days 59-66	5 (5.2%)	6 (6.3%)	6 (6.3%)	1 (1.0%)
At least one report during the 42-day portion of the observation period	20 (20.8%)	20 (20.8%)	23 (24.0%)	11 (11.5%)

Participants were not significantly more likely to report diarrhea ($p=0.08$), vomiting or nausea ($p=0.66$) (Table 2.3), any enteric illnesses ($p=0.71$) or seeking medical care ($p=0.13$) (Table 2.4) using the time-triggered prompted surveys than the self-reported surveys. The individual agreement in reported illness history based on user- and time-triggered surveys, however, was only fair to poor (based on the categorization suggested by Dohoo et al. (2012)) for: diarrhea ($\kappa=0.21$), vomiting or nausea ($\kappa=0.37$), any enteric illnesses ($\kappa=0.29$) or sought medical care ($\kappa<0.01$).

Table 2.4: Number of participants who reported vomiting, nausea or diarrhea (not-associated with alcohol) through user-triggered or time-triggered surveys and those who sought medical care during the 42 days of the 70-day study period where time-triggered surveys requested illness data.

	Time-triggered surveys		User-triggered surveys	
Time Period	Enteric illness	Sought medical care	Enteric illness	Sought medical care
Days 11-18	9 (9.4%)	1 (1.0%)	6 (6.3%)	0 (0.0%)
Days 19-26	12 (12.5%)	0 (0.0%)	4 (4.2%)	0 (0.0%)
Days 31-38	5 (5.2%)	2 (2.1%)	9 (9.4%)	0 (0.0%)
Days 39-46	9 (9.4%)	2 (2.1%)	6 (6.3%)	0 (0.0%)
Days 51-58	10 (10.4%)	1 (1.0%)	5 (5.2%)	0 (0.0%)
Days 59-66	8 (8.3%)	0 (0.0%)	6 (6.3%)	0 (0.0%)
At least one report during the 42-day portion of the observation period	31 (32.3%)	4 (4.2%)	28 (29.2%)	0 (0.0%)

When the self-reported illness data were combined from both the user-triggered and time-triggered sources, 34.4% (33/96) reported nausea or vomiting, 29.2% (28/96) reported diarrhea, and 55.2% (53/96) reported at least one of nausea, vomiting or diarrhea during the 10-week study period. Four individuals sought medical care for their illnesses during the 42-day period. Three individuals sought care for symptoms that occurred beyond the 42-day observational study. In total, 7 individuals (7.3%) sought medical care for the occurrence of nausea or vomiting.

Undergraduate students were 1.8 times (95%CI 1.1 to 2.7, p=0.01) more likely than graduate students to report enteric illness using either the user-triggered or time-triggered features of the app during the 70-day study period. After accounting for student status, there was no significant association between age (less than 25 years: RR 0.89, 95%CI 0.59 to 1.3, p=0.56) or identified gender (men compared to women: RR 1.2, 95%CI 0.61 to 2.2, p=0.65).

2.4.4 Food History Data

Participant compliance varied among the different data collection methods (Table 2.5). Only 115 audio descriptions were collected over the entire 70-day study period. These audio recordings were submitted by 11.5% (11/96) of participants who were given this option. None (0/96) of participants met the criteria for adequate compliance in recording food histories through audio recordings. Seven of the 11 participants saved an audio file < 10 times.

The highest participation was observed in food microsurveys (Table 2.5); during the initial 10 days of the study, 95% of participants submitted at least two microsurveys per day. Slightly lower compliance was observed for the PhotoFoodDiary option in which 51% of participants took an average of at least two photos per day for the first 10 days of the study. In contrast, only 21% of participants with access to this feature completed an average of at least two written meal description per day during the first ten days of the study.

When compared to the initial 10 days of the study, compliance in capturing digital images and meal descriptions over the entire 70-day study period was much lower (Table 2.5). The percentages of participants who collected an average of at least two digital images per day was 15%, and 13% reported at least two meal descriptions per day. Note that food microsurveys were not administered after day 10. Instead, food safety surveys were administered every four days for the remainder of the study. Of the 96 participants, 93 (96.9%) completed each of the five different surveys at least once of the three times the surveys were distributed.

A more detailed comparison of the compliance with data collection methods across the initial 10-day period versus the full 70-day study period is shown in Table 2.5. Compliance in taking pictures of meals during the first 10 days, taking pictures through the entire study, and completing microsurveys did not significantly differ between males and females, based on age,

undergraduate and graduate students, or those who reported gastrointestinal symptoms during the study and those who did not ($p>0.05$).

Table 2.5: Compliance with three data collection strategies for food consumption data (digital images, meal descriptions and microsurveys).

		Initial 10 days	Full 70 day study period
Digital Images (n=96)	Average/participant/day (SD)	2.3 (1.4)	1.2 (1.0)
	Average ≥ 1 per day	85.4% (82)	44.8% (43)
	Average ≥ 2 per day	51.0% (49)	14.6% (14)
	Average ≥ 3 per day	26.0% (25)	6.3% (6)
Written Meal Descriptions (n=79)	Average/participant/day	1.5 (1.1)	1.0 (1.0)
	Average ≥ 1 per day	56.3% (54)	32.3% (31)
	Average ≥ 2 per day	20.8% (20)	12.5% (12)
	Average ≥ 3 per day	7.3% (7)	3.1% (3)
Microsurveys (n=96)	Average/participant/day	2.7 (0.4)	Not administered after day 10
	Average ≥ 1 per	100% (96)	
	Average ≥ 2 per day	94.8% (91)	
	Average = 3 per day	28.1% (27)	

2.4.5 Food Safety Knowledge and Food Preference Data

The Ethica app also provided a tool to poll participant's food safety knowledge; the number of participants who responded to each survey is shown in Table 2.6, followed by the number of correct responses in Table 2.7 and finally food preferences in Table 2.8.

Table 2.6: Number of responses to food safety knowledge questions administered at three different points in time (n=96).

	Responded 0/3 times	Responded 1/3 times	Responded 2/3 times	Responded 3/3 times
Food Safety Survey 1	2 (2.1%)	10 (10.4%)	42 (43.8%)	42 (43.8%)
Food Safety Survey 2	4 (4.2%)	10 (10.4%)	18.8 (10.4%)	64 (66.7%)
Food Safety Survey 3	3 (3.1%)	8 (8.3%)	22 (22.9%)	63 (65.6%)
Food Safety Survey 4	7 (7.3%)	4 (4.2%)	22 (22.9%)	63 (65.6%)
Food Safety Survey 5	5 (5.2%)	9 (9.4%)	23 (24.0%)	59 (61.5%)

Table 2.7: Number of correct responses to food safety knowledge questions administered at three different points in time (n=96).

	Correct 3/3 times	Correct 2/3 times	Correct 1/3 times	Correct 0/3 times	# Participants that responded 3 times (n)
Washing a kitchen sponge with soap will get rid of all the bacteria. *(False)*	47 (74.6%)	10 (15.9%)	2 (3.2%)	4 (6.3%)	63
Raw meat should be washed in the sink before cooking. *(False – due to risk of cross contamination)*	26 (41.3%)	5 (7.9%)	6 (9.5%)	26 (41.3%)	63
Food should be allowed to sit on the counter before putting it in the fridge for storage. *(False)*	39 (61.9%)	6 (9.5%)	6 (9.5%)	12 (19.0%)	63
Prebagged 'ready to eat' greens should not be washed again before eating. *(True – due to risk of cross contamination)*	2 (3.4%)	12 (20.3%)	11 (18.6%)	34 (57.6%)	59
Hamburgers should be cooked until they are 'piping hot' and there is no visible pink left. *(False – the temperature should be checked with a thermometer)*	2 (3.4%)	8 (13.6%)	6 (10.2%)	43 (72.9%)	59
It is important to use a thermometer to check the temperature of meat before serving but you don't need a thermometer for precooked leftovers. *(False)*	14 (23.7%)	11 (18.6%)	8 (13.6%)	26 (44.1%)	59

The relative frequency of correct responses to a series of six questions about food safety knowledge are summarized in Table 2.7 for those who answered each question all three times it was asked. Note that individuals in columns labelled `Correct 2/3 times` and `Correct 1/3 times` changed their responses when asked the same question again at a later date.

When prompted to compare the safety of organic foods with foods from commercial production systems 14.6% (14/96) always indicated that they thought there was no difference. Conversely, 13.5% (13/96) always said that organic foods were the safer choice. It is notable that 17.7% (17/96) changed their responses between the times when the question was administered. Similarly, when prompted to compare the safety of local foods with foods from commercial production systems 11.5% (11/96) always indicated no difference. On the other hand, 20.8% (20/96) always said that local foods were the safer choice. Notably, 15.6% (15/96) changed their responses when asked at a later date.

The responses to a number of food preference questions indicate that on the three days which the surveys were administered, 38% ate at a restaurant (Table 2.8). Furthermore, approximately 17% of participants were conscientious of eating organic foods, 31% of participants reported having eaten locally grown foods and 31% reported eating and hormone-/steroid-free foods on the days on which the surveys were completed (Table 2.8).

When prompted 5.7% of participants indicated that they were vegetarian or vegan and 9.1% of participants in this study indicated that halal or kosher diets influenced their food choices.

Table 2.8: Number of 'Yes' responses to questions about food preferences administered on the different occasions throughout the study period (n=96).

	Yes 3/3 times	Yes 2/3 times	Yes 1/3 times	Yes 0/3 times	# Participants that responded 3 times (n)
Did you eat organic food today?	2 (3.1%)	3 (4.6%)	6 (9.2%)	54 (83.1%)	65
Did you eat food identified as being locally grown today?	0 (0.0%)	7 (10.8%)	13 (20.0%)	45 (69.2%)	65
Did you eat food identified as being raised without steroids/hormones today?	3 (4.7%)	6 (9.4%)	11 (17.2%)	44 (68.8%)	64
Did you eat food identified as being raised without antibiotics today?	2 (3.1%)	3 (4.6%)	11 (16.9%)	49 (75.4%)	65
Did you eat in a restaurant today?	2 (4.8%)	5 (11.9%)	9 (21.4%)	26 (61.9%)	42
Did you eat take out / ready to eat food today?	2 (4.8%)	10 (23.8%)	17 (40.5%)	13 (31.0)	42
Do you eat purchased ready to eat foods and store them at room temperature for more than 1 hour before eating?	0 (0.0%)	6 (14.3%)	10 (23.8%)	26 (61.9%)	42
Do you take leftovers home from restaurants to eat later?	3 (7.10%)	12 (28.6%)	12 (28.6%)	15 (35.7%)	42

2.4.6 Focus Groups

A total of 28 study participants took part in focus group discussions. Key themes obtained pertaining to the advantages, disadvantages and suggestions for improvement of each data collection strategy that were obtained from the two focus groups are shown in Table 2.9.

Table 2.9: Themes obtained from focus group discussions pertaining to the advantages, limitations and suggested changes for three data collection methods - digital-images, food descriptions and microsurvey responses (n=28).

	Advantages	Limitations	Suggestions for improvement
Digital Images	-Efficiency	-Difficulty remembering to complete task -Snack foods often missed -Limited use for wrapped foods	-Push notifications -Include images of grocery receipts -Option to upload images at later time
Text Descriptions	-High level of detail	-Time consuming -Requires manual data input	-Enable retrospective reporting (Backlogging)
Audio Recordings	-Efficiency -Minimal burden on user	-Recording voice made some users feel self-conscious/uneasy	-Incorporate dictation capabilities into the PhotoFoodDiary feature
Food Microsurveys	-Comprehensive -Prompts increased compliance	-Not all foods assessed -Does not require ongoing attention	-Take dietary restrictions into account to provide tailored sets of questions
Illness Surveys	-Timely -Easy to access	-Increased false positives due to accidentally triggering survey	-Follow up illness survey with 48 hour food history questionnaire
Overall	-Relatively small burden on user -Increased food safety awareness	-Requires time to become familiar with new technology	-Link to existing fitness, financial and food tracking applications

The technology was described as “easy to use” (male participant of focus group 2) and “pretty straightforward” (female participant, focus group 2). General themes of focus group discussions centered around convenience of the application in tracking what a person eats with minimal disruption to everyday routines. Participants were not bothered by the idea of continuous data collection under the conditions that informed consent had been given and the

purpose of the study had been disclosed. Several potential applications of the smartphone technology were brought up during focus group discussions. Such detailed food history reports would be an asset to individuals seeking to identify specific food allergies. A common thread through discussions was the potential of the data to identify vulnerable populations, high risk food consumption behaviours and foodborne illness hotspots.

2.5 Discussion

While the underreporting of foodborne illness is widely acknowledged, the extent and implications of underreporting due to cases that do not seek medical care are not as well agreed upon. This study is unique in that by tracking allowing participants to easily self-report symptoms across a 70-day period, it allowed for a more sensitive method of surveillance in a targeted population. The occurrence of enteric illness described here using the user-triggered recording option appeared to be higher than previous estimates of foodborne illness reported as 12.5% - 1 in 8 per year (Thomas et al., 2013). In this study, 34% of participants reported having experienced an episode of vomiting or nausea and 29% reported an episode of diarrhea that was not believed to be caused by alcohol consumption in a 10-week period. However, it is important to recognize that the case definition for acute diarrheal illness in the previously reported study was ≥ 3 loose stools in 24 h with duration lasting > 1 day (Thomas et al., 2013). In the present feasibility study, participants were simply asked to report the occurrence of symptoms; information on frequency within a day, duration, or severity was not requested. In future studies the app questions could be made more specific such that information was requested on vomiting or diarrhea with 3 or more episodes in 24 hours.

Only 7% of individuals who described enteric symptoms on the user-triggered or time-triggered feature reported their symptoms to a health practitioner. The percentage of acute cases

of foodborne illness that seek medical care obtained from retrospective studies – 14% (Thomas et al., 2013), 18% (Scallan et al., 2011) and 22% (Sargeant et al., 2008) were higher, but might also be impacted by the definition of enteric illness in these studies and the relative severity to that in the current study. Discrepancies in estimates of under-reporting may be attributed to differences in the sensitivities of methods used to detect cases. More sensitive methods are able to capture mild cases of foodborne illness that would be missed by other traditional surveillance strategies. Mild cases are less likely to seek medical attention when compared to severe cases (Scallan et al., 2006). Therefore, as the proportion of mild cases to severe cases increases, the percentage of cases that seek medical care will drop accordingly.

Prompting users to complete microsurveys describing symptoms in the previous week may potentially increase the sensitivity of the data collection methods, although the differences observed in this study were not significant. These findings provide evidence regarding the extent of underreporting of enteric illness in a group of university students and the potential for innovative strategies to address this issue.

The complexity of human behavior in regards to schedule, food preferences and risk-benefit perceptions make it unlikely that any single data collection method will provide an accurate account of food history. Instead, tradeoffs must be made to reach a balance between feasibility and level of detail that can be obtained. While some studies have attempted to create passive methods of tracking food intake (Arab et al., 2011) a certain level of disruption to everyday activities is unavoidable. Data obtained from food descriptions was rich in detail and provided information on ingredients that were not evident in digital images. During the initial ten days of the study, participants recorded an average of 1.5 meal descriptions per day. As the study continued, compliance dropped, resulting in an overall rate of approximately 1.0 meal

descriptions per participant per day for the full 70-day study period. The time and effort required to manually input food descriptions is believed to have played an important role in the observed drop in compliance. During focus group discussions, ease and efficiency were described as a key advantages of the option to describe foods through audio recordings. However, only a small number of participants chose to report audio descriptions rather than written descriptions of meals. A recurring disadvantage that participants associated with this data collection method was self-consciousness and a feeling of unease brought about by recording one's own voice.

In focus group discussions, participants perceived stated efficiency as an important advantage for the image capture technique. However, as a consequence of the minimal effort and time required by the user, a degree of ambiguity was introduced into the analysis of the images. Similar to challenges faced in other studies (Arab et al., 2011), there was uncertainty in the identification of foods and whether the user actually consumed all foods seen in the image. In some cases, images were not suitable for analysis because the foods had been partially eaten or the images were too blurry to interpret. Participants highlighted situations in which it was not socially acceptable to use a smartphone device or to take pictures of the meal that was being served (i.e. at restaurants, dinner parties). As the novelty of the task diminished, burn-out and forgetfulness on the part of the participants began to negatively impact compliance. These factors are important consideration when planning the length and intensity of surveillance initiatives.

Food microsurveys were found to have the highest rate of compliance for the initial 10-days of the study, averaging 2.7 survey responses per participant per day. This comparatively high rate of compliance is believed to be partly due to the prompts sent by the application system to indicate to users. The short (1-4 question) design of the surveys, consisting of straightforward, closed-ended multiple choice questions, minimized the time and effort burden placed on

participants. The limitation of such surveys is the loss of data on foods not assessed by the survey questions. In summary, aspects of the food history data collection system that increase compliance include automated reminders, easy access, and most importantly, minimal time and effort requirements. The most relevant pitfall of the user-triggered food history data collection strategies was reliance on the memory of participants. A potential solution for this moving forward is the evolving capability of the Ethica app to issue location or context specific prompts for users who are willing to share location data in addition to time-triggered messages.

When asked the same questions on food safety knowledge surveys at three different points in time, between 17 and 39% of respondents changed their response between surveys. This finding illustrates the flexibility of the app for obtaining different types of information and room for improvement in regards to food safety knowledge among the millennial generation. The simple strategy of repeated assessments proved effective in eliciting improvements in many of the study participants. During focus group discussions, participants emphasized the benefit that providing references for further information on food safety could have on behaviours, further illustrating a strong interest and willingness to learn more about food safety.

Looking forward, this technology has the potential to fill gaps in existing surveillance systems and to provide a foundation for new and innovative strategies. By harnessing the capabilities of smartphone devices, detailed health related data can be collected from sentinel surveillance cohorts in a feasible and sustainable manner. Particular advantages such as the low burden placed on the user, the option to send reminders and prompts and ability to collect a wide range of data without regular access to the internet may alleviate many limitations of existing participatory web-based surveillance systems (Wójcik et al., 2014). While Ethica does require occasional internet access to upload data and download updates, the Ethica app can function and

store data for several days or weeks if needed between access to Wi-Fi or cellular networks. This advantage makes this data collection option feasible even for target groups in rural and somewhat remote areas. Overcoming some of the challenges of traditional surveillance methods creates new opportunities for targeted surveillance strategies to better understand health related behaviours in previously understudied segments of the Canadian population including rural and rural and remote communities (Wójcik et al., 2014).

2.6 Conclusion

Self-reports of enteric symptoms provide valuable insight on the extent, and risk factors of foodborne disease. Smartphone apps, such as Ethica, harness the capabilities of mobile devices to facilitate the collection of data that would be otherwise challenging to acquire from the general population. Results from self-reported illness data indicate that the level of underreporting might be greater than previously described. Collecting an accurate and comprehensive food history is an equally difficult challenge. Features associated with higher compliance include automated reminders, easy access and minimal requirements for time and effort in data entry. Further research is warranted to obtain detailed information on the occurrence of disease, food safety behaviors and the feasibility of new data collection tools.

2.7 References

Arab, L., Estrin, D., Kim, D. H., Burke, J., & Goldman, J. (2011). Feasibility testing of an automated image-capture method to aid dietary recall. *Eur J Clin Nutr, 65*(10), 1156-1162. doi:10.1038/ejcn.2011.75

Bahk, C. Y., Scales, D. A., Mekaru, S. R., Brownstein, J. S., & Freifeld, C. C. (2015). Comparing timeliness, content, and disease severity of formal and informal source outbreak reporting. *BMC infectious diseases, 15*(1), 135.

Brownstein, J. S., Freifeld, C. C., & Madoff, L. C. (2009). Digital disease detection—harnessing the Web for public health surveillance. *New England Journal of Medicine, 360*(21), 2153-2157.

Centers for Disease Control and Prevention. (2006). Foodborne Active Surveillance Network (FoodNet) population survey atlas of exposures. *Atlanta, Georgia: US Department of Health and Human Services, Centers for Disease Control and Prevention, 2007*.

Centres for Disease Control and Prevention. (2017). Burden of Illness: Questions and Answers.

Clarke, J. (2016). Difficulty accessing health care services in Canada.

Harris, J. K., Mansour, R., Choucair, B., Olson, J., Nissen, C., Bhatt, J., Control, C. f. D., & Prevention. (2014). Health department use of social media to identify foodborne illness-Chicago, Illinois, 2013-2014. *MMWR Morb Mortal Wkly Rep, 63*(32), 681-685.

Herikstad, H., Yang, S., Van Gilder, T., Vugia, D., Hadler, J., Blake, P., Deneen, V., Shiferaw, B., Angulo, F., & Group, F. W. (2002). A population-based estimate of the burden of diarrhoeal illness in the United States: FoodNet, 1996–7. *Epidemiology & Infection, 129*(1), 9-17.

Jones, T., McMillian, M., Scallan, E., Frenzen, P., Cronquist, A., Thomas, S., & Angulo, F. (2007). A population-based estimate of the substantial burden of diarrhoeal disease in the United States; FoodNet, 1996–2003. *Epidemiology and Infection, 135*(02), 293-301.

Keusch, G. T. (2013). Perspectives in Foodborne Illness. *Infectious disease clinics of North America, 27*(3), 501-515.

Kikunaga, S., Tomoe, T., Ishibashi, G., & Da-Hong, W. (2007). The application of a handheld personal digital assistant with camera and mobile phone card (Wellnavi) to the general population in a dietary survey. *Journal of nutritional science and vitaminology, 53*(2), 109-116.

Majowicz, S., Dore, K., Flint, J., Edge, V., Read, S., Buffett, M., McEwen, S., McNab, W., Stacey, D., & Sockett, P. (2004). Magnitude and distribution of acute, self-reported gastrointestinal illness in a Canadian community. *Epidemiology and Infection, 132*(04), 607-617.

Mawudeku, A., & Blench, M. (2006). Global public health intelligence network (GPHIN). *7th Conference of the Association for Machine Translation in the Americas;*.

Munro, D., Le Vallée, J.-C., & Stuckey, J. (2012). *Improving food safety in Canada: Toward a more risk-responsive system*.

O'Neil, J., Reading, J., & Leader, A. (1998). Changing the relations of surveillance: the development of a discourse of resistance in Aboriginal epidemiology. *Human organization, 57*(2), 230-237.

Sadilek, A., Kautz, H. A., DiPrete, L., Labus, B., Portman, E., Teitel, J., & Silenzio, V. (2016). *Deploying nEmesis: Preventing Foodborne Illness by Data Mining Social Media.* Paper presented at the AAAI.

Sanmartin, C., & Ross, N. (2006). Experiencing difficulties accessing first-contact health services in Canada. *Healthcare Policy, 1*(2).

Sargeant, J., Majowicz, S., & Snelgrove, J. (2008). The burden of acute gastrointestinal illness in Ontario, Canada, 2005–2006. *Epidemiology and Infection, 136*(04), 451-460.

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., Jones, J. L., & Griffin, P. M. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis, 17*(1).

Scallan, E., Jones, T. F., Cronquist, A., Thomas, S., Frenzen, P., Hoefer, D., Medus, C., & Angulo, F. J. (2006). Factors associated with seeking medical care and submitting a stool sample in estimating the burden of foodborne illness. *Foodbourne Pathogens & Disease, 3*(4), 432-438.

Six, B. L., Schap, T. E., Zhu, F. M., Mariappan, A., Bosch, M., Delp, E. J., Ebert, D. S., Kerr, D. A., & Boushey, C. J. (2010). Evidence-based development of a mobile telephone food record. *Journal of the American Dietetic Association, 110*(1), 74-79.

Tam, C. C., Rodrigues, L. C., & O'Brien, S. J. (2003). The study of infectious intestinal disease in England: what risk factors for presentation to general practice tell us about potential for selection bias in case-control studies of reported cases of diarrhoea. *International Journal of Epidemiology, 32*(1), 99-105.

Thomas, M. K., Majowicz, S. E., Sockett, P. N., Fazil, A., Pollari, F., Doré, K., Flint, J. A., & Edge, V. L. (2006). Estimated numbers of community cases of illness due to Salmonella, Campylobacter and verotoxigenic Escherichia coli: pathogen-specific community rates. *Canadian Journal of Infectious Diseases and Medical Microbiology, 17*(4), 229-234.

Thomas, M. K., Murray, R., Flockhart, L., Pintar, K., Pollari, F., Fazil, A., Nesbitt, A., & Marshall, B. (2013). Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne pathogens and disease, 10*(7), 639-648.

Wang, L., Xu, Y., Wang, Y., Dong, S., Cao, Z., Zhou, W., Sun, H., Huo, D., Zhang, H., Sun, Y., Huang, L., Yuan, Z., Zeng, D., & Song, H. (2010). The Epidemiological Investigation and Intelligent Analytical System for foodborne disease. *Food Control, 21*(11), 1466-1471. doi:10.1016/j.foodcont.2010.04.015

Wójcik, O. P., Brownstein, J. S., Chunara, R., & Johansson, M. A. (2014). Public health for the people: participatory infectious disease surveillance in the digital age. *Emerging themes in epidemiology, 11*(1), 7.

CHAPTER 3 – FOODBORNE OUTBREAK INVESTIGATION: THE EFFECT OF RECALL BIAS ON FOOD HISTORIES COLLECTED A WEEK OR MORE AFTER CONSUMPTION

3.1 Abstract

Collecting accurate and comprehensive food histories has been recognized as an obstacle in foodborne illness investigation. Recall bias is a key limitation in a foodborne outbreak investigation. The ability to associate an outbreak of enteric illness with a common food exposure is dependent on the accuracy of data collected from food history questionnaires. The purpose of this study was to investigate the effect of recall bias on the validity of food history data in a context comparable to outbreak investigations, and to characterize the food exposures of a previously understudied segment of the Canadian population. The food consumption of 96 university students was collected using Ethica, a smartphone-based data acquisition system. Comprehensive food histories were captured through a combination of digital images, meal descriptions, and short food exposure surveys. This real-time data was used as a reference to measure the sensitivity and specificity of food history questionnaires administered after an average of 7 or 18 days (2.5 weeks) after consumption (n=86). The questionnaires and time intervals used in this study were designed to resemble a range of plausible local, provincial and national enteric outbreak investigations conducted by public health officials in Canada. The validity of food history data collected after 7 days was low for many foods with sensitivities with a median of 54.5% for peanut butter and ranged from 14.3% for sprouts to 100% for leafy greens. The median of the observed specificities was 71.4% for sprouts and ranged from 30.4% for beef to 80.4% for peanuts. Similarly, the sensitivity of data collected after 18 days had a median of 46.2% for melons and ranged from 15.8% for sprouts to 77.8% for tomatoes. The

specificities after 18 days had a median for 72.2% of nuts and ranged from 21.2% for leafy greens to 92.1% for melons. The impact of recall bias on the accuracy of food history data was found to vary with food type. This study serves as a first step to quantifying the implications of recall bias so that recall can be accounted for in future outbreak investigation strategies.

3.2 Introduction

Foodborne illness is a global concern with many practical challenges. In 2010, the consumption of contaminated foods caused approximately 600 million illnesses worldwide (WHO, 2015a). In Canada, 4 million (1 in 8) individuals are thought to be affected by domestically acquired foodborne illness each year (Thomas et al., 2013). When reports of unusual cases are received or a sudden and unexpected increase in illness is detected, an outbreak investigation may be necessary to identify the source of illness and to prevent further harm (Vik and Hexemer, 2014). If a review of the preliminary information – which may include laboratory test results and basic epidemiological data such as person, place and time – suggests that the outbreak may be linked to a common food exposure then a foodborne outbreak investigation may be initiated. Advances in laboratory molecular-typing techniques allow investigators to distinguish between closely related pathogenic strains. Pulse Field Gel Electrophoresis (PFGE), Multiple-Locus Variable number tandem repeat Analysis (MLVA) and Whole Genome Sequencing (WGS) are techniques used to match cases infected with the same outbreak strain and to test isolates from food samples.

If food samples are not available or no match can be found the ability to generate an accurate hypothesis relies largely on the validity of data collected from food history questionnaires (Vik et al., 2014). While the timeframe of investigation varies depending on the circumstances of each outbreak, public health practitioners in Canada and the US traditionally

aim to administer these questionnaires within one week after the onset of symptoms as part of local or regional outbreak investigations, and after to 2-3 weeks for national investigations (CIFOR, 2014). In practice, these targets often need to be extended due to limitations in the timeliness of reporting, laboratory testing and outbreak detection. In an analysis of metrics of enteric outbreaks investigations, the British Columbia Centre for Disease Control (BCCDC) reported that the median time to initiate an outbreak investigation was 36 days (Fong et al., 2017). The incubation period, including the time from consumption to clinical signs, for many of the most common foodborne pathogens ranges from hours to several days (CIFOR, 2014). For the purposes of this study, time intervals refer to the period between consumption of the food item and when the participant was prompted to recall the exposure.

Recall bias occurs when people search for explanations for their disease and may assign more significance to some exposures than others (Mausner and Kramer, 1985). Thus far, public health practice has accepted recall bias as an inherent limitation to outbreak investigation. Progress in designing surveys in such a way as to minimize this recall bias has plateaued in recent years. As the complexity of the food supply network in Canada continues to grow, so does the need for new investigation tools and to address the impact that recall bias has on outbreak investigations. Quantifying the effect of recall bias is an important first step to addressing this issue.

In order to examine the sensitivity and specificity of food history questionnaires administered at different points in time, researchers would need an accurate account of what each individual ate during the time period of interest. Obtaining such a detailed record of food consumption continues to be a challenge in nutritional research. Real-time records of food exposure can provide accurate accounts of food exposures. Capturing digital images of foods has been shown to be a feasible way to aid dietary recall (Arab et al., 2011). Web-based surveys have

been implemented in a number of outbreaks (Beatty et al., 2009; Srikantiah, 2005) and have been shown to increase the timeliness of investigations without compromising response rates or participant satisfaction (Ghosh, 2008). However, data collected from self-reported, retrospective dietary histories will always be limited by the quality of the memories from which the information was retrieved.

Limitations in human memory have implications for the accuracy of food histories collected during outbreak investigations. Memories of similar and frequently repeated events are particularly susceptible to memory decay (Wirfält, 1998). When an individual is unable to retrieve details of an event such as a meal, a plausible inference is made based on general knowledge about their diet (Smith, 1993). Such compromises may result in the omission of rare food exposures as well as the intrusion of more commonly consumed foods (Smith, 1993). The longer the delay between exposure and recall, the more likely the event is to be misattributed to an earlier or later date. This phenomenon, known as the telescoping effect (Janssen et al., 2006), is problematic for outbreak investigations, which assess specific periods of time and place importance on the temporality of exposure and the onset of illness. Recall strategies used by respondents to retrieve such memories have been found to differ with food type (Johnson-Kozlow et al., 2006), suggesting that certain food exposures may be more susceptible to recall bias than others.

The impact of recall bias on the validity of self-reported food history data during outbreak investigations remains an area of uncertainty as few studies have been conducted in the area. Mann (1981) and Decker et al. (1986) investigated dietary recall of meals served at a single event after 3 and 5 days, respectively. Both studies involved observing individuals at a buffet meal and found sensitivities ranging from 81.2% to 95.2% and specificities ranging from 75% to 98.5%. While these studies provide insight into local point source outbreak investigations, they

offer limited generalizability to national outbreak investigations that involve complex food distribution networks and longer times between consumption and recall.

The Public Health Agency of Canada (Thomas et al., 2013) and the US Centers for Disease Control and Prevention (Scallan et al., 2011) have recognized recall bias, in addition to under-reporting of enteric illness, as a key limitation to the study of foodborne illness. The purpose of this study was to investigate the effect of recall bias on the validity of food history data in the context timeframes reported for both local and national outbreak investigations and to characterize the food exposure of a previously understudied segment of the Canadian population at potentially high risk of food-borne disease.

3.3 Methods

3.3.1 Study Sample

Participants were recruited at the University of Saskatchewan through a collaboration with the Social Sciences Research Laboratory which actively recruits and maintains a cohort of active volunteers. All interested participants with access to an Android phone version 4.0 or greater were included in the study. Advertisements were also placed on the campus-wide Personalized Access to Web Services (PAWS) online platform and announcements were made in a number of classrooms. Upon completion of the study, each participant received compensation of up to $100 CAD depending on the study components they had completed. Ethics approval was obtained from the University of Saskatchewan Behavioural Ethics Review Committee (BEH #15-187).

3.3.2 Data Collection

Data collection took place across three time periods, each lasting 10 weeks. Data were collected from the first cohort (n=17) between September 21st and November 29, 2015, from the second cohort (n=40) between January 18th and March 27, 2016, and from the third cohort (n=39) between January 25nd and April 3, 2016. Basic demographic data were collected using an initial online survey through a link emailed to study participants.

Comprehensive records of food exposures were collected from each group during the initial 10 days of the study through a combination of real-time images of meals, food descriptions, and daily food exposure mini-survey questions.

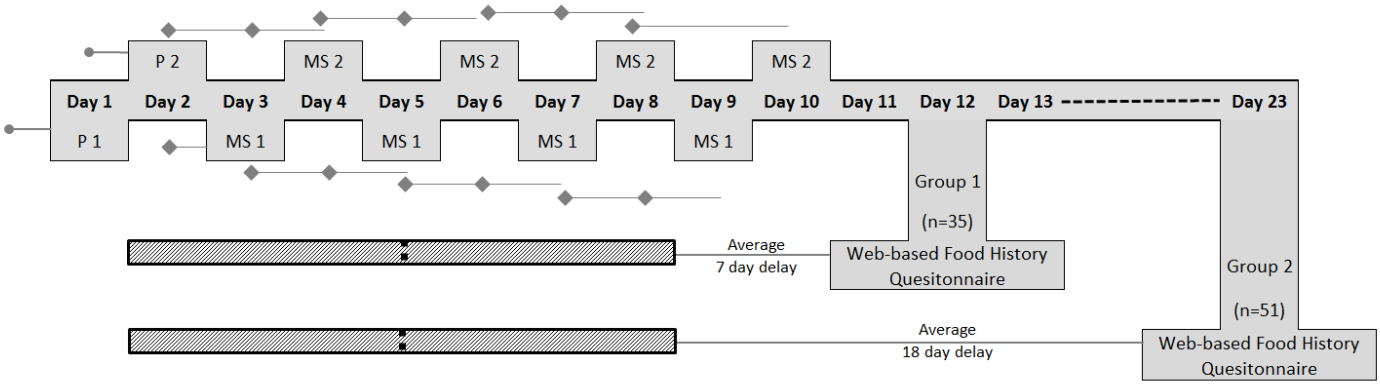

Figure 3.1: Timeline of data collection for the recall bias study relative to the 7-day focus period between day 2 and day 8 of the study (rectangular block). P1 and P2 were practice microsurveys. M1 represents Microsurvey Type 1 and M2 represents Microsurvey Type 2. Microsurveys administered on day 1, 2, 3 and day 10 assessed food exposures of only one previous day. The remainder of the microsurveys assessed food exposures over the 2 days prior. Diamonds (♦) indicate the target for food intake data collection.

Food exposure data were collected via Ethica smartphone technology developed at the University of Saskatchewan. The application can be used to acquire, store, and analyze data on human behavior and mobility (Hashemian et al., 2012). Although Ethica had not previously been used to study foodborne illness, the application has proven useful in collecting related food consumption data (Hashemian et al., 2012). In this study, Ethica provided a convenient and cost

effective method by which to collect real-time data in the form of digital images and brief user-triggered and time-triggered surveys. When this study initiated, Ethica was available for the Android operating system version 4.0 or greater. Data were encrypted and securely transmitted to Ethica servers at the University of Saskatchewan when the smartphones were connected to networks.

This study utilized specific features of the application custom designed to capture information on food consumption behavior and illness. Participants were asked, using a button followed by a multiple choice question, to self-report when they were eating and the source of the food: eating food prepared at home, eating at a restaurant, eating ready-to-eat food purchased off campus, and each food purchased on campus. The participants were then given the option to take a picture of the food. The PhotoFoodDiary feature of the application was used to capture real-time food consumption information. Participants were asked to capture a digital image of each meal or snack consumed over the 10-day period. Each image was automatically labeled with the date, time and geographic coordinates. All participants had the option to provide an audio description of the meal or snack. Participants in groups 2 and 3 were also given the opportunity to provide written descriptions of the components as well as a list of ingredients as desired.

The survey feature of the application was used to administer microsurveys as an additional means of acquiring information about recent food exposures. Microsurveys were administered at breakfast, lunch and dinner. These microsurveys consisted of 1 to 4 short multiple-choice questions regarding food exposures from the previous 24 or 48 hours (Figure 3.1). Participants were given a timeframe of 4 hours to complete each short survey. The microsurveys consisted of short questions modeled after a subset of questions from the subsequent detailed web-based food history survey. This follow-up detailed electronic food

history was administered after the initial data collection period (Figure 3.1). Representative questions were included on the microsurveys for the following food types: fruits, nutrition bars, snacks, nuts, salads, sprouts and fish (Microsurvey 1) and nutritional supplements, cereals, eggs and dairy products, cheese, nut butters, noodle mixes, tofu, poultry and vegetables (Microsurvey 2) (see Appendix 1 for examples of the two sets of microsurveys). The surveys were structured to ensure that all participants would be asked about their consumption of the target food groups for days 2 through 8 of the study.

Participants from cohorts 2 and 3 were randomly assigned to complete a detailed electronic food history questionnaire administered either 7 or 18 days (2.5 weeks) after the midpoint of the data collection period using an online survey platform (Qualtrics Survey Software, Washington, DC). All participants in cohort 1 completed the survey at 2.5 weeks. This questionnaire required participants to recall food exposures between day 2 and day 8 of the study (midpoint – day 5) and was sent to participants at either day 12 (day 5 plus 7 days) or 23 (day 5 plus 18 days). The detailed food history questionnaire was modeled on the Foodbook questionnaire (MacDonald, 2016) with some minor adaptations in consent wording required by local ethics and to provide appropriate context for the participants. The questionnaires and time intervals used in this study were designed to resemble those used by public health officials in Canada.

Self-reported real-time data from days 2 through 8 were summarized for comparison to follow up surveys (Figure 3.1). Data on the occurrence of illness was also recorded through a user-triggered survey which assessed information on symptoms, whether a health-care professional was consulted, and association with alcohol consumption.

3.3.3 Data Analysis

Foods items commonly implicated in enteric outbreaks in Canada were identified based on expert opinion and publications as targets for evaluation of recall bias (Greig and Ravel, 2009; Painter et al., 2009; Pires et al., 2009). Foods that were reported as consumed by between twenty and eighty percent of the study population on the detailed food history questionnaire were deemed eligible for analysis based on having sufficient variability in intake to evaluate the impact of recall on both sensitivity and specificity. Foods that were not easily distinguishable in images were excluded. Based on these criteria, fifteen food items of interest were selected for analysis including lettuce or leafy greens, poultry, nuts, sprouts, cabbage, berries, cucumber, melons, peanut butter, peanuts (not including peanut butter), tomatoes and breaded chicken (see Appendix 2 for a complete list of foods that were selected eligible for analysis).

The analysis of recall bias within the detailed food history questionnaire focused on food exposures between day 2 and day 8 of the study. A total of 1645 images, 984 text descriptions, 18 audio recordings, and 2127 microsurvey responses were collected during this 1-week period. Information obtained from each data collection method was assessed for evidence of exposure to each of the fifteen foods of interest. Foods were only considered present in an image if the food item was visible and clearly distinguishable. Audio recordings were transcribed to text and common misspellings were accounted for during the interpretation of each text description. Data from microsurveys collected on days 3 through 10 reflected the food consumption on days 2 to 8 and were also extracted for comparison to the follow up surveys (Figure 3.1).

A true exposure was defined as the presence of a food item in any image, description, or microsurvey response collected between day 2 and day 8 of the study. Accordingly, if a food item did not appear in any of the three data collection methods between days 2 and 8 of the study, it was assumed the individual did not consume the food item during that period of time.

The potential for under-reporting among these three options was acknowledged and is discussed later in this paper. However, for the purpose of this analysis, data collected using the Ethica app immediately and shortly after exposure was assumed to be the working "gold standard" or reference value for food history.

Cohen's kappa (κ) is a widely accepted method by which to assess the extent of agreement between two or more tests and estimates the level of agreement above that expected by chance alone (Landis and Koch, 1977). Previously accepted guidelines for the interpretation of levels of agreement, as proposed by Landis and Koch (1977) are as follows: less than or equal to 0 = poor; 0.01-0.2 = slight; 0.21-0.4 = fair; 0.41-0.6 = moderate; 0.61-0.8 = substantial; 0.81-1.0 almost perfect. In this study, the level of agreement between the reference standard and data subsequently obtained from the food history questionnaire was assessed by κ values. In this context, a κ statistic of 1.0 would indicate perfect agreement between the reference standard (real-time data) and the results of the food history questionnaire. On the other extreme, a κ statistic of 0.0 would indicate no agreement between the two data sets, suggesting a profound effect of recall bias. This analysis was conducted for each of the 15 food items of interest.

Recall bias was examined by estimating the sensitivity and specificity of the web-based detailed food history questions administered at days 12 and 23 relative to the reference standard information collected on or before day 10. In the context of this study, sensitivity refers to the proportion of participants exposed to a food item (as determined by the reference standard) who reported exposure to that food item on the food history questionnaire. Specificity refers to the proportion of participants not exposed to a food item (as determined by the reference standard) who did not report exposure to that food item on the food history questionnaire. The sensitivity and specificity of exposure to the 15 foods of interest reported on the detailed food history questionnaire were determined for both the 7-day and 18-day average delay period. Furthermore,

the positive predictive value and negative predictive value for exposure to each food item from the food history questionnaire was calculated given the proportion of people who had indicated eating the food item using the reference standard reporting method.

Generalized estimating equations (GEE) with a logit link function and assuming a binomial distribution were used to estimate the effect of factors of interest on the sensitivity of food history questions while accounting for correlations among data on different foods collected from the same participant. Bivariate or unconditional analysis were completed to assess the individual effects of three potential risk factors for sensitivity and specificity: different time delays between consumption and data collection, the presence of gastrointestinal symptoms, and for whether there were any prompts to aid in recall of food exposures. Time delay refers to either the 7-day or 18-day period of delay between the midpoint of the one-week data collection period (day 5 of the day 2 to day 8 focus period) and completion of the food history questionnaire on days 12 or 23 (Figure 3.1). Gastrointestinal illness was defined as diarrhea, vomiting, abdominal pain and cramps, or nausea experienced between day 2 and day 8 of the study. Illnesses due to the consumption of alcohol were excluded from the analysis.

Prompts that might have aided in recall of food exposures refer to questions about the particular food of interest on the microsurveys which preceded the later food history questionnaires. Only four of the 15 food items of interest for assessing recall bias (beef, pork, spinach and strawberries) from the food history questionnaire did not appear on the microsurveys. Participants were prompted to report exposure to the remaining 11 food items on the microsurveys up to day 10 before subsequently completing the food history questionnaires on days 12 or 23. Using only data collected from digital images and food descriptions as the reference standard for this targeted analysis, the accuracy of recall was compared between food items, where responses on the detailed survey might have been prompted by the earlier

microsurvey questions and food items queried on the detailed food history survey where no comparable questions had been asked on the earlier microsurveys.

To determine the generalizability of the food history results to the larger Canadian population, the relative frequency of exposure to the fifteen foods of interest was compared to results obtained from the national Foodbook survey conducted in 2014 by the Public Health Agency of Canada (MacDonald, 2016). Study results were compared to values for the Canadian population as a whole, the Saskatchewan population, the Canadian population for the month of January, and the segment of the Canadian population between 20 and 64 years of age. Databases were created and managed using Microsoft Access 2016 and Microsoft Excel 2016. Data were analyzed using a commercial statistical software program (Stata/SE 14.1 for Windows, StataCorp LP, College Station, TX).

3.4 Results

3.4.1 Study Sample

A total of 96 participants enrolled in the study and completed the initial phase of data collection. Nine participants were lost to follow up before the web-based food history questionnaires on days 12 and 23. One participant failed to consent to completing the food history questionnaire. Data from the remaining 86 participants was included in this analysis. The mean age of the participants providing data for the analysis of recall bias was 26.6 (SD=7.9). The study population contained 66% (57/86) females; 37.8% (31/82) of participants were newcomers to Canada (within the last two years). When prompted 4.8% (4/83) of participants indicated that they were vegetarian or vegan and 8.4% (7/83) of participants in this study indicated that halal or kosher diets influenced their food choices.

Based on the data available from participants that were lost to follow up, 6/7 were female, 71.3% (5/7) were between the 18 to 25 years of age, 20% (1/5) was a newcomer to Canada (moved to Canada within the last two years), 20% (1/5) was vegetarian or vegan and 20% (1/5) indicated that halal or kosher diets influenced their food choices.

During the 7 days of interest (Day 2 to Day 8 of the study) 18.6% (16/86) of participants experienced at least one symptom of foodborne illness and one participant sought medical treatment for their illness. More specifically, 10.5% (9/86) of participants experienced vomiting or diarrhea; none sought medical care for their symptoms during this period.

3.4.2 Inter-rater Agreement between Food History Questionnaires and Real-Time Data

Inter-rater agreement between reported exposure from the reference standard and detailed food history data after 7 days (Table 3.1a) and after 18 days (Table 3.1b) was summarized for each food of interest. The highest observed agreements for lettuce and greens as reported at 7 days (κ=0.548) and for peanut butter at 18 days (κ=0.530) were considered moderate. Most κ values were classified as fair, slight or even poor.

With the exception of leafy greens, no significant differences were apparent based on the 95%CI of levels of agreement between the reference standard and data collected after either 7 days or 18 days. Inter-rater agreement between the reference standard and leafy greens exposure measured on the web-based questionnaire was greater after 7 days, κ=0.548 [0.303 to 0.794] than when compared to that measured at 18 days, κ=-0.052 [-0.250 to 0.146].

Table 3.1: (a) Inter-rater agreement (κ and 95% confidence intervals) between food history data collected after 7 days and the reference standard. (b) Inter-rater agreement between food history data collected after 18 days and the reference standard exposure history collected on the day of consumption or within 48 hours.

a)	**7-Day Delay (n=35)**		
	κ	95% CI	
Lettuce and Leafy Greens	0.548	0.303	0.794
Nuts	-0.091	-0.339	0.158
Sprouts	-0.066	-0.235	0.104
Cabbage	0.043	-0.236	0.321
Berries	-0.071	-0.356	0.215
Cucumber	0.288	-0.031	0.608
Melon	0.364	0.02	0.707
Peanut Butter	0.388	0.059	0.718
Peanuts	0.189	-0.126	0.505
Tomatoes	0.418	0.099	0.736
Beef	0.043	-0.203	0.289
Pork	0.199	-0.125	0.524
Spinach	0.283	-0.005	0.572
Strawberries	0.252	-0.102	0.606
Breaded Chicken	0.149	-0.122	0.420

b)	**18-Day Delay (n=51)**		
	κ	95% CI	
Lettuce and Leafy Greens	-0.052	-0.250	0.146
Nuts	0.180	-0.057	0.417
Sprouts	-0.088	-0.254	0.078
Cabbage	0.187	-0.019	0.393
Berries	0.100	-0.142	0.342
Cucumber	0.172	-0.099	0.442
Melon	0.426	0.134	0.718
Peanut Butter	0.530	0.298	0.762
Peanuts	0.184	-0.048	0.416
Tomatoes	0.305	0.026	0.584
Beef	0.108	-0.133	0.349
Pork	0.036	-0.236	0.308
Spinach	-0.123	-0.379	0.134
Strawberries	0.014	-0.273	0.245
Breaded Chicken	0.058	-0.132	0.247

3.4.3 Accuracy of the Food History Questionnaire

The specificity, sensitivity, positive predictive value (PPV), and negative predictive value (NPV) of food history questionnaires administered after 7 days and after 18 days are shown in Table 3.2. Sensitivity with a 7-day delay to the detailed food history questionnaire varied from a minimum of 14.3% for sprouts to maximum of 100% for leafy greens. Similarly, sensitivity with an 18-day delay to the detailed food history questionnaire varied from a minimum of 15.8% for sprouts to maximum of 77.8% for tomatoes. Specificity with a 7-day delay to the detailed food history questionnaire varied from a minimum of 30.4% for beef to maximum of 80.4% for peanuts. While, specificity with an 18-day delay to the detailed food history questionnaire varied from a minimum of 21.2% for leafy greens to maximum of 92.1% for melons.

Table 3.2: Diagnostic test measures of the validity and predictive values of data collected on exposure after 7 days (a) and after 18 days (b) in regards to 15 food items commonly implicated in foodborne outbreaks. Values in bold represent the medians.

a) 7-day Delay (n=35)

		Point	95% CI	
Lettuce and Leafy Greens	Sensitivity (%)	100	80.5	100
	Specificity (%)	55.6	30.8	78.5
	PPV (%)	68.0	46.5	85.1
	NPV (%)	100	69.2	100
Nuts	Sensitivity (%)	21.7	7.46	43.7
	Specificity (%)	66.7	34.9	90.1
	PPV (%)	55.6	21.2	86.3
	NPV (%)	30.8	14.3	51.8
Sprouts	Sensitivity (%)	14.3	4.03	32.7
	Specificity (%)	**71.4**	**29.0**	**96.3**
	PPV (%)	66.7	22.3	95.7
	NPV (%)	17.2	5.85	35.8
Cabbage	Sensitivity (%)	33.3	14.6	57.0
	Specificity (%)	71.4	41.9	91.6
	PPV (%)	63.6	30.8	89.1
	NPV (%)	41.7	22.1	63.4
Berries	Sensitivity (%)	46.2	26.6	66.6
	Specificity (%)	44.4	13.7	78.8
	PPV (%)	70.6	44.0	89.7
	NPV (%)	22.2	6.41	47.6
Cucumber	Sensitivity (%)	61.5	31.6	86.1
	Specificity (%)	68.2	45.1	86.1
	PPV (%)	53.3	26.6	78.7
	NPV (%)	75.0	50.9	91.3
Melon	Sensitivity (%)	66.7	22.3	95.7
	Specificity (%)	79.3	60.3	92.0
	PPV (%)	40.0	12.2	73.8
	NPV (%)	92.0	74.0	99.0
Peanut Butter	**Sensitivity (%)**	**54.5**	**23.4**	**83.3**
	Specificity (%)	83.3	62.6	95.3
	PPV (%)	60.0	26.2	87.8
	NPV (%)	80.0	59.3	93.2
Peanuts	Sensitivity (%)	30.8	9.09	61.4
	Specificity (%)	86.4	65.1	97.1
	PPV (%)	57.1	18.4	90.1
	NPV (%)	67.9	47.6	84.1

b) 18-day Delay (n=51)

		Point	95% CI	
Lettuce and Leafy Greens	Sensitivity (%)	72.2	46.5	90.3
	Specificity (%)	21.2	8.98	38.9
	PPV (%)	33.3	19.1	50.2
	NPV (%)	58.3	27.7	84.8
Nuts	Sensitivity (%)	48.5	30.8	66.5
	Specificity (%)	**72.2**	**46.5**	**90.3**
	PPV (%)	76.2	52.8	91.8
	NPV (%)	43.3	25.5	62.6
Sprouts	Sensitivity (%)	15.8	6.02	31.3
	Specificity (%)	69.2	38.6	90.9
	PPV (%)	60.0	26.2	87.8
	NPV (%)	22.0	10.6	37.6
Cabbage	Sensitivity (%)	39.4	22.9	57.9
	Specificity (%)	83.3	58.6	96.4
	PPV (%)	81.3	54.4	96
	NPV (%)	42.9	26.3	60.6
Berries	Sensitivity (%)	37.9	20.7	57.7
	Specificity (%)	72.7	49.8	89.3
	PPV (%)	64.7	38.3	85.8
	NPV (%)	47.1	29.8	64.9
Cucumber	Sensitivity (%)	54.2	32.8	74.4
	Specificity (%)	63.0	42.4	80.6
	PPV (%)	56.5	34.5	76.8
	NPV (%)	60.7	40.6	78.5
Melon	**Sensitivity (%)**	**46.2**	**19.2**	**74.9**
	Specificity (%)	92.1	78.6	98.3
	PPV (%)	66.7	29.9	92.5
	NPV (%)	83.3	68.6	93
Peanut Butter	Sensitivity (%)	74.1	53.7	88.9
	Specificity (%)	79.2	57.8	92.9
	PPV (%)	80.0	59.3	93.2
	NPV (%)	73.1	52.2	88.4
Peanuts	Sensitivity (%)	34.6	17.2	55.7
	Specificity (%)	84.0	63.9	95.5
	PPV (%)	69.2	38.6	90.9
	NPV (%)	55.3	38.3	71.4

Table 3.2: Diagnostic test measures of the validity and predictive values of data collected on exposure after 7 days (a) and after 18 days (b) in regards to 15 food items commonly implicated in foodborne outbreaks. Values in bold represent the medians (continued).

a)	**7-day Delay (n=35)**				**b)**	**18-day Delay (n=51)**			
		Point	95% CI				Point	95% CI	
Tomatoes	Sensitivity (%)	82.6	61.2	95.0	Tomatoes	Sensitivity (%)	77.8	60.8	89.9
	Specificity (%)	58.3	27.7	84.8		Specificity (%)	53.3	26.6	78.7
	PPV (%)	79.2	57.8	92.9		PPV (%)	80.0	63.1	91.6
	NPV (%)	63.6	30.8	89.1		NPV (%)	50.0	24.7	75.3
Beef	Sensitivity (%)	75.0	42.8	94.5	Beef	Sensitivity (%)	70.0	45.7	88.1
	Specificity (%)	30.4	13.2	52.9		Specificity (%)	41.9	24.5	60.9
	PPV (%)	36.0	18.0	57.5		PPV (%)	43.8	26.4	62.3
	NPV (%)	70.0	34.8	93.3		NPV (%)	68.4	43.4	87.4
Pork	Sensitivity (%)	61.1	35.7	82.7	Pork	Sensitivity (%)	31.6	12.6	56.6
	Specificity (%)	58.8	32.9	81.6		Specificity (%)	71.9	53.3	86.3
	PPV (%)	61.1	35.7	82.7		PPV (%)	40.0	16.3	67.7
	NPV (%)	58.8	32.9	81.6		NPV (%)	63.9	46.2	79.2
Spinach	Sensitivity (%)	54.5	32.2	75.6	Spinach	Sensitivity (%)	48.6	31.4	66
	Specificity (%)	76.9	46.2	95.0		Specificity (%)	37.5	15.2	64.6
	PPV (%)	80.0	51.9	95.7		PPV (%)	63.0	42.4	80.6
	NPV (%)	50.0	27.2	72.8		NPV (%)	25.0	9.77	46.7
Strawberries	Sensitivity (%)	44.4	13.7	78.8	Strawberries	Sensitivity (%)	18.8	4.05	45.6
	Specificity (%)	80.8	60.6	93.4		Specificity (%)	80.0	63.1	91.6
	PPV (%)	44.4	13.7	78.8		PPV (%)	30.0	6.67	65.2
	NPV (%)	80.8	60.6	93.4		NPV (%)	68.3	51.9	81.9
Breaded Chicken	Sensitivity (%)	38.1	18.1	61.6	Breaded Chicken	Sensitivity (%)	18.5	6.3	38.1
	Specificity (%)	78.6	49.2	95.3		Specificity (%)	87.5	67.6	97.3
	PPV (%)	72.7	39.0	94.0		PPV (%)	62.5	24.5	91.5
	NPV (%)	45.8	25.6	67.2		NPV (%)	48.8	33.3	64.5

3.4.4 The Effects of Time, Illness and Multiple Assessments on the Accuracy of Dietary Recall

There were no differences in either sensitivity or the specificity of the reported exposures based on the detailed food history questionnaires for the 15 foods of interest regardless of whether the questionnaire was administered after the 7- or 18-day delay (Table 3.3). Similarly, there was no difference among the sensitivities and specificities of the reported exposures for the 15 foods of interest depending on whether or not the participant reported any illness during the study period (Table 3.3).

Table 3.3: Odds ratios for the bivariate analysis of effect of time delay, enteric illness and multiple assessments on the sensitivity and specificity of dietary recall for 15 foods of interest (n=86).

	Sensitivity				Specificity			
	OR	95% CI		P	OR	95% CI		P
Time Delay 18-day delay (In reference to 7-day delay)	0.99	0.96	1.02	0.43	1.00	0.96	1.03	0.83
Enteric Illness Symptomatic (In reference to non-symptomatic)	1.20	0.77	1.87	0.42	1.08	0.71	1.66	0.72
Memory Priming Assessed on Microsurveys AND FHQ (In reference to assessed only on FHQ)	1.18	0.83	1.68	0.37	.638	.473	.859	.003

Multiple assessments had a negative effect on the specificity and positive predictive value of food history questionnaires (Table 3.3). In other words, repeatedly prompting participants to recall a food exposure on microsurveys increased the likelihood of reporting a false positive on food history questionnaires administered later. The sensitivity of food history questionnaires was not significantly affected.

3.4.5 Comparing Apparent Prevalences to Foodbook Exposure Frequencies

The apparent prevalence of each food item of interest based on the combined results of the detailed food history questionnaire at 7 and 18 days was compared to results obtained from the national Foodbook Survey conducted in 2014 by the Public Health Agency of Canada (MacDonald, 2016). This national initiative interviewed participants over the telephone to assess food exposures over the past 7 days. Results were compared to national exposure rates, as well as to exposure rates specific to the province of Saskatchewan, and to the corresponding age demographic.

Table 3.4: Food exposures reported by the study population of university students based on a food history questionnaire completed an average of at least 1 week after exposure (n=86) compared to reference values from the Foodbook report (MacDonald, 2016).

	Study Population			Foodbook Report (MacDonald, 2016)		
	Prevalence (%)	95% CI		Canada (%)	Saskatchewan (%)	20-64 Years (%)
Lettuce/Leafy Greens	41	30.0	51.8	82.4	80.7	86.3
Melons	22	14.0	32.3	39.7	41.2	38.3
Tomatoes	69	58.0	78.2	72.9	66.4	77.9
Breaded Chicken	56	45.0	66.9	16.5	18.7	14.2
Pork	43	32.0	54.2	55.1	62.7	55.1
Spinach	66	55.0	76.1	28.4	28.9	29.7
Strawberries	29	20.0	39.9	49.6	49.7	46.3
Peanut Butter	44	33.0	55.3	55.0	58.8	54.5
Peanuts	45	35.0	56.5	33.6	36.3	35.7
Nuts	65	54.0	75.1	65.4	64.6	68.9
Sprouts	77	66.0	85.2	12.9	11.7	13.0
Beef	37	27.0	48.3	78.4	86.3	79.7
Cabbage	63	52.0	73.3	30.0	32.0	30.9
Berries	64	53.0	74.4	65.2	61.1	62.5
Cucumber	43	32.0	54.2	62.9	59.0	61.7

To account for seasonality, data gathered in January (groups 1 and 2) were also specifically compared to Foodbook reference values corresponding to the month of January (MacDonald, 2016).

Table 3.5: Food exposures reported by the study population of university students based on a food history questionnaire completed an average of at least 1 week after exposure (n=86) compared to reference values from the Foodbook report reference values for the month of January (MacDonald, 2016).

	Study Population			Foodbook Report (MacDonald, 2016)
	Prevalence (%)	95% CI		January (%)
Lettuce/Leafy Greens	41	30.0	51.8	76.2
Melons	22	14.0	32.3	22.0
Tomatoes	69	58.0	78.2	63.0
Breaded Chicken	56	45.0	66.9	18.0
Pork	43	32.0	54.2	54.7
Spinach	66	55.0	76.1	29.8
Strawberries	29	20.0	39.9	42.3
Peanut Butter	44	33.0	55.3	62.3
Peanuts	45	35.0	56.5	36.6
Nuts	65	54.0	75.1	61.7
Sprouts	77	66.0	85.2	13.2
Beef	37	27.0	48.3	79.9
Cabbage	63	52.0	73.3	22.8
Berries	64	53.0	74.4	55.2
Cucumber	43	32.0	54.2	51.8

3.5 Discussion

The fallibility of human memory has long been recognized as a key limitation to foodborne outbreak investigation (Decker et al., 1986; Mann, 1981). Misclassifications in self-reported exposure status reduces the power of epidemiological studies to detect meaningful associations between exposures and the development of illness. While the limitations associated with recall bias are commonly reported (Dechet et al., 2006; Schenkel et al., 2006), little is known about the extent and implications of the problem. This study serves as a first step in quantifying recall bias in a context comparable to how cases and controls might be questioned for outbreak investigations so that recall bias can be accounted for and mitigated in public health practice.

The validity of food history data collected after 18 days was found to be generally low with sensitivities ranging from 15.8% to 77.8% for 15 foods of interest, and specificities ranging from 21.2% to 92.1%. There was no observed significant improvement in sensitivity or specificity observed when recall time was reduced to 7 days from the midpoint of the period of interest. Similarly, agreement with the gold standard did not differ between the time intervals for most food items, with leafy greens being the lone exception. The comparably low accuracy of dietary memory after both time intervals examined in this study suggests there is a substantial potential for bias for most food types following the first week after consumption.

Although this study utilized a multitude of data collection methods in an effort to capture comprehensive food histories, imperfections in the reference standard will still exist. Compliance measured as an average of at least one report per day in the data collection methods used to create the gold standard of food history in the first ten days of the study were 100% (96/96) for food microsurveys, 85.4% (82/96) for digital images and 68.4% (54/79) for meal descriptions (Chapter 2). It is possible that participants failed to record a food exposure in all three data collection methods used to create the reference standard with Ethica (false negative), but remembered the exposure when later completing the Food History Questionnaire.

Such false negatives in the Ethica data could have inflated the negative denominator derived from the reference standard and have potentially lead to an underestimate of the specificity of Food History Questionnaires. These same false negative results in the Ethica data could have resulted in an underestimate of the denominator for the questionnaire sensitivity calculation and an inflated estimate of sensitivity. It is much less likely that participants would have said that they ate something that they did not when using Ethica (false positives), given most reporting was done shortly after the meal was consumed.

The sensitivity of recall based on the Ethica reference standard was highest for common foods such as leafy greens and tomatoes (77.2%-100%). These same foods were associated with relatively low specificities (21.2%-58.3%). Recognizing that the low specificity estimates might be impacted by incomplete reporting, this finding may still indicate a tendency for participants to use knowledge of habitual diet to answer questions about common food exposures rather than recalling specific exposure events. On the other hand, specificities were found to be highest (67.7%-72.2%) for sprouts and nuts.

The sensitivities of 18-day dietary recall were compared to the findings of past relevant outbreak investigations reported by PHAC and the US CDC. This comparison included only solved outbreaks investigations in which cases were laboratory confirmed, and where cases were infected with the same outbreak strain (PFGE, MLVA or WGS) that was isolated from the contaminated food source. Infection with the same outbreak strain suggests consumption of the same contaminated food source. Therefore, the percentage of laboratory confirmed cases that failed to report consuming the implicated food item might provide an estimate of the false negative rate and then the sensitivity of food history questionnaires. While cross-contamination and multiple source outbreaks may account for some differences, a substantial proportion of the discrepancies where cases did not report eating the food of interest was assumed to be a consequence of recall bias for this comparison.

The sensitivity of dietary recall to tomato consumption computed in this study (77.8% [60.8-89.9]) is consistent with the relative frequency of exposure reported among confirmed cases in previous tomato outbreaks: 75% (Reller et al., 2006), 81% (Behravesh et al., 2012), 70% (Greene et al., 2008), and 69% (Donnan et al., 2012). Comparable results were found between the computed sensitivity estimated for reporting consumption of cucumbers (61.6 [31.6-86.1]) and previous cucumber outbreaks: 62% (Angelo et al., 2015), 75% (Centers for Disease Control

and Prevention, 2015) and 69% (Donnan et al., 2012). Finally, the sensitivity of dietary memory for consumption of leafy greens (72.2% [46.5-90.3]) was also congruent with the exposure frequencies observed among confirmed cases in past leafy green outbreaks: 71% (Nuorti et al., 2004), 70% (Ackers et al., 1998), 85% (Centres for Disease Control and Prevention, 2013), and 64.9% (Slayton et al., 2013). The similarity of these findings indicates that recall bias may serve as a plausible explanation for the lack of exposure reported among confirmed cases during outbreak investigations.

In addition to time after consumption, one of the other factors considered as a potential influence on recall was illness during the observation period. The dietary recall of participants who reported enteric illness during the study were not found to be more sensitive or specific than that of asymptomatic individuals. Similarly, prompting participants to recall food exposures on more than one occasion (i.e. on microsurveys and again on food history questionnaires) did not have a significant effect on the sensitivity of dietary recall. However, food exposures that were assessed on multiple occasions (i.e. on microsurveys and again on Food History Questionnaires) were found to have a lower specificity than food exposures that were assessed only on the food history questionnaire. In other words, having been previously asked about a food exposure increases the likelihood of reporting a false positive.

Hupbach et al. (2009) describe the retrieval of episodic memories as a reactivation process in which memories become fragile before they are reconsolidated. Each time a memory is retrieved, it becomes susceptible to the intrusion of information from similar events that took place at a later date (Hupbach et al., 2009). In this way, details of memory may become lost, misattributed or inadvertently falsified. Jacoby and Whitehouse (1989) also found that priming often resulted in a misattribution of familiarity. Memories of recalling whether or not an event occurred may be misclassified as having actually experienced the event (Janssen et al., 2006). In

the context of this study, priming from microsurveys is believed to have caused a misattribution of familiarity leading to a higher rate of false positives. Foods that were not assessed on microsurveys experiences no such increase.

Advances in smartphone technology continue to give rise to new opportunities for the acquisition of data on human behavior. The Ethica app has itself advanced and is now available for iPhones as well as Android phones. As new data collection and analysis tools become available, they allow for a more accurate representation of dietary history. While a true gold standard of dietary histories may not be attainable for direct comparison, future studies may employ methods such as Bayesian latent class analysis to investigate the sensitivity and specificity of data collection methods that are currently available in the absence of a true gold standard. Formal incorporation of prior knowledge regarding measurement uncertainty may offer a means by which to account for, rather than aim to eliminate, recall bias in foodborne outbreak investigations.

The magnitude of the effect that recall bias had on the accuracy of dietary memory was not uniform across food types. This type of differential error has considerable implications on outbreak investigations as it may diminish some associations while inflating others. Quantifying the effects of recall bias on the accuracy of self-reported exposure to different foods is the first step toward accounting for this bias. Future research into the effects of recall bias at different points in time and on different commonly implicated foods is warranted. Not only will this research provide insight on strategies to mitigate recall bias during data collection but it may lead to data analysis strategies to adjust for these inherent biases.

3.6 Conclusion

Limitations in human memory present practical implications to the investigation of foodborne disease outbreaks. This study serves as a first step towards quantifying these implications so that they can be accounted for and mitigated in future outbreak investigation strategies. The sensitivity and specificity of dietary recall of different food items will allow for a better understanding of the implications of recall bias on outbreak investigation strategies. Through multidisciplinary collaborations, advances in areas of psychology may be applied to epidemiological studies to inform best practices in foodborne outbreak investigation.

3.7 References

Ackers, M. L., Mahon, B. E., Leahy, E., Goode, B., Damrow, T., Hayes, P. S., Bibb, W. F., Rice, D. H., Barrett, T. J., Hutwagner, L., Griffin, P. M., & Slutsker, L. (1998). An outbreak of Escherichia coli O157:H7 infections associated with leaf lettuce consumption. *J Infect Dis, 177*(6), 1588-1593.

Angelo, K. M., Chu, A., Anand, M., Nguyen, T. A., Bottichio, L., Wise, M., Williams, I., Seelman, S., Bell, R., Fatica, M., Lance, S., Baldwin, D., Shannon, K., Lee, H., Trees, E., Strain, E., Gieraltowski, L., Centers for Disease, C., & Prevention. (2015). Outbreak of Salmonella Newport infections linked to cucumbers--United States, 2014. *MMWR Morb Mortal Wkly Rep, 64*(6), 144-147.

Arab, L., Estrin, D., Kim, D. H., Burke, J., & Goldman, J. (2011). Feasibility testing of an automated image-capture method to aid dietary recall. *Eur J Clin Nutr, 65*(10), 1156-1162. doi:10.1038/ejcn.2011.75

Beatty, M., Shevick, G., Shupe-Ricksecker, K., Bannister, E., Tulu, A., Lancaster, K., Alexander, N., Zellner, D., Lyszkowicz, E., & Braden, C. (2009). Large Salmonella Enteritidis outbreak with prolonged transmission attributed to an infected food handler, Texas, 2002. *Epidemiology and Infection, 137*(03), 417-427.

Behravesh, C. B., Blaney, D., Medus, C., Bidol, S. A., Phan, Q., Soliva, S., Daly, E. R., Smith, K., Miller, B., Taylor, T., Jr., Nguyen, T., Perry, C., Hill, T. A., Fogg, N., Kleiza, A., Moorhead, D., Al-Khaldi, S., Braden, C., & Lynch, M. F. (2012). Multistate outbreak of Salmonella serotype Typhimurium infections associated with consumption of restaurant tomatoes, USA, 2006: hypothesis generation through case exposures in multiple restaurant clusters. *Epidemiol Infect, 140*(11), 2053-2061. doi:10.1017/S0950268811002895

Centers for Disease Control and Prevention. (2015). Multistate outbreak of Salmonella poona infections linked to imported cucumbers.

Centres for Disease Control and Prevention. (2013). Multistate outbreak of E. coli O157: H7 infections linked to romaine lettuce.

CIFOR. (2014). Guidelines for foodborne disease outbreak response [Internet]. Atlanta: Council of State and Territorial Epidemiologists.

Dechet, A. M., Scallan, E., Gensheimer, K., Hoekstra, R., Gunderman-King, J., Lockett, J., Wrigley, D., Chege, W., Sobel, J., & Group, M. W. (2006). Outbreak of multidrug-resistant Salmonella enterica serotype Typhimurium definitive type 104 infection linked to commercial ground beef, northeastern United States, 2003–2004. *Clinical infectious diseases, 42*(6), 747-752.

Decker, M., Booth, A., Dewey, M., Fricker, R., Hutcheson, R., & Schaffner, W. (1986). Validity of food consumption histories in a foodborne outbreak investigation. *American journal of epidemiology, 124*(5), 859-863.

Donnan, E. J., Fielding, J. E., Gregory, J. E., Lalor, K., Rowe, S., Goldsmith, P., Antoniou, M., Fullerton, K. E., Knope, K., Copland, J. G., Bowden, D. S., Tracy, S. L., Hogg, G. G., Tan, A., Adamopoulos, J., Gaston, J., & Vally, H. (2012). A multistate outbreak of hepatitis A associated with semidried tomatoes in Australia, 2009. *Clin Infect Dis, 54*(6), 775-781. doi:10.1093/cid/cir949

Fong, D., Otterstatter, M., Taylor, M., & Galanis, E. (2017). Analysis of enteric disease outbreak metrics, British Columbia Centre for Disease Control, 2005-2014. *Canada Communicable Disease Report, 43*(1), 1.

Ghosh, T. S. (2008). Internet-versus Telephone-based Local Outbreak Investigations-Volume 14, Number 6—June 2008-Emerging Infectious Disease journal-CDC.

Greene, S. K., Daly, E. R., Talbot, E. A., Demma, L. J., Holzbauer, S., Patel, N. J., Hill, T. A., Walderhaug, M. O., Hoekstra, R. M., Lynch, M. F., & Painter, J. A. (2008). Recurrent multistate outbreak of Salmonella Newport associated with tomatoes from contaminated fields, 2005. *Epidemiol Infect, 136*(2), 157-165. doi:10.1017/S095026880700859X

Greig, J. D., & Ravel, A. (2009). Analysis of foodborne outbreak data reported internationally for source attribution. *Int J Food Microbiol, 130*(2), 77-87. doi:10.1016/j.ijfoodmicro.2008.12.031

Hashemian, M., Knowles, D., Calver, J., Qian, W., Bullock, M., Bell, S., Mandryk, R., Osgood, N., & Stanley, K. (2012). iEpi: an end to end solution for collecting, conditioning and utilizing epidemiologically relevant data. *Proceedings of the 2nd ACM international workshop on Pervasive Wireless Healthcare*.

Hupbach, A., Gomez, R., & Nadel, L. (2009). Episodic memory reconsolidation: updating or source confusion? *Memory, 17*(5), 502-510.

Janssen, S. M., Chessa, A. G., & Murre, J. M. (2006). Memory for time: How people date events. *Memory & Cognition, 34*(1), 138-147.

Johnson-Kozlow, M., Matt, G. E., & Rock, C. L. (2006). Recall strategies used by respondents to complete a food frequency questionnaire: an exploratory study. *J Am Diet Assoc, 106*(3), 430-433. doi:10.1016/j.jada.2005.12.007

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *biometrics*, 159-174.

MacDonald, D. (2016). *Foodbook: The Canadian Food, Water and Animal Exposure Study*. Paper presented at the IAFP 2016 Annual Meeting.

Mann, J. M. (1981). A prospective study of response error in food history questionnaires: implications for foodborne outbreak investigation. *American journal of public health, 71*(12), 1362-1366.

Mausner, J., & Kramer, S. (1985). Epidemiology: An introductory text. Philadelphia, PA: WB Saunders: Harcourt Brace Jovanovich.

Nuorti, J. P., Niskanen, T., Hallanvuo, S., Mikkola, J., Kela, E., Hatakka, M., Fredriksson-Ahomaa, M., Lyytikainen, O., Siitonen, A., Korkeala, H., & Ruutu, P. (2004). A widespread outbreak of Yersinia pseudotuberculosis O:3 infection from iceberg lettuce. *J Infect Dis, 189*(5), 766-774. doi:10.1086/381766

Painter, J. A., Ayers, T., Woodruff, R., Blanton, E., Perez, N., Hoekstra, R. M., Griffin, P. M., & Braden, C. (2009). Recipes for foodborne outbreaks: a scheme for categorizing and grouping implicated foods. *Foodborne pathogens and disease, 6*(10), 1259-1264.

Pires, S. M., Evers, E. G., van Pelt, W., Ayers, T., Scallan, E., Angulo, F. J., Havelaar, A., & Hald, T. (2009). Attributing the human disease burden of foodborne infections to specific sources. *Foodborne pathogens and disease, 6*(4), 417-424.

Reller, M. E., Nelson, J. M., Mølbak, K., Ackman, D. M., Schoonmaker-Bopp, D. J., Root, T. P., & Mintz, E. D. (2006). A large, multiple-restaurant outbreak of infection with Shigella flexneri serotype 2a traced to tomatoes. *Clinical infectious diseases, 42*(2), 163-169.

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., Jones, J. L., & Griffin, P. M. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis, 17*(1).

Schenkel, K., Bremer, V., Grabe, C., Van Treeck, U., Schreier, E., Höhne, M., Ammon, A., & Alpers, K. (2006). Outbreak of hepatitis A in two federal states of Germany: bakery products as vehicle of infection. *Epidemiology and Infection, 134*(06), 1292-1298.

Slayton, R. B., Turabelidze, G., Bennett, S. D., Schwensohn, C. A., Yaffee, A. Q., Khan, F., Butler, C., Trees, E., Ayers, T. L., Davis, M. L., Laufer, A. S., Gladbach, S., Williams, I., & Gieraltowski, L. B. (2013). Outbreak of Shiga toxin-producing Escherichia coli (STEC) O157:H7 associated with romaine lettuce consumption, 2011. *PLoS One, 8*(2), e55300. doi:10.1371/journal.pone.0055300

Smith, A. (1993). Cognitive psychological issues of relevance to the validity of dietary reports. *European journal of clinical nutrition, 47*, S6-18.

Srikantiah, P. (2005). Web-based Investigation of Multistate Salmonellosis Outbreak-Volume 11, Number 4—April 2005-Emerging Infectious Disease journal-CDC.

Thomas, M. K., Murray, R., Flockhart, L., Pintar, K., Pollari, F., Fazil, A., Nesbitt, A., & Marshall, B. (2013). Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne pathogens and disease, 10*(7), 639-648.

Vik, J., & Hexemer, A. (2014). Summary: Canada's Food-borne Illness Outbreak Response Protocol. *Canada Communicable Disease Report, 40*(14), 306.

Vik, J., Hexemer, A., & Farber, J. (2014). Summary: Weight of evidence-Factors to consider when investigating a food-borne illness outbreak. *Canada Communicable Disease Report, 40*(14), 303.

WHO. (2015a). WHO estimates of the global burden of foodborne diseases. *WHO: Geneva, Switzerland*.

Wirfält, E. (1998). Cognitive aspects of dietary assessment. *Food & Nutrition Research, 42*, 56-59.

CHAPTER 4 – SHORT COMMUNICATION: APPLYING BAYESIAN LATENT CLASS ANALYSIS TO DETERMINE THE SENSITIVITY AND SPECIFICITY OF A REAL-TIME DATA COLLECTION METHODS, MICROSURVEYS, AND A FOOD HISTORY QUESTIONNAIRE

4.1 Abstract

One of the most pertinent obstacles to foodborne outbreak investigation continues to be the lack of a true gold standard by which to measure the accuracy data collected from food history questionnaires. Even data collected in real-time, may involve misclassifications of exposure status due to issues in compliance. Bayesian latent class analysis incorporates previous knowledge of a parameters and does not require a gold standard test. The purpose of this study was to determine the accuracy and validity of data collected in real-time from digital images and meal descriptions, microsurveys and the food history questionnaire. Using Bayesian analysis, the sensitivities and specificities of the three data collection methods were assessed. Data was obtained from 51 students at the University of Saskatchewan that captured their food consumption through digital images and meal descriptions, microsurveys and completed a food history questionnaire that was administered 2.5 weeks after the midpoint of a 1 week window of interest. The five food exposures that were assessed included tomatoes, lettuce and leafy greens, cucumbers, berries and nuts. The sensitivities of the three data collection methods ranged from 53.0-79.1% for images/meal descriptions, 52.4-75.1% for microsurveys and 59.1-78.6% for the food history questionnaire. The specificities ranged from 84.3-90.9% for images/meal descriptions, 83.3-89.3% for microsurveys and 47.8-88.4% for the food history questionnaire. Overall, the sensitivities, specificities and true prevalences computed by Bayesian latent class analysis did not differ significantly from findings of the frequentist approach used in Chapter 3.

However, the previously assumed gold standard was found to be imperfect with sensitivities from 85.3-91.0% and specificities from 73.4-80.6% for the parallel interpretation of images/meal descriptions and microsurveys. This study provides an innovative approaches to adjusting for recall bias in estimating food exposures with imperfect tests.

4.2 Introduction

Food recall is influenced by various individual characteristics, such as food preferences, memory performance and attention to detail. While we consume a wide variety of foods on a daily basis, the attention that our minds devote to processing and storing this information varies depending on context and food type (Johnson-Kozlow et al., 2006; Smith, 1993). Although information on when which foods were eaten may seem of little importance in our daily lives, these details are crucial to foodborne outbreak investigations.

The epidemiological tools used in outbreak investigations are susceptible to a variety of biases. While biases introduced by the interviewer and the instruments used to collect data are at the focus of methodological studies, error introduced by the interviewee is often overlooked. It is well known that dietary memory fades with time (Wirfält, 1998), yet data collected from food history questionnaires is assumed to be accurate regardless of time delays. Although the fallibility of human memory may be unavoidable, it may be possible to mitigate the error it introduces through careful quantification and calibration of epidemiological tools.

Health studies are often faced with the challenge of assessing the accuracy of a diagnostic test when the true disease status cannot be determined with certainty. Latent class modelling has been introduced as a method by which to obtain valid estimates of sensitivities, specificities, and true prevalence even when no gold standard test is available (Formann and Kohlmann, 1996; Rindskopf and Rindskopf, 1986; van Smeden et al., 2013). The Bayesian approach to latent class

modelling has some important advantages when compared to the frequentist approach. Bayesian modelling allows researchers to incorporate prior knowledge from published articles and expert opinion. This approach allows new data to be assessed in the context of previous work that has been done in the area. If minimum prior knowledge is incorporated into the Bayesian model, then it is likely to arrive at a similar result as the frequentist model. An illustrative example of the application of Bayesian latent class modelling in diagnostic testing is found in the work of Schumacher et al. (2016) on the evaluation of diagnostic tests for childhood pulmonary tuberculosis. By simultaneously assessing the results of different tests, and inputting prior knowledge on the measurement accuracy of each test, Schumacher et al. (2016) determined the sensitivity and specificity of radiography, microscopy, Xpert MTB/RIF, tuberculin skin test, and liquid culture in diagnosing the disease.

During outbreak investigations, data collected from food history questionnaires is assumed to be an accurate and complete account of dietary history. Since these questionnaires are typically administered anywhere from several days to several weeks after the onset of symptoms, errors in memory may result in incomplete and unreliable data. By asking participants to capture images of foods, describe meals and complete microsurveys, using the Ethica application, the study reported in Chapter 3 attempted to create a reference standard for dietary history. However, less than perfect compliance was observed; this may have had consequences on the validity of the reference standard. The purpose of this study was to determine sensitivity and specificity without assuming a gold standard for the analysis for the data collected in real-time from digital images and meal descriptions, from the microsurveys and from a food history questionnaire similar to those currently being used to conduct national foodborne outbreak investigations in Canada. These results will be used to measure the validity

of the gold standard used in the previous study, to estimate the true prevalence of exposure to foods of interest, and to gain a more informative estimate of the effect of recall bias on dietary memory.

4.3 Methods

This study utilized data collected as part of previous investigation (Chapter 3) to assess the validity of food history techniques that are currently being implemented to investigate foodborne outbreaks. Data on the food histories were collected from 51 students at the University of Saskatchewan through three different methods – images/meal descriptions, microsurveys and food history questionnaires – to assess how accurately participants could recall food exposures.

Data was collected over three different time periods; the first group began data collection on September 21, 2015 and completed the food history questionnaire on October 13, 2015 (n=15); the second group began data collection on January 18, 2016 and completed and the February 9, 2016 (n=22); the third group began data collection on January 25, 2016 and completed the food history questionnaire on February 16, 2016 (n=14). See Figure 4.1 for the timeline of data collection.

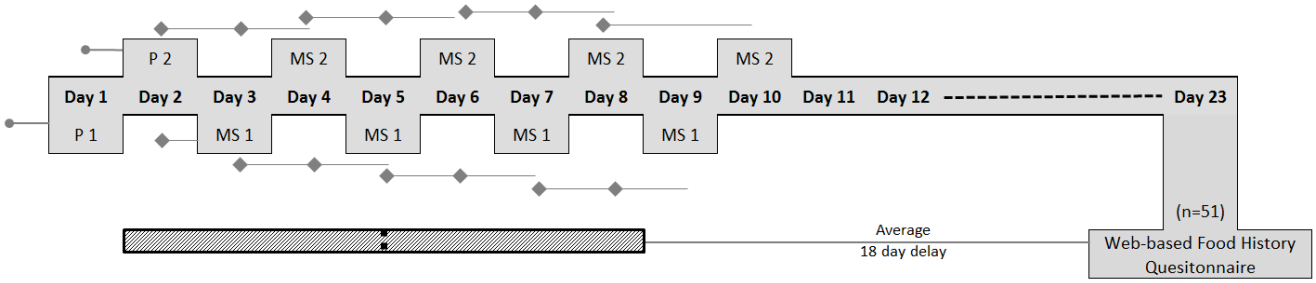

Figure 4.1: Timeline of data collection for microsurveys and the food history questionnaire relative to the 7-day period of interest between day 2 and 8 of the study (rectangular block). P1 and P2 were practice microsurveys. M1 represents Microsurvey Type 1 and M2 represents Microsurvey Type 2. Microsurveys administered on day 1, 2, 3 and day 10 assessed food exposures of only one previous day. The remainder of the microsurveys assessed food exposures over the 2 days prior. Diamonds (♦) indicate the target for food intake data collection.

Refer to Chapter 3 for a complete description of the timeline, tools and methods that were used to collect the relevant data.

4.3.1 Bayesian Latent Class Analysis

Bayesian latent class analysis was conducted to determine the sensitivity and specificity of foodborne history questionnaires for five foods that contained non-zero values in each cell of the contingency tables. The foods included in the Bayesian analysis were tomatoes, nuts, lettuce and leafy greens, berries, and cucumbers. The Bayesian simulations were run assuming 3 tests and 1 population (Branscum et al., 2005). The model also allowed for correlation between the images/meal descriptions and microsurvey data collected using Ethica.

Sokal (1989) suggests that the burn-in period should be less than 5% of the length of the simulation to ensure that the parameter estimates are minimally biased by any data generated from the non-stationary segment of the distribution. A burn-in period of 100,000 was discarded based on the inferences made from the Brook-Gelman-Rubin diagnostic plots (Gelman and

Rubin, 1992). A total run length of 2,500,000 was used; the burn in period comprised 4.2% of the total number of simulations. After excluding the 100,000 iterations of the burn-in period, the remaining 2,400,000 iterations were monitored and analyzed. To avoid correlation between sampling points, only every 100th sample in the simulation was used. This method of analysis combined data from three tests to determine the accuracy of each test as well as the prevalence of exposure to each food item. Bayesian latent class modelling was conducted in OpenBUGS version 3.2.3 (Spiegelhalter et al., 2007). See Appendix 10 for an example of the OpenBUGS code that was used.

4.3.1.1 Test 1 – Digital Images and Meal Descriptions

The first test consisted of data extracted from digital images and meal descriptions collected in real-time. As part of the previous study, participants captured a digital image of each meal or snack consumed between day 2 and day 8 of the study through the PhotoFoodDiary feature of the Ethica application. All participants had the option to accompany each image with an audio description of the meal or snack. Participants in groups 2 and 3 were also given an additional option to provide written descriptions of the components as well as a list of ingredients as desired.

4.3.1.2 Test 2 – Microsurveys

The second test was comprised of data extracted from microsurveys. In order to allow participants to become familiar with the Ethica Survey Tool, two practice surveys were administered on days 1 and 2 of the survey (Figure 4.1). After this brief introductory period, two microsurveys were administered on alternating days during the days of interest. Each

microsurvey consisted of up to four multiple answer questions that assessed specific food exposures of interest that may have taken place on the previous day or two. Microsurvey 1 was administered on days 3, 5, 7, and 9 and assessed recent exposures to foods such as fruits, nutrition bars, snacks, nuts, salads, sprouts and fish (Figure 4.1). Microsurvey 2 was administered on days 4, 6, 8 and 10 and assessed recent exposures to foods such as nutritional supplements, cereals, eggs and dairy products, cheese, nut butters, noodle mixes, tofu, poultry and vegetables. A time limit of 4 hours was given to complete each microsurvey (refer to Appendix 1 for microsurveys questionnaires).

4.3.1.3 Test 3 – Food History Questionnaire

The third test was comprised of information provided on food history questionnaires modeled on the Foodbook questionnaire (MacDonald, 2016) with some minor adaptations in consent wording required by local ethics and to provide appropriate context for the participants. This questionnaire required participants to recall food exposures between day 2 and day 8 of the study (midpoint – day 5) and was sent to participants on day 23 (day 5 plus 18 days) (Figure 4.1).

4.3.1.4 Priors and Assumptions

Priors on the sensitivities of the food history questionnaire for each food item were informed by published reports of past outbreak investigations. Although data on the sensitivity of food history questionnaires in the context of national outbreak investigations is not directly available, indirect measures of this information can be deduced from past outbreak reports. With the exception of situations of cross-contamination and multiple source outbreaks, when a

pathogen is isolated from a particular food item, it may be assumed that cases infected with that particular outbreak strain have consumed the contaminated product before falling ill. Therefore, the proportion of laboratory confirmed cases that misclassified their exposure status to the implicated food item serves as a proxy for the sensitivity of the food history questionnaire. Using this line of deductive reasoning, priors for the sensitivities were inferred from past outbreak reports from PHAC and the US CDC. Only published reports of outbreaks in which cases were confirmed through PFGE, MLVA or WGS and the outbreak strain was successfully isolated from the contaminated food source were used to inform priors. For a list of the publications that were used to inform priors see Appendix 11. No previous knowledge was available for the specificities of food history questionnaires. Therefore, Jeffrey's priors (non-informative) were used for specificities of the food history questionnaire in the Bayesian analysis.

In regards to data collected from digital images and meal descriptions, no previous knowledge was available for sensitivities so again Jeffrey's priors were used. Since this data was collected in real-time, it is unlikely that it contains many false positives – the priors for specificity were based on 95% certainty that the true value was above 80% with a point estimate of 95%. Due to a lack of previous knowledge and relevant literature, Jeffrey's priors were used for the sensitivity of microsurveys. Since these were completed up to 48 hours after the foods were consumed, the specificities were expected to be slightly less than those of real-time data. Priors for specificities of the microsurvey were based on 95% certainty that the true value was above 80% with a point estimate of 90% (slightly less than real-time data). The true prevalence of each exposure was unknown so Jeffrey's priors were used. The beta distributions for each of the priors was calculated using epiR beta buster (Stevenson et al., 2015). For a complete list of priors used and corresponding beta parameters see Appendix 12.

Convergence diagnosis and output analysis (CODA) was conducted to ensure that Markov chains were sampling from stationary distributions. Four diagnostic tests were conducted on the Coda of each measure to verify that convergence had been reached including Gelman and Rubin (1992), Geweke (1991), Raftery and Lewis (1992) and Heidelberger and Welch (1981). Conditional correlations were assessed by applying the conditional independent Bayesian model. Coda diagnostic tests were conducted using packages "coda" and "mcmcplots" in the R software package (R Core Team, 2013). For an example of the R code that was used to assess convergence, see Appendix 13.

4.3.1.5 Comparing findings from Bayesian Analysis to Frequentist Analysis (used in Chapter 3)

The sensitivity and specificity of test 1 and test 2 run in parallel was computed using Ausvet EpiTools epidemiological calculator to estimate the accuracy of the reference standard used in Chapter 3 (calculator available at www.epitools.ausvet.com.au/content.php?page=2Tests).

4.4 Results

A total of 51 participants recorded their food consumption through digital images/meal descriptions and microsurveys during the time period of interest and completed the food history questionnaire 2.5 weeks later. See Appendix 14 for the contingency tables of the number of participants that reported each food exposure using each of the three data collection methods.

4.4.1 Convergence and Conditional Independence

The results of the CODA diagnostic tests suggested that the Markov chains had reached their stationary distributions. The Gelman and Rubin (1992) diagnostic indicated that the

potential scale reduction were approaching 1. The dependence factor values computed by the Raftery and Lewis (1992) diagnostic all fell below 5. The p-values in the Geweke (1991) diagnostic were all greater than 0.05. Finally, the p-values in the Heidelberger and Welch (1981) diagnostic were also all greater than 0.05. The conditional correlation estimates – rhoD for sensitivities and rhoDc for specificities – were found to be near zero and did not include the value 1, indicating that the tests were conditionally independent (see Appendix 15 for a complete list of rhoD and rhoDc values).

4.4.2 Test 1 – Digital Images and Meal Descriptions

The sensitivities of the real-time self-reported data pertaining to each food item ranged from 53.0% [24.8-91.2] for cucumbers to 79.1% [64.9-92.3] for tomatoes (Table 4.1).

Table 4.1: Sensitivities and 95% credible intervals (CrI) for test 1 (digital images and meal descriptions) for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts (n=51).

	Sensitivity	95% CrI	
Tomatoes	79.1	64.9	92.3
Lettuce and Leafy Greens	69.1	51.2	90.1
Cucumbers	53.0	24.8	91.2
Berries	71.5	47.3	95.0
Nuts	69.8	37.5	99.8

The specificities of real-time self-reported data pertaining to each food item ranged from 84.3% [65.5-97.8] for berries to 90.9% [76.1-98.8] for tomatoes (Table 4.2).

Table 4.2: Specificities and 95% credible intervals (CrI) for test 1 (digital images and meal descriptions) for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts (n=51).

	Specificity (%)	95% CrI	
Tomatoes	90.9	76.1	98.8
Lettuce and Leafy Greens	90.4	74.7	98.7
Cucumbers	90.3	79.1	98.3
Berries	84.3	65.5	97.8
Nuts	88.1	72.9	98.3

4.4.3 Test 2 – Microsurveys

The sensitivities of microsurveys pertaining to each food item ranged from 52.4% [36.6-70.5] for lettuce and leafy greens to 75.1% [40.3-99.9] for cucumbers (Table 4.3).

Table 4.3: Sensitivities and 95% credible intervals (CrI) for test 2 (microsurveys) for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts (n=51).

	Sensitivity (%)	95% CI	
Tomatoes	57.0	42.8	71.2
Lettuce and Leafy Greens	52.4	36.6	70.5
Cucumbers	75.1	40.3	99.9
Berries	54.1	31.0	93.3
Nuts	65.4	37.1	98.9

The specificities of microsurveys pertaining to each food item ranged from 83.3% [71.6-93.6] for nuts to 89.3% [79.9-96.0] for cucumbers (Table 4.4).

Table 4.4: Specificities and 95% credible intervals (CrI) for test 2 (microsurveys) for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts (n=51).

	Specificity (%)	95% CrI	
Tomatoes	87.9	77.1	95.5
Lettuce and Leafy Greens	88.1	77.4	95.7
Cucumbers	89.3	79.9	96.0
Berries	88.9	79.2	95.9
Nuts	83.3	71.6	93.6

4.4.4 Test 3 – Food History Questionnaire

The sensitivities of the food history questionnaire pertaining to each food item ranged from 59.1% [55.0-63.2] for berries to 78.6% [68.1-87.8] for lettuce and leafy greens (Table 4.5).

Table 4.5: Sensitivities and 95% credible intervals (CrI) for test 3 (food history questionnaire) for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts.

	Sensitivity (%)	95% CrI	
Tomatoes	70.3	61.8	78.3
Lettuce and Leafy Greens	78.6	68.1	87.8
Cucumber	68.9	61.5	75.9
Berries	59.1	55.0	63.2
Nuts	70.7	54.4	86.4

The sensitivity for lettuce and leafy greens, 78.6% [68.1-87.8], was significantly higher than the sensitivity for berries 59.1% [55.0-83.2].

The specificities of the food history questionnaire pertaining to each food item ranged from 47.8% [0.96-99.6] for lettuce and leafy greens to 88.4% [66.9-100] for berries (Table 4.6).

Table 4.6: Specificities and 95% credible intervals (CrI) for test 3 (food history questionnaire) for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts.

	Specificity (%)	95% CrI	
Tomatoes	59.7	1.21	99.8
Lettuce and Leafy Greens	47.8	0.96	99.6
Cucumbers	72.9	51.3	99.3
Berries	88.4	66.9	100
Nuts	72.7	48.4	99.5

4.4.5 True Prevalence

The true prevalence of exposure to each food item ranged from 38.3% [15.4-69.5] for cucumbers to 92.4% [76.5-100] for tomatoes (Table 4.7).

Table 4.7: The true prevalence and 95% credible intervals (CrI) of exposure to for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts.

	Prevalence (%)	95% CrI	
Tomatoes	92.4	76.5	100
Lettuce and Leafy Greens	86.2	59.4	100
Cucumbers	38.3	15.4	69.6
Berries	54.0	25.2	77.7
Nuts	46.4	15.9	84.0

4.4.6 Comparison between Bayesian Analysis Results and Frequentist Analysis

The sensitivity and specificity of the food history questionnaire computed through Bayesian Latent class analysis are shown in Table 4.8 directly next to the corresponding frequentist results from Chapter 3.

Table 4.8: Sensitivities of food history questionnaire for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts for Bayesian approach compared to the sensitivities computed through the frequentist approach.

	Bayesian Approach Food History Questionnaire Sensitivity (%) [95%CrI]	Frequentist Approach Food History Questionnaire Sensitivity (%) [95%CI]
Tomatoes	70.3 [61.8-78.3]	77.8 [60.8-89.9]
Lettuce and Leafy Greens	78.6 [68.1-87.8]	72.2 [46.5-90.3]
Cucumbers	68.9 [61.5-75.9]	54.2 [32.8-74.4]
Berries	59.1 [55.0-63.2]	37.9 [20.7-57.7]
Nuts	70.7 [54.4-86.4]	48.5 [30.8-66.5]

A comparison between the specificities of the food history questionnaire computed through Bayesian Latent class analysis and the results of specificity results from the frequentist analysis are shown in Table 4.9.

Table 4.9: Specificities of food history questionnaire for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts for Bayesian approach (used in this chapter) compared to the sensitivities computed through the frequentist approach.

	Bayesian Approach Food History Questionnaire Specificities (%)[95%CrI]	Frequentist Approach Food History Questionnaire Specificities (%)[95%CI]
Tomatoes	59.7 [1.21-99.8]	53.3 [26.6-78.7]
Lettuce and Leafy Greens	47.8 [0.96-99.6]	21.2 [8.98-38.9]
Cucumbers	72.4 [51.3-99.3]	63.0 [42.4-80.6]
Berries	88.4 [66.9-100]	72.7 [49.8-89.3]
Nuts	72.7 [48.4-99.5]	72.2 [46.5-90.3]

The accuracy parameters of the combination of test 1 and test 2 run in parallel are shown in Table 4.10. This combination of tests represents the reference standard used in chapter 3.

Table 4.10: Sensitivities and specificities of test 1 (digital images and meal descriptions) and test 2 (microsurveys) run in parallel for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts.

	Sensitivity (%)	Specificity (%)
Tomatoes	91.0	79.9
Lettuce and Leafy Greens	85.3	79.6
Cucumbers	88.3	80.6
Berries	86.9	74.9
Nuts	89.6	73.4

A direct comparison between the true prevalances calculated above and the prevalances computer through Bayesian Latent Class analysis are shown in Table 4.11.

Table 4.11: The prevalence of exposure to for tomatoes, lettuce and leafy greens, cucumbers, berries and nuts as determined based on the food history questionnaire data compared to the true prevalence calculated using Bayesian latent class analysis.

	Bayesian Approach Prevalence (%) [95%CrI]	Apparent Prevalence Food History Questionnaire (%) (n=51)	Apparent Prevalence Image/ Descriptions (%) (n=51)	Apparent Prevalence Microsurvey Data (%) (n=51)
Tomatoes	92.4 [76.5-100]	68.6 (35/51)	74.5 (38/51)	88.2 (45/51)
Lettuce and Leafy Greens	86.2 [59.4-100]	76.5 (39/51)	60.8 (31/51)	76.5 (39/51)
Cucumbers	38.3 [15.4-69.6]	45.1 (23/51)	25.5 (13/51)	45.1 (23/51)
Berries	54.0 [25.2-77.7]	33.3 (17/51)	49.0 (25/51)	60.8 (31/51)
Nuts	46.4 [15.9-84.0]	41.2 (21/51)	39.2 (20/51)	72.5 (37/51)

4.5 Discussion

The overall objective of this study was to determine the accuracy and validity of data collected in real-time from digital images and meal descriptions, microsurveys and the food history questionnaire. By incorporating sensitivity information derived for questionnaires of previous outbreak investigations, as well as the results of each of the diagnostic tests into Bayesian latent class models, the sensitivity and specificity of each test was determined for the classification of exposures to tomatoes, cucumbers, nuts, lettuce/leafy greens and berries as well as the true prevalences of exposure.

The sensitivities of the three data collection methods were fairly similar – 53.0% to 79.1% for images/meal descriptions, 52.4% to 75.1% for microsurveys and 59.1% to 78.6% for the food history questionnaire. The range of specificities for images/meal descriptions and microsurveys were also similar to one another – 84.3% to 90.9% for images/meal descriptions and 83.3% to 89.3% for microsurveys. The specificities of the food history questionnaire were

found to vary more than the other tests, ranging from 47.8% to 88.4%. When looking at each food type individually, the credible intervals of the sensitivity of the tests were found to overlap – indicating no significant difference in the sensitivities of the tests. The credible intervals specificities of each test also overlapped for each of the five foods.

The sensitivity and specificities for individual foods were not significantly different from each other with the exception of the microsurvey data for tomatoes and lettuce/leafy greens, where the specificity was significantly higher than the sensitivity. While the credible intervals were relatively wide for most estimates of sensitivity and specificity, many of the estimates did contain useful information both on most probably value and the extent of uncertainty in the estimates even with this relatively small sample size. However, the credible intervals for the sensitivity of the food history questionnaires were narrower than for the frequentist approach outlined in Chapter 3, while the specificity intervals were wider. Two particularly notable results were the specificity estimates for tomatoes and lettuce/leafy greens based on the food history questionnaire. In both cases the specificity estimate was close to 50% with credible intervals that extended from 1% to >99%. The interpretation is that there is so much uncertainty that the risk of false positives could not be meaningfully estimated given the available data.

The true prevalence of exposure to tomatoes 92.4% [76.5-100] calculated via Bayesian Latent Class analysis was significantly higher than the true prevalence for cucumbers 38.5% [15.4-69.6]. The same observation was made using the frequentist approach in Chapter 3 where the prevalence of tomatoes was 69% [58.0-78.2] and the prevalence of cucumbers was 43% [32.0-54.2]. This finding is also in agreement of the findings of the Foodbook Report (MacDonald, 2016) in which the prevalence of exposure to tomatoes was 72.9% [66.4-77.9] in the Canadian population and the prevalence of exposure to cucumbers was 62.9% [59.0-61.7].

The lower boundary of the 95% credible intervals for the estimated true prevalence of tomatoes (76.5%) calculated using Bayesian latent class analysis was, however, greater than the apparent prevalence of consumption of tomatoes (68.6%) based on the food history questionnaire. This could be a combination of the low sensitivity of the food history questionnaire and the limited information on the specificity of the food history questionnaire for this food group. All other reported apparent prevalences fell within the credible intervals for the estimated true prevalences, suggesting the apparent and true prevalences were not significantly different from each other for the remaining tests and foods.

These results of this study provide context to the findings of the previous study (Chapter 3) that implemented a frequentist approach to quantifying recall bias. The credible intervals of the sensitivity and specificity of the food history questionnaire for each food item computed in Chapter 4 were found to overlap between the corresponding results from Chapter 3. In other words, no significant differences were found in the measures of accuracy computed by the methods of analysis.

The previously used reference standard was based on parallel interpretation of the real-time data from images/meal descriptions and the microsurveys administered with 48 hours. The results of the Bayesian analysis indicated that this combination of tests had sensitivities ranging from 85.3-91.0% and specificities ranging from 73.4-80.6%. These findings indicate that the reference test used in Chapter 3 was as suspected not a perfect gold standard. The Bayesian latent class analysis, therefore, allows for a better approximation of both sensitivity and specificity of the food history questionnaire and the true prevalence of exposure to each food item.

The Bayesian latent class analysis approach has several important advantages over the more traditional frequentist approach. This method provides a more comprehensive assessment of the uncertainty associated with sensitivity and specificity of all three diagnostic tests and provides a better assessment of the true prevalence of exposure to each food group. This served to reduce the impact of the limitations of each individual data collection method. Furthermore, by formally incorporating prior knowledge regarding measurement uncertainty, more informed estimations of sensitivities and specificities of the questionnaire were obtained for each of the food items given the limitations of the study sample size.

This approach to quantifying recall bias may have important implications on the field of outbreak investigation. By determining the sensitivity and specificity of different tools, data can be adjusted to account for errors in human memory yielding a more accurate representation of the food consumption histories. Given the pretest probability of exposure and the sensitivity and specificity of the test, the positive and negative predictive values can be determined to adjust questionnaire results for the potential recall bias.

This method does, however, require a strong sample size to ensure that the final answers are not influenced too strongly by prior information. In this example, there were only five foods of interest where there were data from all three tests and no zero cells. The estimated true prevalences had very wide confidence intervals for all foods but tomatoes emphasizing the uncertainty associated with this relatively small sample. The list of foods that were assessed in this study are a subset of the foods commonly implicated in outbreak investigations. Future studies may consider conducting similar analyses on a wider range of food items to gain a more comprehensive perspective on the effect of recall bias on dietary memory. By recruiting larger sample sizes and building on the knowledge from this work, future studies can also narrow the

confidence intervals of the estimated parameter and determine whether some food items are less susceptible to recall bias than others.

4.6 Conclusion

Bayesian latent class modelling is an effective strategy for determining the sensitivity and specificity of food history questionnaires in the absence of a true gold standard. Using this approach, recall bias can be quantified and adjusted for foodborne outbreak investigations. By incorporating prior knowledge and avoiding the need for a true gold standard, this approach allows researchers to quantify biases that would otherwise be difficult to measure. The sensitivities and specificities computed via the Bayesian approach did not differ significantly from the corresponding parameters found by using the frequentist approach used in Chapter 3. Similarly, findings regarding prevalence of exposures did not differ significantly between the two methods of analysis. This study serves as a first step to a new and innovative approach to quantifying biases of human memory to inform public health practice.

4.7 References

Branscum, A., Gardner, I., & Johnson, W. (2005). Estimation of diagnostic-test sensitivity and specificity through Bayesian modeling. *Preventive veterinary medicine, 68*(2), 145-163.

Formann, A. K., & Kohlmann, T. (1996). Latent class analysis in medical research. *Statistical methods in medical research, 5*(2), 179-211.

Gelman, A., & Rubin, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statistical science*, 457-472.

Geweke, J. (1991). *Evaluating the accuracy of sampling-based approaches to the calculation of posterior moments* (Vol. 196): Federal Reserve Bank of Minneapolis, Research Department Minneapolis, MN, USA.

Heidelberger, P., & Welch, P. D. (1981). A spectral method for confidence interval generation and run length control in simulations. *Communications of the ACM, 24*(4), 233-245.

Johnson-Kozlow, M., Matt, G. E., & Rock, C. L. (2006). Recall strategies used by respondents to complete a food frequency questionnaire: an exploratory study. *J Am Diet Assoc, 106*(3), 430-433. doi:10.1016/j.jada.2005.12.007

MacDonald, D. (2016). *Foodbook: The Canadian Food, Water and Animal Exposure Study.* Paper presented at the IAFP 2016 Annual Meeting.

R Core Team. (2013). R: a language and environment for statistical computing. *R Foundation for Statistical Computing*.

Raftery, A. E., & Lewis, S. M. (1992). [Practical Markov Chain Monte Carlo]: comment: one long run with diagnostics: implementation strategies for Markov Chain Monte Carlo. *Statistical science, 7*(4), 493-497.

Rindskopf, D., & Rindskopf, W. (1986). The value of latent class analysis in medical diagnosis. *Statistics in medicine, 5*(1), 21-27.

Schumacher, S. G., van Smeden, M., Dendukuri, N., Joseph, L., Nicol, M. P., Pai, M., & Zar, H. J. (2016). Diagnostic test accuracy in childhood pulmonary tuberculosis: a Bayesian latent class analysis. *American journal of epidemiology*.

Smith, A. (1993). Cognitive psychological issues of relevance to the validity of dietary reports. *European journal of clinical nutrition, 47*, S6-18.

Sokal, A. D. (1989). Monte Carlo methods in statistical mechanics: foundations and new algorithms: Troisieme cycle de la physique en Suisse Romande.

Spiegelhalter, D., Thomas, A., Best, N., & Lunn, D. (2007). OpenBUGS user manual, version 3.0. 2. *MRC Biostatistics Unit, Cambridge*.

Stevenson, M., Nunes, T., Heuer, C., Marshall, J., Sanchez, J., Thornton, R., Reiczigel, J., Robison-Cox, J., Sebastiani, P., & Solymos, P. (2015). epiR: Tools for the analysis of epidemiological data. vol R package version 0.9–62: R.

van Smeden, M., Naaktgeboren, C. A., Reitsma, J. B., Moons, K. G., & de Groot, J. A. (2013). Latent class models in diagnostic studies when there is no reference standard—a systematic review. *American journal of epidemiology*, kwt286.

Wirfält, E. (1998). Cognitive aspects of dietary assessment. *Food & Nutrition Research, 42*, 56-59.

CHAPTER 5 – CONCLUSION

It has long been recognized that limitations in human memory introduce errors and biases into the data collected during outbreak investigations (Decker et al., 1986; Mann, 1981). More specifically, as cases misclassify their own exposure status, the power of epidemiological tools to detect meaningful associations between exposures and the development of illness decreases significantly. This may increase the time it takes to identify the source of the outbreak and result in more cases of illness. While these problems are widely acknowledged (Dechet et al., 2006; Schenkel et al., 2006), the studies presented here are among the first to directly measure the implications of recall bias and provide a way by which to overcome the challenge. In doing so, various other important barriers to outbreak surveillance were also addressed including collecting accurate data on the incidence of foodborne illness and food histories for affected individuals.

Bayesian latent class analysis can provide estimates of the accuracy of different test and surveillance methods, even when gold standards do not exist. Knowledge of the measures of accuracy of a test allow researchers and public health practitioners to convert apparent prevalances to true prevalances, as described by Rogan and Gladen (1978). In the context of this study, accuracy of data collected from food history questionnaires reflected the effect that recall bias had on participants' abilities to remember different food exposures. When accounting for the sensitivity and specificity of the questionnaire used, the true prevalence was found to slightly decrease from the apparent prevalence reported from food history questionnaires. This additional step of calibrating data analysis methods to account for errors in human memory is expected to yield more accurate representation of the food consumption histories and allow public health officials to generate more informed hypotheses about the source of outbreaks.

5.1 Summary of Key Findings

A combination of innovative approaches and technological tools has provided new insights into the surveillance and investigation of foodborne illness. Ninety-six students at the University of Saskatchewan were recruited and asked to report their food consumption over a 10-week period. These studies have provided information on the frequency of symptoms of enteric illness and the food consumption habits of a previously understudied segment of the Canadian population. In addition, the three studies collected comprehensive food consumption histories, quantified recall bias and provided a method of accounting for such bias in outbreak investigations.

Chapter 2 focused on obtaining an accurate and comprehensive account of self-reported enteric illness and tracking food consumption over the course of 10 weeks. The option to trigger foodborne illness surveys at the touch of a button gave participants a quick and convenient way to report information on the occurrence of illness. Ethica provided a feasible way to collect detailed information from a large group of people over a long period of time. Digital images provided information about food consumption behaviours and an accurate and objective account of each meal. Meal descriptions allowed participants to provide information such as specific ingredients and other details that may not be evident in the images. Microsurveys helped to ensure that even foods missed in pictures and descriptions, were still recorded by prompting participants for responses.

When asked the same questions on food safety knowledge at three different time points, between 17% and 39% of respondents changed their response between surveys. Thus, illustrating the room for improvement in regards to food safety knowledge among the millennial generation, and the willingness to learn more about food safety, as supported by findings from focus group discussions. Participants were found to be reasonably compliant with data collection protocols;

for example, during the first 10 days of the study 100% (96/96) of participants completed at least 1 food microsurvey per day, 85.4% (82/96) submitted at least one digital image per day and 68.4% (54/79) provided at least one meal description per day. Overall, the convenience of the application in tracking what a person eats with minimal disruption to everyday routines, and the vast amounts of data that were reported made this an effective and feasible strategy for capturing food histories.

In Chapter 3 the real-time data collected through digital images, meal descriptions, and microsurveys were used as a reference to measure the sensitivity and specificity of traditional 7 or 18-day (2.5 weeks) food history questionnaires. The validity of food history data collected after 7 days was found to be consequentially low with the validity of data varying by food type. Sensitivities ranged from 14.3% for sprouts to 100% for leafy greens while specificities ranged from 30.4% for beef to 80.4% for peanuts. The findings were similar for the questionnaires administered after 18 days; sensitivities ranged from 15.8% for sprouts to 77.8% for tomatoes and specificities ranged from 21.2% for leafy greens to 92.1% for melons. Similarly, agreement between data collected in real-time and data collected from food history questionnaires did not differ between the time intervals for most food items, with leafy greens being the lone exception. Dietary recall among participants who reported enteric illness during the study did not differ from that of individuals who experienced no such symptoms. Foods exposures that were assessed on more than one occasion (i.e. on microsurveys and again on food history questionnaires) were also not found to be more sensitive or specific than food exposures that were assessed only once (on only food history questionnaires). Interestingly, having been previously asked about a food exposure decreased the specificity and increased the likelihood of reporting a false positive – the rate of false negatives was unaffected. The key findings of this chapter were that dietary memory

has comparably low accuracy after 7 and 18 days and that the magnitude of the effect that recall bias had on the accuracy of dietary memory was not uniform across food types.

In Chapter 4 Bayesian analysis was applied to determine the specificity and sensitivity of food history questionnaire data collected at 2.5 weeks and data from the Ethica app (self-reported images with descriptions and time-triggered microsurveys) in the absence of a true gold standard. The sensitivities and specificities of all three tests for 5 different foods were determined using this method. The findings of this study suggest that data collected from digital images and descriptions alone did not provide a true gold standard of dietary history. Similarly, data collected from microsurveys alone does not serve as a true gold standard. While neither method is perfect, the combination of these strategies interpreted in parallel provided a relatively sensitive assessment of food exposures. Overall, the sensitivities and specificities of the food history questionnaire calculated using Bayesian analysis did not significantly differ from those calculated in Chapter 3. Bayesian latent class modelling proved as an effective strategy for determining the sensitivities and specificities of each of the data collection methods. In this sample, the true prevalences based on the Bayesian analysis adjusted for test sensitivities and specificities and sample size were not significantly different from the apparent prevalences with one exception. The Bayesian estimate of the true prevalence for exposure to tomatoes was higher than that based on the food history questionnaire.

5.2 The Broader Context of these Findings

By directly measuring the symptoms of enteric illness, these studies provided a comparison for previous estimates of foodborne illness in Canada. Among the 96 study participants, 34% reported having experienced an episode of vomiting or nausea and 29% reported an episode of diarrhea that was not believed to be caused by alcohol consumption in a

10-week period. These findings are considerably higher than previous estimates of 12.5% - 1 in 8 per year (Thomas et al., 2013). Differences in the case definition of acute diarrheal illness may contribute to this discrepancy. The study by Thomas et al., 2013 defined cases of acute diarrheal illness as having ≥ 3 loose stools in 24 hours with duration lasting > 1 day. In the present feasibility study (Chapter 2), participants were simply asked to report the occurrence of symptoms. Including information such as duration, frequency, and severity on app questionnaires in future studies will allow for more direct comparisons to be made.

Accurately measuring the percentage of cases of acute foodborne illness that seek medical care is an important component to measuring the extent and implications of foodborne illness in Canada. In this study, 7% of individuals who described enteric symptoms on the user-triggered or time-triggered feature reported their symptoms to a health practitioner. Previous estimates of the percentage of cases of foodborne illness that seek medical care include 14% (Thomas et al., 2013), 18% (Scallan et al., 2011) and 22% (Sargeant et al., 2008). Discrepancies in these estimates may be attributed to differences in the definition of enteric illness that was used in each study and the sensitivities of methods used to detect cases. Severe cases of foodborne illness are more likely to seek medical care than mild or moderate cases (Scallan et al., 2006). As a result, surveillance methods that utilize only data from health records often fail to capture less severe cases of illness.

The sensitivity of food history questionnaires has rarely been directly studied. Mann (1981) and Decker et al. (1986) investigated dietary recall after 3 and 5 days, respectively, in the context of local point source outbreak investigations. Both studies observed individuals as they ate at a buffet dinner and measured the accuracy of dietary recall of a few specific food items. Both studies found sensitivities ranging from 81.2% to 95.2% and specificities ranging from

75% to 98.5%. The analysis conducted in Chapter 3 included a greater time lag between exposure and the questionnaire and a larger number of food items. The effect of recall bias of dietary memory was found to vary considerably with food type. Therefore, the broader range of sensitivities (14.3-100% after 7 days and 15.8-77.8% after 18 days) computed in Chapter 3 was to be expected. The same can be said for the specificities which ranged from 30.4-80.4% after 7 days and 21.2-92.1% after 18 days.

The variety of food preferences and risk-benefit perceptions that characterize the Canadian population make it challenging to create data collection strategies that could be applied across contexts and situations. New technologies such as Ethica will play an important role in collecting and analyzing public health intelligence and identifying risk factors to prevent disease. This technology has the potential to collect detailed health related data from sentinel surveillance cohorts in a feasible and sustainable manner. Particular advantages of the app include the wide range of options to collect data, the low burden placed on the user, the option to prompt and remind users throughout the data collection process, and the ability to collect data without regular access to the Internet. These advantages create new opportunities for targeted surveillance strategies to better understand health related behaviours.

5.3 Limitations

Although considerable efforts were made and different data collection methods were utilized, limitations still exist in the methodology and analysis of the study. Because compliance was not perfect, Bayesian estimates of sensitivity and specificity suggest that some participants may have failed to report a particular food exposure in digital images, meal descriptions and microsurveys, but remembered the exposure when later completing the food history questionnaire. The resulting false negatives within the collection of real-time data could have

caused a decrease in the specificities for the retrospective questionnaire calculated in Chapter 3. It is considered less likely that participants would falsely report having consumed a food item shortly after the meal was consumed. These limitations were examined by implementing different data analysis methodologies to calculate sensitivities and specificities in Chapters 3 and 4. Chapter 3 used real-time data as a reference to measure the accuracy of data collected on food history questionnaires. Chapter 4 implemented Bayesian latent class modelling to analyze self-reported data, microsurvey results, and food history questionnaire results to calculate sensitivities and specificities in the absence of a true gold standard. The results of the two different approaches were not significantly different.

While the true prevalence of exposure was estimated in Chapter 4 after adjusting for errors associate with sensitivity and specificity, the relatively small sample size limited the analysis to five foods and resulted in relatively wide credible intervals. Differences were found between the frequency of food exposures observed in this study and those published in the national Foodbook report Canada (MacDonald, 2016). These differences remained even after adjusting for age, province and time of year for 9 of 15 foods that were investigated. This may be attributed to differences in the method of data collection, but also emphasizes the importance that social, cultural and other demographics that characterize sub-populations have in outbreak investigations.

5.4 Tying It All Together

To solve a challenge that has always faced the field of foodborne illness investigation, these studies used a combination of new tools, advanced data analysis methodologies and creative innovations. A comprehensive record of the food consumption of each participant was collected using Ethica smartphone technology. The effect of recall bias on dietary history was quantified

by using data collected in real-time as a reference standard to measure the accuracy of data collected from food history questionnaires administer later in the study. Bayesian latent class analysis was conducted to measure the sensitivity and specificity of food history questionnaires in the absence of a true gold standard. By adjusting for sensitivity and specificity, an example of how data analysis methods can then be calibrated to account for the effect of recall bias was illustrated. This three step process– collect, quantify and mitigate – guided the process of quantifying and mitigate recall bias in foodborne outbreak investigations. To ensure that the results of these studies would have practical value for public health officials in Canada, the questionnaires and time intervals used were designed to resemble a range of plausible local, provincial and national enteric outbreak investigations conducted. It is important to note that during outbreak investigations, data collected from food history questionnaires is interpreted in combination with information gathered from food testing, food/environmental investigations and laboratory results. These other sources of evidence can be used to verify and complement food history data. Decisions regarding public health action are guided by the weight of evidence approach that incorporates each piece of available information (Vik et al., 2014).

Future research into the effects of recall bias at different points in time and on different commonly implicated foods is warranted. The list of foods commonly implicated in outbreak investigations that were assessed in this study is by no means meant to be comprehensive. By including a wider range of food items as well as a larger sample size, more comprehensive perspective may be gained on the effect of recall bias on dietary memory. Such investigations may unearth trends and insights into why the recollection of some food exposures is better than others. Capturing information from grocery receipts may be a beneficial strategy that can be incorporated into these types of studies. Although images of such receipts would not guarantee

that the individual consumed all foods that were purchased, this strategy has the potential to verify food history reports. Furthermore, this type of data may be particularly useful in prompting participants' memories during outbreak investigations that involve long windows of time.

As technologies and methodologies continue to advance, so will opportunities for the acquisition of data on human behavior. For example, the Ethica app has itself advanced substantially since the time of this study and is now available for iPhones as well as Android phones. By harnessing these new technologies new data collection and analysis tools can continue to be developed to allow for more accurate representations of diet history. Several potential applications of the Ethica smartphone technology were brought up during focus group discussions including the identification of specific food allergies as well as risk factors and foodborne illness hotspots.

While a true gold standard of dietary histories may never be attained, data analysis methods such as Bayesian latent class analysis can be used to determine sensitivities and specificities of a wide range of epidemiological tests and estimate true prevalence of foodborne exposures after correcting for recall bias. Formal incorporation of prior knowledge regarding measurement uncertainty may offer a means by which to account for, rather than aim to eliminate, recall bias in foodborne outbreak investigations. It is important to note that these studies serve as a first step to a new and innovative approach to accounting for errors in human memory during epidemiological investigations. Through multidisciplinary approaches, advances in areas of psychology can continue to be applied to epidemiological studies to inform best practices in public health.

5.5 References

Dechet, A. M., Scallan, E., Gensheimer, K., Hoekstra, R., Gunderman-King, J., Lockett, J., Wrigley, D., Chege, W., Sobel, J., & Group, M. W. (2006). Outbreak of multidrug-resistant Salmonella enterica serotype Typhimurium definitive type 104 infection linked to commercial ground beef, northeastern United States, 2003–2004. *Clinical infectious diseases*, *42*(6), 747-752.

Decker, M., Booth, A., Dewey, M., Fricker, R., Hutcheson, R., & Schaffner, W. (1986). Validity of food consumption histories in a foodborne outbreak investigation. *American journal of epidemiology*, *124*(5), 859-863.

MacDonald, D. (2016). *Foodbook: The Canadian Food, Water and Animal Exposure Study*. Paper presented at the IAFP 2016 Annual Meeting.

Mann, J. M. (1981). A prospective study of response error in food history questionnaires: implications for foodborne outbreak investigation. *American journal of public health*, *71*(12), 1362-1366.

Rogan, W. J., & Gladen, B. (1978). Estimating prevalence from the results of a screening test. *American journal of epidemiology*, *107*(1), 71-76.

Sargeant, J., Majowicz, S., & Snelgrove, J. (2008). The burden of acute gastrointestinal illness in Ontario, Canada, 2005–2006. *Epidemiology and Infection*, *136*(04), 451-460.

Scallan, E., Hoekstra, R. M., Angulo, F. J., Tauxe, R. V., Widdowson, M.-A., Roy, S. L., Jones, J. L., & Griffin, P. M. (2011). Foodborne illness acquired in the United States—major pathogens. *Emerg Infect Dis*, *17*(1).

Scallan, E., Jones, T. F., Cronquist, A., Thomas, S., Frenzen, P., Hoefer, D., Medus, C., & Angulo, F. J. (2006). Factors associated with seeking medical care and submitting a stool sample in estimating the burden of foodborne illness. *Foodbourne Pathogens & Disease*, *3*(4), 432-438.

Schenkel, K., Bremer, V., Grabe, C., Van Treeck, U., Schreier, E., Höhne, M., Ammon, A., & Alpers, K. (2006). Outbreak of hepatitis A in two federal states of Germany: bakery products as vehicle of infection. *Epidemiology and Infection*, *134*(06), 1292-1298.

Thomas, M. K., Murray, R., Flockhart, L., Pintar, K., Pollari, F., Fazil, A., Nesbitt, A., & Marshall, B. (2013). Estimates of the burden of foodborne illness in Canada for 30 specified pathogens and unspecified agents, circa 2006. *Foodborne pathogens and disease*, *10*(7), 639-648.

Vik, J., Hexemer, A., & Farber, J. (2014). Summary: Weight of evidence-Factors to consider when investigating a food-borne illness outbreak. *Canada Communicable Disease Report*, *40*(14), 303.

APPENDICES

Appendix 1: Microsurvey Questionnaires (Type 1 and Type 2)

Microsurvey Type 1

Survey A **8:00**
Breakfast **Time sent to smartphone:** **AM**

You are invited to participate in the following study. Thank-you for considering to support this research to investigate new technology that we believe will help to improve current methods of understanding enteric illness and foodborne diseases.
By completing and submitting the questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the conditions of participation in this study.

Food Group: Fruit

1 Q1 On _____, Did you eat any servings of fresh fruit?
Please do not include frozen, canned/jarred or dried fruits.
Please do include any fruits in salads, smoothies, desserts or as a garnish.

Q2 What fruit (s)

- Bananas
- Mangoes
- Kiwi
- Pomegranate (not powdered)
- Pineapple
- Avocado (including guacamole)
- Olives
- Melons
- Berries
- Unpasteurized fruit juice
- Fruit Smoothies
- Pre-cut fruit mix or fruit platter

Survey B Lunch **Time sent to smartphone:** **12:45 PM**

Food Group: Nutrition bars

Q1: On _____, did you eat any nutrition bars for a snack or meal 2 replacement?
- granola bar
- energy bar
- protein bar
- other

Food Group: Meal replacements

Q2: On _____, did you eat any meal replacements?
- fruit based drink
- vegetable based drink
- fruit + vegetable
- protein powder
- dairy based drink
- full meal replacement (e.g Boost, Ensure)

Food Group: Snacks

Q3: On _____, did you eat any convenience snacks when you were hungry?
- chips, nachos, popcorn or pretzels
- chocolate or chocolate containing candy
- non chocolate candy
- pizza or pizza pops.
- dips (e.g. salsa, hummus, prepared or store-bought)

Food Group: Nuts

Q4: On _____, did you eat any nuts as snacks when you were hungry (whole, pieces or ground nuts)?
- Almonds
- Cashews
- Peanuts (not including peanut butter)
- Walnuts
- Pecans
- Hazelnuts (filberts)

Survey C Supper **Time sent to smartphone:** **6:30 PM**

Food Group: Salads

3 Q1: On _____ did you eat any store bought or homemade salads?

Lettuce or leafy greens salad
with tomatoes?
Coleslaw
Egg salad (as own dish or prepared on sandwich or pita)
Potato salad
Pasta salad
Fruit salad/pre-cut
Fish or seafood salad (e.g tuna, mixed seafood)

Q2: Did you use a salad dressing?

processed salad dressing
mayonnaise-containing homemade salad dressing
oil and/or vinegar-containing homemade salad dressing

Food Group: Sprouts

Q3: On _____, did you eat any sprouts?

alfalfa
mung bean
onion
radish
mustard
broccoli

Did you cook the sprouts?

Food Group: Fish

Q4: On _____ did you eat any fish products?

finned fish (e.g. tuna, bass, sole, cod, salmon, trout, pickerel, perch, jackfish, whitefish etc.)

Q5: what type of product?

canned
fresh fish (caught in wild or bought fresh, not frozen)
smoked
salted
frozen fish fillets
raw (e.g. sushi, sashimi, ceviche, tartare)

Thank you for your time and contribution

Microsurvey Type 2

Survey A Breakfast

Time sent to smartphone: **8:00 AM**

You are invited to participate in the following study. Thank-you for considering to support this research to investigate new technology that we believe will help to improve current methods of understanding enteric illness and foodborne diseases.

By completing and submitting the questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the conditions of participation in this study.

Food Group: Supplements

Q1: On MONDAY did you eat any dietary or nutritional supplements?

4

What type of supplement?

Chia seeds or powder
Hemp seeds or hearts
Whey powder
Protein mix
Flax seeds (whole or ground)
Herbs (dried or fresh)

Food Group: Cereals

Q2: On _____ did you eat cereal for breakfast?

Cold breakfast cereal
Hot breakfast cereal (e.g oatmeal, cream of wheat, porridge)

Food Group: Eggs and Dairy

Q3: On _____ did you eat eggs or dairy products for breakfast?

Any eggs (e.g. scrambled eggs, omelets, put into drinks or homemade salad dressing)

raw or undercooked (runny or over-easy) eggs

Dairy milk (Pasteurized or Unpasteurized (raw)dairy milk

Any dairy substitutes or non-dairy milk (e.g soy, almond, coconut, oat, or rice milk)

Survey B Lunch **Time sent to smartphone:** **12:45 PM**

Food Group: Cheese

5 Q1: On _____ did you eat any cheese products?

blue-veined cheese (e.g. gorgonzola, blue cheese)

fresh cheese (e.g , feta, buffalo mozzarella sold in water, goat or sheep cheese)

processed cheese (e.g. sliced, string or tube cheese)

hard cheese (e.g cheddar, gouda, mozzarella, Monterey jack)

soft cheese (e.g. camembert)

cottage cheese or ricotta

parmesan cheese (fresh grated or as dried grated product)

Food Group: Sandwiches

Q2: On _____ did you prepare any sandwiches, pitas, crepes, snacks or drinks using peanut butter or other nut pastes or spreads?

peanut butter

almond butter

chocolate hazelnut spread

Q3: did you sweeten it with honey or other sugar replacements

honey

stevia

prepared jam

Q4: On _____ did you eat any dried soup or noodle mixes?

purchased as prepared

made it myself (e.g. instant rice noodles)

Q5: On _____ did you eat any tofu or tofu-containing products?

tofu purchased in package

meal or snack containing tofu

Survey C Supper **Time sent to smartphone:** **6:30 PM**

Food Group: Poultry

6 Q1: In _____, did you eat any poultry products?

chicken
turkey
duck
partridge
goose

Q2: what type of product?

fast food purchase or made at home from frozen ready-cook meal (e.g. chicken nuggets/strips)
ground meat
pieces (e.g thighs, breast, neck, wings)
whole
deli meat
used to make soup

Food Group: Vegetables

Q3: On _____ did you eat any fresh, frozen or canned vegetables?

tomatoes
peppers
carrots
broccoli
cauliflower
onions
leeks
celery
mushrooms
cucumber
peas
beans
corn
cabbage
vegetable juices
fresh squeezed or processed

Thank you for your time and contribution

Appendix 2: Selection process for Foods of Interest

Eligible Foods (Relative frequency = 20-80%)*	Visually Distinguishable	Literature Review	Selected for Study
Any lettuce or leafy greens	Yes	Yes	Yes
Poultry	Yes	Yes	No
Whole, pieces or ground nuts	Yes	Yes	Yes
Sprouts	Yes	Yes	Yes
Cabbage(including coleslaw)	Yes	Yes	Yes
Berries	Yes	Yes	Yes
Cucumber	Yes	Yes	Yes
Melons	Yes	Yes	Yes
Peanut butter	Yes	Yes	Yes
Peanuts (not including peanut butter)	Yes	Yes	Yes
Any tomatoes	Yes	Yes	Yes
Breaded Chicken	Yes	Yes	Yes
Any cheese products	Yes	Yes	-
Mushrooms	Yes	-	-
Peas (shelled or in pods)	Yes	-	-
Broccoli	Yes	-	-
Onions	Yes	-	-
Nutrition bars	Yes	-	-
Peppers	Yes	-	-
Carrots	Yes	-	-
Bananas	Yes	-	-
Avocado (including guacamole)	Yes	-	-
Breakfast cereal	Yes	-	-
Cold breakfast cereal	Yes	-	-
Convenience snacks	Yes	-	-
Chips, nachos, popcorn, pretzels	Yes	-	-
Chocolate	Yes	-	-
Non-chocolate candy	Yes	-	-
Almonds	Yes	-	-
Dried soup or noodle mixes	-	-	-
Dairy milk	-	-	-
Dairy substitutes or non-dairy milk	-	-	-
Hot breakfast cereal	-	-	-
Parmesan cheese	-	-	-
Fruit smoothies	-	-	-
Chicken (not including deli-meat)	-	-	-
Chicken pieces or parts in soups, or as part of a dish, not including deli-meat)	-	-	-

Frozen fish (e.g. breaded or non-breaded fillets of cod, haddock, sole fish, basa, tilapia)	-	-	-
Any dietary or nutritional supplements (e.g. meal replacements, protein powder, vitamins, herbs)	-	-	-
Frozen vegetables	-	-	-
Finned-fish	-	-	-

*Based on Food History Questionnaire

Appendix 3: Food Safety Survey 1

You are invited to participate in the following study. Thank-you for considering to support this research to investigate new technology that we believe will help to improve current methods of understanding enteric illness and foodborne diseases.

By completing and submitting the questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the conditions of participation in this study.

Set 1

Q1. Did you eat in a restaurant today?

Yes

No

Q2. Did you eat take out / ready to eat food today?

Yes

No

Q3. Do you eat purchased ready to eat foods and store them at room temperature for more than 1 hour before eating?

Yes

No

Q4. Do you take left overs home from restaurants to eat later?

Yes

No

Thank you for your time and contribution

Appendix 4: Food Safety Survey 2

You are invited to participate in the following study. Thank-you for considering to support this research to investigate new technology that we believe will help to improve current methods of understanding enteric illness and foodborne diseases.

By completing and submitting the questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the conditions of participation in this study.

Q1. Did you eat organic food today?

Yes
No
Not sure

Q2. Did you eat food identified as being locally grown today?

Yes
No
Not sure

Q3. Did you eat food identified as being raised without antibiotics today?

Yes
No
Not sure

Q4. Did you eat food identified as being raised without steroids/hormones today?

Yes
No
Not sure

Q5. Did you feel sick in the last week?
Branch if yes / If no then done

5a. Did you feel nauseous or vomit in the last week

5b. Did you have diarrhea in the last week?
Branch if yes to either / If no then done

5c. Did you consult a health care professional regarding this illness?

5d. Do you suspect your illness might be related to consumption of alcoholic beverages?

Thank you for your time and contribution

Appendix 5: Food Safety Survey 3

You are invited to participate in the following study. Thank-you for considering to support this research to investigate new technology that we believe will help to improve current methods of understanding enteric illness and foodborne diseases.

By completing and submitting the questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the conditions of participation in this study.

Q1. Are organic foods safer than foods from conventional production systems?

Yes
No
Not sure

Q2. Are locally grown foods safer than foods from commercial production systems?

Yes
No
Not sure

Thank you for your time and contribution

Appendix 6: Food Safety Survey 4

You are invited to participate in the following study. Thank-you for considering to support this research to investigate new technology that we believe will help to improve current methods of understanding enteric illness and foodborne diseases.

By completing and submitting the questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the conditions of participation in this study.

Q1. Washing a kitchen sponge with soap will get rid of all the bacteria.
TRUE
FALSE

Q2. Raw meat should be washed in the sink before cooking.
TRUE
FALSE

Q3. Food should be allowed to sit on the counter before putting it in the fridge for storage.
TRUE
FALSE

Appendix 7: Food Safety Survey 5

You are invited to participate in the following study. Thank-you for considering to support this research to investigate new technology that we believe will help to improve current methods of understanding enteric illness and foodborne diseases.

By completing and submitting the questionnaire, YOUR FREE AND INFORMED CONSENT IS IMPLIED and indicates that you understand the conditions of participation in this study.

Q1. Prebagged "ready to eat" greens should not be washed again before eating.

TRUE

FALSE

Q2. Hamburgers should be cooked until they are "piping hot" and there is no visible pink left.

TRUE

FALSE

Q3. It is important to use a thermometer to check the temperature of meat before serving, but you don't need a thermometer for precooked leftovers.

TRUE

FALSE

Thank you for your time and contribution

Appendix 8: User-triggered Foodborne Illness Survey

Did you consult a health-care professional regarding this illness?

Did you suspect your illness might be related to consumption of alcoholic beverages?

I've got…

- Diarrhea
- Vomiting
- Abdominal pain and craps
- Nausea
- Fever
- Other

Thank you for your time and contribution

Appendix 9: Time-triggered Foodborne Illness Questions

Did you feel sick in the last week?

Branch if yes / If no then done

4a. Did you feel nauseous or vomit in the last week

4b. Did you have diarrhea in the last week?

Branch if yes to either / If no then done

4c. Did you consult a health care professional regarding this illness?

4d. Do you suspect your illness might be related to consumption of alcoholic beverages?

Thank you for your time and contribution

Appendix 10: Example of OpenBUGS code for 3 Tests and 1 Population

Berries:

```
model;
# FBQ = foodbook , MicS=microsurvey, SelfR = self reported data
{
# Multinomial Model
y[1:2, 1:2, 1:2] ~ dmulti(p[1:2, 1:2, 1:2], n)
# Prob all 3 tests positive
p[1, 1,1] <- pi*SeFBQ*(SeSelfR*SeMicS+covDp) + (1-pi)*(1-SpFBQ)*((1-SpSelfR)*(1-SpMicS)+covDn)
# SelfR+, MicS-,  FBQ+
p[1,2,1] <- pi*SeFBQ*(SeSelfR*(1-SeMicS)-covDp) + (1-pi)*(1-SpFBQ)*((1-SpSelfR)*SpMicS-covDn)
# SelfR+,MicS+, FBQ-
p[1,1,2] <- pi*(1-SeFBQ)*(SeSelfR*SeMicS+covDp) + (1-pi)*SpFBQ*((1-SpSelfR)*(1-SpMicS)+covDn)
# SelfR+,MicS-, FBQ-
p[1,2,2] <- pi*(1-SeFBQ)*(SeSelfR*(1-SeMicS)-covDp) + (1-pi)*SpFBQ*((1-SpSelfR)*SpMicS-covDn)
# SelfR-,MicS+, FBQ+
p[2,1,1] <- pi*SeFBQ*((1-SeSelfR)*SeMicS-covDp) + (1-pi)*(1-SpFBQ)*(SpSelfR*(1-SpMicS)-covDn)
# SelfR-, MicS-,  FBQ+
p[2,2,1] <- pi*SeFBQ*((1-SeSelfR)*(1-SeMicS)+covDp) + (1-pi)*(1-SpFBQ)*(SpSelfR*SpMicS+covDn)
# SelfR-,MicS+, FBQ-
p[2,1,2] <- pi*(1-SeFBQ)*((1-SeSelfR)*SeMicS-covDp) + (1-pi)*SpFBQ*(SpSelfR*(1-SpMicS)-covDn)
# SelfR-,MicS-, FBQ-
p[2,2,2] <- pi*(1-SeFBQ)*((1-SeSelfR)*(1-SeMicS)+covDp) + (1-pi)*SpFBQ*(SpSelfR*SpMicS+covDn)

# Constraints for covDp and covDn
ls <- (SeSelfR-1)*(1-SeMicS)
us <- min(SeSelfR,SeMicS) - SeSelfR*SeMicS
lc <- (SpSelfR-1)*(1-SpMicS)
```

```
uc <- min(SpSelfR,SpMicS) - SpSelfR*SpMicS


# Correlations between Ethica tests, conditional on food status
rhoD <- covDp / sqrt(SeSelfR*(1-SeSelfR)*SeMicS*(1-SeMicS))
rhoDc <- covDn / sqrt(SpSelfR*(1-SpSelfR)*SpMicS*(1-SpMicS))


# Priors
pi ~ dbeta(0.5, 0.5) ## Jeffrey's Prior
SeSelfR ~ dbeta(0.5, 0.5) ## Jeffrey's Prior
SpSelfR ~ dbeta(21.2019, 2.0633) ## Mode=0.95, 95% sure >0.80
SeMicS ~ dbeta(0.5, 0.5) ## Jeffrey's Prior
SpMicS ~ dbeta(42.5732, 5.6192) ## Mode=0.95, 90% sure >0.80
SeFBQ ~ dbeta(52.4511, 22.5292) ## Mode=0.705, 95% sure > 0.61
SpFBQ ~ dbeta(0.5, 0.5) ## Jeffrey's Prior
covDn ~ dunif(lc, uc)
covDn ~ dunif(lc, uc)
covDp ~ dunif(ls, us)
}
# Data
#     (test1=+,test2=+,test3=+), (+,+,-), (+,-,+), (+,-,-), (-,+,+), (-,+,-), (-,-,+), (-,-,-)#
#     (SelfR=+,MicS=+,FBQ=+), (+,+,-), (+,-,+), (+,-,-), (-,+,+), (-,+,-), (-,-,+), (-,-,-)#
list(n=51, y=structure(.Data=c(6,4,4,11,4,3,3,16),.Dim=c(2,2,2)))
END
#  Initial Values
list(pi=0.75, SeSelfR=0.75, SpSelfR=0.75, SeMicS=0.75, SpMicS=0.75, SeFBQ=0.75, SpFBQ=0.75)
list(pi=0.7, SeSelfR=0.7, SpSelfR=0.7, SeMicS=0.7, SpMicS=0.7, SeFBQ=0.7, SpFBQ=0.7)
list(pi=0.65, SeSelfR=0.65, SpSelfR=0.65, SeMicS=0.65, SpMicS=0.65, SeFBQ=0.65, SpFBQ=0.65)
```

Appendix 11: Priors for sensitivities and specificities used for Bayesian Latent Class Modelling.

	Food Item	Prior for Sensitivity (%)	Prior for Minimum Sensitivity (%)	Beta Distribution
Food History Questionnaire	Tomatoes	70.5	61.0	52.4511, 22.5292
	Lettuce and Leafy Greens	79.7	64.9	24.6045, 7.0122
	Cucumbers	68.7	62.0	100.0694, 46.1364
	Berries	59.5	56.0	326.0973, 222.2847
	Nuts	75.0	70.0	23.3398, 5.1763
Microsurveys	Tomatoes	Jeffrey's Proir		0.5, 0.5
	Lettuce and Leafy Greens	Jeffrey's Proir		0.5, 0.5
	Cucumbers	Jeffrey's Proir		0.5, 0.5
	Berries	Jeffrey's Proir		0.5, 0.5
	Nuts	Jeffrey's Proir		0.5, 0.5
Real-Time Data (Images and Descriptions)	Tomatoes	Jeffrey's Proir		0.5, 0.5
	Lettuce and Leafy Greens	Jeffrey's Proir		0.5, 0.5
	Cucumbers	Jeffrey's Proir		0.5, 0.5
	Berries	Jeffrey's Proir		0.5, 0.5
	Nuts	Jeffrey's Proir		0.5, 0.5

*pi – uninformative Jeffrey's prior (0.5, 0.5)

	Food Item	Prior for Specificity (%)	Prior for Minimum Specificity (%)	Beta Distribution
Food History Questionnaire	Tomatoes	Jeffrey's Prior		0.5, 0.5
	Lettuce and Leafy Greens	Jeffrey's Prior		0.5, 0.5
	Cucumbers	Jeffrey's Prior		0.5, 0.5
	Berries	Jeffrey's Prior		0.5, 0.5
	Nuts	Jeffrey's Prior		0.5, 0.5
Microsurveys	Tomatoes	95.0	80.0	42.5732, 5.6192
	Lettuce and Leafy Greens	95.0	80.0	42.5732, 5.6192
	Cucumbers	95.0	80.0	42.5732, 5.6192
	Berries	95.0	80.0	42.5732, 5.6192
	Nuts	95.0	80.0	42.5732, 5.6192
Real-Time Data (Images and Descriptions)	Tomatoes	90.0	80.0	21.2019, 2.0633
	Lettuce and Leafy Greens	90.0	80.0	21.2019, 2.0633
	Cucumbers	90.0	80.0	21.2019, 2.0633
	Berries	90.0	80.0	21.2019, 2.0633
	Nuts	90.0	80.0	21.2019, 2.0633

*pi – uninformative Jeffrey's prior (0.5, 0.5)

Appendix 12: Priors of Sensitivities of Food History Questionnaire for tomatoes, lettuce/leafy greens, nuts, berries and cucumbers used for Bayesian Latent Class Modelling

	Sensitivity	Reference	Average	Minimum
Tomatoes	61	Gupta, S. K., Nalluswami, K., Snider, C., Perch, M., Balasegaram, M., Burmeister, D., Lockett, J., Sandt, C., Hoekstra, R. M., & Montgomery, S. (2007). Outbreak of Salmonella Braenderup infections associated with Roma tomatoes, northeastern United States, 2004: a useful method for subtyping exposures in field investigations. Epidemiol Infect, 135(7), 1165-1173. doi:10.1017/S0950268807007911	70.5	61.0
	70	Greene, S. K., Daly, E. R., Talbot, E. A., Demma, L. J., Holzbauer, S., Patel, N. J., Hill, T. A., Walderhaug, M. O., Hoekstra, R. M., Lynch, M. F., & Painter, J. A. (2008). Recurrent multistate outbreak of Salmonella Newport associated with tomatoes from contaminated fields, 2005. Epidemiol Infect, 136(2), 157-165. doi:10.1017/S095026880700859X		
	81	Behravesh, C. B., Blaney, D., Medus, C., Bidol, S. A., Phan, Q., Soliva, S., Daly, E. R., Smith, K., Miller, B., Taylor, T., Jr., Nguyen, T., Perry, C., Hill, T. A., Fogg, N., Kleiza, A., Moorhead, D., Al-Khaldi, S., Braden, C., & Lynch, M. F. (2012). Multistate outbreak of Salmonella serotype Typhimurium infections associated with consumption of restaurant tomatoes, USA, 2006: hypothesis generation through case exposures in multiple restaurant clusters. Epidemiol Infect, 140(11), 2053-2061. doi:10.1017/S0950268811002895		
	69	Donnan, E. J., Fielding, J. E., Gregory, J. E., Lalor, K., Rowe, S., Goldsmith, P., Antoniou, M., Fullerton, K. E., Knope, K., Copland, J. G., Bowden, D. S., Tracy, S. L., Hogg, G. G., Tan, A., Adamopoulos, J., Gaston, J., & Vally, H. (2012). A multistate outbreak of hepatitis A associated with semidried tomatoes in Australia, 2009. Clin Infect Dis, 54(6), 775-781. doi:10.1093/cid/cir949		
	67	Carvalho, C., Thomas, H., Balogun, K., Tedder, R., Pebody, R., Ramsay, M., & Ngui, S. (2012). A possible outbreak of hepatitis A associated with semi-dried tomatoes, England, July-November 2011. Euro Surveill, 17(6).		
	75	Reller, M. E., Nelson, J. M., Mølbak, K., Ackman, D. M., Schoonmaker-Bopp, D. J., Root, T. P., & Mintz, E. D. (2006). A large, multiple-restaurant outbreak of infection with Shigella flexneri serotype 2a traced to tomatoes. Clinical infectious diseases, 42(2), 163-169.		

Lettuce/ Leafy Greens	64.9	Slayton, R. B., Turabelidze, G., Bennett, S. D., Schwensohn, C. A., Yaffee, A. Q., Khan, F., Butler, C., Trees, E., Ayers, T. L., Davis, M. L., Laufer, A. S., Gladbach, S., Williams, I., & Gieraltowski, L. B. (2013). Outbreak of Shiga toxin-producing Escherichia coli (STEC) O157:H7 associated with romaine lettuce consumption, 2011. PLoS One, 8(2), e55300. doi:10.1371/journal.pone.0055300	79.7	70.0
	71	Nuorti, J. P., Niskanen, T., Hallanvuo, S., Mikkola, J., Kela, E., Hatakka, M., Fredriksson-Ahomaa, M., Lyytikainen, O., Siitonen, A., Korkeala, H., & Ruutu, P. (2004). A widespread outbreak of Yersinia pseudotuberculosis O:3 infection from iceberg lettuce. J Infect Dis, 189(5), 766-774. doi:10.1086/381766		
	70	Ackers, M. L., Mahon, B. E., Leahy, E., Goode, B., Damrow, T., Hayes, P. S., Bibb, W. F., Rice, D. H., Barrett, T. J., Hutwagner, L., Griffin, P. M., & Slutsker, L. (1998). An outbreak of Escherichia coli O157:H7 infections associated with leaf lettuce consumption. J Infect Dis, 177(6), 1588-1593.		
	93	Centers for Disease Control and Prevention. (2016). Multistate outbreak of listeriosis linked to packaged salads produced at Springfield, Ohio Dole processing facility (final update). Centers for Disease Control and Prevention, Atlanta, GA: http://www. cdc. gov/listeria/outbreaks/bagged-salads-01-16/index. html		
	81	Centers for Disease Control and Prevention. (2012). Multistate Outbreak of Shiga Toxin-producing Escherichia Coli O157: H7 Infections Linked to Organic Spinach and Spring Mix Blend. Retrieved from www.cdc.gov/ecoli/2012		
	85	Centres for Disease Control and Prevention. (2013). Multistate outbreak of E. coli O157: H7 infections linked to romaine lettuce.		
	93	Tataryn, J., Morton, V., Cutler, J., McDonald, L., Whitfield, Y., Billard, B., Gad, R., & Hexemer, A. (2014). Outbreak of E. coli O157: H7 associated with lettuce served at fast food chains in the Maritimes and Ontario, Canada, Dec 2012. Canada Communicable Disease Report, 40(S1), 2.		
Nuts	70	Centers for Disease Control and Prevention. (2014). Multistate outbreak of human Salmonella Enteritidis infections linked to Turkish pine nuts. Final update.	75	70.0
	80	Centers for Disease Control and Prevention. (2013). Multistate outbreak of Salmonella Montevideo and Salmonella Mbandaka infections linked to tahini sesame paste (final update).		

Berries	56	Calder, L., Simmons, G., Thornley, C., Taylor, P., Pritchard, K., Greening, G., & Bishop, J. (2003). An outbreak of hepatitis A associated with consumption of raw blueberries. Epidemiology & Infection, 131(1), 745-751.	59.5	56.0
	63	Severi, E., Verhoef, L., Thornton, L., Guzmán Herrador, B. R., Myrmel, M., Stene-Johansen, K., & Vold, L. (2015). Large and prolonged food-borne multistate hepatitis A outbreak in Europe associated with consumption of frozen berries, 2013 to 2014.		
Cucumber	62	Angelo, K. M., Chu, A., Anand, M., Nguyen, T. A., Bottichio, L., Wise, M., Williams, I., Seelman, S., Bell, R., Fatica, M., Lance, S., Baldwin, D., Shannon, K., Lee, H., Trees, E., Strain, E., Gieraltowski, L., Centers for Disease, C., & Prevention. (2015). Outbreak of Salmonella Newport infections linked to cucumbers--United States, 2014. MMWR Morb Mortal Wkly Rep, 64(6), 144-147.	68.7	62.0
	75	Centers for Disease Control and Prevention. (2015). Multistate outbreak of Salmonella poona infections linked to imported cucumbers.		
	69	Donnan, E. J., Fielding, J. E., Gregory, J. E., Lalor, K., Rowe, S., Goldsmith, P., Antoniou, M., Fullerton, K. E., Knope, K., Copland, J. G., Bowden, D. S., Tracy, S. L., Hogg, G. G., Tan, A., Adamopoulos, J., Gaston, J., & Vally, H. (2012). A multistate outbreak of hepatitis A associated with semidried tomatoes in Australia, 2009. Clin Infect Dis, 54(6), 775-781. doi:10.1093/cid/cir949		

Appendix 13: Example of R Code

```
# Installing CODA files
install.packages("coda")
install.packages ("mcmcplots")
#Save environment
save.image("C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\BerriesRData")
#Load CODA environment
load ("C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\BerriesRData")
#Analyses of estimates
#read.coda ESTIMATES "Berries"
SeSelfRCodaChain1 <-
read.coda("C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\SeSelfRBerriesCodaChain1.txt",
"C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\SeSelfRBerriesCodaIndex.txt", 10000,
300000, 1, quiet=F)
SeSelfRCodaChain2 <-
read.coda("C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\SeSelfRBerriesCodaChain2.txt",
"C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\SeSelfRBerriesCodaIndex.txt", 10000,
300000, 1, quiet=F)
SeSelfRCodaChain3 <-
read.coda("C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\SeSelfRBerriesCodaChain3.txt",
"C:\\Users\\pjs147\\Desktop\\CODAFiles\\Berries\\SeSelfRBerriesCodaIndex.txt", 10000,
300000, 1, quiet=F)
SeSelfRCodaChains <- mcmc.list(list(SeSelfRCodaChain1, SeSelfRCodaChain2,
SeSelfRCodaChain3))
```

```
#Descriptives

summary (SeSelfRCodaChains)

#visual diagnostics

plot(SeFBQCodaChains)

autocorr.plot(SeFBQCodaChains)

mcmcplot(SeFBQCodaChains)

rmeanplot (SeFBQCodaChains)

#Convergence Diagnostics, Patrick Lam's syntax

#1. Gelman and Rubin Multiple Sequence Diagnostic

gelman.diag(SeSelfRCodaChains, confidence = 0.95, transform=FALSE, autoburnin=TRUE,

multivariate=TRUE)

gelman.plot(SeFBQCodaChains, bin.width = 10, max.bins = 50, confidence = 0.95, transform =
FALSE, autoburnin=TRUE, auto.layout = TRUE)


#2. Geweke Diagnostic

geweke.diag(SeSelfRCodaChains)

geweke.plot(SeFBQCodaChains)

#3. Raftery and Lewis Diagnostic

raftery.diag(SeSelfRCodaChains, q = 0.025, r = 0.005, s = 0.95)

 #4. Heidelberger and Welch Diagnostic

heidel.diag(SeSelfRCodaChains)
```

Appendix 14: Contingency tables for 5 foods used in Bayesian Latent Class Analysis

Lettuce and Leafy Greens

		Food History Questionnaires	
Microsurveys	Self-Report	No	Yes
No	No	4	8
	Yes	5	10
Yes	No	1	7
	Yes	2	14

Nuts

		Food History Questionnaires	
Microsurveys	Self-Report	No	Yes
No	No	13	8
	Yes	4	4
Yes	No	7	3
	Yes	6	6

Berries

		Food History Questionnaires	
Microsurveys	Self-Report	No	Yes
No	No	16	3
	Yes	11	4
Yes	No	3	4
	Yes	4	6

Cucumber

		Food History Questionnaires	
Microsurveys	Self-Report	No	Yes
No	No	21	9
	Yes	2	2
Yes	No	2	6
	Yes	3	6

Tomatoes

		Food History Questionnaires	
Microsurveys	Self-Report	No	Yes
No	No	3	3
	Yes	6	11
Yes	No	3	4

Appendix 15: Measures of conditional independent for tomatoes, lettuce/leafy greens, nuts, berries and cucumbers used for Bayesian Latent Class Modelling

	rhoD (for Sensitivity)	95% CI	
Tomatoes	-0.070	-0.343	0.222
Lettuce and Leafy Greens	-.017	-0.381	0.308
Cucumbers	0.059	-0.393	0.542
Berries	-0.139	-0.513	0.285
Nuts	0.064	-0.370	0.541

	rhoDc (for Specificity)	95% CI	
Tomatoes	0.272	-0.113	0.809
Lettuce and Leafy Greens	0.272	-0.118	0.811
Cucumbers	0.318	-0.077	0.810
Berries	0.172	-0.175	0.709
Nuts	0.154	-0.169	0.649