# An Investigation of Integrating Program- and System-level Datasets for Examining Relationships Among Variables of a Complex Health System

by

Alvin Yapp

A thesis submitted in partial fulfillment of the requirements for the degree of

Master of Education

in

Measurement, Evaluation and Cognition

Department of Educational Psychology
University of Alberta

## Abstract

The potential for integrated program- and system-level datasets for generating previously inaccessible evaluative insights remains untapped. This study, situated within a yearlong evaluation of a lifestyle intervention program for patients with metabolic syndrome (MetS), provides an illustrative example of using an integrated dataset to identify relationships among the different levels of healthcare and examining the statistical findings with key informant interviews to extended interpretations beyond statistical significance towards clinical significance. The analysis of the program-level dataset revealed that the lifestyle intervention program was an effective treatment for MetS, while the additional analysis of the system-level dataset found an increase in patients' ED utilization post-enrolment. Further, the new evaluative insight from the integrated dataset of no significant relationship among the clinical indicators and ED utilization begins to situate the program within the larger system context, suggesting some relationship between lifestyle intervention programs and ED utilization unrelated to clinical indicators. In so doing, this study advances a design for evaluating health programs with implications for enhancing program-level evaluations within complex health systems through integration of system-level datasets, and for increasing clinical validity of evaluation findings through inclusion of clinical experts as members of the evaluation team.

## Preface

This thesis is an original work by Alvin Yapp. The research project, of which this thesis is a part, received research ethics approval from the University of Alberta Health Research Ethics Board – Health Panel, Project Name "Integrating Administrative Datasets with Program Datasets: The Effect of a Lifestyle Intervention Program on Emergency Department Utilization of Patients with Metabolic Syndrome", Pro00054995, March 25, 2015.

*"All collected data had come to a final end. Nothing was left to be collected.*

*But all collected data had yet to be completely correlated and put together in all possible relationships.*

*A timeless interval was spent in doing that."*

- *Issac Asimov, The Last Question*

**Acknowledgements**

After what seems like “a timeless interval”, here I am at the end, and with a completed document! I am only here because of the important contributions of the wonderful people I have in my life. First and foremost, my parents, you taught me the value of struggle through your own hardships. The foundation of my education began with post-dinner sessions at the kitchen table, and I will be forever grateful for the opportunities your sacrifices have provided me. To my brother, Adrian, you kept me grounded and provided welcome, if unproductive, distractions.

To my colleagues in the Centre for Research in Applied Measurement and Evaluation, academic writing can be an isolating and lonely experience with long hours spent agonizing over seemingly important minutia. Thank you all for sharing in this experience and being alone together! To Katya Chudnovsky and Krista Brower, thank you for your advice on working with the datasets, and for your feedback on this document, respectively.

To my colleagues at the Edmonton Oliver Primary Care Network, and the CHANGE Alberta research team, thank you for providing the opportunity and resources to conduct this thesis research. To my colleagues at Alberta Health Services, Drs. Oddie, and Krajnak, thank you for your advice on navigating the request process for accessing system-level datasets. To the patients and my interview participants, thank you for your participation, your contributions have been invaluable and led to findings that would not have been possible otherwise.

To my committee members, Drs. Bulut and Walton, thank you for taking the time to serve on my thesis committee. Your feedback and expertise are greatly appreciated! To my thesis supervisor, Dr. Cheryl Poth, thank you for your ongoing support throughout this journey and your endless patience as we charted new territory academically and logistically. I would not

have gotten here without you. Your mentorship has helped me become a better academic and evaluator. I will carry your lessons into all future endeavors. Thank you.

Finally, to my wife, Jinglan. Thank you for your unconditional support and encouragement, for putting up with my late nights and absent weekends, and for your clinical input and unique perspectives as you suffered through my early drafts. I couldn't have asked for someone better to be at my side during this adventure. This is for you.

## Table of Contents

## List of Figures

## List of Tables

## Chapter 1: Introduction

There exists untapped potential for the integration of program-level and system-level datasets to be used for generating new evaluative insights about the relationships among the different levels of healthcare to inform both policy and program decisions. I observed this possibility first-hand in my evaluator roles with the health service authority of Alberta, Alberta Health Services (AHS) from 2009-2014, and currently with a primary care network (PCN). Common across both roles was my work with healthcare datasets examining patterns of change and outcomes. What differed across the roles was the type of databases I had access to; whereas previously, I used the system-level datasets to examine the changes in utilization patterns across the provincial health system; currently I have access to program-level clinical datasets for examining individual patient outcomes of participants in health programs. I have discovered that while it is common practice for individual datasets at the program- and system-levels to be analyzed separately, opportunities are limited for integrating these datasets to examine the relationship among the variables across different levels of healthcare.

My familiarity with the available datasets led me to believe their integration held strong potential for uncovering new and useful information that would potentially influence both program- and system- level decisions. However, through experience, I recognized the integration of program- and system- level datasets are often prevented by logistical challenges beyond the control of individual researchers. In addition, I also noticed that the interpretation of the statistical findings is often left to folks like me with no clinical training. This thesis study provides an illustrative example through my journey, in dual roles as an evaluator and researcher to: (1) access and integrate datasets from different levels of healthcare, (2) analyze the resulting integrated dataset, and (3) interpret the statistical findings using a clinical interpretive lens.

Through this illustration, I will demonstrate how the integration of data across two levels of health care can lead to the generation of new evaluative insights that would not have been available otherwise, and how these new evaluative insights might be used to inform program-level and system-level decisions.

In this introductory chapter, I tell the story of how I came to focus my thesis research on integrating program- and system-level datasets. I begin by describing my dual roles as a researcher and an evaluator, describe each type of dataset (program-level dataset from the PCN and system-level dataset from AHS), and outline what I recognized as limitations inherent to each of the datasets. I then situate the thesis within an ongoing evaluation of a lifestyle intervention program, the Canadian Health Advanced by Nutrition and Graded Exercise (CHANGE) program. To do this, I provide background information on the clinical impact of metabolic syndrome (MetS); and describe the CHANGE program, the Program Outcomes Dataset, and the System Outcomes Dataset. Finally, I describe the organization of the thesis into its six chapters.

**Describing my Dual Roles as Researcher and Evaluator**

As I conducted my thesis research in my role as a thesis researcher, I was concurrently conducting the evaluation of the CHANGE program as an internal evaluator with a Primary Care Network (PCN). A PCN is an organization that helps local physicians provide comprehensive patient care to Albertans through coordination of multidisciplinary healthcare professionals (Primary Care Network Project Management Office, 2016). The role of a PCN evaluator involves working with the management team and the clinicians to develop and administer evaluation plans across the PCN. This role involved a relationship of trust with management and

clinicians that facilitated the process of obtaining permission to conduct this thesis study with PCN data. At the beginning of the thesis research, I had been working in this role for 12-months.

The dual roles of a researcher and evaluator served numerous benefits for me as a researcher, for my colleagues at the PCN, and for the CHANGE program personnel. As a researcher, it provided access to the data of the CHANGE program (hereafter referred to as the Program Outcomes Dataset), a familiarity to the CHANGE program, and access to practicing clinicians involved with CHANGE for the clinical interpretation portion of this study. For my colleagues at the PCN, working with someone on this research that they are already familiar may give them greater comfort, which could potentially help them gain new insights to inform program decisions. For the CHANGE program personnel, the ability to integrate the program-level dataset with the previously inaccessible system-level dataset from AHS containing emergency department utilization data (hereafter referred to as the System Outcomes Dataset) allows for additional dimensions of the CHANGE program to be evaluated. Due to my familiarity with the components of System Outcomes Dataset, I could see the potential of how the integration of system-level data with program-level data may be beneficial to enhancing the evaluation of a health program, such as the CHANGE program.

In addition to the benefits of my dual role, I faced two challenges. As a researcher, my role as an internal evaluator led to closeness with the program and clinicians involved, resulting in a bias that needed to be mitigated. As an internal evaluator, my role as a researcher meant that my request for data was processed as an external researcher, which is very stringent and time consuming, resulting in a limited number of datasets being available for integration.

**Limited Use of Integrated Datasets for Program- and System-Level Evaluations**

In this study, I will refer to two types of datasets: system-level datasets, which are generated by and retained by Alberta Health Services (AHS), and program-level datasets, which are generated by individual health programs. In this section, I will describe my experience working with these datasets, how these datasets are currently used, highlight the limitations inherent to each of the datasets, and point to how these limitations can be mitigated through integrating the datasets together.

AHS is the provincial health authority of Alberta and as such is responsible for the "planning and delivery of health supports and services [to Albertans]" (Alberta Health, 2016). AHS is the source of many system-level health sector datasets. During my 5 years employed as an evaluator with AHS, I worked with an evaluation unit that had access to a wealth of system-level data. This system-level data was used to examine the performance of the health system. The ability to properly evaluate the performance of system-level initiatives was attributable to the large-scale and quantitative nature of these datasets.

Many of these system-level datasets contain administrative data, which is "data collected for the purpose of carrying out various non-statistical programs" (Statistics Canada, 2015). Administrative data is generated alongside clinical data with each patient encounter with a healthcare provider and kept in system-level datasets. A patient encounter is defined as "contact between a patient and a healthcare provider, occurring at a given time and place, where one or more services or produces are provided to maintain or restore the patient's health" (Alberta Health, 2015, p. 2). The type of administrative data generated varies for each dataset, but typically includes billing, time spent, and location of service. Administrative data is collected to facilitate the delivery of care, and to build patient records informing health care providers of the

patient's medical history. However, administrative health data has also been found to be valuable for answering system-wide questions such as patterns of emergency department utilization for different groups of patients (Lee, 2013; Nguyen & DeJesus, 2010) and health costs (Hemmila et al., 2008). In order to facilitate use of this resource, AHS maintains a repository of administrative data and has staff dedicated to helping with its use for quality improvement, research, and evaluation of system-wide issues.

The limitation of using system-level data is the lack of insight into underlying contributors to the findings, such as the existence of health programs (i.e., lifestyle intervention programs) in primary care contributing to improved patient outcomes. For example, if emergency department utilization increased in the past year, there is no readily available information about the factors causing that increase in these system-level reports. In contrast, program-level reports measure outcomes deemed important by the program in question, and is limited solely to data collected from the measured program outcomes. Integrating system-level data to program-level data, and assessing the effect of specific program-level interventions on the system-level outcome indicators may allow for the additional examination of factors contributing to system-level outcomes. I was aware of this potential, but I was unable to access program-level data whose custodians are primary care physicians. However, I suspected that if program-level datasets were integrated with the system-level datasets, it would be possible to provide insights into the role of these smaller scale programs on system-wide measures, which may in turn influence system-wide initiatives.

The program-level datasets available at the PCN are descriptive of program activities and of enrolled patients, containing clinical indicators relevant to each program; this was the type of program-level data that I aimed to integrate with system-level data stored separately at AHS to

identify what was driving system-level performance indicators at the program-level. As the evaluation specialist at the PCN, I used this data to provide evaluation support for programs underway at the PCN. The program-level data from the PCN measures how utilization of PCN resources affects patient health; evaluation of this data often informs program- and patient-improvement initiatives. However, this dataset is limited to primary care services; it does not provide any information about patient utilization of health resources at AHS, specifically, emergency department utilization data. I had seen the type of data available from both sides of the access barrier, but I was not yet able to put them together. Without integrating program-level datasets together with system-level datasets, a more complete picture of a patient's journey through the health system cannot be obtained. In this regard, the effect of a program on the wider health system remains unexplored in the current evaluation environment.

The integration of program- and system-level datasets would allow for the exploration of primary health care programs effects on the wider health system. Across my roles in AHS and with a PCN, I observed few instances of integrations of datasets across the program- and system-levels. The integration of health sector datasets holds strong potential for generating evaluative insights about complex relationships between health programs and the broader health system. However, there remain substantial barriers to accessing and integrating datasets across the program- and system-levels. In my current evaluator role at the PCN, I do not have direct access to the AHS administrative data. This limits the extent to which I can examine patient utilization of non-PCN resources, such as emergency department utilization, and therefore limits my ability to examine the effect of the PCN activities on the wider health system. Ironically, I experienced the opposite dilemma while employed by AHS; I did not have access to System Outcomes Datasets rather than the program-level datasets. The current thesis study presents an opportunity

to integrate program-level and system-level datasets within the evaluation of the CHANGE program to demonstrate the potential benefits of integrating datasets for generating evaluative insights that would have otherwise been unavailable.

## Situating CHANGE Program as the Evaluation Context

To understand the important contributions of integrated datasets for informing the evaluation of the CHANGE program, some background information on the health impact of metabolic syndrome (MetS) is necessary. The high prevalence of MetS in the population, the debilitating effect it has on people's health, and its position as a unique early indicator of deteriorating health highlights the potential importance of lifestyle change programs to create better health outcomes at both the program- and system-levels. Although important to the study, the diagnosis and treatment of MetS is not the focus of this thesis; therefore, this is provided as background information in the introduction rather than in the literature review.

### Impacts of Metabolic Syndrome on Patients and Health Resources

MetS is a condition of insulin resistance associated with an increased risk for developing chronic conditions such as cardiovascular disease (CVD), hypertension, type 2 diabetes (T2DM), chronic kidney disease, and dyslipidemia (Rao, Dai, Lagacé, & Krewski, 2014). MetS has been identified both as a national and global public health problem (Riediger & Clara, 2011; Zimmet, Magliano, Matsuzawa, Alberti, & Shaw, 2005). According to the Canadian Diabetes Association Clinical Practice Guidelines Expert Committee, a diagnosis of MetS requires three of the following five criteria: (a) elevated waist circumference (ranges differ based on ethnicity), (b) elevated triglycerides (≥1.7mmol/L), (c) reduced high-density lipoprotein cholesterol (<1.0mmol/L for males, <1.3mmol/L for females), (d) elevated blood pressure (≥130/85 mm/hg), and (e) elevated fasting blood glucose (≥5.6 mmol/L) (Goldenberg & Punthakee, 2013).

Progression of a patient from MetS to chronic disease increases the patient's risk of requiring emergency department services, inpatient admissions, and reliance on prescription drugs (Mensah & Brown, 2007).

The predisposition of patients with MetS towards developing serious chronic conditions such as T2DM and CVD is well documented. Ford, Li, & Sattar (2008) completed a quantitative review of 16 prospective studies examining the association of MetS with T2DM, concluding that MetS is a strong predictor of T2DM. In a population-based prospective cohort study, Lakka et al., (2002) found that Finnish men aged 42 to 62 with MetS were more likely to develop CVD at the 11-year follow-up mark, even if the patient did not have CVD or diabetes at baseline testing. This association between MetS and chronic diseases may in part be due to similar criteria required for both diagnosis of MetS and diagnosis of certain chronic diseases. Furthermore, MetS is a cluster of risk factors related to increased incidence and mortality of CVD (Grundy, 2004; Yoon, Yoo, & Kim, 2014).

The link between MetS and the progression towards certain chronic diseases makes a diagnosis of MetS a unique warning sign. This presents clinicians and patients with an opportunity to alter the patient's health trajectory, stopping and potentially reversing progression towards serious chronic diseases. It is important to seize this opportunity to decrease the prevalence of not only MetS, but also potentially the prevalence of T2DM and CVD in our population. This can be achieved through the implementation of lifestyle intervention programs designed to support patients in adopting healthy diets and exercise regimens, such as the CHANGE program. Lifestyle interventions have been shown to stop the progression of MetS to chronic diseases, and are the recommended treatment for MetS (Grundy, 2004). Any decrease in the prevalence of chronic diseases may also result in a corresponding decrease in emergency

department utilization related to these diseases. This makes MetS an ideal condition for this study; it can be treated within a primary care system, and treatment impacts can be measured in the broader health system. This conveniently makes evaluation of the CHANGE program an opportunity to demonstrate integration of program-level data with system-level administrative data, and how this can improve understanding of outcomes at both the program and the broader health system levels.

**CHANGE Program and Program Outcomes Dataset**. The PCN management team provided the initial opportunity for this study by requesting evaluation building upon the previous work of the CHANGE researchers. The purpose of this evaluation was to assist clinical and policy decision-making by examining the effectiveness of the CHANGE program in managing MetS, and to improve future lifestyle intervention programs. A description of the CHANGE program provides some context in which the current study takes place, in terms of the patient population enrolled in the program, the nature of data available, and the current efforts to assess program outcomes. Understanding the goals of CHANGE also establishes the need for integrating the program-level data of CHANGE with the system-level data from AHS.

An initial research project focused on the CHANGE effects on clinical indicators generated the program-level. The Canadian Health Advanced by Nutrition and Graded Exercise (CHANGE) program is a 12-month behavioral change intervention program developed through a collaborative effort from physicians and researchers located at five Canadian institutions (St. Joseph's Healthcare Hamilton, Polyclinic, University of Guelph, Laval University, and Royal Alexandra Hospital Edmonton). This program is targeted at patients with Metabolic Syndrome (MetS). The three primary objectives of CHANGE are to: (1) promote healthy lifestyle modification, (2) reduce patient reliance on prescription drugs, and (3) reduce progression of

chronic diseases (CHANGE Alberta, 2015). The program aims to achieve these objectives through lifestyle interventions, namely validated nutrition and exercise programs. In 2012, the national research team chose three "demonstration sites", with 300 patients (100 patients at each site), to pilot the CHANGE program. A program coordinator at each site managed referrals to the CHANGE program and provided administrative support. Registered dietitians and kinesiologists were responsible for providing clinical services. The PCN for which I worked was one of these demonstration sites, and administered the program as follows.

Family physicians (FPs) affiliated with this PCN referred patients with Metabolic Syndrome (MetS) to the CHANGE program, where they were screened by a program coordinator to ensure they meet inclusion/exclusion criteria (Table 1). Screening involved a one-on-one session with the patient to discuss program requirements, an interview to determine the patient's stage of readiness to make lifestyle changes, bloodwork, measurements of weight and waist circumference. Eligible patients were enrolled in the program for 12 months throughout which they received dietary coaching from a registered dietitian and exercise coaching from a kinesiologist on a weekly basis for 3 months, then monthly for 9 months. The first CHANGE patient was enrolled on October 18, 2012. Together, the project coordinator, registered dietitian, and kinesiologist were known as the CHANGE team.

As a part of program requirements, patients had regular visits every 3 months with their family physician to monitor the patients' progress through the program. During these visits, patients underwent bloodwork, resulting in clinical data being collected at five time points (baseline, 3, 6, 9, and 12 months). The clinical indicators collected were related to MetS or chronic diseases associated with MetS: blood pressure (BP), weight, waist circumference, fasting blood glucose (FBG), hemoglobin A1C (HgA1c), total cholesterol, triglycerides (TC), high-

density lipoprotein (HDL), low-density lipoprotein (LDL), and creatinine. These clinical indicators, as a dataset, will be referred to hereafter as the Program Outcomes Dataset. The baseline and 12-month individual clinical indicators were previously assessed by CHANGE researchers, allowing them to identify the program as an effective treatment for MetS.

Table 1
*Inclusion/Exclusion Criteria of the CHANGE Program*

| Inclusion | Exclusion |
|---|---|
| 3 of 5 of…<br>Blood Pressure at least one of:<br>● ≥130/85 mmHg<br>● Receiving pharmacotherapy for elevated blood pressure<br>Fasting Blood Glucose at least one of:<br>● Fasting Blood Glucose ≥ 5.6mmol/L<br>● Receiving pharmacotherapy for elevated blood glucose levels<br>Triglyceride (at least one of...)<br>● Triglycerides of ≥ 1.7 mmol/L<br>● Receiving pharmacotherapy for elevated triglycerides<br>HDL-C<br>● <1.0mmol/L (males)<br>● <1.3mmol/L (females)<br>Abdominal circumference > ranges specified below:<br>• Abdominal circumference > ranges specified below<br>● Europids, Whites, sub-Saharan Africans, Mediterranean, middle east (Arab) ≥ 94 cm males, 80 cm females<br>● Asian and South Central Americans ≥ 90 cm males and 80 cm females<br>● US and Canadian Whites ≥ 102 cm males, 88 cm females<br>● Ethnicity unclear (use range for Europids) | Inability to speak, read or understanding English and/or French<br><br>Having a medical or physical condition that makes moderate intensity physical activity (like a brisk walk) difficult or unsafe.<br><br>Diagnosis of Type 1 Diabetes Mellitus<br><br>Type 2 Diabetes Mellitus only if any one of the following are present (select all that apply)<br>● Proliferative diabetic retinopathy<br>● Nephropathy (suggested parameters: serum creatinin > 160)<br>● Clinically manifest neuropathy defined as absent ankle jerks<br>● Severe fasting hyperglycemia > 11 mmol/L<br><br>Peripheral vascular disease<br><br>Significant medical co-morbidities, including uncontrolled metabolic disorders (e.g., thyroid, renal, liver), heart disease, stroke and ongoing substance abuse.<br><br>Clinically significant renal failure |

This current thesis study extends the initial CHANGE research program focused on program effects on clinical indicators related to MetS to examine program effects on emergency department utilization through integrated datasets. The effect of primary care programs, such as CHANGE, on emergency department (ED) utilization is of interest due to the latter's high costs in the health system; there is potential for primary care interventions to divert patient trajectories away from the ED visits. Since the Program Outcomes Dataset does not include measures of ED

utilization, the Systems Outcome Dataset needs to be integrated with the Program Outcomes Dataset in order to examine the effect of the CHANGE program on patient ED utilization rates.

**System Outcomes Dataset**. The Systems Outcome Dataset consists of data from one administrative dataset, the National Ambulatory Care Report System (NACRS), which tracks emergency department activity. Data from emergency departments are cleaned and aggregated before being stored by the health authority in this repository. Several variables are available through this dataset (Appendix A), but for this study, the variable of interest is the number of ED visits each patient had before and after enrolment in the CHANGE program. An additional factor in the selection of this dataset is the feasibility of accessing the datasets in a timely manner for the purposes of a thesis study; including more data elements would have required other organizations or permissions which were not feasible for this study. The data from NACRS is limited to the system-level measure of ED utilization; this is only one aspect of a complex health system, but integrating the System Outcomes Dataset with the Program Outcomes Dataset will allow for an initial look at the effect of health programs on system-level indicators like ED utilization.

**Integrated Dataset**. The System Outcomes Dataset will be integrated with the Program Outcomes Dataset to examine the effect of this lifestyle intervention program on patients' pattern of emergency department utilization. The resulting dataset will be referred to as the Integrated Dataset. Using the Integrated Dataset, the relationship between clinical indicators of MetS and emergency department utilization can then be explored. A summary of the three datasets and their sources is presented in Table 2.

Table 2
*Datasets and Data Sources*

| Dataset | Data Source |
|---|---|
| Program Outcomes Dataset | CHANGE Program (Electronic Medical Records) |
| System Outcomes Dataset | Alberta Health Services (National Ambulatory Care Reporting System) |
| Integrated Dataset* | Integrated Program Outcomes Dataset and System Outcomes Datasets |

*Integration done by Alberta Health Services using Patient Health Numbers. Datasets were returned de-identified.

**Thesis Purpose Statement and Organization**

The purpose of this study is to illustrate the usefulness of evaluative insights for a complex health system generated from the examination of nuanced relationships among program- and system-level variables of an integrated dataset. This thesis is organized into six chapters. In chapter 2, I describe the results of a literature review focused in health care, illustrating the need for integration of program- and system-level datasets, the current state of evaluation designs and use, and the challenges inherent to dataset access, analysis, and interpretation. In chapter 3, I present the study approach and dataset descriptions, analysis and interpretation procedures, and ethical considerations and efforts. In chapter 4, I report on the main findings for the three datasets, and the clinical interpretations of the statistical results by clinicians. In chapter 5, I discuss the study findings, using three points to illustrate the usefulness of integrated datasets for generating evaluative insights and one to point to the importance of the clinical perspective. In Chapter 6, I highlight three implications for health care evaluation practices: emerging standards of integrating datasets and including clinical perspectives to explore a complex health system, mitigating strategies for accessing data, and

broadening perspectives by engaging clinicians in evaluations. I conclude the thesis by linking the future directions for research to the study limitations and offer final thoughts on the integration of datasets across different levels of healthcare.

## Chapter 2: Literature Review

There is a need for understanding patient journey as they access and utilize health care to determine how to provide patients with "…the right care, at the right time, from the right provider, and in the right place…" (Alberta Health Services, 2015, p.67) and to enhance resource use in health care. The purpose of this literature review is to illustrate the need for the current thesis study in generating new evaluative insights for assessing program- and system-level outcomes from the integration of program- and system-level datasets. This chapter begins by describing how evaluation is currently used within a complex health care system to assess the impacts of programs at the program- and system-levels, and points towards the potential of integrated data to examine nuanced relationships among variables of a complex health system. Next, I describe the barriers to accessing datasets at both the program- and system -levels, and propose that these barriers limit the types of data available to researchers and evaluators, and thus limit the evaluation designs that can be conducted at each level. Lastly, I highlight the important role of clinical experts in an evaluation of a complex health system, specifically the importance of interpreting statistical results through a clinical interpretive lens by interviewing clinicians with informed clinical and contextual expertise. Together, this review of the current published literature highlights the need for an illustrative example of a study that explores the evaluative insights gained from integrated program- and system-level datasets, and clinical interpretation of the statistical findings by experts.

### Evaluation Designs and Use within Complex Health Systems

Evaluations in health organizations have long been conducted (Marjoua & Bozic, 2012). How healthcare evaluations are conducted and used are dependent upon the data available at program- and system-levels and the challenges inherent to dataset integration, analysis, and

interpretation. A review of the literature reveals the prevalent use of certain research designs and approaches to evaluations (Forbes et al., 2015). Some designs are better than others but what becomes clear is that the availability of data influences evaluation design and subsequent evaluation use. Yet what remains to be investigated is the effect of programs within the larger systems in which they are situated.

**Use and limitations of randomized control trial designs** The randomized control trial (RCT) is touted as the most powerful experimental design and the gold standard in clinical research (Stolberg, Norman & Trop, 2004). At its simplest, the RCT involves participants receiving one of several clinical interventions allocated at random (Jadad, 1998 via Stolberg, Norman & Trop, 2004). Thus, RCTs are conducted in environments that control for as many variables as possible, allowing researchers to make causal inferences about clinical interventions. A RCT is especially useful when the variables related to the condition and the clinical intervention are well defined and measurable. This design is used throughout the health sector to assess clinical interventions ranging from program-level interventions, such as the effectiveness of pharmacological drugs on blood pressure, to system-level interventions, such as the impact of a type of treatment on hospital length of stay.

RCTs are considered the gold standard for explanatory research, but have limited usefulness for assessing effectiveness of a complex healthcare system due to the large number of variables in the system that are difficult to control (Campbell et al., 2000; Boon et al., 2007). A complex healthcare system is defined as a collection of "complex interventions to improve or enhance health and well-being as well as to prevent disease" (Boon et al., 2007, p.2). In this definition, "complex" refers to the "entangled interrelationships among multiple 'active' components of the intervention" (Boon et al., 2007, p.2). Due to these complex relationships

amongst variables, RCTs alone may not be an appropriate study design for evaluation of a complex healthcare system.

Evaluation plays an important role within health organizations in addressing the deficiency of RCT to examine complex relationships, by providing a rigorous process to seek evidence that are "useful, feasible, ethical, and accurate" (Koplan, Milstein, & Wetterhall, 1999, p1) to inform decision making in a complex healthcare system. There are a variety of evaluation designs that examine factors outside the independent variables using a variety of methods. To encourage the use of evaluation in healthcare, health authorities in the United States and Canada have developed guides to introduce the field of Evaluation to program managers and clinicians (Koplan, Milstein, & Wetterhall, 1999; Bowen, 2012; Alberta Health, 2013; Public Health Ontario, 2015), and evaluation has seen increased use at both the program and system-levels.

The increase in the use of evaluation in healthcare is promising, but for results of program evaluations to be beneficial, they must be used. Evaluation use is one of the most researched areas in the evaluation field (Christie, 2007). In the literature, the terms evaluation "use" and "utilization" are used interchangeably, and a reasonable working definition is: "the effect the evaluation has on the evaluand – the 'thing' being evaluated – and those connected to the evaluand" (Christie, 2007, p. 8). Alkin and Taut (2002) identified two major types of evaluation use: "process use" and "use of evaluation findings". Process use is defined as "cognitive, behavioural, program, and organizational changes resulting, either directly or indirectly, from engagement in the evaluation process and learning to think evaluatively" (Patton, 2008, p. 108). Use of evaluation findings is more complicated, being divided into three types: instrumental, conceptual, or symbolic (Alkin, 2011; King & Pechman, 1986; Leviton & Hughes, 1981 via Johnson et al., 2009). Instrumental use refers to cases where evaluation findings are

directly applied. Conceptual use refers to cases where evaluation findings have not resulted in a change in the program, but resulted in a change in understanding. Symbolic use refers to an evaluation whose results will not be used, but rather, the evaluation itself was used to fulfill some requirement.

The type of data used in evaluations can have an impact on the utilization value of an evaluation. Christie (2007) examined the influence of different types of evaluation data (large-scale, case study, and anecdotal) on the actions of 131 decision makers, and found that most decision makers were somewhat, or greatly influenced, by all types of evaluation data. This finding suggested that decision makers are willing to use different types of information while making decisions. The author went on to posit that it may be advantageous for evaluators to identify the types of data the decision makers would like to use in informing their decisions. In addition, the author also found that certain types of evaluation data influenced certain groups of decision makers more in particular situations, "[pointing] to the decision maker's desire for both broad and more in-depth information when making decisions about programs, thus reflecting the need for evaluation studies that are designed to yield both types of information" (Christie, 2007, p. 20). There appears to be value in identifying the types of data that would be most valuable to stakeholders in the current context of the evaluation; understanding the type of data that would be most persuasive is an important aspect of evaluations that could influence evaluation use.

Looking at the types of evaluation data available at the different levels of healthcare reinforces how the data made available influences the evaluation designs, which then ultimately influences evaluation use. At the program-level, it is the patient clinical indicators data and experience measures of patients for a program that are easier to access, and local program-level changes are easier to implement. Thus, improvement measures are used more at the program-

level. At the system-level of a health authority, the data that is most readily available are quantitative administrative data that spans the health authority. Evaluations at the system-level are generally commissioned by managers interested in performance measures of the whole health system (i.e., emergency department wait times), especially because system-wide changes are difficult to implement. Thus, performance indicators are used more at the system-level. These differences in data availability drives the differences in evaluation designs at the program-level to look more towards formative measures of experience for improvement use (process/conceptual use), and at the system-level to look at more summative quantitative measures for accountability (instrumental use). I will provide examples that highlight the current methods that are prevalent within program-level and system-level evaluations in the next two sub-sections, focusing in on evaluations of lifestyle intervention programs and emergency departments. I point out limitations of these evaluations, and propose how integrated datasets can address these limitations.

**Evaluating impacts at the program-level**. Current program evaluations of health programs consist of examining the clinical indicators of the patient, measures of the patient experience (i.e., satisfaction, self-reported outcome measures), and measures of the providers of care (i.e., performance, perspective) (Wozniak et al., 2014). The effect of health program on system-level measures, such as the decrease in use of healthcare services (i.e., emergency department utilization), has been identified as an important indicator to measure (Hollander et al., 2010); however, measurement requires the challenging task of integrating program and system-level datasets. Consequently, most program-level evaluations are unable to integrate system-level datasets, and therefore focus on the program itself, and the evaluations of current lifestyle intervention programs are no exception. For the purposes of this literature review, I will focus on research directed on efficacy of lifestyle interventions to examine the current state of lifestyle intervention programs, and the methods used in these evaluations of lifestyle programs.

The diet intervention used in the CHANGE program is the gold standard diet intervention taught to patients with MetS: the Mediterranean diet. It is characterized by a "high intake of olive oil, fruits, nuts, vegetables, and cereals; a moderate intake of fish and poultry; a low intake of dairy products, red meat, processed meats, and sweets; and wine in moderation, consumed with meals" (Willett et al., 1995 as cited in Estruch et al., 2013, pp.1). This diet has been shown to improve the clinical indicators associated with MetS (Kastorini et al., 2011) and has even been shown to facilitate reversion to normal from a baseline status of MetS (Babio et al., 2009; Babio et al., 2014; Estruch et al., 2013). Studies assessing the effectiveness of this diet intervention did so through pre-post analyses of the clinical indicators associated with MetS. The effect of these diet interventions and improved clinical indicators on the wider health care system was not investigated in any published studies.

As established previously, RCTs are used to determine the causal relationship between variables, but may not be appropriate for investigating outcomes in the complex health system. RCTs are still important tools for evaluators in the health context, as demonstrated by an RCT by Bo et al., (2007) who found that an intervention of five sessions of individualized recommendations from nutritionists, endocrinologists and internal medicine specialists covering diet, exercise and behaviour modifications over a one-year period led to a decrease in BMI and waist circumference, as well as substantially decreasing the prevalence of MetS within the treatment population. They also found that the control group, which did not receive the individualized recommendations, had a worsening in cardiovascular risk factor clinical indicators. Similarly, Gomez-Huelgas et al., (2015) examined the effects of a lifestyle intervention program on the long-term management and prevalence of patients with MetS. The authors assessed the success of this program through pre-post measures of the clinical indicators related with MetS, as well as inclusion of a control group. They found that intensive lifestyle interventions conducted by primary care physicians over a three-year period substantially improved clinical indicators related to MetS, finding significant differences in abdominal obesity, HDL-cholesterol and blood pressure. They also found that lifestyle interventions reduced the number of components of MetS more than conventional treatments. However, these studies are limited by the specific procedures RCT methods, and are ultimately not able to incorporate secondary health data into their analysis to further investigate the effect of interventions on the patients' broader health care utilization.

Looking to the investigation of program outcomes, Rubenfire et al. (2011) examined a diet and exercise education program (Med Fit) targeted at patients with MetS at the University of Michigan Health Care System. Med Fit focused on teaching the Mediterranean diet and

increasing the activity of the patients to the recommended level of 150 to 200 minutes per week including moderate intensity aerobic (50-75% of max heart rate), resistance training and stretching (Rubenfire et al., 2011). A weekly session of 45 minutes of diet education and another 45 minutes of exercise was held over 12 weeks, and resulted in significant decreases in five clinical indicators of MetS (Triglycerides, HDL-C, Blood Pressure and Fasting glucose). Again, there was no inclusion of secondary health datasets to place the program within a broader health system, and this study was not able to investigate its effects on other health care system utilization of the participants.

**Evaluating impacts at the system-level**. Evaluations of a system-level service, such as emergency departments, are generally focused around the emergency department themselves, not accounting for the influences of the programs or initiatives in other levels of health care.

An example of this focus inward is from a study by Saunders et al. (2015) that looked at the relationship between family immigration status and revisit rate to the emergency department. The study team focused on the demographic factors of emergency department utilization. This required the linking of two datasets, each from different sources: an administrative datasets of the emergency department (NACRS) from the health authority, and a demographic dataset (PRDS) from Citizenship and Immigration Canada. The sample size was large (n = 3,332,901 records) and the authors could identify some areas of possible improvement to emergency department care. However, the study was limited in that it did not include additional datasets that would have allowed examination of the effects of their other health services utilization on emergency department utilization. Additionally, a review of emergency department performance indicators found five categories of indicators: outcome, process, satisfaction, equity and structural/organizational measures (Madsen et al., 2015). These indicators are appropriate for

instrumental use, but are limited, as they do not directly examine the wider health system, or the role of the emergency department within that system. Integrating the emergency department performance indicators with program-level datasets may address this limitation by providing additional data on context of the patients' previous interactions with the health system before visiting the emergency department. This would include measures of the patients' health outside the emergency department and provide a more complete picture of the patients' reason for visiting the emergency department.

System-level indicators can be integrated to provide insights into other datasets. For example, Silverman et al., (2015) screened 256 patients aged 18-85 in the Long Island Jewish Medical Center observation unit for dysglycemia (irregular HbA1C levels). The authors found 52% of patients were subsequently diagnosed with pre-diabetes, and 9% were diagnosed with full diabetes. Their study suggested that, without the benefit of primary care, many of these patients would have dysglycemia remain undiagnosed, with a potential to progress towards more serious chronic disease. However, they note that "it is not clear why the frequency of undiagnosed dysglycemia was higher in the observation unit than in the United States outpatient population." (p. 8). Although they explored some possible explanations (high prevalence of African-Americans and Asian patients, uninsured patients, income, and health literacy), the authors ultimately point to a lack of social, economic and health service utilization data that leaves this question unanswered in their current study. This demonstrates that the data from the emergency departments, without integration with other datasets, are limited in their ability to examine the nuanced relationships among variables in a complex health system.

In addition to determining the effect of a program on emergency department utilization, there is also an opportunity to explore clinical indicators that are related to emergency

department utilization. Due to the high cost of emergency department visits (Tran, 2016), there is incentive in determining the variables associated with emergency department utilization (Porter et al, 2015). Initially, efforts were made to reduce costs to emergency departments through improving efficiencies within acute care, for example, replacing mid-level providers with emergency medicine residents (Clinkscales et al., 2015). However, if a patient can avoid the emergency department visits altogether through primary care interventions, this would also reduce costs (Sidorov et al., 2002). Intensive lifestyle intervention primary care programs, like the CHANGE program in this thesis study, have the potential to increase the quality and quantity of interactions between patients and caregivers. In literature, increased care continuity has been shown to decrease the likelihood of emergency department visits for congestive heart failure (Hussey et al., 2014). Although decreasing costs to the health system would be an important outcome measure of the CHANGE program, the effect of the CHANGE program on emergency department utilization has not yet been examined until this thesis study.

**Integration across program- and system-level data**. Integrating secondary datasets into the study of patient health outcomes through interventions programs does not seem to be a common occurrence. An attempt example was demonstrated by Forbes et al., (2015) who conducted a literature review of exercise programs for people with dementia. One of their secondary objectives was to examine the utilization of healthcare services of patients with dementia after completion of the exercise program; this would require the integration of secondary datasets from other health care services with the primary dataset. However, after screening papers and reviewing in-depth 18 studies, they noted that none of the included studies contained measures of healthcare service utilization, highlighting the challenges of integrating program- and system-level datasets to investigate the relationships among the variables between the two different levels of healthcare.

The studies reviewed by Forbes, et al., (2015) highlight possible methods of evaluating a lifestyle intervention program at the program-level; by examining clinical indicators related to MetS, quality of life indicators, and specific health outcomes. The studies reviewed were program-level clinical studies, and so the findings were tailored towards instrumental use. None of these studies examined the links between the program and the larger health system in which it exists. The current thesis study begins to address this lack of literature further examining emergency department utilization of patients with MetS who are enrolled in a lifestyle intervention program.

In this section, I established the current practices and methods in program-level and system-level evaluations. I also demonstrated how the type of data available could influence evaluation use, as in; an integrated dataset can potentially lead to an expansion in different types of evaluation use at different levels of healthcare. In the next section, I will describe the barriers

to accessing and integrating system-level and program-level datasets in the Alberta healthcare context, and the potential healthcare system benefits of overcoming these barriers.

**Barriers to Integrating Datasets across Levels of Healthcare**

Health data is generated from each patient encounter, and this data is mainly used for treatments and building a record of the patient's health care. In Alberta, health data is held by the different organizations, each of which provides different patient care. For example, an emergency department visit would be recorded in one dataset and a family physician visit would be held in another dataset. These datasets describe the same set of Albertan patients, and can potentially be brought together in an integrated dataset to help program evaluators provide a more complete picture of the effects of a program on both patient outcomes and the wider health system.

Integrated datasets can provide more information about patients by investigating the effect of the health care system on patients; a domain of health inquiry called health impacts (Hollander et al., 2010). The investigation of health impacts could, in turn, benefit from an increased use of integrated datasets. Integrated datasets would allow for examination of the broader effects on the patient and their utilization of other health resources, rather than focusing solely on the effect of a single health resource on the clinical indicators of the patient. Outcome measures can now include utilization rates of other health resources.

An example of the potential of integrated datasets was illustrated in a study about the relationship between primary care appointment no-show rates and emergency departments utilization by Nguyen & DeJesus (2010). To examine this relationship, the authors needed to integrate data from the primary care providers with data from emergency departments. In doing so, they could find that that an increase of patient no-shows in primary care clinics is associated

with an increase in emergency department visits for patients who more often no-showed to primary care appointments. Due to this discovery of this relationship between a program-level measure and a system-level measure, the authors could offer potential new strategies for improving the appropriate use of emergency departments through initiatives at the primary care level to decrease appointment no-shows.

Secondary use of health data is the use of health data for a purpose beyond what the data was initially collected for. Secondary use of health data is valuable for academic research, the evaluation of health programs, and the evaluation of the health system (van Panhuis et al., 2014). Recently, there has been a call to increase the sharing of health data for secondary purposes (Walport, M., & Brest, P., 2011). Being able to link different types of secondary health data allows researchers to "study the effects of health determinants on health services utilization and health outcomes, study the effects of health events and health care utilization on factors such as employment, income and psychosocial characteristics, and [to] make regional and inter-provincial comparison." (Kephart, 2002, p. 10). This would be beneficial in creating additional evaluation insights in program- and system- level evaluation efforts.

Accessing and integrating system-level and program-level datasets containing health utilization data and patient clinical indicators has proven to be very challenging (van Panhuis et al., 2014; Kephart, 2002). In my experience, this has resulted in isolated silos of work resulting in incomplete pictures of the program, the system, and the patient experience. In my view, evaluations of health programs in Alberta currently lack examination of the relationship between the patient, and their system-level and program-level datasets; there are individual lines of communication between each patient and the system or program, but no line of communication between the system- and program levels. Integrating program and system-level datasets can help

to bridge this gap of communication. From my work with these datasets, I experienced the five challenges noted by Kephart (2002) in accessing and integrating program- and system-level data: (a) approval process and contractual agreements, (b) standardization of administrative data, (c) data extraction and linkage procedures, (d) derived variables measuring health services utilization, and (e) data security and access. I will outline how these challenges apply to the Alberta healthcare data context.

**Approval process and contractual agreements**. Administrative databases are held by different data custodians, who are defined as those responsible for the secure collection and safe-keeping of patient-identifiable datasets in accordance with Alberta's Health Information Act (Government of Alberta, 2016). Data custodians range from organizations, such as Primary Care Networks, to individual physicians. To link datasets together, each data custodians must enter into data sharing agreements, and consent to the use of their patients' data for research purposes on a study-by-study basis. Obtaining permissions from all relevant parties can be challenging and time consuming.

**Standardization of administrative data**. There is no standardized central EMR system in Alberta; each clinic is free to choose a different EMR for use. As consequence, sharing and comparing of data from different EMRs require additional time standardizing fields and gathering data from multiple systems. Thus, integrating data is not a simple process of linking patient identifiers, but ensuring variables of all datasets meet some standard before being integrated.

**Data extraction and linkage procedures**. Pathways for data access are not apparent or easy to navigate, particularly for external researchers. AHS has gone through several rounds of restructuring in the past years, leading to dead ends and out-of-date information for data access; it is difficult to find the contacts that will be able to provide the requested datasets. Additionally, some data may be in paper charts, which will need to be digitized prior to storage and linkage. This adds another layer to the challenges of integrating data.

**Derived variables measuring health services utilization**. Different datasets are collected in different ways. A lot of the datasets are derived from calculations of other data. This makes it difficult to understand exactly where data has originated from, and which data definitions are used, which could affect the way that data is analyzed, and how the results are interpreted.

**Data security and access**. Data with patient identifiers, which are required to link patients across different datasets, is even more difficult to obtain due to strict requirements to protect patient privacy, laid out in the Health Information Act (Government of Alberta, 2016). This barrier decreases the number of researchers and evaluators who would have access to data for integration.

Due to the five identified barriers, most system-level evaluations often do not examine individual programs that contribute to the functioning of the system. Conversely, program-level evaluations focus on the program itself, but often do not examine the system in which the program exists. There is a gap between the program-and system-level evaluations, especially when programs are developed to help address system issues, such as health impacts. Integrating these datasets and examining the effect of the program at the system level will be helpful in

creating a fuller picture of the patient experience, the effect of a program within the greater health system, and further inform both clinical and policy decision-making.

These barriers to data access may have also limited the amount of published research that integrated program and system-level datasets (Kephart, 2002; van Panhuis, 2014). Furthermore, the existing barriers limit findings in studies that would require integration (Forbes et al., 2015). A main barrier to data integration is the requirement of obtaining the agreements between the multiple data custodians. Under current legislation and AHS guidelines, a full ethics approval and set of research agreements between the custodians of the data need to be signed for research projects to proceed; this includes any evaluations that will be published in peer-reviewed journals. Many projects examining the effectiveness of health programs are completed in a program evaluation context for internal quality improvement, without a need for external publication. The findings of many program evaluations will not be published nor made publicly available when done in this setting. For example, most the projects I have been involved in while working as an evaluator with Alberta Health Services and the Primary Care Networks have not been published outside of internal reports, and are not available to the general public. Few evaluators working in the health sector of Alberta have the resources to pursue publication of their findings in this environment, leading to a dearth in published articles integrating health data from multiple sources outside of a research context.

Addressing this publication gap through this thesis study could increase evaluation use in health care in three ways: 1) providing a road map for accessing, integrating and analyzing data, 2) increase stakeholder involvement in the use of system-level data through using them at the program-level, and 3) increase access to broad and in-depth evaluation data that inform decision-makers. There is currently no resource that outlines the datasets available to health researchers,

or the process of accessing these datasets. Thus, this study may also have some implications for future process and use; the process of attempting to access and integrate these datasets can drive interest and further innovations into how these datasets can be accessed and used. The resulting road map to the datasets can identify areas for improvement for easier access. Stakeholder involvement is also important for increasing the use of evaluation findings (Alkin, 2011; Johnson et al., 2009). Currently, stakeholder involvement is sorely lacking in system-level evaluations due to researchers working solely with quantitative analysis of secondary datasets. Bringing this data into the context of programs or health initiatives has the potential to make it more meaningful at the program-level, and may result in an increased use of evaluation findings through of clinically actionable recommendations. In addition, providing decision-makers access to broad (i.e., system-level measures) and in-depth (i.e., program-level measures) evaluation data that describes the same intervention can result in them feeling better informed and more likely to use evaluation findings (Christie, 2007). Also important to the process is improving the validity of findings from these datasets by engaging front-line clinicians who have the clinical and contextual expertise to interpret the statistical findings from analyses of integrated datasets through a clinical interpretive lens.

**Engaging the Clinical Perspective as an Interpretive Lens in Health Evaluation**

Clinicians are an important stakeholder group to engage in the processes of health evaluations to improve the validity of interpretations and to access their perspective in the evaluation process. Engaging clinicians can improve the validity of the findings in two ways: by providing appropriate clinical interpretations of statistical findings, and by providing important contextual information that may be missed through routine data collection. Including clinicians in the evaluation process also gives a key stakeholder group a voice in the evaluation process.

This participatory approach can lead to "better data, better understanding of the data, more appropriate recommendations, [and] better uptake of findings" (Guijt, 2014, p.2). Being engaged in the evaluation process increases the likelihood that the clinicians would use the results from the evaluation to inform their practice; it is important to include them so that they are helping *to do* evaluation rather than having evaluation *done to* them.

Statistical significance is often used in health evaluation as evidence of the success of a clinical intervention and its measurement limitations are overlooked. In truth, statistical significance represents a single type of evidence to be used in the evaluation of clinical interventions; the overinterpretation of statistical significance results in an incomplete look at the program outcomes. Bothe & Richardson (2011) outlined and described three different constructs of significance used when assessing the effectiveness of clinical interventions: (a) statistical significance, whether there is a difference between groups; (b) practical significance, the magnitude of the difference between groups; and (c) clinical significance, what the difference between groups mean. Bothe & Richardson (2011) also introduced a fourth type of significance that was previously unnamed: (d) personal significance, whether the difference solved a person's problem. Here, I will present a summary of each type of significance as described by Bothe & Richardson (2011), and how the inclusion of a clinical interpretive lens can help address the limitations of statistical and practice significance.

Statistical significance is an important part of determining the effectiveness of clinical interventions. It provides necessary evidence about the typical effect of interventions at a group-level: the likelihood of obtaining the observed difference between the samples if the populations from which the samples were drawn were identical. In other words, the probability of the differences between groups being real, with the conventional threshold for a real difference being

set at 95%. A weakness of statistical significance is that it does not provide any additional interpretive information about the nature of that difference. For example, the magnitude of the difference, or whether the difference resulted in a change in the groups, cannot be determined through significance testing. To answer these additional questions, measures of practical and clinical significance must be used.

Practical significance is a measure of the magnitude of the difference between groups. The most common measures of practice significance are those of effect size (Cohen, 1988), and a common measure of effect size is Cohen's *d*, which calculates the standardized difference between two means by finding the difference between the means and dividing the result by the pooled standard deviation (Cohen, 1988). The reporting of effect sizes often accompanies statistical tests to provide information about the magnitude of any statistically significant differences, but like statistical significance, it is important not to overinterpret the results. Effect size still does not provide any information about the clinical significance of the differences; what the difference means.

Bothe & Richardson (2011) touch on the complexity of defining, describing and identifying clinical significance; clinical significance is a loaded term with formal and informal usage across the literature. They point out that clinical significance is often conflated with practical significance, in that the use of the word "practical" does not refer to practical significance, but to clinical significance, such as defining clinical significance as "the practical or applied value of a treatment effect" (Houle and Stump, 2008, p. 5 via Bothe & Richardson, 2011) or referring to "clinical or practical significance of results" (Johnson, 2006, p. 21 via Bothe & Richardson, 2011) as one construct. They note a good working definition in the context of clinical interventions in that clinical significance is defined as an intervention "which makes a

real (e.g., genuine…noticeable) difference in everyday life to the clients or to others with whom the clients interact" (Kazdin, 1999, p. 332 via Bothe & Richardson, 2011) and as a "recognizable treatment change that is valued by the clinician, client, and relevant others" (Finn, 2003, p. 215 via Bothe & Richardson, 2011). Bothe & Richardson (2011) further differentiate clinical significance from practical significance by reinforcing that practical significance refers to the *magnitude* of a change (Kirk, 2001) and that clinical significance refers to the *meaning* of a change (Ogles, Lunnen, & Bonesteel, 2001). Determining the meaningfulness of a change is a challenging undertaking, and there are both formal and informal methods doing this.

Some work has been done to try and formalize the process of determining clinical significance through mathematical and statistical methods (Follette & Callaghan, 2001; Kazdin, 1999; Ogles et al., 2001). Bothe & Richardson (2011) pointed out one attempt by Jacobson, Follette, & Revenstorf (1984) that stood out as influential. Jacobson et al., (1984) defined a clinically significant (i.e., meaningful) change as (a) beyond what could be expected from measurement error alone, and (b) represents movement from a dysfunctional range to a normal range. In other words, a clinically significant change is one that is statistically reliable and involves movement from a value in the dysfunctional range to a value in the normal range. A reliability change index (RCI) score is used to measure whether the observed change is statistically reliable, and there are published guidelines for dysfunctional/normal ranges for clinical indicators. This definition allows for the identification of real-world outcomes, but only does so at the group-level and using distributional measures. The limitation of this formal process of determining clinical significance is that clinicians' values judgements about the meaningfulness of treatment results are only included through the guideline dysfunctional-normal ranges. Bothe & Richardson (2011) note that the clinicians' informed judgements are

important for determining clinical significance, and suggest "recognizing and, indeed incorporating as definitive this professional-level aspect of clinical significance" (p. 3). However, they do not go into detail about how to incorporate the clinician-perspective into the process for determining the clinical significance of a change. Identifying the clinical significance is a good measure of change as judged by a clinician or by clinical standards, but does not provide any information about how this change has affected the patients themselves. This is addressed by personal significance.

Personal significance "refers to whether individual patients report feeling improvements that matter to them in the context of their own lives and, more importantly, whether they demonstrate functioning in ways that reflect improvements" (Bothe & Richardson, 2011, p. 3) Bothe & Richardson (2011) go on to differentiate personal significance from clinical significance "by the fact that clinical significance reflects or may incorporate the clinician's judgments, decisions or values, whereas personal significance reflects the client's" (p. 3). Further, they note the importance of drawing this distinction and including a measure of personal significance due to the role of the patient in determining the success of the intervention; they are the ones who are best able to judge the meaningfulness of a change due to the intervention. Thus, for a measurement of clinical outcomes to be considered complete, it should have measures from the clinician-perspective as well as measures from the patient-perspective.

**Need for the Study**

Current approaches to healthcare evaluations limit the ability to examine the relationships among program- and system-level variables. These relationships have the strong potential to provide insights that are otherwise inaccessible. At the system-level, measures are focused on performance indicators and results that lend itself for accountability purposes (e.g., Clinkscales

et al., 2015). At the program-level, measures are focused on patient clinical indicators appropriate for the program and are used for improvement purposes (e.g., Kastorini et al., 2011). What is missing is linking the program-level measures to the system-level measures that would be available through the integration of the datasets. This study responds to three pressing needs: (1) generate new insights about the complex relationships between programs and the broader health system they reside in (van Panhuis et al., 2014; Kephart, 2002), (2) offer procedures for mitigating barriers to integrating these datasets (Kephart, 2002; Nguyen & DeJesus, 2010), and (3) involve the clinical perspective in evaluations to increase the validity of interpretations of statistical findings (Bothe & Richardson, 2011) and uptake of findings (Guijt, 2014) in healthcare. This study provides an illustrative example of an evaluation approach generating an initial understanding of the complexities of health program and system outcomes. Specifically, this study will identify the limitations of individual program- and system- datasets and new evaluative insights that are possible with integrated dataset.

**Research Questions**

The overall research question is: What is the impact of integrating datasets and adding the clinical perspective on the new evaluative insights? This research question will be answered through answering three sub-research questions: (1) from the interpretation of the Program Outcomes Dataset, which clinical indicators are improved by the participation of patients with metabolic syndrome in a lifestyle intervention program (CHANGE)? (2) From the interpretation of the System Outcomes Dataset, does participation of patients with metabolic syndrome in a lifestyle intervention program decrease the number of emergency department visits? (3) From the interpretation of the Integrated Dataset, to what extent can the clinical indicators related to metabolic syndrome predict the number of emergency department visits of MetS patients?

## Chapter 3: Methods

This chapter outlines the study approach and dataset descriptions, analysis and interpretation procedures, and ethical considerations and efforts. First, I introduce the quantitative approach and justify its use for the study design to address the research questions. Second, I describe the datasets in terms of variables and preparation for analysis. Third, I outline the procedures involved in analyzing the datasets separately and integrated, followed by the procedures for clinical interpretations of the statistical findings. Finally, to conclude this section, I explain the ethical considerations of this study and my efforts to mitigate ethical concerns.

### Quantitative Study Approach and Study Designs

A quantitative approach is appropriate when considering the purpose of the study, which is to illustrate evaluative insights required for understanding complex relationships among program- and system-level outcomes from separate existing datasets. The Program Outcomes Dataset, System Outcomes Dataset and Integrated Dataset all contain quantitative variables, which are examined for patterns by performing (1) separate analyses, looking at variables pre- and post-intervention, and (2) an integrated statistical analysis where the relationships among variables will be explored. The findings from these two analyses will be compared to identify new evaluative insights possible through an integrated dataset.

The initial, statistical interpretation of the dataset analyses will be followed by a clinical interpretation, to overcome the limitations inherent with statistically significant differences in a healthcare setting. Statistically significant differences may not imply a clinically significant difference; for example, there may be no statistically significant difference in a patient's BMI from baseline to 12-months, but a dietitian would consider a patient maintaining a constant BMI over one year to be a clinically significant outcome. The author of this study is not a medical

clinician who can make judgments regarding clinical significance; to enhance validity of the study findings, there is a need for expert opinion on the clinical interpretations of the findings to explore clinical significance and implications (Bothe & Richardson, 2011).

To answer the overall research question, the evaluative insights that are generated by analysis of the separate datasets will be compared with the evaluative insights generated by the analysis of the integrated dataset. Due to different types of analyses possible with an integrated dataset, there are two separate designs for evaluation of this study: one for the separate analyses of the individual datasets and another for the integrated dataset.

The separate analysis of the individual datasets employs a quasi-experimental single-group interrupted time-series design (Creswell, 2009). This is appropriate for the analysis of separate datasets, since the data contains variables measured from patients of CHANGE at two different points in time: baseline and 12-months. The data from the two datasets are analyzed separately to answer SQ1 and SQ2. CHANGE patient recruitment included a level of self-selection, as all patients who meet the inclusion criteria could participate in CHANGE if they wished; there was no random assignment and no control group.

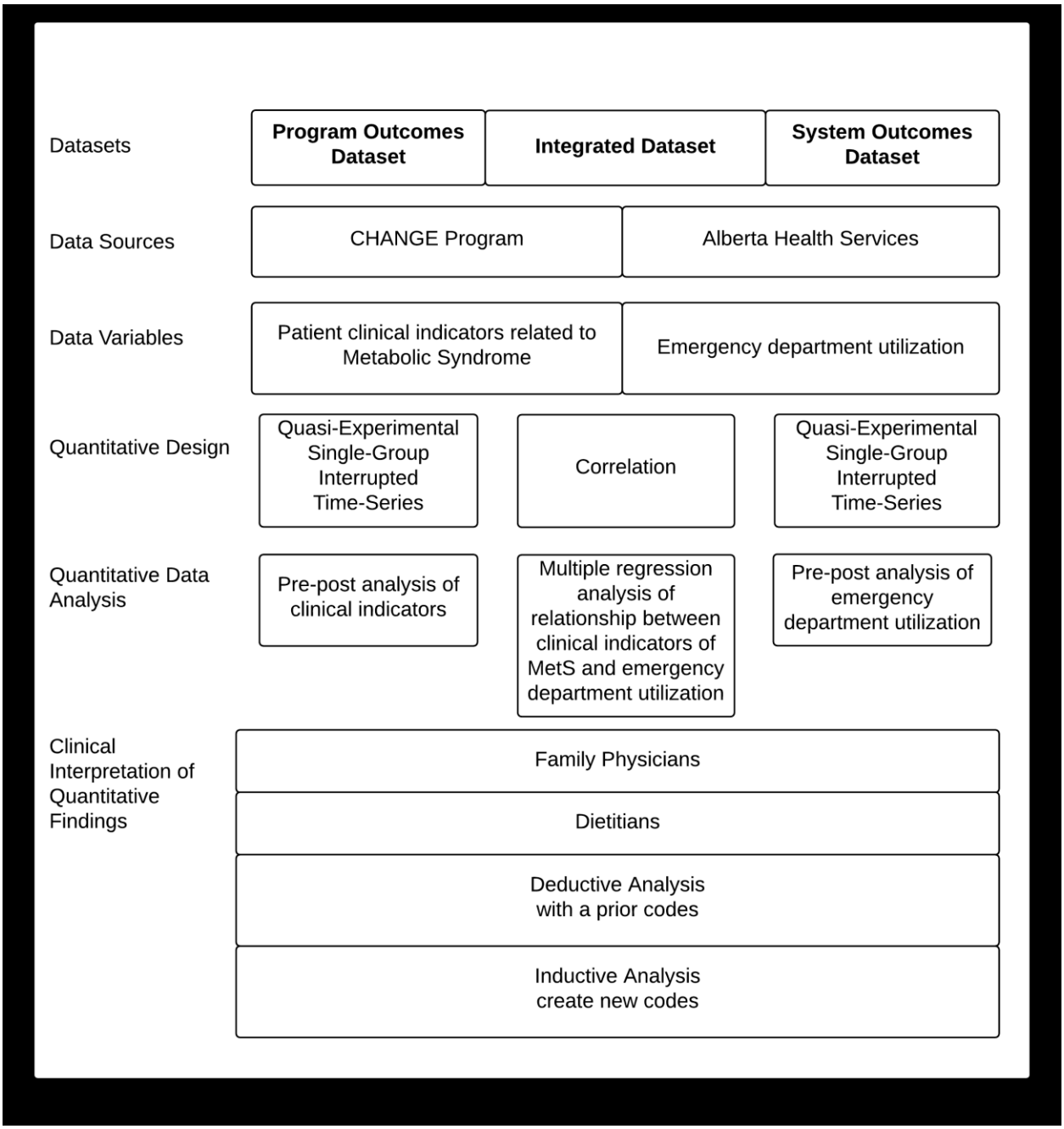
The integrated analysis employed a correlational design (Creswell, 2009) to answer SQ3 by determining the relationship between MetS clinical indicators and emergency department utilization of patients with MetS. This design was chosen to determine the relationship between the clinical indicators collected at the program-level with the system-level measure of ED utilization. This component of the study demonstrates an analysis that is only possible with an integrated dataset.

The analysis of the datasets separately and then after integration will demonstrate what initial analysis is possible with the separate datasets, and what new evaluative insights are

generated by analysis of an integrated dataset. Specifically, a regression analysis will be possible with an integrated dataset; this will identify the relationships among the variables of the program-level and system-level outcomes.

In addition to analysis of datasets, an additional procedure will need to be taken to address the author's lack of clinical knowledge. Other research teams which study health interventions usually involve a clinical expert to provide clinical expertise required for clinical interpretation of statistical results. Although there may be statistically significant findings, these findings may not be clinically significant (Bothe & Richardson, 2011). This may be due to the extent and size of the differences, or due to the relevance of the findings in actual clinical care. Formal clinical significance will be calculated using reliability change index (RCI) (Jacobson & Truax, 1991). In this study, to further increase validity of findings within the context of CHANGE and MetS, dietitians, kinesiologists, and family physicians involved with the CHANGE program were recruited to participate in key informant interviews (Bothe & Richardson, 2011). In these interviews, the clinicians were asked to apply their clinical perspective and expertise in interpretation of the statistical findings. The overall study design is summarized in Figure 1.

Figure 1
*Overall Study Design*



**Dataset Description and Access Procedures**

The data used in this study was collected from patients with MetS attached to PCN physicians; the patients completed the full 12-months of the CHANGE program (n=85) between October 18, 2012 and December 31, 2014. The present study used this previously collected data in a secondary manner. These patients were identified by their respective PCN physicians as candidates for the CHANGE program, and underwent an intake process with the CHANGE program coordinator to screen for the inclusion/exclusion criteria (Table 1). Some patients who have scores close to, but not quite meeting, the inclusion/exclusion criteria may also have been admitted into the program at the discretion of the project coordinator or physician; thus, some indicators of patients admitted into the program may fall outside of the strict inclusion/exclusion criteria.

A total of 85 patient identifiers from the Program Outcomes Dataset were sent to Alberta Health Services to integrate said dataset with the System Outcomes Dataset. Of these 85 patients, 50.6% were male, and had ages ranging from 37 to 75, with a mean age of 60.4 (S.D. = 8.7). This high mean age of participants is likely influenced by the CHANGE program being offered during regular office hours, limiting participants to those who had the time to attend sessions during the day, and introducing a barrier to entry for those who work regular hours.

**Variable description**. There are two datasets accessed: The Program Outcomes Dataset and Systems Outcomes Dataset. The Program Outcomes Dataset consisted of data collected by a program coordinator as a part of the CHANGE program; this data includes patient age, sex, and 10 clinical indicators related to MetS (Table 3). Clinical indicators requiring lab work were gathered into the dataset directly from the family physicians' electronic medical records

Table 3
*Clinical Indicators and Guidelines of Metabolic Syndrome*

| Clinical Indicators | Clinical Guidelines |
|---|---|
| Body Mass Index (BMI) (kg/m$^2$) | Underweight: <18.5 |
| | Normal: 18.5 – 24.9 |
| | Overweight: 25 – 29.9 |
| | Obesity class I: 30.0 – 34.9 |
| | Obesity class II: 35.0 – 39.9 |
| | Obesity class III: ≥40.0 |
| Waist Circumference (WC) (cm) | -- |
| Blood Pressure (BP) (Systolic / Diastolic) (mmHg) | Normal: <140 / 90 |
| Fasting Blood Glucose (FBG) (mmol/L) | Normal: 3.3 – 5.8 |
| Hemoglobin A1C (HgA1C) (%) | Normal: 4 – 6 |
| Triglycerides (TG) (mmol/L) | Normal: < 2.20 |
| Creatinine (Crea) (µmol/L) | Normal: 70 – 120 |
| Total Cholesterol (TC) (mmol/L) | Desirable: <5.2 |
| | Borderline high: 5.2 – 6.2 |
| | High >6 |
| Low-Density Lipoprotein (LDL) (mmol/L) | High risk patients: <2.0 |
| | Low risk patients: <3.37 |
| High Density Lipoprotein (HDL) (mmol/L) | Low: <1.00 |

*Adapted from Council for Continuing Pharmaceutical Education (2013)

Administrative data related to ED utilization in the time period between January 1, 2011 and December 31, 2014, by patients in the CHANGE program (System Outcomes Dataset) was pulled to capture 1-year pre- and 1-year post- enrolment in the program for all patients. A data request was sent to AHS to obtain this specified System Outcomes Dataset. The AHS analyst tasked with the data pull required the Patient Health Numbers (PHNs) of patients in the Program Outcomes Dataset, as well as the Program Outcomes Dataset itself, to link the two datasets to create the Integrated Dataset. Once linking was completed and confirmed, the analyst replaced the PHNs with generic patient identifiers. A list of the variables pulled is available in Appendix A.

**Dataset preparation**. Accessing the datasets for this study required consent and data sharing agreements with the various custodians of all data. For the Program Outcomes Dataset, consent was required from the CHANGE national research team. For the Systems Outcomes Dataset, consent and a data sharing agreement between the current study team and AHS was required. Once all agreements were signed, the Program Outcomes Dataset was encrypted and sent to AHS to integrate with the System Outcomes Dataset. The AHS analyst performed integration of the datasets and removed all patient identifiers, replacing them with generic unique identifiers before sending the datasets to the current study team.

Once datasets were obtained, the data then needed to be prepared and cleaned before analysis. All datasets were imported into SPSS 17.0 for analysis. Descriptive statistics of age, sex and the 10 clinical indicators measured will provide a clinical profile of patients entering the CHANGE program. The data was examined for completeness and entry errors through analysis of frequencies and descriptive statistics (e.g., means, standard deviations and ranges). Although these 85 patients were considered to have the most complete datasets, there were still some data

missing at baseline and 12-month post intervention. Analysis that required both baseline and 12-months post-intervention data excluded those cases involving incomplete data; the n reported in the results reflects the number of cases which had complete baseline and 12-month post measures.

***Assumption testing for t-test***. There are four assumptions that must be met to perform a valid t-test. Here I list the required assumptions and the procedures for how they were addressed: (a) Dependent variable (DV) is on a continuous scale: The DVs were all continuous variables; (b) Independent variable (IV) consists of two categorical matched pairs: The IV is time period (baseline and 12-months) and cases were matched by patient; (c) There are no significant outliers in the differences between the matched pairs: Outliers were identified using Median Absolute Deviation (MAD) (Leys, Ley, Klein, Bernard, & Licata, 2013) and removed from the statistical tests to meet the required assumptions for the statistical tests; (d) The distribution of the differences in the DVs is approximately normally distributed: After outliers were removed, normality plots of each DV were created to confirm normal distribution.

***Assumption testing for multiple linear regression***. There are six assumptions that must be met to perform a valid multiple linear regression: (a) ratio of cases to IVs, (b) absence of outliers among the IVs and on the DV, (c) absence of multicollinearity and singularity, (d) normality, linearity, homoscedasticity of residuals, (e) independence of errors, and (f) absence of outliers in the solution. These assumptions were investigated and were met by the datasets; outliers were already removed for this analysis by initial CHANGE researchers.

## Statistical Analysis Procedures

There are three types of statistical analyses to be done, one for each of the quantitative datasets. I will describe the statistical analyses procedures for conducting the paired t-tests with the Program Outcomes Dataset, conducting the paired t-tests for the System Outcomes Dataset, and conducting the multiple linear regression for the Integrated Dataset.

**Analysis of program outcomes dataset**. With the Program Outcomes Database, the MetS clinical indicator values at baseline and 12-months served as the pre- and post-enrolment values compared to determine the program's effect on each enrolled patient. Repeated-measures t-test (Tabachnick & Fidell, 2007) was used with patients' baseline and 12-month clinical indicator values to check for differences in each of the clinical indicators pre- and post- program enrolment. Since each of the clinical indicators measured a different component of MetS, it was appropriate to examine the effect of the program on each clinical indicator separately. Effect size was calculated using Cohen's *d* (Cohen, 1988). Formal clinical significance was calculated using reliable change index (RCI) (Jacobson & Truax, 1991).

In addition to these statistical tests, a descriptive analysis of the patients' MetS status was examined to provide some clinical context to the statistical results. The patient's MetS status at baseline and 12-months were identified to determine how many patients still meet the criteria for MetS after 12-month program participation.

**Analysis of system outcomes dataset**. The System Outcomes Dataset will provide the number of emergency department visits between one-year pre- to one-year post-enrolment in the program. To identify any differences in ED utilization pre-and post-program enrolment, another set of dependent-measures t-tests were completed on patients' CHANGE ED utilization rates to identify any statistically significant differences. Similar to the Program Outcomes Dataset, effect size was calculated using Cohen's *d*(Cohen, 1988), and formal clinical significance was calculated using reliable change index (RCI) (Jacobson & Truax, 1991).

**Analysis of integrated outcomes dataset**. To determine the relationship between MetS indicators and ED utilization, a standard multiple linear regression (Tabachnick & Fidell, 2007) was performed. Due to existing clinical relationships between variables related to MetS, the correlations between these variables were calculated to identify which variables would be most appropriate to enter into a standard regression.

A new set of 10 composite variables were calculated by taking the difference between baseline and 12-month scores, creating a gain score to incorporate into a single multiple regression (Gelman & Hill, 2006). Although there is currently no underlying theory that explains the relationship between the clinical indicators of MetS and the rate of ED utilization, it was decided based on calculated correlations that the variables entered into the regression model would be the five clinical indicators used as inclusion criteria into the study: (1) Blood Pressure (both systolic and diastolic), (2) Fasting Blood Glucose, (3) Triglycerides, (4) High-Density Lipo-Protein and (5) Waist Circumference. In addition to these variables, age and sex were added to the regression model to determine how much those two additional demographic factors were related to emergency department utilization.

## Clinical Interpretations of Statistical Findings

Clinical interpretation is necessary to increase the validity of the study findings. Key informant interviews were undertaken with clinicians involved with CHANGE who could clinically interpret the statistical findings. Including clinicians in this way develops a clinical interpretive lens through which the statistical findings can be examined which increases the validity of the study findings and allows for examination beyond statistical significance and towards clinical significance (Bothe & Richardson, 2011).

**Participant sampling**. Participants were recruited per their ability to (a) apply their clinical knowledge to interpret the quantitative findings from analysis of the three datasets, (b) properly contextualize the findings within the context of the CHANGE program and the treatment of MetS, and (c) provide insight into the implications of these findings. Due to their clinical expertise with MetS and their familiarity with the CHANGE program, family physicians, kinesiologists, and dietitians involved with the CHANGE program were recruited as key informants. They undertook 60-minute key informant interviews in which they provided their expert opinion on the validity and implications of the statistical findings. An effort was made to recruit at least one clinician in each field to capture a broader clinical perspective across different practices and contexts; however, only 3 family physicians and 3 dietitians consented to being interviewed. Recruitment was completed through a recruitment email, which was sent to potential participants by the executive director of the PCN. This email contained details of the study and contact information of the author, who the clinicians directly contacted to participate.

**Key informant interview procedures**. Initial development of the interview protocol was performed by examining the author's limitations in clinical interpretation of results from the statistical analyses, and identifying appropriate questions obtain clinicians' interpretations of the findings in both the separate and integrated analyses. Participants were presented the preliminary findings from the statistical analyses of the three datasets, and asked about (a) the clinical significance of the findings from the statistical analyses, from their perspective, for clinical practice, (b) unanticipated results from these findings, (c) how these results might inform their practice, and (d) the next steps for research in this area. Specific probes targeted elaboration based on the statistical results that were notable, and additional probing questions were asked based on participant responses.

Potential CHANGE clinician key informants were contacted via email sent by the management team of the PCN (Appendix B). The clinicians interested in participating in the study then contacted the author directly, after which a meeting was at a mutually convenient time and location where private conversation was possible. Participants were reminded that their participation was voluntary and anonymous to encourage their providing of unrestricted expert opinions prior to obtaining informed consent (Appendix C).

One-on-one key informant interviews with clinicians were then conducted following a pre-determined protocol (Appendix D). At the end of each interview, a summary of about the interview was provided by the interviewer; this member checking provided an opportunity for the clinician to correct any misinterpretations, thereby decreasing the likelihood of comments being misrepresented (Shenton, 2004). The key informant interviews were recorded with an audio recorder and transcribed for analysis. After each interview, I listened to and reflected on the session, highlighting key points such as the perceived impact of specific clinical indicators.

Preliminary findings from these early reflections were used to inform future interviews; for example, if there was a discrepancy between responses from interviewees, this was explored further in future interviews to determine the reasoning behind the discrepancy (Shenton, 2004). Transcripts of the key informant interviews were imported into QDAMiner4 Lite for analysis. The participants' responses to the questions were analyzed using deductive and inductive thematic analysis, with the goal of finding themes and codes validating and exploring the clinical and policy implications of findings. A deductive analysis (Braun & Clarke, 2006) was completed using a list of anticipated codes that were developed a priori. This analysis was used to determine which clinical indicators showed changes that were clinically significant, whether these findings were surprising, and whether these findings would alter current clinical practice. Following the deductive analysis, additional codes were created through inductive examination of the transcripts; the codes were modified as required, providing further depth and context to the deductive analysis (Braun & Clarke, 2006).

**Ethical Considerations**

The author requested and received approval from the Health Research Ethics Board at the University of Alberta for this study. Research agreements were made with Alberta Health Services, which outlined restrictions in use of the requested datasets. There were three key ethical issues that were considered, and efforts made in which to mitigate for this study: (a) confidentiality of participants, (b) security of data, and (c) dual role of researcher.

**Confidentiality of participants**. Confidentiality was particularly important due to the potential for harm of the participants through unintended access to identifying data. In this study, there were two groups of participants that required consideration – patients whose health data were involved in the analysis of the datasets, and clinical experts involved in the key informant interviews. For the patients, Patient Health Numbers (PHNs) were required to integrate the Program Outcomes Dataset and the System Outcomes Dataset. Since the PHN uniquely identifies patients within the Alberta health care system, any communication of PHNs had the potential for a breach of confidentiality unless performed securely. To minimize risk, a data sharing agreement was required for the sharing of any data containing PHNs; thus, the present study team and AHS entered into a data sharing agreement which outlined patient variables that would be shared (Appendix A) and the responsible parties. The PHNs were replaced with generic identifiers during data integration before being sent to the present study team to anonymize the data. This meant that although the data was linked to the same patient, it was no longer possible to identify the patient based on this dataset received by the researcher. As this research involved secondary analysis of de-identified patient data, with no direct patient contact, this was determined to be a low-risk study for which a waiver of informed patient consent was granted.

Clinical experts who participated in key informant interviews were not identified in this study, and precautions were taken so that they remained anonymous. To maintain the anonymity of the participants and to help the reader to easily distinguish between the two groups of clinicians, pseudonyms were assigned as follows: names starting with D are dietitians, those starting with P are physicians. Audio recordings of the interviews were kept in encrypted files on secure hard drives, accessible only by the author and his supervisor. Due to the small number of

clinicians involved with the CHANGE program, careful consideration was made in the reporting of quotes to minimize risk of the informants being identified.

**Security of data**. As individual physicians were the custodians of patient data contained in the Program Outcomes Dataset, any breech of the datasets could have resulted in a breach of patient and physician confidentiality, which ultimately was the responsibility of the custodian. Access to the program-level data thus required physicians involved with CHANGE to consent to usage of their patients' data for research. The data also had to be stored securely. The datasets obtained for this study were stored on encrypted hard drives kept in secure locations to mitigate the risk of data breach.

**Dual role of researcher**. A possible conflict of interest was that the author of this study was also employed by the PCN. The two roles remained separate by clearly outlining the responsibilities of each role, that of an employee and a researcher, prior to the onset of this study. The potential for bias during key informant interviews was mitigated through member-checking of interview results and using statistical methods for eliminating outliers. A thesis supervisor was involved to help mitigate any conflicts of interest by providing oversight over the research and an external view over how some relationships and findings may be perceived.

## Chapter 4: Study Findings

This chapter is organized into three sections corresponding to the three major findings from the analysis of the three datasets. For each major finding, I present the evidence from the statistical analyses that served as the basis for clinical interpretations that are presented subsequently. The first finding examined the impacts of participation in Canadian Health Advanced by Nutrition and Graded Exercise (CHANGE) participation on clinical indicators related to metabolic syndrome (MetS) from the Program Outcomes Dataset. The second finding examined the impacts of CHANGE participation on emergency department (ED) utilization from the System Outcomes Dataset. The third finding reports on the relationships between clinical indicators of MetS and ED utilization from the Integrated Dataset.

### Finding 1: Impacts of CHANGE Participation on Clinical Indicators

Statistical analysis of the Program Outcomes Dataset revealed statistically significant improvements in 8 of 11 clinical indicators of MetS after CHANGE participation. Statistical significance suggests that the patients' clinical indicators improved from baseline to 12-months, but it does not provide any information as to the clinical significance of this change. To identify the clinical significance of these improvements, the statistical analysis was followed by the calculation of formal clinical significance through reliable change index (RCI) (Jacobson & Truax, 1991) and clinical interpretation of the statistical findings by dietitians and physicians through one-on-one key informant interviews. Whereas RCI provides a quantitative measure of clinical significance, the key informant interviews include the clinical context and perspective into identifying clinical significance.

The key informant interviews revealed clinically significant improvements in 5 of 11 clinical indicators of MetS after CHANGE participation. The clinicians also provided some

information about how they use the threshold values in published clinical guidelines (Table 4) and their own clinical experience to identify clinical significance, brought forth an interpretive concept of the clinical significance of maintaining indicators at baseline levels, and highlighted the treatment-focused approach as the emerging standard to treating MetS.

All clinicians to cite clinical guidelines for the normal values of each MetS indicator, which guided their choice of important clinical target thresholds, and how they assessed clinically significant changes related to MetS. Some of the clinical indicators measure a similar indicator of health (e.g., LDL and HDL are both measures of the lipid profile), thus the clinical indicators are grouped together into four clinical categories: body mass, blood pressure, lipid profile, and blood glucose. Some notable trends about how the clinicians applied guidelines emerged from the analysis of these descriptions: (1) the clinicians did not comment on the clinical significance of waist circumference, instead looking to body mass index (BMI) as an indicator of body mass, (2) the dietitians did not use blood pressure in their clinical practice, (3) Hemoglobin A1C (HbA1C) is the preferred long-term indicator of glucose levels, and (4) High-density lipoprotein-Cholesterol ratio is the preferred long-term indicator of the lipid profile. The clinical guidelines for the normal values of each indicator, and a summary of how the clinicians described applying the guidelines are presented in Table 4.

Table 4

*Clinical Guidelines for Clinical Indicators of MetS and how Clinicians Determine Clinical Significance*

| Clinical Category | Clinical Indicators | Clinical Guidelines | Clinical Usage |
|---|---|---|---|
| Body Mass | Body Mass Index (BMI) ($kg/m^2$) | Underweight: <18.5<br>Normal: 18.5 – 24.9<br>Overweight: 25 – 29.9<br>Obesity class I: 30.0 – 34.9<br>Obesity class II: 35.0 – 39.9<br>Obesity class III: ≥40.0 | Dietitians: 3-5% improvement<br>Physicians: 1 point, or 5% improvement. |
| | Waist Circumference (WC) (cm) | -- | |
| Blood Pressure | Blood Pressure (BP) (Systolic / Diastolic) (mmHg) | Normal: <140 / 90 | Physicians: Decrease of 7.5 in hypertensive patients. Decrease of 5 in normal patients. |
| Lipid Profile | Triglycerides (TG) (mmol/L) | Normal: < 2.20 | Prefer to use HbA1C. |
| | Total Cholesterol (TC) (mmol/L) | Desirable: <5.2<br>Borderline high: 5.2 – 6.2<br>High >6 | Prefer to look at HDL-Cholesterol ratio. |
| | Low-Density Lipoprotein (LDL) (mmol/L) | High risk patients: <2.0<br>Low risk patients: <3.37 | Prefer to look at HDL-Cholesterol ratio. |
| | High Density Lipoprotein (HDL) (mmol/L) | Low: <1.00 | Prefer to look at HDL-Cholesterol ratio. |
| Blood Glucose | Fasting Blood Glucose (FBG) (mmol/L) | Normal: 3.3 – 5.8 | Prefer to use HbA1C. |
| | Hemoglobin A1C (HbA1C) (%) | Normal: 4 – 6 | |

*Adapted from Council for Continuing Pharmaceutical Education (2013)

One of the dietitians, Drew, brought forth an emergent concept for interpreting the statistical results, baseline maintenance. She talked about how maintaining the indicators at baseline level may be considered a clinical success because of the natural progression of MetS towards chronic diseases, and the eventual worsening of clinical indicators in the absence of intervention. Therefore, being able to demonstrate that the indicators have either improved or stabilized could potentially be clinically significant for understanding the impact of the CHANGE program. This idea, that maintenance of indicators may be a clinically significant result, was presented to the physicians during their interviews. The physicians agreed with this idea, with one physician noting: "…the natural process for diabetes is to worsen with time…" (Presley). Applying this emerging concept of baseline maintenance to the interpretation of the statistical findings, we cannot assume that stability in the indicators (that is non-significance) means that there is no impact of CHANGE participation. In fact, stability in clinical indicators may signify a halt in the progression of MetS, and would be considered a clinically significant outcome from CHANGE participation.

Another emergent concept, treatment-focused approach, was brought forth by two of the physicians, Presley and Parker. These two physicians identified two divergent approaches of using clinical thresholds of the lipid profile and treatment goals: (1) guideline-focused, and (2) treatment-focused. Whereas being guideline-focused is an older standard involving focusing treatment on improving clinical indicators until they meet specific guideline values, being treatment-focused is the emerging standard that is currently being taught to resident physicians involving emphasizing the importance of ongoing treatment of lipid indicators, without the primary focus being on target thresholds.

Presley elaborated that the practice of being guideline-focused is falling out of favor, as newly graduated family physicians have been trained to be treatment-focused. Patrice is a recently graduated family physician who confirmed being trained with the treatment-focused approach; he commented on the advantage of being treatment-focused, stating that cholesterol regulation is still an ongoing field of study, with the relationship between treatment, clinical indicator values, and outcomes still unclear, which lead him not to focus on numeric values:

> …when we see patients improve their lifestyle, you can see that improvement [translate into] increase in HDL and reduction LDL, and so both of those go in the equation for total cholesterol…but I think that our body's cholesterol regulation and the importance of these numbers is being found out more and more in medicine, and that's why I guess they're getting [physicians] away from the surrogate markers and numbers…as long as the intervention is happening then that's when we're generally happy` in terms of total lipids and cardiac.

Treatment-focused physicians, while they acknowledge there are target thresholds in published guidelines, do not focus too closely at the numeric values when adjusting treatment or determining improvement. Rather, they consider patient compliance with their treatment plan as an indicator of success. In the CHANGE program, their treatment plan includes compliance with their medication, and taking steps to improve lifestyle factors such as diet and exercise, as recommended by the dietitians and kinesiologists. Thus, there may not be a clinically significant improvement in the indicators, but there may be improvements in lifestyle factors; a treatment-focused physician would consider this a clinically significant change. Applying this emerging concept of a treatment-focused approach to the interpretation of the statistical findings, we know that we ought to look beyond the impact of the CHANGE program on values of the clinical

indicators, and towards the impact that the CHANGE program had on patient compliance with their treatment plans, especially the diet and exercise components, which are designed to improve those clinical indicators.

To delve deeper into the impacts of CHANGE participation on the clinical indicators of MetS, I will first report the statistical evidence supporting the finding that CHANGE participation resulted in a statistically significant improvement in 8 of the 11 clinical indicators related to MetS. Then, I will present the evidence from the key informant interviews for each group of indicators that support the finding that CHANGE participation resulted in a clinically significant improvement in 5 of the 11 clinical indicators related to MetS, and that the maintenance of the remaining 6 clinical indicators at baseline levels is also a clinically significant event.

**Statistical findings of the effect of CHANGE on MetS**. The statistical analysis of the Program Outcomes Dataset revealed 8 of the 11 clinical indicators related to MetS improved after participation in the CHANGE program. There were statistically significant differences found for 8 of 11 clinical indicators of MetS when comparing values at 12-months to baseline; the exceptions to this trend are one of the glucose measures, Fasting Blood Glucose, and two of the lipid profile Triglycerides and Total Cholesterol. Further, the large effect size values ($d$ =.73 to $d$ =1.51) suggest a high level of practical significance for all 8 of the statistically significant differences (see Table 5). This is especially true for the indicators of body mass and blood pressure, further suggesting that the CHANGE program was particularly effective at improving these two aspects of MetS within a 12-month time period. Body mass and blood pressure indicators were also found to be formally clinically significant (RCI≥1.96)

Table 5
*Clinical Indicators at Baseline and 12-months Post-Enrolment*

| | | | Baseline | | 12-Months | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| Clinical Category | Clinical Indicators | N | M | SD | M | SD | T | P | d | RCI |
| Body Mass | Weight*** | 75 | 89.13 | 14.67 | 85.83 | 15.06 | 5.64 | < .000 | 1.31 | 7.04 |
| | BMI*** | 75 | 31.21 | 3.19 | 30.05 | 3.67 | 5.54 | < .000 | 1.29 | 2.42 |
| | WC*** | 72 | 106.08 | 7.63 | 101.56 | 10.42 | 6.37 | < .000 | 1.51 | 7.14 |
| Blood Pressure | SBP*** | 70 | 129.50 | 13.62 | 124.93 | 11.31 | 3.14 | .003 | .76 | 23.58 |
| | DBP*** | 75 | 78.63 | 8.64 | 75.32 | 9.18 | 3.12 | .003 | .73 | 14.33 |
| Lipid Profile | TG | 76 | 1.727 | 0.66 | 1.727 | 0.68 | 0.002 | .998 | - | 1.13 |
| | LDL* | 78 | 2.57 | 0.82 | 2.16 | 0.098 | 5.41 | < .000 | 1.23 | 0.89 |
| | HDL* | 75 | 1.11 | 0.19 | 1.17 | 0.21 | -3.30 | .001 | -.77 | 0.23 |
| | TC | 80 | 4.22 | 0.97 | 4.20 | 0.95 | 0.204 | .839 | - | 1.17 |
| Glucose | HbA1C* | 70 | 6.34 | 0.65 | 6.12 | 0.61 | 4.54 | < .000 | 1.09 | 0.93 |
| | FBG | 73 | 6.54 | 1.18 | 6.57 | 1.27 | -.294 | .769 | - | 1.17 |

* denotes statistically significant difference ($p < .05$)
** denotes clinically significant improvement (RCI $\geq 1.96$)
*** denotes personally significant improvement (clinicians)

In addition to looking at the clinical indicators separately, I also examined the prevalence of MetS as a diagnosis. At the end of the CHANGE program, there were fewer patients meeting clinical criteria for MetS based on their clinical indicators. Specifically, 44.7% of patients (n=21) who met the clinical criteria for MetS at baseline no longer met the criteria at 12-months post-enrolment (Table 6). These patients had, with regards to clinical indicators, reversed their MetS status. Conversely, 5 patients who did not meet the diagnostic criteria for MetS at baseline ended up meeting the diagnostic criteria at 12-months post-enrolment.

Table 6
*Number of Patients Meeting Clinical Criteria for MetS at Baseline and 12-months Post-Enrolment*

| | | MetS (M12) | | Total |
|---|---|---|---|---|
| | | Not MetS | MetS | |
| MetS (Baseline) | Not MetS | 12 | 5 | 17 |
| | MetS | 21 | 26 | 47 |
| | Total | 33 | 31 | 64 |

**Clinical significance and important thresholds of clinical indicators**. The results of the statistical tests in Table 5 and 4.3 were presented to the clinicians, and they were asked (1) which of the changes in clinical indicators stood out to them as unexpected, and (2) based on their clinical expertise and perspective, which of the changes in clinical indicators from baseline to 12-months represented a clinically significant change. Two themes related to clinical significance emerged from analysis of the interview responses: (1) classification of the differences between baseline and 12-month values as clinically significant, and (2) identification of important clinical thresholds of each clinical indicator. These two themes will be presented sequentially for each of the clinical indicators, grouped by the clinical category that each indicator measures: (1) body mass, (2) blood pressure, (3) lipid profile, (4) blood glucose, and (5) overall MetS status.

***Body mass***. The body mass category of indicators includes weight, body mass index (BMI), and waist circumference (Table 5). Formal clinical significance was found for all three measures of body mass, which converged with the consensus among clinicians that there was a clinically significant improvement in all the body mass measures after CHANGE. The dietitians reported that the improvements in weight and BMI were considered clinically significant when following published guidelines of 3-5% of weight loss being a significant improvement. The physicians agreed that there was a clinical significance with one saying, "…the weight and BMI…it's great to see these numbers. I think the numbers are excellent." (Patrice). Importantly, Patrice also conveyed that there is clinical value in patients moving out of the obesity range (defined as BMI≥30.0), as this decreases risk for chronic diseases:

> I'm looking at the [BMI] here that you've got, going from 31 down to 30, and we use 30 as our higher risk number, 30 and up, and these people were able to really bring themselves out of the risk range there.

None of the clinicians identified any target thresholds for waist circumference, yet all the clinicians interviewed reported that the difference in waist circumference measured (4.52cm) was clinically significant. Patrice elaborated that a decrease in waist circumference helps with managing blood sugar levels, an important step in managing MetS. Similarities between the groups in this regard is not surprising since there appear to be agreement upon guidelines for body mass.

There is evidence of the idea of baseline maintenance being important from the dietitians who voiced that any improvement in BMI could be considered significant, with Devon noting that "changing BMI is hard" and Drew remarking "…any decrease in BMI I think is significant."

There is also evidence of adopting a more treatment-focused approach in their consideration, in that, despite the mean percentage decrease in weight and BMI (3.7) not reaching this 5% target threshold cited by the physicians, the physicians still noted that it is significant to see this much body mass improvement in just one year.

***Blood pressure***. The blood pressure category of indicators includes systolic (SBP) and diastolic blood pressure (DBP) (Table 5). Formal clinical significance was found for both measures of blood pressure. Differences were reported in how the clinical significance of the changes seen in blood pressure was determined among physicians, and between how dietitians and physicians used blood pressure in their practice.

Two physicians reported that the changes seen in blood pressure were clinically significant while the remaining physician (Patrice) reported that it was clinically significant, but

with some qualifiers. Patrice commented that the mean baseline level for both systolic and diastolic blood pressures were already within a normal range, and that the small improvements seen would not be considered clinically significant for a patient whose baseline blood pressure was in the abnormal, hypertensive range (>140/90). The two other physicians aligned with this line of thinking, arguing that the decrease seen in patients with lower baseline blood pressure (from 130/79 at baseline to 125/75 at 12-months) is considered a clinically significant improvement. This viewpoint is noted by Presley, "the blood pressures aren't very high to begin with…so to drop 5 in a fairly normal blood pressure is…a significant effect". For patients with higher blood pressure, a bigger proportionately bigger change will be necessary before it is considered clinically significant. For example, Patrice would like to:

> …see [a decrease of] 7.5 or something…3.3 would be pretty small for somebody who you're trying to reduce them if they're in the hypertensive [high blood pressure] range.

Since the baseline blood pressure was not in the hypertensive range, the physicians considered the decrease in blood pressure to be clinically significant.

It is noteworthy that two dietitians did not feel they were qualified to comment on the clinical significance of changes in blood pressure (Dakota, Devon). The remaining dietitian, Drew, noted that a decrease in blood pressure is a positive outcome, as it "reduces your risk of having strokes and heart attacks". She did not elaborate further on clinical significance. This resulted in a difference in the way that the dietitians responded about blood pressure and how the physicians did. The dietitians did not comment on target thresholds or guidelines for blood pressure, reporting that blood pressure was not in their scope of practice. The physicians reported being experienced in working with blood pressures in their practice, and provided more details on this clinical indicator.

Medication was not explicitly examined in this study, and thus, clinicians were not asked about medications during the interview, however, two of the physicians specifically mentioned the use of pharmaceutical agents for blood pressure management. Parker commented that the changes seen in blood pressure from CHANGE participation approach the changes that would be seen from pharmaceutical interventions. Presley mentioned that the CHANGE program was focused on reducing pharmacological interventions, so an increase in medications was not likely for these patients. In fact, there is anecdotal evidence from Presley that patients decreased medication use throughout the duration of the program:

> …the idea behind the project was that putting the intervention in would enable you to pull off the medications and stay normal…I know from my personal patients, there were some patients that I ended up pulling medication off during the project…

It may not be surprising given the differences in scopes of practice that the dietitians did not specifically mention any medications in detail, but they did mention that decreasing medications would be considered a clinically significant event.

***Lipid profile***. The lipid profile consists of four clinical indicators that describe the cholesterol levels of a patient: low-density lipoprotein (LDL), high-density lipoprotein (HDL), total cholesterol (TC) and triglycerides (TG) (Table 5). These four indicators are related to each other; LDL, HDL and TG are used in the calculation of TC. It is important to note the direction of clinical improvements of these indicators: a decrease in LDL, TC and TC, and an increase in HDL are considered clinical improvements. No formal clinical significance was found for any of the measures of lipid profile. This converged with most of the clinicians' judgements.

Most clinicians reported that the improvements seen in lipid profile from CHANGE patients were not clinically significant. One physician, Parker, had a contrary view to the other

clinicians, commenting that the change in LDL appeared clinically significant; however, he did not elaborate further on the reasons behind it being clinically significant. The interviews focused in on reasons that there may not have been significant improvements in the lipid profile indicators as well as limitations in the analysis of the lipid profile variables.

The clinicians provided two ideas for the non-significance of the lipid profile indicators. The first is that the lipid profile indicators were already at a relatively low baseline value and thus are more difficult to improve significantly. This is evidenced by the clinicians' comments that they expected an improvement in TG from the diet interventions in CHANGE, but that TG values were already low at baseline and may not have further room to improve within a year. Secondly, patient motivation could have decreased over the yearlong program, resulting in a decrease in diet and exercise compliance. There were some comments indicating that since CHANGE was a yearlong intervention, some patients may have lost motivation and reverted to old eating habits. It was also noted by all clinicians that TG is a sensitive test that can be easily impacted by recent meals; recent meals could have affected results by skewing the short-term indicators away from long-term results.

Two limitations to the current lipid profile analysis emerged from the analysis of the interviews: (1) A cholesterol ratio variable was not calculated, and (2) gender differences in lipid profile were not examined. The clinicians all noted that the relationship between indicators of the lipid profile made it difficult to tease out the clinically significant improvements of each individual indicator. One of the dietitians, Drew suggested looking at the ratio between LDL and HDL, which may be more useful than looking at the indicators separately. This is an important alternate indicator because when beneficial HDL increases and the detrimental LDL decreases, TC could remain stable; thus, the beneficial redistribution of cholesterol is not captured when

only examining individual indicators. This sentiment was echoed by a physician, Presley, who pointed out the importance of looking at the redistribution of cholesterol (a decrease in LDL and an increase in HDL), rather than total cholesterol:

> Total cholesterol…I didn't actually think that was going to change, so I think it's actually more the redistribution…

Presley also pointed out a gender difference in LDL values, and suggested looking at a gender breakdown. He conveyed that the decrease in LDL seen in the study is would be considered clinically significant if collected from a female patient. Although the clinicians do not consider changes in the lipid profile to be a clinically significant improvement, applying our first emergency interpretive concept of baseline maintenance, they do consider that the maintenance of the lipid profile at baseline levels over a yearlong program to be a clinically significant event for a patient with MetS.

***Blood glucose***. The indicators of blood glucose levels consist of two clinical indicators that describe a patients' blood sugar: fasting blood glucose (FBG) and hemoglobin A1C (HbA1C). No formal clinical significance was found for any of the measures of blood glucose. There was some disagreement between the dietitians and the physicians regarding their clinical interpretations of the change in blood glucose.

The dietitians reported a clinically significant improvement in HbA1C, but not in FBS, whereas the physicians did not consider the change in either indicator have shown a clinically significant improvement. This difference is due to the dietitians considering the maintenance of HbA1C at the baseline level in the MetS population to be clinically significant. The clinicians cited guidelines that a value of 4-6% for HbA1C is considered normal, with values greater than 7% being considered high and increasing the risk for developing chronic diseases. One of the

dietitians, Drew, specifically points out the low HbA1C at 12-months (6.12), which is below the threshold for an abnormal value (≥7.0):

> …The fact that we saw a low [HbA1C], I mean that is more significant than anything…because then we can see that…over three months. So that's actually very significant.

The clinicians also noted that the baseline value of HbA1C (6.34) was already at a normal level, and thus had little room left for improvement over 12-months of the CHANGE program; the dietitians noted maintaining HbA1C below abnormal levels to be clinically significant.

Fasting blood glucose was an indicator that did not show a statistically significant change, and clinicians were asked about why they thought this was the case. The clinicians all noted that the mean FBG level at baseline, like HbA1C, while not at normal levels, were still relatively low. Thus, there was little room left for improvement over 12-months of the CHANGE program. The physicians noted it was interesting that the FBG did not improve even though there was an improvement in HbA1C because "…they're both linked together quite closely in terms of cause and effect." (Patrice). The clinicians did not provide any guidelines for the FBG thresholds, and noted that they prefer to use HbA1C instead of FBG as a measure of long-term blood glucose levels due to the sensitive nature of FBG. All the clinicians noted that FBG, like triglycerides, is a sensitive indicator that can be affected by events of the day; therefore, it is more suitable for looking at short-term changes. FBG is not usually looked at when examining longer-term changes, or trends in clinical indicators. HbA1C is a more stable indicator of blood sugar, and is more suitable for looking at long-term changes, such as determining clinical significance of interventions.

Although the physicians do not consider changes in the blood glucose to be a clinically significant improvement, like the lipid profile, we can apply our first emergency interpretive concept of baseline maintenance. All the clinicians do consider that the maintenance of blood glucose at baseline levels over a yearlong program to be a clinically significant event for a patient with MetS.

***Overall MetS status***. To contextualize the finding that 44.7% of patients in the CHANGE program had reversed their MetS status, clinicians were asked about (1) the typical rate of reversal for MetS that they see in their practice and (2) the clinical significance of reversing MetS status.

The two clinician groups differed in the typical rate of reversal for MetS that they see in their practice based on current standards of care. The dietitians cited a higher rate of reversal (0-60%) than the family physicians (0-5%). Given this low rate of reversal expected by the physicians, all physicians reported being impressed by the 44.7% reversal rate that we saw with the CHANGE patients. The clinicians all agreed that reversal of overall MetS status is a clinically significant outcome.

A table of the clinicians' responses to whether there was a clinically significant difference after the CHANGE program for each of the clinical indicators, and the number of responses that align with the statistical findings is presented in Table 7

Table 7

*Key Informant Interviewee Identification of Clinically Significant Differences Among Clinical Indicators of MetS 12-months Post-Enrolment in CHANGE*

| Clinical Category | Clinical Indicators | Statistically Significant | Formally Clinically Significant | Dakota | Devon | Drew | Patrice | Parker | Presley | # Converge |
|---|---|---|---|---|---|---|---|---|---|---|
| Body Mass | Weight*** | Yes | Yes | Unsure | Yes | Yes | Yes | Yes | Yes | 5 |
| | BMI*** | Yes | Yes | Unsure | Yes | Yes | Yes | Yes | Yes | 5 |
| | WC*** | Yes | Yes | Yes | Yes | Yes | Yes | Yes | Yes | 6 |
| Blood Pressure | SBP*** | Yes | Yes | Unsure | Yes | Yes | Unsure | Yes | Yes | 4 |
| | DBP*** | Yes | Yes | Unsure | Unsure | Yes | Unsure | Yes | Yes | 4 |
| Lipid Profile | TG | No | No | No | No | No | No | No | No | 6 |
| | LDL* | Yes | No | Unsure | Unsure | No | No | Yes | Yes | 2 |
| | HDL* | Yes | No | Unsure | Unsure | Unsure | No | No | Yes | 1 |
| | TC | No | No | No | No | No | No | No | No | 6 |
| Glucose | HbA1C* | Yes | No | Yes | Yes | Yes | No | No | No | 3 |
| | FBG | No | No | No | No | No | No | No | No | 6 |

* denotes statistically significant difference ($p < .05$)

** denotes clinically significant improvement (RCI $\geq 1.96$)

*** denotes personally significant improvement (clinicians)

**Finding 2: Impacts of CHANGE Participation on ED Utilization**

There is a statistically significant increase in ED utilization during participation in the CHANGE program. Clinicians also agree that the increase in ED utilization is clinically significant, however, two important interpretive considerations emerged from the key informant interviews. First, the dietitians did not have experience working with patients' ED visits, and thus deferred to the physicians when asked about the clinical significance of patients' ED utilization. Second, the physicians suggest that more datasets, providing more variables over a longer timeframe, are required to provide a meaningful interpretation of the increase in ED utilization of CHANGE patients in the context of MetS.

**Statistical findings of the effect of CHANGE on ED utilization**. In the baseline period, 8 patients had an ED visit (range: 1-4), whereas in the 12-month period, 21 patients had an ED visit (range:1-9). The evidence of increased patient ED utilization over the first year of participation in the CHANGE program is the statistically significant difference ($t(84) = -2.581$, $p = .012$, $d = -.563$) in emergency department utilization when comparing the patients' utilization at baseline and at 12-months, but that this difference is not formally clinically significant (RCI = 1.25) (see Table 8). This timing of this difference suggests a link between participation in CHANGE and an increase in ED utilization. Despite the increase, the mean number of ED visits of CHANGE participants ($M$=.153, $M$=.529) remains lower than the mean ED visits of the Alberta population ($M$=.557) (Statistics Canada, 2011; Alberta Health Services, 2013).

Table 8
*Mean Number of Emergency Department Visits at Baseline and 12-months Post-Enrolment*

| | | Baseline | | 12-Months | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | N | M | SD | M | SD | t | p | *d* | RCI |
| ED visits | 85 | .153 | .567 | .529 | 1.33 | -2.581 | .012 | -.563 | 1.25 |

**Clinical significance of increase in ED utilization**. Two themes emerged from the key informant interviews regarding the increase in ED utilization: (1) the increase in ED utilization was clinically significant from a numeric perspective, and (2) the increase in ED utilization was likely not a result of CHANGE or MetS. The physicians noted that the numeric increase in ED utilization was clinically significant, but that they were unable to provide any further interpretations about the role of CHANGE in this increase without knowing the reasons for the patients' visit to the ED. Therefore, they are unable to comment on whether a visit to the ED was related to MetS, or related to the CHANGE program, or due to some other factor. However, both clinician groups were willing to lend their expertise and provide some potential reasons for why there was an increase in ED utilization and pointed to additional data sources that could help answer this question. The clinical interpretations of the increase in ED utilization differed between the physicians and dietitians. All the physicians reported that this increase in ED utilization was clinically significant, whereas the dietitians reported the opposite. Common to both was a reference in general to number of visits; for example, Presley, a physician stated: "Yeah, if you look at the numbers, going from 1 in 7 to 1 in 2 people visiting the ER, that sounds significant." Whereas Drew, a dietitian stated, "…to me, that's not significant…" What this means is that the two different clinician groups have different opinions about the number emergency department visits. It is important to note that the physicians are more experienced working with patients who have ED visits since these visits are often followed-up with by the family physicians, whereas follow-up with dietitians after an ED visit is not common. Thus, when looking at the clinical significance of the increase in ED utilization, the physician perspective takes priority; the increase in ED utilization is clinically significant.

The clinicians described a few reasons for an increase in ED utilization that could be related to participation in CHANGE: (a) patients became more active and therefore are at higher risk for injuring themselves, (b) patients had increased awareness of their own health, and are therefore utilizing the health care system more, or (c) patients were not following-up with their family physicians for medication adjustments, leading to an increase in adverse medication effects. These are components that were not explored in the scope of this study, and points to areas with valuable information to guide future studies.

Two physicians, Presley, and Parker, commented that the increase in ED utilization was likely unrelated to the CHANGE program, or MetS. Although patients are with MetS are at higher risk for developing chronic diseases, they have not actually developed the chronic diseases which would lead to a visit to the emergency department. These chronic diseases take 10-30 years to manifest, as noted by Presley: "…the thing it's going to have an impact on …it may decrease the development of diabetes…[but] your emergency visit related to your diabetes, is probably 20 years down the road."

Parker commented that the CHANGE program should have led to decreased ED utilization, rather than the increase that was observed. Parked explained that since the program is providing a lot of support to the patients through increased frequency of family physician, dietitian, and kinesiologist visits, this increased attachment of a patient to a set of clinicians. This attachment should theoretically increase the likelihood that patients would contact the program clinicians with health concerns related to MetS, rather than visit the emergency department.

**Finding 3: Relationships Among Clinical Indicators of MetS and ED Utilization**

No statistically or clinically significant relationship was found between the clinical indicators of MetS and ED utilization. I will first present the statistical analyses, including the calculation of composite variables and the regression model from the Integrated Dataset. Then I will present the clinical interpretation of these findings by the clinicians which suggests that ED visits may not be captured within the current variables or timeframe of this study

**Statistical findings among clinical indicators of MetS and ED utilization**. Composite variables of baseline and 12-month variables were calculated to create a single set of variables to enter into the regression model. These composite variables were calculated by taking the difference between the clinical indicators (baseline – 12-months) and the ED visits (pre – post). The correlations of baseline (Table 9), 12-month (Table 10) and the composite variables (Table 11) were calculated to confirm that the relationships among the variables remained similar after the creation of the composite variables. Additionally, the relationships among these variables were examined and were used to identify which variables would be entered into the regression model. Based on the significant correlations between variables, six variables that were least related to each other were chosen to be entered into the regression model: SBP, DBP, FBG, TG, HDL, and WC.

Table 9
*Correlation Matrix of Baseline Variables*

| Baseline | Age | Sex | W | BMI | SBP | DBP | FBS | TG | LDL | HDL | AC | TC | HbA1C | ED Visits |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 1 | | | | | | | | | | | | | |
| Sex | .185 | 1 | | | | | | | | | | | | |
| W | -.162 | -.565** | 1 | | | | | | | | | | | |
| BMI | -.040 | -.079 | .711** | 1 | | | | | | | | | | |
| SBP | .102 | -.028 | .158 | .245* | 1 | | | | | | | | | |
| DBP | -.259* | -.295** | .196 | .076 | .490** | 1 | | | | | | | | |
| FBG | .073 | -.124 | .135 | .135 | .131 | .064 | 1 | | | | | | | |
| TG | -.269* | -.168 | .005 | -.114 | .147 | .305** | .018 | 1 | | | | | | |
| LDL | -.171 | .077 | -.011 | -.008 | .188 | .258* | -.190 | .351** | 1 | | | | | |
| HDL | .063 | .344** | -.161 | .084 | -.139 | -.251* | -.163 | -.593** | -.101 | 1 | | | | |
| AC | -.075 | -.366** | .799** | .712** | .050 | .040 | .212 | -.051 | -.005 | -.050 | 1 | | | |
| TC | -.243* | .073 | -.047 | -.037 | .183 | .283** | -.187 | .584** | .942** | -.098 | -.040 | 1 | | |
| HbA1C | .072 | -.109 | .087 | .031 | -.006 | -.103 | .655** | .051 | -.178 | -.119 | .156 | -.152 | 1 | |
| ED Visits | .149 | .149 | -.047 | .075 | .026 | -.073 | -.014 | -.076 | -.102 | .116 | .098 | -.088 | .038 | 1 |

**Correlation is significant at the .05 level (2-tailed).*
***Correlation is significant at the .01 level (2-tailed).*

Table 10
*Correlation Matrix of 12-month Variables*

| 12-Month | Age | Sex | Weight | BMI | SBP | DBP | FBS | TG | LDL | HDL | AC | TC | HbA1C | ED Visits |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 1 | | | | | | | | | | | | | |
| Sex | .185 | 1 | | | | | | | | | | | | |
| Weight | -.191 | -.486** | 1 | | | | | | | | | | | |
| BMI | -.053 | -.017 | .770** | 1 | | | | | | | | | | |
| SBP | .063 | -.094 | .240 | .285* | 1 | | | | | | | | | |
| DBP | -.326** | -.333** | .235 | .031 | .371** | 1 | | | | | | | | |
| FBG | .052 | -.039 | -.001 | .117 | .140 | .007 | 1 | | | | | | | |
| TG | -.270* | -.187 | .194 | .132 | .144 | .264* | .013 | 1 | | | | | | |
| LDL | -.102 | .057 | -.023 | -.120 | .020 | .018 | -.308** | .063 | 1 | | | | | |
| HDL | .020 | .308** | -.276* | -.173 | -.038 | -.245* | .015 | -.269* | .171 | 1 | | | | |
| AC | .024 | -.283* | .841** | .828** | .260* | .035 | .073 | .122 | -.109 | -.280* | 1 | | | |
| TC | -.100 | .035 | -.045 | -.120 | .080 | .043 | -.211 | .320** | .929** | .299** | -.153 | 1 | | |
| HbA1C | .116 | -.018 | -.082 | .091 | .109 | -.054 | .787** | .022 | -.248* | .005 | .053 | -.104 | 1 | |
| ED Visits | .031 | .014 | .034 | .166 | .136 | -.088 | .034 | .169 | -.057 | .062 | .083 | -.029 | -.123 | 1 |

**Correlation is significant at the .05 level (2-tailed).*
***Correlation is significant at the .01 level (2-tailed).*

Table 11
*Correlation Matrix of Difference Variables*

| DIFF | Age | Sex | Weight | BMI | SBP | DBP | FBS | TG | LDL | HDL | AC | TC | HbA1C | ED Visits |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Age | 1 | | | | | | | | | | | | | |
| Sex | .185 | 1 | | | | | | | | | | | | |
| Weight | .015 | -.136 | 1 | | | | | | | | | | | |
| BMI | .035 | -.060 | .988** | 1 | | | | | | | | | | |
| SBP | .067 | .022 | .308* | .279* | 1 | | | | | | | | | |
| DBP | .086 | .040 | .411** | .391** | .700** | 1 | | | | | | | | |
| FBG | -.008 | -.106 | .163 | .129 | .138 | .026 | 1 | | | | | | | |
| TG | -.205 | -.119 | .102 | .111 | -.096 | .079 | .163 | 1 | | | | | | |
| LDL | -.007 | .096 | .025 | .026 | .179 | .078 | .023 | .094 | 1 | | | | | |
| HDL | .059 | .082 | -.113 | -.142 | .023 | -.016 | -.172 | -.736** | -.106 | 1 | | | | |
| AC | -.182 | .049 | .666** | .675** | .145 | .108 | .125 | .225 | .043 | -.303** | 1 | | | |
| TC | -.204 | .055 | .054 | .054 | .097 | .126 | .111 | .578** | .879** | -.309** | .119 | 1 | | |
| HbA1C | .000 | -.111 | .176 | .156 | .087 | -.002 | .596** | .111 | .035 | -.099 | .119 | .080 | 1 | |
| ED Visits | .032 | .049 | .102 | .112 | .088 | .085 | .037 | .045 | .226* | -.018 | .024 | .167 | -.035 | 1 |

**Correlation is significant at the .05 level (2-tailed).*
***Correlation is significant at the .01 level (2-tailed).*

A regression model was created with the age and sex variables. This did not result in a significant regression model ($F(2, 82) = .123$, $p = .885$) with an adjusted $R^2 = -.021$, suggesting that age and sex are not significant predictors of ED utilization in patients with MetS. The six clinical indicators (SBP, DBP, FBG, TG, HDL, WC) were entered into the model, and did not result in a significant regression model ($F(8, 55) = .391$, $p = .921$) with an adjusted $R^2 = -.084$, suggesting this model is not a significant predictor of ED utilization in patients with MetS.

These findings from the regression analysis were presented to the clinicians to determine their clinical significance, and the impact these findings would have on clinical practice. Two themes emerged: (1) that they expect the clinical indicators of MetS to be unrelated to ED utilization, and (2) that there is value in identifying the relationships among variables to identify predictors of ED utilization.

**Clinical significance of relationships among clinical indicators of MetS and ED utilization**. The clinicians confirmed the statistical finding that there is no clinically significant relationship among the clinical indicators of MetS and ED utilization. The clinicians reported that this the result that they expected. The reasoning of one of the physicians, Presley, is that although MetS is associated with a higher risk of chronic disease, the ED visit associated with that chronic disease is decades away from the initial diagnosis of MetS. One of the risk assessment calculators, the Framingham Risk Score, is used to estimate a patient's 10-year risk of developing cardiovascular disease, and includes some indicators of MetS (SBP, TC, HDL), but is also a long-term risk calculator. Thus, the clinicians did not expect the clinical indicators of MetS to be good predictors of patients' ED utilization within the timeframe of this study.

The lack of a statistically or clinically significant relationships between the clinical indicators of MetS and ED utilization suggests that some other factor that isn't measured in this

study is related to the increase in patients' ED utilization. The clinicians point towards other indicators that could explain the increase in ED utilization within the timeframe of CHANGE: (a) other patient-level clinical indicators (e.g., medication, sleep, mental health), and (b) other system-level indicators (e.g., cost, outpatient visits). One of the physicians, Patrice, noted the importance of looking at patient medication use and the cost of delivering the program:

> I think knowing the number of meds the patients start and stop on…[and] the cost of the program [would be beneficial]

Some of these datasets reside at the patient-level, some at the program-level, and some at the system-level, highlighting the importance of working towards overcoming data integration barriers.

Clinicians reported that there is clinical value in looking for relationships among program- and system-level variables to inform their practice and identify predictors of ED utilization that can be addressed. However, they also noted that reducing ED utilization is a long-term goal that may not be measurable within the short timeframe of this study. The clinicians made suggestions for additional measures that may be affected within the timeframe of this CHANGE to be examined.

**Summary of Findings**

Three major findings emerged from the statistical analyses of each dataset and clinical interpretation of those findings. The first finding from the Program Outcomes Dataset, demonstrated program effectiveness, evidenced by the statistically significant improvements in 8 indicators, 5 of which were also clinically significant, and the maintenance of the remaining clinical indicators at baseline levels. The second finding of statistically significant increased ED utilization by patients participating in CHANGE, revealed by the System Outcomes Dataset, was

surprising to clinicians because of the improvements seen in MetS indicators. The third finding, from the Integrated Dataset, revealed a lack of either a statistically or clinically significant relationship between the clinical indicators of MetS and ED utilization. The clinical interpretation suggests that ED visits may not be captured in the existing variables, and that ED visits related to MetS would be decades away from initial diagnosis and would not be captured in the time frame of this study.

# Chapter 5: Discussion

This chapter discusses the study findings and addresses the guiding research question, "What is the impact of integrating datasets and adding the clinical perspective on the new evaluative insights?" The chapter is organized around four discussion points: First, the enhanced understanding of clinical indicators of program effectiveness for treating metabolic syndrome (MetS). Second, the emerging evidence of program effects on the health system through examining ED utilization. and specifically, the divergence of the findings, and the usual expectations informed by clinicians and the literature. Third, how integrated datasets situates programs within the health system, and the usefulness of integrated datasets for generating evaluative insights. Finally, the importance of a clinical interpretation of statistical findings for enhancing the validity of evaluative insights.

## Enhanced Understanding of Clinical Indicators of Program Effectiveness for Treating MetS

This study established the effectiveness of the CHANGE program, as a lifestyle intervention program providing education around the Mediterranean diet and exercise coaching, as an effective treatment for MetS. Participation in the CHANGE program improved patients' clinical indicators related to MetS after 12-months, evidenced by convergence in clinical with statistical significance for body mass and blood pressure indicators. The CHANGE program outcomes are considered consistent with the statistical results studies from similar lifestyle intervention programs using the Mediterranean diet and guided exercise sessions (Bo et al., 2007; Gomez-Huelgas et al., 2015; Rubenfire et al., 2011). Furthermore, at the end of the program, 44.7% of patients no longer met clinical criteria for MetS based on their clinical indicators, an outcome that is considered clinically significant by the clinicians and converges

with the results from previous studies of diet-only interventions (Babio et al., 2009; Babio et al., 2014; Estruch et al., 2013).

Less evident, because it was statistically non-significant, is the effectiveness of CHANGE participation on blood sugar and the lipid profile. Improvements to body mass indicators and blood pressures are usually found concurrently with improvements in blood sugar and cholesterol levels (Case et al., 2002), but this was not the case in this study. Instead, the clinical interpretations from this study noted that the blood sugar and lipid indicators remained at baseline levels after 12-months of CHANGE participation. The clinicians identified that blood sugar and lipid indicators take longer to show improvement, and that a significant improvement would not have been expected after only 12-months in the CHANGE program. Further, they noted that maintaining the blood sugar and lipid profile indicators at baseline levels after 12-months could be a clinically significant event since the natural progression of MetS is for the blood sugar and lipid profile indicators to worsen and thus increase the risk of cardiovascular disease. This is supported by Bo et al., (2007) who found that a control group of patients with MetS who did not receive intervention saw their risk of cardiovascular incidence increase.

The results from the analysis of the Program Outcomes Dataset in this study aligns with the published literature of other lifestyle intervention programs, and represents the current standard for evaluation lifestyle interventions. There is a common limitation in this method that is addressed by this study: the effects of the lifestyle intervention program on the wider health system is not examined through integration of program-level data with system-level data. This is an important limitation because without examining the effect of a program on system-level data, the scope and usefulness of study results to evaluating the effectiveness of the program on patients (Bo et al., 2007; Gomez-Huelgas et al., 2015; Rubenfire et al., 2011). This is an

important limitation to address because system-level measures are starting to be used as indicators of program success; in some cases, reporting on the impact of the program on these system-level indicators is a requirement for program funding. System-wide analysis requires obtaining data from health authorities recording the patients' utilization of other health services within the wider health system, and integrating this data with program-level datasets. Without integrating these datasets, the evaluation of the CHANGE program would have been limited to a single evaluative insight: that patients had a statistically significant improvement on 8 of 11 clinical indicators of MetS. This conclusion is useful for examining the effect of the CHANGE program on the patients' clinical indicators, but does not provide any insight into the effect of the program on system-level variables. The integration of datasets in this study allowed for the identification of the impact of the CHANGE program on ED utilization, and represents an emerging type of study being done.

**Emerging Evidence of Program Effects on the Health System**

The study findings suggest some association between CHANGE participation and increased patients' ED utilization, despite improvements in patients' clinical indicators. This finding is puzzling as it runs counter to the expectations of study clinicians and previous studies (Gillespie et al., 2016); both which notes that patients' ED utilization should decrease after improvements were made to clinical indicators, which are indicators of improved general health. This finding may be explained by a missing variable not currently measured in this study that is driving the increase in ED utilization because of a lacking statistically significant relationship between the clinical indicators of MetS and ED utilization. The clinicians provided some support to this hypothesis by offering potential reasons for the increase in ED utilization that are

related to CHANGE, but none of the reasons they provided could be explored in within the scope of this study without additional datasets.

The clinicians' speculations require integration of additional datasets to further investigate the relationship between MetS indicators, ED utilization, and a lifestyle intervention program. Two physicians commented that the increase in ED utilization was unlikely to be related to participation in CHANGE since the ED visit related to MetS would be decades away from initial diagnosis. This is an important idea to consider, because this suggests that a longitudinal study, one that spans decades, is required to examine the effect of MetS on ED utilization. This further suggests that there is some other variable that is not measured that is related to the increase in ED utilization by CHANGE patients.

**Integrated Datasets Situates Programs Within the Health System**

This study provides important insights about the relationships among variables across datasets that are accessible only when those datasets are integrated and useful for policy-makers and program developers. That CHANGE was found to be an effective treatment for MetS is helpful for making the case for funding decisions yet it is limited in its usefulness for policy-makers to determine the success of the CHANGE program on the wider health system. What it does highlight for policy-makers are the potential ripple effects of lifestyle intervention programs on the rest of the healthcare system and identify specific demands at other levels of health care. This type of information about the effect of a program on system-level measures, such as ED utilization, while not a current reporting requirement for health programming, is increasingly becoming a priority for measurement and evaluation initiatives. There reflects an increasing understanding that it is possible, and perhaps more efficient, to improve system-level performance indicators at the program-level. Alberta Health (2013) specifically notes in their

Health System Outcomes and Measurement Framework that system-level outcomes are dependent on the attainment of program-level outcomes. Thus, it is important to examine the effect of the CHANGE program on a system-level measure like ED utilization to ensure the programs are having positive effects on the overall health system, but doing so requires additional data. Accessing the Systems Outcome Dataset through the health authority allowed for an examination of the effect of the program on ED utilization; analysis of the Systems Outcome Dataset revealed that there was an increase in ED utilization after enrolment in CHANGE.

Identifying an increase in ED utilization after enrolment in the CHANGE program is useful for program administrators, as it provided some information to guide future quality improvement efforts at the program-level. This is because there may be some aspect of the CHANGE program that is increasing ED utilization for patients. This highlights that assumptions around the effect of CHANGE on system-level measures need to be re-examined, and points to an area where the program could improve. Most importantly, this finding posits a new question to the study administrators: is there some causal relationship between CHANGE and this increase in ED utilization after enrolment in the CHANGE program?

An attempt was made at investigating the relationship between CHANGE and the increase in patients' ED utilization by looking for relationships among the clinical indicators of MetS and ED utilization. No significant relationship was found, which suggested that factors outside of the MetS clinical indicators were driving the increase in ED utilization observed. This is supported by a physician who commented that the ED visits due to complications of MetS should be decades after initial diagnosis, once chronic diseases manifest. The additional unknown factors associated with increased ED utilization could be related to the CHANGE

program itself; if so, these factors will need to be identified and improved to decrease the risk of ED utilization in future CHANGE participants. This is useful information for program administrators; they now know that some aspect of the CHANGE program, unrelated to measured clinical indicators, may be associated with increased ED utilization. The clinicians identified certain variables that could have led to increased ED utilization, such as changes in medication use, increased physical activity leading to increased risk of injury, and aggravation of previous injuries. These postulated variables can be used to guide future studies in identifying factors associated with ED utilization of CHANGE patients.

Integrating the System Outcomes Dataset with the Program Outcomes Dataset allowed for a deeper exploration into the nuanced relationships among variables of both datasets. This addressed a limitation of similar studies of lifestyle intervention programs (Kramer et al., 2011; Rao et al., 2014; Wang et al., 2014) and evaluations of emergency departments (Porter et al., 2015; Nguyen & DeJesus, 2010) to identify similar relationships among different levels in healthcare. The Integrated Dataset provided access to the patients' ED utilization data, and allowed me to explore the effect of the program on patients' ED utilization pre- and post-enrolment in CHANGE. This deeper exploration yielded two additional evaluative insights into the effect of the CHANGE program on ED utilization that would have been inaccessible with a singular dataset. Thus, this answers the research question of this study: there were two evaluation insights that were generated from a clinical interpretation of an integrated dataset of program-level and system-level data that would not have been possible with a statistical approach of each dataset alone: (1) ED utilization of patients increased after enrolment in CHANGE, and (2) that there was no relationship between the clinical indicators of MetS and ED utilization. The analysis of integrated datasets was therefore useful for expanding the scope of this evaluation

beyond program-level clinical indicators (Kastorini et al., 2011; N.Banio et al., 2009), to include the effect of the program on an important system-level outcome measure, ED utilization. This highlights the potential of integrated datasets to explore the relationships among variables at multiple levels of healthcare, leading to a deeper and more nuanced understanding of these relationships that can be used to inform the decision-making process. Another important component to this study was the inclusion of clinicians to interpret the statistical findings using their clinical and contextual expertise.

**Clinical Perspectives Enhance Validity of Evaluative Insights**

This study contributes understandings about the importance of clinical interpretations for enhancing validity. This is evidenced by the findings that demonstrated areas of divergence and convergence with traditional statistical significance. Bothe & Richardson (2011) described four types of significance: statistical, whether there is a difference between groups; practical, the magnitude of that difference; clinical, what that difference means; and personal, whether that difference solved the problem. Personal significance was not examined in this study, but would require interviews with patients to determine the personal significance of the CHANGE program.

The previous studies of lifestyle intervention programs focused in on statistical and practical significance of their findings (Rubenfire et al., 2011; Gomez-Huelgas et al., 2015), but did not include measures of clinical significance, such as RCI (Jacobson & Truax, 1991) or clinical interpretations of the statistical findings (Bothe & Richardson, 2011). There two important differences found between this study that was not found in the previous studies; both of them which involve the inclusion of a clinical interpretive lens: (1) that some of the statistically significant improvements were not clinically significant (LDL, HDL, HbA1C) in the context of MetS, and that (2) the maintenance of blood sugar and lipid profile indicators at

baseline could be considered a clinically significant event, due to the natural progression of MetS to worsen without intervention.

There is considerable similarity in the statistical findings of this study and other studies of lifestyle intervention programs (Rubenfire et al., 2011; Gomez-Huelgas et al., 2015), especially in the finding that measures of the blood sugar and lipid profile are maintained at baseline levels. However, whereas those studies stopped at statistical significance, this study went further to look at the clinical significance of the statistical results through a clinical interpretive lens (Bothe & Richardson, 2011). These previous studies identified a difference between groups (statistical significance), and the magnitude of that difference (practical significance), but they did not go further to examine what this difference means (Bothe & Richardson, 2011). In this study, the inclusion of clinicians allowed me to explore the meaning behind the difference found (clinical significance). Although was no statistical significance found for measures of blood sugar and lipid profile, the clinicians noted that maintaining clinical indicators at a baseline level is an important first step at treating MetS, especially because blood sugar and lipid profile indicators take a long time to improve. Thus, the clinicians consider the maintenance of blood sugar and lipid profile at baseline levels to be a clinically significant event. This is a finding that would not have been possible without the clinical and contextual expertise gained through applying a clinical interpretive lens.

**Summary of Discussion**

The effectiveness of the CHANGE program for treating MetS was established in this study. while the clinical interpretation captured the positive outcomes of all the clinical indicators, the usual method of examining statistical significance alone (Bo et al., 2007; Gomez-Huelgas et al., 2015; Rubenfire et al., 2011) did not capture the positive outcomes in blood sugar

and lipids. This difference reinforces the need to take timeframe and clinical expertise into account in the statistical interpretation of clinical findings. Further, integration of system-level ED utilization datasets allowed for the investigation of the effect of the program on ED utilization.

The usual system-level indicator of ED utilization is an emerging measured indicator of program effectiveness, but it can't currently be used to assess program effectiveness because the relationships between programs and ED utilization is not yet understood. The increase in ED visits by patients during CHANGE cannot be directly attributable to CHANGE. There is a danger of making erroneous decisions based on an incomplete look at the effects of individual programs on the health system; more data is needed.

With the integrated datasets, we begin to examine the relationships among program- and system-level variables beyond the direct effects of the program on the patient. Specifically, the integration of datasets led to two new evaluative insights: (1) ED utilization of patients increased after enrolment in CHANGE, and (2) that there was no relationship between the clinical indicators of MetS and ED utilization. These new evaluative insights begin to situate the program within the larger system context, which is currently lacking in health evaluation designs, and represents emerging methods in evaluating health programs.

Clinical perspectives allow evaluators to interpret findings beyond statistical significance, which identifies that there is a difference after treatment, and towards clinical significance, which identifies if the difference identified is meaningful in the context of the condition. The clinical perspective is also valuable throughout the study, but particularly during the design phase where their input can be used to strengthen the clinical relevance and validity of the study design.

## Chapter 6: Implications, Limitations, and Final Thoughts

This study advances three implications for consideration in future studies of the health care system. The first is a practice implication of how researchers and evaluators can increase the type of data used in studies to enhance studies of complex health systems. The second is a methodological implication of mitigating data access barriers to increase the type of clinical data available to enable examination of the nuanced relationships among patient-, program-, and system-level variables of health care. The third is a clinical implication of how researchers and evaluators can apply a clinical interpretive lens to increase the validity and clinical usefulness of statistical findings.

### Enhancing Evaluations of Complex Health Systems Through Increasing Data Types

In this study, I introduced two new types of data into a program-level evaluation: system-level data and clinical interpretation of statistical findings. The system-level data of patients' ED utilization led to additional findings around patients' ED utilization, but also to findings about the relationship between the program-level variables, the clinical indicators of MetS, and ED utilization. Increasing the types of data available to examine a health program increases the different kinds of analyses that can be done to examine the many aspects of that program. The inclusion of system-level data expands the scope of the evaluation from looking just at the program-level, and towards examining the program within the context of the wider health system. Without integrated datasets, it becomes impossible to look at the nuanced relationships among variables of a complex health system; integrating datasets allows for more information to inform program decision-making by situating the program within the system. The clinical interpretation of the statistical findings allowed for examination of these relationships beyond statistical significance and towards clinical significance.

Looking at these nuanced relationships through integrated datasets is an emerging method of evaluating complex health systems. Examining the relationships among variables across different levels of health care is the only way to investigate the effect of clinical interventions on system-level measures. This study has presented an example of integrating just one additional dataset, which led to two additional findings, but also led to raising additional questions. These additional questions can be answered through the integration of even more system-level datasets, such as medication data. The future of evaluation in health care will be in the integration of appropriate datasets to explore the nuanced relationships among variables of a complex health system and using a clinical interpretive lens to examine the statistical findings. This study provided a glimpse at that future, but more work needs to be done in the mitigation of barriers to data access and refinement of the analysis methods before dataset integration becomes commonplace and widely used.

**Strategies to Increase Data Types by Mitigating Barriers to Data Access**

This study points to the challenges faced by researchers and evaluators in their design of studies and evaluations to investigate the health system due to the availability of data types. The primary focus of current designs to system-level evaluations on quantitative accountability measures due to system-level data being available; conversely, program-level evaluations are focused on quality improvement and patient experience due to program- and patient-level data being available. Integrating these two datasets in this study led to new evaluative insights about the relationships among program-level and system-level measures, which can be used to guide future improvements at the program- and system-levels. Rather than having studies focused on either program- or system-level indicators, studies utilizing integrated datasets could investigate the indicators at both levels and the relationships among these indicators. This would lead to a

more nuanced understanding of the patients' journey in accessing health care in a complex system.

This study points to 4 strategies that guide how researchers and evaluators can overcome each of the barriers to the integration of datasets, as described by Kephart (2002); I will expand on Kephart's (2002) work by providing some guidance to mitigate the barriers he described. The barriers to dataset integration need to be mitigated if integrated dataset analyses are to become commonplace (Silverman et al., 2015). I will also present one strategy that the health authority could implement to help mitigate these barriers. I provide these strategies in the hopes of increasing the amount of dataset integration that occurs in the evaluation of health initiatives.

**Establish rapport with gatekeepers of data pathways**. In the initial stages of this study, I intended to include the pathway for accessing system-level data from the Alberta health authority (AHS); however, the pathway has since changed and is no longer useful. This is an ongoing challenge in health care, and the important strategy for users trying to access system-level data is to identify and develop relationships with the gatekeepers of this data. In addition to gaining access to the data, the gatekeepers may have in-depth knowledge of the datasets, and may be able to provide some important contextual information.

**Get to know all aspects of the target variables**. When I first received the data, I needed to examine the variables from both datasets to fully understand how they were calculated and collected. Data being collected may not be standardized, and it is particularly important to identify the origin and description of each variable collected when integrating datasets. A variable may look the same as another, but may be collected in a different way. This is particularly important when using derived variables measuring health services utilization, where the variables that are being accessed were not collected directly, but derived from other variables. For example, if age is calculated rather than recorded, the researcher or evaluator would need to know how it is calculated: is it the age at referral, enrolment, or first visit? Researchers and evaluators will need to be particularly careful if the custodians they are working with have not standardized the data, and if they are working with several custodians; each custodian could have a different convention for collecting each variable. Thus, they should consider whether the data accessed follow standardized practices before they integrate the datasets together.

**Develop data extraction and linkage protocols**. When I began to integrate the datasets, a common patient-identifying variable is required. In the Alberta context, this will be the Patient Health Number (PHN), but other provincial health authorities will have a similar identifier. The data that you send in to the health authority to be matched will require PHNs, therefore necessitating the need for encryption. Make sure you are familiar with and diligently follow data security protocols of all the custodians involved. Typically, this includes encrypting data when transmitting and storing data, and working with a de-identified dataset. In general, the encryption built into office suites (AES-128) is strong enough for this purpose. It cannot be stressed enough that PHNs must be securely guarded, as all health information for that patient is linked to their PHN.

**Promote/enhance database security**. When negotiating the agreement, privacy and security of all data should be taken seriously. Demonstrating an understanding of how to securely transmit and store data increases the confidence of the health authority to release data to you. A breach of data security is harmful not only to the researchers/evaluators involved, but to the health authority, the custodians of the data, and the patients. The security of patient-identifying health data is particularly important, which is the reason it has been emphasized throughout this study. Where possible, as it was in this study, I would recommend working with de-identified data. The original raw data and master list should be stored securely and encrypted, with only encrypted files being transmitted.

**Development of accessible system data repository**. When I accessed the system-level datasets, it was through a province-wide data repository. Access to this data repository is relatively closed due to the previously mentioned barriers. Although researchers have some level of control over those barriers to data integration, and can my suggested steps to mitigate some access issues, the custodians of the administrative datasets are ultimately responsible for the safeguarding of health information. Generally, this responsibility falls upon the health authority. With this responsibility, the health authorities become the gatekeepers of a lot of administrative datasets that can be used for secondary purposes, particularly in examining the effects of programs on the entire health system. This situation puts Alberta Health Services in a good position to create and maintain a repository of health information, for managing access to this data and for developing some expertise in the utilization of these datasets for research purposes. If such a repository were in place, the data access and integration process for this study could have been more easily and expediently facilitated. In addition to the system-level datasets of which AHS are already custodians, this new repository could contain data from other studies that have been completed and can be further used for a secondary purpose.

Data collected for the evaluation of specific programs results in the creation of datasets which are used for a specific purpose; however, after this purpose has been fulfilled, the data is no longer used. If this data can be stored in a repository, it can be used by other researchers in a secondary analysis (Hollander et al., 2010; van Panhuis et al., 2014). In healthcare, maintaining a repository of identifiable data allows researchers to integrate datasets together to explore unforeseen relationships in a variety of situations. Developing and publishing a provincial pathway to more easily access secondary health information would potentially increase the number of studies using these datasets. Integrating datasets together will provide more variables

to explore more areas of health care, and will allow researchers to explore additional impacts of a health initiative on both patients involved and the health system as a whole.

Improving access to these datasets could allow data integration to happen more readily and across multiple studies. Pooling these datasets in research could lead to breakthroughs in the understandings of the many relationships in a complex health system. As demonstrated through findings of this study, integrating data sets together across levels of the health care system allows for development of a more nuanced understanding of the patient journey in the health care system.

The collection of data for program evaluation can be resource intensive; therefore, data collection is always focused on how to measure the achievement in the primary goals of a program. This approach limits the study team to using primary data collected, which is guided by what the original collecting team foresees as affecting original targets of the program. For evaluation design, this limits studies from exploring the relationships between the program and the wider system in which it resides. This limitation can be mitigated through the integration of additional secondary data sources to complement the primary data collection. Finding and integrating data beyond the program will allow the study team to find new ways of interpreting existing data and findings through additional information that situates the program within a larger system.

In order increase access to these secondary datasets, additional resources are required to develop and maintain appropriate infrastructure. The pathways to accessing these datasets must be clearly laid out, open and available for researcher use. An important part of this infrastructure will be the creation of a role with the expertise to identify which variables may be applicable to

each study context, and who can lead researchers down the right path to accessing datasets for integration.

**Including Clinicians' Interpretive Lens to Enhance Validity and Usefulness**

This study highlights the critical role of subject-matter experts as an interpretive lens to increase the validity of statistical findings, which has implications for future evaluations of healthcare programs in the Canadian context. Including clinicians in this study helped achieve three Canadian Evaluator Competencies; specifically, technical competencies related to evaluation design, analyzing and interpreting data, and drawing conclusions and making recommendations based on the data (Canadian Evaluation Society, 2010). I will describe how to include clinicians in evaluations of health programs to validate findings, and how this helps to achieve each of the technical evaluation competencies.

In this study, I recruited dietitians and family physicians to provide their clinical expertise in interpreting statistical findings. Had I included them earlier in the study process, their clinical perspective would have also been helpful in designing this study. Thus, the recruitment of clinicians to aid in the evaluation should occur early in the process so that they may guide the evaluation process throughout. Clinicians that are working within program are those who will be able to provide contextual information to guide the types of data that can and should be incorporated in the evaluation. In addition, they can provide ideas for how the evaluation can identify areas of program improvement, and highlight areas of impact and influence of the program.

Including the clinical perspective in this study allowed me to look beyond statistical significance and towards clinical significance. It is important to differentiate between the sequential types of significance: statistical, practical, clinical, and personal significance (Bothe &

Richardson, 2011). Statistical significance refers to whether there is a difference between groups, which can be identified using statistical tests; practical significance refers to the magnitude of the differences found, which can be identified using effect size measures; clinical significance refers to what the difference means, which can be identified using RCI, but better if contextualized by content experts (i.e., clinicians); and personal significance which refers to whether the difference solved the patients' problem, which can be identified by the patient. This last component of personal significance was not addressed in the scope of this study.

Once I had collected the data, I had to analyzed and interpreted it appropriately, keeping in mind the four types of significance. Data coming from a specialized field, such as health, and the resultant findings can be difficult to analyze and interpret in a valid manner without the necessary expertise; thus, it is important for experts to assist with the interpretation of clinical significance (Bothe & Richardson, 2011). The clinicians and other experts that are recruited should be ones who can provide interpretations of findings beyond statistical significance, and to explore meaningful findings that would not have otherwise been found. Specifically, the experts should be able to provide contextual information about the condition being studied and the program designed to treat it, and assist in identifying clinically significant results that would be missed by relying solely on statistical findings. The inclusion of the clinicians improves validity of findings by providing more meaningful interpretations through the lens of clinical and contextual experts.

In this study, I was fortunate enough to interview both dietitians and family physicians to incorporate their perspectives into my findings. Using multiple clinicians from multiple clinical areas provides additional perspectives and allowed me to examine how each clinical area differed in the level of expertise and adjust accordingly. For example, the dietitians noted that they did

not have much experience working with blood pressure, thus, I put more weight on the family physicians' interpretation of the changes in blood pressure. This allows evaluators to triangulate the results, thus minimizing the risk of bias while also exploring additional perspectives of an issue. It is critically important to recruit the right experts for the study's context and to properly identify the background of the participants so that the data they provide through their clinical perspectives is properly interpreted (Marhsall, 1996). It is also important to identify clinician roles and perspectives that may be missing; I was unable to recruit kinesiologists into the study and therefore had to consider that I am missing some expert perspectives on the exercise component of the program.

I have highlighted how clinical experts are an important source of contextual and clinical expertise, and how they play a vital role in evaluations for identifying the nuanced impacts of a program on patient outcomes (Bothe & Richardson, 2011) and on the health system. Including clinical experts in evaluations can be difficult, but successful recruitment is very worthwhile. When recruiting clinicians for a study, it is important for evaluators to reach out and include them in the development of the evaluation. This is limited by how engaged the clinicians are willing to be, but clinician involvement at any stage of the evaluation should be considered an advantage, as their unique perspective are used to improve the validity and therefore quality of the evaluation.

Engaging these important clinician stakeholders in the evaluation process may also improve the utilization of evaluation results (Alkin, 2011; Johnson et al., 2009; Patton, 2008). In the current study, the clinicians lent their clinical expertise to contextualize the statistical findings to determine clinically significant results, as well as provide insights into possible reasons for some findings, such as how injuries resulting from exercise that could contribute to the increased

ED utilization. Hearing the clinician perspective on the findings could lead to a higher level of trust in the validity of the study.

**Limitations and Directions for Future Research**

There were two areas of limitations in this study that should be considered in the interpretation of results and suggests directions for future research: datasets and key informants. The study datasets were limited to what I could access through AHS and the CHANGE program; this resulted in limited variables, data points, and a small sample size. The participants in key informant interviews were limited to convenience sampling; this resulted in limited perspectives of clinical areas, absence of the patient perspective, and a small sample size

Additional variables were identified by the physicians as important for looking at the relationships among program- and system-level variables were not included in the dataset accessed for this study: medication information, inpatient admissions, outpatient visits, patient reported outcome measures (i.e., quality of life measures), and physician billing data. Access to these variables could have facilitated the identification of additional relationships among the program- and system-levels of healthcare and provided more contextual information around study findings. Thus, to address the limited variables in this study, additional datasets could be accessed and integrated with data from a lifestyle intervention program. Including cost data would allow for an exploration of how cost efficient and financially sustainable this program can be. Specifically related to medication data, the clinicians interviewed in this study have noted that decreased medication dosages could be considered clinically significant. The clinicians have provided anecdotal examples of their own patients who have decreased medication requirements because of the CHANGE program. A future study which includes medication data could look at the effect of this lifestyle intervention program on the medication dosage and report back on the

effectiveness of specific patient activities for decreasing medication usage. Similar to medication usage, future studies could examine and integrate other system-level measures, such as inpatient admissions, to further examine how lifestyle intervention programs could affect other areas of health system usage.

There is also a limitation in how some of the lipid profile variables were examined in this study. Some of the physicians noted that the lipid ratio is an important measure that was not calculated, and one of the physicians went on to note that there is a gender difference in how lipid profile is used. As such, there may be limitations to the conclusions we could make about the effect of the CHANGE program on the lipid profile. Future studies could calculate the lipid ratio and examine gender differences in the lipid profile.

The data points available in the Program Outcomes Dataset was also limited temporally, to a single year of data collected from the 85 patients enrolled in the CHANGE program. The System Outcomes Dataset only included data for these same patients from one year pre- and post- enrolment in CHANGE. The physicians noted that the 2-year timeframe of ED utilization data included in the study may be too short to fully examine the relationship between the clinical indicators of MetS and ED utilization, elaborating that ED visits due to complications from MetS are decades away from the initial diagnosis of the condition. To address the limited data points in this study, a longitudinal study could be conducted which examines the long-term effects of a lifestyle intervention program by examining the extent that treating MetS decreases risk of developing chronic diseases and how that affects health system utilization.

The small sample size (n=85) decreased the robustness of the statistical tests and the generalizability of the results. The patients in this study are not selected randomly from the general population, but from patients who met the inclusion/exclusion criteria for CHANGE and

decided to voluntarily enroll in the program. This non-randomized sample limits the generalizability of the results and the robustness of the statistical analyses. The clinical findings of this study may only be generalizable to patients with MetS who fit the inclusion/exclusion criteria for CHANGE. The small sample size of the study is because CHANGE program was intended as a demonstration project and a pilot, which led to a limited number of patients being selected and enrolled in the program. The CHANGE program was also only offered during regular office hours, which limited the number of patients to those who could attend an intensive 12-month program during regular hours. This limited availability could have led to the relatively older population of participants in CHANGE ($M$=60.4). The clinical findings can be strengthened through randomly sampling participants from the general population in future studies and offering the CHANGE program outside of regular hours. To address the limitations of a small sample size, this study could be replicated with either a larger population or a larger sample size. This would increase the robustness of the statistical tests and improve generalizability to a wider population. Greater access to longitudinal data covering more variables and a greater sample size over a longer time frame would permit more extensive investigation of program impacts on a larger number of system-level variables over a longer-term. This would provide a richer description of the patient journey and improve the generalizability of the data.

The clinical perspectives included in the key informant interviews were limited to dietitians and family physicians due to convenience sampling. This was particularly limiting because the kinesiologist could have been able to provide more information around the exercise component of the CHANGE program. Also due to the convenience sampling is that the clinicians were only brought in at the end of the study to examine the clinical significance of the

statistical findings. Including the clinicians throughout the study would have been beneficial because they could provide feedback around the design of the study overall, such as suggesting additional datasets to integrate and providing feedback around the statistical analysis of the clinical indicators, such as the lipid profile. To address the limitation of the limited number of clinician perspectives in this study, future studies could include kinesiologists. Going one step further, there could be open the recruitment of all clinicians to increase the number of clinical perspectives that are included. For example, a chronic disease management nurse may be able to provide some insight into how the patient is managing their chronic diseases, and a pharmacist may be able to provide information around medication use. Additionally, if a specific perspective is missing, but vital for the interpretation of results, clinicians outside the specific program may be recruited to access this clinical perspective. To address the limited engagement of the clinicians in this study, future studies could include clinicians throughout the study process to include the clinical perspective throughout the study.

The study was also limited by the small number of clinicians (n=6) recruited for the key informant interviews, which decreased the number of clinical perspectives available for the interpretation of the statistical findings. Attempts were made to recruit clinicians from each of the three clinical groups involved with the CHANGE program: dietitians, kinesiologist, and physicians. However, attempts to recruit kinesiologists for an interview were not successful, and their perspective was therefore not included in this study. The kinesiologists could have provided more insight into the exercise component of CHANGE, as they are actively involved in the exercise component of the program. Options to mitigate this limitation would include increasing the time frame of study, or considering the more user-friendly option of a survey to obtain feedback from clinicians.

The patient perspective is missing from this study, thus the personal significance (Bothe & Richardson, 2011) of the improvements in clinical indicators was not investigated. Bothe & Richardson (2011) argue that a complete examination of a clinical intervention must include assessment of personal significance, which is currently missing from this, and many other studies. To address the missing patient perspective, future studies could include patients in key informant interviews, which would allow for their perspective of how CHANGE has affected them and their health in their lives outside of the clinic.

**Final Thoughts**

This study highlights the potential of integrated datasets to be used for exploring relationships among variables in a complex health system. Specifically, the study contributes to the literature in two ways: First the analysis of the integrated dataset generated two findings that were not accessible without an integrated dataset. Second, this study highlights the importance of including clinical experts when working with clinical indicators to incorporate their clinical perspective into the interpretation of the statistical findings.

The study findings have three important implications: (1) for promoting use of integrated datasets by encouraging researchers and health authorities alike to mitigate barriers in health data access for research and evaluation, (2) advocating for the inclusion of clinical perspectives in interpreting quantitative health findings, and (3) enhancing evaluations of complex health systems through increasing the types of data that are used in program- and system-level evaluations. Taken together, an emerging standard of health program evaluation is beginning to take shape, where system-level datasets are integrated with program-level datasets to explore the nuanced relationships between health programs and the wider health system.

## References


Alberta Health. (2013). *Health System Outcomes and Measurement Framework*. Retrieved from http://www.health.alberta.ca/documents/PMIS-Outcomes-Measurement-Framework-2014.pdf

Alberta Health Services. (2013). *Alberta Health Services Annual Report 2012-2013*. Retrieved from https://www.albertahealthservices.ca/Publications/ahs-pub-2012-2013-annual-report.pdf

Alberta Health. (2015). *Alberta Netcare Portal Glossary Quick Reference*. Retrieved from http://www.albertanetcare.ca/learningcentre/documents/EMR-Glossary.pdf

Alberta Health. (2016). *Health services in Alberta*. Retrieved July 30, 2016, from http://www.health.alberta.ca/health-services.html

Alberta Health Services. (2015*). Alberta Health Services' 2015-2018 Health Plan and Business Plan*. Retrieved from http://www.albertahealthservices.ca/assets/about/publications/ahs-pub-health-business-plan-2015-18.pdf

Alkin, M. C. (2011). *Evaluation essentials from A to Z*. New York: Guilford Press.

Alkin, M. C., & Taut, S. M. (2002). Unbundling evaluation use. *Studies in Educational Evaluation*, *29*(1), 1–12. https://doi.org/10.1016/S0191-491X(03)90001-0

Babio, N., Bulló, M., Basora, J., Martínez-González, M. A., Fernández-Ballart, J., Márquez-Sandoval, F., … Salas-Salvadó, J. (2009). Adherence to the Mediterranean diet and risk of metabolic syndrome and its components. *Nutrition, Metabolism and Cardiovascular Diseases*, *19*(8), 563–570. https://doi.org/10.1016/j.numecd.2008.10.007

Babio, Nancy, Toledo, E., Estruch, R., Ros, E., Martínez-González, M. A., Castañer, O., … others. (2014). Mediterranean diets and metabolic syndrome status in the PREDIMED

randomized trial. *Canadian Medical Association Journal*, *186*(17), E649–E657.

Bo, S., Ciccone, G., Baldi, C., Benini, L., Dusio, F., Forastiere, G., … Pagano, G. (2007). Effectiveness of a Lifestyle Intervention on Metabolic Syndrome. A Randomized Controlled Trial. *Journal of General Internal Medicine*, *22*(12), 1695–1703. https://doi.org/10.1007/s11606-007-0399-6

Boon, H., MacPherson, H., Fleishman, S., Grimsgaard, S., Koithan, M., Norheim, A. J., & Walach, H. (2007). Evaluating Complex Healthcare Systems: A Critique of Four Approaches. *Evidence-Based Complementary and Alternative Medicine*, *4*(3), 279–285. https://doi.org/10.1093/ecam/nel079

Bothe, A. K., & Richardson, J. D. (2011). Statistical, practical, clinical, and personal significance: definitions and applications in speech-language pathology. *American Journal of Speech-Language Pathology*, *20*(3), 233–242. https://doi.org/10.1044/1058-0360(2011/10-0034)

Bowden, S. C., Harrison, E. J., & Loring, D. W. (2014). Evaluating Research for Clinical Significance: Using Critically Appraised Topics to Enhance Evidence-based Neuropsychology. *The Clinical Neuropsychologist*, *28*(4), 653–668. https://doi.org/10.1080/13854046.2013.776636

Bowen, S. (2012). *A Guide to Evaluation in Health Research*. Retrieved from http://www.cihr-irsc.gc.ca/e/documents/kt_lm_guide_evhr-en.pdf

Braun, V., & Clarke, V. (2006). Using thematic analysis in psychology. *Qualitative Research in Psychology*, *3*(2), 77–101. https://doi.org/10.1191/1478088706qp063oa

Campbell, M., Fitzpatrick, R., Haines, A., Kinmonth, A. L., Sandercock, P., Spiegelhalter, D., & Tyrer, P. (2000). Framework for design and evaluation of complex interventions to

improve health. *BMJ (Clinical Research Ed.)*, *321*(7262), 694–696. https://doi.org/10.1136/bmj.321.7262.694

Case, C. C., Jones, P. H., Nelson, K., O'Brian Smith, E., & Ballantyne, C. M. (2002). Impact of weight loss on the metabolic syndrome. *Diabetes, Obesity and Metabolism*, *4*(6), 407–414.

Cecchini, M., Sassi, F., Lauer, J. A., Lee, Y. Y., Guajardo-Barron, V., & Chisholm, D. (2010). Tackling of unhealthy diets, physical inactivity, and obesity: health effects and cost-effectiveness. *The Lancet*, *376*(9754), 1775–1784. https://doi.org/10.1016/S0140-6736(10)61514-0

Christie, C. A. (2007). Reported Influence of Evaluation Data on Decision Makers' Actions: An Empirical Examination. *American Journal of Evaluation*, *28*(1), 8–25. https://doi.org/10.1177/1098214006298065

Clinkscales, J. D., Fesmire, F. M., Hennings, J. R., Severance, H. W., Seaberg, D. C., & Patil, N. (2016). The Effect of Emergency Medicine Residents on Clinical Efficiency and Staffing Requirements. *Academic Emergency Medicine*, *23*(1), 78–82. https://doi.org/10.1111/acem.12834

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed). Hillsdale, N.J: L. Erlbaum Associates.

Creswell, J. W. (2010). *Research design: qualitative, quantitative, and mixed methods approaches* (3. ed., [Nachdr.]). Los Angeles: SAGE Publ.

Donnelly, J. E., Blair, S. N., Jakicic, J. M., Manore, M. M., Rankin, J. W., & Smith, B. K. (2009). Appropriate Physical Activity Intervention Strategies for Weight Loss and Prevention of Weight Regain for Adults: *Medicine & Science in Sports & Exercise*, *41*(2), 459–471.

https://doi.org/10.1249/MSS.0b013e3181949333

Estruch, R., Ros, E., Salas-Salvadó, J., Covas, M.-I., Corella, D., Arós, F., … Martínez-González, M. A. (2013). Primary Prevention of Cardiovascular Disease with a Mediterranean Diet. *New England Journal of Medicine*, *368*(14), 1279–1290. https://doi.org/10.1056/NEJMoa1200303

Finn, P. (2003). Evidence-based treatment of stuttering: *Journal of Fluency Disorders*, *28*(3), 209–218. https://doi.org/10.1016/S0094-730X(03)00039-1

Follette, W. C., & Callaghan, G. M. (2001). The evolution of clinical significance. *Clinical Psychology: Science and Practice*, *8*(4), 431–435.

Forbes, D., Forbes, S. C., Blake, C. M., Thiessen, E. J., & Forbes, S. (2015). Exercise programs for people with dementia. In The Cochrane Collaboration (Ed.), *Cochrane Database of Systematic Reviews*. Chichester, UK: John Wiley & Sons, Ltd. Retrieved from http://doi.wiley.com/10.1002/14651858.CD006489.pub4

Ford, E. S., Li, C., & Sattar, N. (2008). Metabolic Syndrome and Incident Diabetes: Current state of the evidence. *Diabetes Care*, *31*(9), 1898–1904. https://doi.org/10.2337/dc08-0423

Gelman, A., & Hill, J. (2009). *Data analysis using regression and multilevel/hierarchical models* (11. printing). Cambridge: Cambridge Univ. Press.

Gillespie, S. M., Shah, M. N., Wasserman, E. B., Wood, N. E., Wang, H., Noyes, K., … McConnochie, K. M. (2016). Reducing Emergency Department Utilization Through Engagement in Telemedicine by Senior Living Communities. *Telemedicine and E-Health*. https://doi.org/10.1089/tmj.2015.0152

Goldenberg, R., & Punthakee, Z. (2013). Definition, Classification and Diagnosis of Diabetes, Prediabetes and Metabolic Syndrome. *Canadian Journal of Diabetes*, *37*, S8–S11.

https://doi.org/10.1016/j.jcjd.2013.01.011

Gomez-Huelgas, R., Jansen-Chaparro, S., Baca-Osorio, A. J., Mancera-Romero, J., Tinahones, F. J., & Bernal-López, M. R. (2015). Effects of a long-term lifestyle intervention program with Mediterranean diet and exercise for the management of patients with metabolic syndrome in a primary care setting. *European Journal of Internal Medicine*, *26*(5), 317–323. https://doi.org/10.1016/j.ejim.2015.04.007

Government of Alberta. Health Information Act (2016). Retrieved from http://www.qp.alberta.ca/documents/Acts/H05.pdf

Grundy, S. M. (2004). Definition of Metabolic Syndrome: Report of the National Heart, Lung, and Blood Institute/American Heart Association Conference on Scientific Issues Related to Definition. *Circulation*, *109*(3), 433–438. https://doi.org/10.1161/01.CIR.0000111245.75752.C6

Guijt, I. (2014). Participatory approaches. *Methodological Brief: Impact Evaluation*, *5*. Retrieved from http://betterevaluation.org/sites/default/files/Participatory_Approaches_ENG.pdf

Hemmila, M. R., Jakubus, J. L., Maggio, P. M., Wahl, W. L., Dimick, J. B., Campbell, D. A., & Taheri, P. A. (2008). Real money: Complications and hospital costs in trauma patients. *Surgery*, *144*(2), 307–316. https://doi.org/10.1016/j.surg.2008.05.003

Hollander, M. J., Miller, J. A., & Kadlec, H. (2010). Evaluation of healthcare services: asking the right questions to develop new policy and program-relevant knowledge for decision-making. *Healthcare Quarterly (Toronto, Ont.)*, *13*(4), 40–47.

Houle, T. T., & Stump, D. A. (2008). Statistical Significance Versus Clinical Significance. *Seminars in Cardiothoracic and Vascular Anesthesia*, *12*(1), 5–6. https://doi.org/10.1177/1089253208316440

Hussey, P. S., Schneider, E. C., Rudin, R. S., Fox, D. S., Lai, J., & Pollack, C. E. (2014). Continuity and the Costs of Care for Chronic Disease. *JAMA Internal Medicine*, *174*(5), 742. https://doi.org/10.1001/jamainternmed.2014.245

Jacobson, N. S., & Truax, P. (1991). Clinical significance: a statistical approach to defining meaningful change in psychotherapy research. *Journal of Consulting and Clinical Psychology*, *59*(1), 12–19.

Jacobson, Neil S., Follette, W. C., & Revenstorf, D. (1984). Psychotherapy outcome research: Methods for reporting variability and evaluating clinical significance. *Behavior Therapy*, *15*(4), 336–352. https://doi.org/10.1016/S0005-7894(84)80002-7

Jadad, A. R. (2002). *Randomised controlled trials: a user's guide* (Reprint). London: BMJ Books.

Johnson, C. J. (2006). Getting started in evidence-based practice for childhood speech-language disorders. *American Journal of Speech-Language Pathology*, *15*(1), 20–35.

Johnson, K., Greenseid, L. O., Toal, S. A., King, J. A., Lawrenz, F., & Volkov, B. (2009). Research on Evaluation Use: A Review of the Empirical Literature From 1986 to 2005. *American Journal of Evaluation*, *30*(3), 377–410. https://doi.org/10.1177/1098214009341660

Kastorini, C.-M., Milionis, H. J., Esposito, K., Giugliano, D., Goudevenos, J. A., & Panagiotakos, D. B. (2011). The Effect of Mediterranean Diet on Metabolic Syndrome and its Components. *Journal of the American College of Cardiology*, *57*(11), 1299–1313. https://doi.org/10.1016/j.jacc.2010.09.073

Kazdin, A. E. (1999). The meanings and measurement of clinical significance. Retrieved from http://psycnet.apa.org/journals/ccp/67/3/332/

Kephart, G., Canadian Institute for Health Information, & Canadian Population Health Initiative. (2002). *Barriers to accessing & analyzing health information in Canada*. Ottawa: Canadian Institute for Health Information.

King, J. A., & Pechman, E. M. (1984). Pinning a Wave to the Shore: Conceptualizing Evaluation Use in School Systems. *Educational Evaluation and Policy Analysis*, *6*(3), 241–251. https://doi.org/10.3102/01623737006003241

Kirk, R. E. (2001). Promoting Good Statistical Practices: Some Suggestions. *Educational and Psychological Measurement*, *61*(2), 213–218. https://doi.org/10.1177/00131640121971185

Koplan, J. P., Milstein, R., & Wetterhall, S. (1999). Framework for program evaluation in public health. *MMWR: Recommendations and Reports*, *48*, 1–40. https://doi.org/10.1037/e547822006-001

Kramer, M. K., McWilliams, J. R., Chen, H.-Y., & Siminerio, L. M. (2011). A Community-Based Diabetes Prevention Program: Evaluation of the Group Lifestyle Balance Program Delivered by Diabetes Educators. *The Diabetes Educator*, *37*(5), 659–668. https://doi.org/10.1177/0145721711411930

Lakka, H.-M., Laaksonen, D. E., Lakka, T. A., Niskanen, L. K., Kumpusalo, E., Tuomilehto, J., & Salonen, J. T. (2002). The metabolic syndrome and total and cardiovascular disease mortality in middle-aged men. *JAMA*, *288*(21), 2709–2716. https://doi.org/10.1001/jama.288.21.2709

Lee, H. S. (2013). The Impact of Emergency Room Utilization by Depression Patients on Medical Treatment Expense in Korea. *Osong Public Health and Research Perspectives*, *4*(5), 240–245. https://doi.org/10.1016/j.phrp.2013.09.007

Leviton, L. C., & Hughes, E. F. X. (1981). Research On the Utilization of Evaluations: A Review and Synthesis. *Evaluation Review*, *5*(4), 525–548. https://doi.org/10.1177/0193841X8100500405

Leys, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, *49*(4), 764–766. https://doi.org/10.1016/j.jesp.2013.03.013

Madsen, M., Kiuru, S., Castrèn, M., & Kurland, L. (2015). The level of evidence for emergency department performance indicators: systematic review. *European Journal of Emergency Medicine*, *22*(5), 298–305. https://doi.org/10.1097/MEJ.0000000000000279

Marjoua, Y., & Bozic, K. J. (2012). Brief history of quality movement in US healthcare. *Current Reviews in Musculoskeletal Medicine*, *5*(4), 265–273. https://doi.org/10.1007/s12178-012-9137-8

Marshall, M. N. (1996). The key informant technique. *Family Practice*, *13*(1), 92–97.

Mensah, G. A., & Brown, D. W. (2007). An overview of cardiovascular disease burden in the United States. *Health Affairs*, *26*(1), 38–48. https://doi.org/10.1377/hlthaff.26.1.38

Nguyen, D. L., & DeJesus, R. S. (2010). Increased Frequency of No-Shows in Residents' Primary Care Clinic Is Associated With More Visits to the Emergency Department. *Journal of Primary Care & Community Health*, *1*(1), 8–11. https://doi.org/10.1177/2150131909359930

Ogles, B. M., Lunnen, K. M., & Bonesteel, K. (2001). Clinical significance: History, application, and current practice. *Clinical Psychology Review*, *21*(3), 421–446.

Olde, D., Alpert, P., & Dalusung-Angosta, A. (2013). Metabolic syndrome: Clinical perspective

for best practice: Metabolic Syndrome: Clinical Perspective for Best Practice. *Journal of the American Association of Nurse Practitioners*, *25*(12), 644–652. https://doi.org/10.1002/2327-6924.12048

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed). Thousand Oks, Calif.: SAGE.

Porter, M., Diaz, V. A., Gavin, J. K., Zacarias, A., Dickerson, L., Hueston, W. J., & Carek, P. J. (2015). Cost and Utilization: Hospitalized Patients on a Family Medicine Service. *Southern Medical Journal*, *108*(6), 364–369. https://doi.org/10.14423/SMJ.0000000000000293

Primary Care Network Project Management Office. (2016). *Alberta PCNs*. Retrieved July 30, 2016, from https://www.pcnpmo.ca/alberta-pcns/Pages/default.aspx

Public Health Ontario. (2015). *At A Glance: The ten steps for conducting an evaluation*. Retrieved from http://www.publichealthontario.ca/en/eRepository/At_A_Glance_Evaluation_2015.pdf

Rao, D. P., Dai, S., Lagacé, C., & Krewski, D. (2014). Metabolic syndrome and chronic disease. *Women*, *886*(50.4), 49–8.

Riediger, N. D., & Clara, I. (2011). Prevalence of metabolic syndrome in the Canadian adult population. *Canadian Medical Association Journal*, *183*(15), E1127–E1134. https://doi.org/10.1503/cmaj.110070

Rubenfire, M., Mollo, L., Krishnan, S., Finkel, S., Weintraub, M., Gracik, T., … Oral, E. A. (2011). The Metabolic Fitness Program: Lifestyle modification for the metabolic syndrome using the resources of cardiac rehabilitation. *Journal of Cardiopulmonary Rehabilitation and Prevention*, *31*(5), 282–289. https://doi.org/10.1097/HCR.0b013e318220a7eb

Saunders, N. R., To, T., Parkin, P. C., & Guttmann, A. (2015). Emergency Department Revisits by Urban Immigrant Children in Canada: A Population-Based Cohort Study. *The Journal of Pediatrics*. https://doi.org/10.1016/j.jpeds.2015.11.037

Shenton, A. K. (2004). Strategies for Ensuring Trustworthiness in Qualitative Research Projects. *Education for Information*, *22*(2), 63–75.

Sidorov, J., Shull, R., Tomcavage, J., Girolami, S., Lawton, N., & Harris, R. (2002). Does Diabetes Disease Management Save Money and Improve Outcomes? : A report of simultaneous short-term savings and quality improvement associated with a health maintenance organization-sponsored disease management program among patients fulfilling health employer data and information set criteria. *Diabetes Care*, *25*(4), 684–689. https://doi.org/10.2337/diacare.25.4.684

Silverman, R. A., Schleicher, M. G., Valente, C. J., Kim, M., Romanenko, Y., Schulman, R. C., … Jornsay, D. L. (2015). Prevalence of undiagnosed dysglycemia in an emergency department observation unit: Undiagnosed Dysglycemia Observation Unit. *Diabetes/Metabolism Research and Reviews*, n/a-n/a. https://doi.org/10.1002/dmrr.2674

Statistics Canada. (2011). *Population, urban and rural, by province and territory (Alberta)* (table). *Summary Tables.* Retrieved April 24, 2017 from http://www.statcan.gc.ca/tables-tableaux/sum-som/l01/cst01/demo62j-eng.htm

Statistics Canada. (2015). *Use of administrative data*. Retrieved July 30, 2016, from http://www.statcan.gc.ca/pub/12-539-x/2009001/administrative-administratives-eng.htm

Stolberg, H. O., Norman, G., & Trop, I. (2004). Randomized controlled trials. *American Journal of Roentgenology*, *183*(6), 1539–1544. https://doi.org/10.2214/ajr.183.6.01831539

Tabachnick, B. G., & Fidell, L. S. (20). *Using multivariate statistics* (5. ed., Pearson internat. ed.,

[Nachdr.]). Boston, Mass.: Pearson/Allyn and Bacon.

Tran, B., Falster, M. O., Girosi, F., & Jorm, L. (2016). Relationship between use of general practice and healthcare costs at the end of life: a data linkage study in New South Wales, Australia. *BMJ Open*, *6*(1), e009410. https://doi.org/10.1136/bmjopen-2015-009410

U.S. Department of Health and Human Services Centers for Disease Control and Prevention Office of the Director, Office of Strategy and Innovation. (n.d.). Introduction to program evaluation for public health programs: A self-study guide. Retrieved from https://www.cdc.gov/eval/guide/cdcevalmanual.pdf

van Panhuis, W. G., Paul, P., Emerson, C., Grefenstette, J., Wilder, R., Herbst, A. J., … Burke, D. S. (2014). A systematic review of barriers to data sharing in public health. *BMC Public Health*, *14*(1), 1144. https://doi.org/10.1186/1471-2458-14-1144

Walport, M., & Brest, P. (2011). Sharing research data to improve public health. *Lancet (London, England)*, *377*(9765), 537–539. https://doi.org/10.1016/S0140-6736(10)62234-9

Wang, Q. C., Chawla, R., Colombo, C. M., Snyder, R. L., & Nigam, S. (2014). Patient-Centered Medical Home Impact on Health Plan Members With Diabetes: *Journal of Public Health Management and Practice*, *20*(5), E12–E20. https://doi.org/10.1097/PHH.0b013e3182a8eb3d

Willett, W. C., Sacks, F., Trichopoulou, A., Drescher, G., Ferro-Luzzi, A., Helsing, E., & Trichopoulos, D. (1995). Mediterranean diet pyramid: a cultural model for healthy eating. *The American Journal of Clinical Nutrition*, *61*(6 Suppl), 1402S–1406S.

Wozniak, L., Soprovich, A., Mundt, C., Johnson, J. A., & Johnson, S. T. (2015). Contextualizing the Proven Effectiveness of a Lifestyle Intervention for Type 2 Diabetes in Primary Care: A Qualitative Assessment Based on the RE-AIM Framework. *Canadian Journal of*

*Diabetes*, *39*, S92–S99. https://doi.org/10.1016/j.jcjd.2015.05.003

Yoon, N.-H., Yoo, S., & Kim, H. (2014). Composition of Metabolic Syndrome Among Korean Adults in a Lifestyle Modification Intervention. *Osong Public Health and Research Perspectives*, *5*(6), 370–377. https://doi.org/10.1016/j.phrp.2014.10.005

Zimmet, P., Magliano, D., Matsuzawa, Y., Alberti, G., & Shaw, J. (2005). The metabolic syndrome: a global public health problem and a new definition. *Journal of Atherosclerosis and Thrombosis*, *12*(6), 295–300. https://doi.org/10.5551/jat.12.295

## Appendix A: Dataset Variables

| Data Source | Dataset | Variables | Description |
|---|---|---|---|
| Alberta Health Services (System Outcomes Dataset) | Alberta Ambulatory Care Report System (AACRS) | | Unique identifier assigned by DIMR |
| | | | Age |
| | | | Gender |
| | | | Diagnosis Code 1-10 |
| | | | Diagnosis Description 1-10 |
| | | | Visit Date |
| | | | Disposition Date |
| | | | CCI Intervention Code 1-10 |
| | | | CCI Intervention Description 1-10 |
| | | | Provider Service 1-8 |
| | | | Provider Type 1-8 |
| | | | Comprehensive Ambulatory Classification System (CACS) Code |
| | | | CACS Description |
| | | | Resource Intensity Weight (RIW) |
| Physicians (Program Outcomes Dataset) | CHANGE Dataset | PHN | |
| | | Sex | Sex |
| | | DoB | Date of Birth |
| | | At baseline, 3, 6, 9, and 12 months: | |
| | | BMI | Body Mass Index |
| | | WC | Waist Circumference |
| | | BP | Blood Pressure |
| | | FBG | Fasting Blood Glucose |
| | | TG | Triglycerides |
| | | Crea | Creatinine |
| | | HbA1C | Hemoglobin A1C |
| | | TC | Total Cholesterol |
| | | LDL | Low-density lipoprotein |
| | | HDL | High-density lipoprotein |

## Appendix B: Key Informant Interview Participant Recruitment Email

My name is Alvin Yapp and I am the Evaluation Specialist with Edmonton Oliver Primary Care Network, but I am also a student from the University of Alberta. I am writing you in my capacity as a student to invite you to participate in my research study investigating the effect the CHANGE program had on patients' emergency department utilization.

Your position as a clinician or policy-maker with the CHANGE program makes you eligible for this study. I obtained your contact information from the CHANGE program.

If you decide to participate in this study, you will be asked to participate in a one-on-one interview with me where you will be asked to provide your expert opinion in order to contextualize current findings regarding the CHANGE program and the effects it may have had on patients' emergency department utilization. I would like to record the audio from this interview which will be transcribed and analyzed.

Participation in this study is completely voluntary. You can choose not to participate without any negative repercussions. If you'd like to participate, or have any questions about the study, please email me at ayapp@ualberta.ca or at 587.408.8426

Thanks for your consideration.

Sincerely,

Alvin Yapp

## Appendix C: Key Informant Interview Participant Consent Form

### PARTICIPANT CONSENT FORM

**Title of Study:** *Integrating Administrative Datasets with Program Datasets: The Effect of a Lifestyle Intervention Program on Emergency Department Utilization of Patients with Metabolic Syndrome*

**Principal Investigator:** Cheryl Poth ( cpoth@ualberta.ca ; 780.492.1144 )
**Co-investigator:** Alvin Yapp ( ayapp@ualberta.ca ; 587.408.8426 )

---

**Why am I being asked to take part in this research study?**

In order to examine the effect of the Canadian Health Advanced by Nutrition and Graded Exercise (CHANGE) program on participants' emergency department utilization, administrative data from Alberta Health Services (AHS) has been linked with data from the CHANGE program. The relationships between the clinical indicators of metabolic syndrome and emergency department utilization were also explored. The statistical analyses have been completed, but the clinical significance of these findings still needs to be identified.

As a clinician or policy-making stakeholder involved with the CHANGE program, your expert feedback is valuable in contextualizing the findings of the statistical analyses. Your contact information was obtained through the CHANGE program in order to invite you to participate in an interview to provide your expert opinion on these findings.

In addition to informing further development of the CHANGE program, the results of this study will be used in support of my thesis, which is a partial requirement of for the degree of Masters of Education in Measurement, Evaluation and Cognition.

Before you make a decision, the interviewer will go over this form with you. You are encouraged to ask questions if you feel anything needs to be made clearer. You will be given a copy of this form for your records.

**What is the reason for doing the study?**

This study will provide an illustrative example of how the Integrated Dataset created by the integration of administrative data from AHS (Emergency Use Dataset) with clinical indicators from a lifestyle intervention program (Program Outcomes Dataset) can address clinical and policy questions that would not have been possible with either data set alone.

**What will I be asked to do?**

The interview will be no longer than 60 minutes and will be conducted in a one-on-one fashion at a mutually convenient location where a private conversation can be had. The interview will be audio recorded and transcribed for analysis. This is a key informant interview where you will provide your expert opinion on questions regarding:

1. The significance of the findings from the statistical analyses, from your perspective, for clinical practice or policy-making.
2. Unanticipated results from these findings.
3. How these results might inform your practice.
4. The next steps for research in this area.

Transcripts and recordings of the interview will not be made available, however, a summary of the findings from this study will be provided upon request.

**What are the risks and discomforts?**

There is a small chance that confidentiality may be breeched, however, every precaution will be taken to mitigate this risk as much as possible. Audio recordings will be transcribed for analysis, and the raw audio files will be stored on an encrypted hard drive when not in use.

You may incur a small cost, such as parking. This will not be reimbursed.

**What are the benefits to me?**

The information we get from doing this study has the strong potential to help us better understand the relationship between participation in a lifestyle intervention program and emergency department utilization for patients with metabolic syndrome, and the relationships between the clinical indicators of metabolic syndrome and emergency department utilization.

Better understanding these relationships will allow for better planning of programs to help prevent the progression of metabolic syndrome to chronic diseases.

**Do I have to take part in the study?**

Being in this study is your choice. If you decide to be in the study, you can change your mind and stop being in the study at any time, and it will in no way affect your relationship with the CHANGE program that you are entitled to.

After the interview is complete, you will be asked one last time if you consent to using the information gathered in the interview. You will be able to withdraw your responses at this time. If you consent to the use of your data, your data will be used in the study without any identifiers.

**Will my information be kept private?**

During the study we will be collecting data about you. We will do everything we can to make sure that this data is kept private. No data relating to this study that includes your name will be released outside of the researcher's office or published by the researchers. Sometimes, but law, we may have to release your information with your name so we cannot guarantee absolute privacy. However, we will make every legal effort to make sure that your information is kept private.

Audio recordings and meeting notes will be kept confidential. Only the principal investigator and co-investigator will have access to these. Participation in this interview will not be identified in the dissemination of the research. Identifiable information will be edited out of any quotes used. Audio recordings will be kept on an encrypted hard drive for a minimum of 5 years following completion of the research project, and appropriate destroyed in a way that ensures privacy and confidentiality.

**What if I have questions?**

*If you have any questions about the research now or later, please contact Alvin Yapp ( 587.408.8426 ) or Cheryl Poth ( 780.492.1144 ).*

*If you have any questions regarding your rights as a research participant, you may contact the Health Research Ethics Board at 780-492-2615. This office has no affiliation with the study investigators.*

**CONSENT**

---

**Title of Study:** Integrating Administrative Datasets with Program Datasets: The Effect of a Lifestyle Intervention Program on Emergency Department Utilization of Patients with Metabolic Syndrome

**Principal Investigator:** Cheryl Poth **Phone Number(s):** 780.492.1144
**Co-Investigator:** Alvin Yapp **Phone Number(s):** 587.408.8426

| | Yes | No |
|---|---|---|
| Do you understand that you have been asked to be in a research study? | ☐ | ☐ |
| Have you read and received a copy of the attached Information Sheet? | ☐ | ☐ |
| Do you understand the benefits and risks involved in taking part in this research study? | ☐ | ☐ |
| Have you had an opportunity to ask questions and discuss this study? | ☐ | ☐ |
| Do you understand that you are free to leave the study at any time, without having to give a reason and without penalty? | ☐ | ☐ |
| Has the issue of confidentiality been explained to you? | ☐ | ☐ |
| Do you understand who will have access to your study records? | ☐ | ☐ |

Who explained this study to you?
_____

I agree to take part in this study:

Signature of Research Participant
_____

(Printed Name)
_____

Date:_____

Signature of Investigator or Designee _____ Date _____

## Appendix D: Key Informant Interview Protocol

**Key Informant Interview Protocol**

This interview will follow these procedures:

1. The location of the interview will be in a place of mutual convenience where a private conversation can be had.
2. The interview will be no longer than one hour in length.
3. The co-investigator (Alvin Yapp) will be conducting the interviews.
4. It will be confirmed that the participants have signed the consent forms before the interview will be started, and that they understand:
    a. their right to refuse to answer any questions
    b. participating or not participating will not affect their standing within the organization
    c. that the interview will be digitally recorded with their permission
5. Thanking participants for their participation and feedback.

The primary purpose of these interviews is to contextualize the quantitative findings from the integrated dataset using the participants' expert opinions, and how these findings may be clinically relevant and/or impact clinical practice.

**Preamble**

Canadian Health Advanced by Nutrition and Graded Exercise (CHANGE) is a demonstration project that provides individualized diet and exercise programs for patients with Metabolic Syndrome. Three sites in Canada have been chosen to participate in this project, one of which is the Edmonton Oliver Primary Care Network (EOPCN). Enrolled patients are supported with weekly customized visits to the PCN Dietitian and Kinesiologist for the first 3 months and then monthly visits for 9 additional months (1 year total).

Patients were screened and those who met the inclusion/exclusion criteria for the program were presented to family physicians. Family physicians would then approve patients for the program, and the patients contacted to discuss participation in the program.

The patient inclusion/exclusion criteria for the CHANGE program are:

[present inclusion/exclusion criteria, highlighting the clinical indicators]

Diet Interventions included:

- Individualized diet plans based on the Mediterranean diet, risk factors and needs
- Stepped approach to weight loss (i.e., lifestyle changes, weight loss, blood sugars, and then hypertension/dyslipidemia)
- Counselling weekly for the first 3 months and then follow-up with the patient's progress monthly for 9 months
- Behavioural strategies

Exercise Intervention:

- Individualized and graded exercise plan

As a part of the program, the following clinical parameters were measured and compared from baseline to 12-months program:

- Routine blood work
    - Fasting blood glucose
    - Triglyceride
    - HDL-C (including LDL-C and total cholesterol)
- CBCD, Creatinine, HbA1C

- Waist circumference
- Height and weight (BMI)
- Blood Pressure

Statistical analyses were done on these indicators to identify any differences in patients at baseline and 12-months. Although there may be statistically significant results, this does not always mean that there are clinically significant results. The purpose of this interview is to get your clinical interpretation of my statistical findings related to the CHANGE program. Your responses will be kept anonymous, so please be candid with your responses. I will be asking some questions about your practice to contextualize your clinical interpretations.

**Introduction**

1. How many patients do you see in your practice?
    a. How many of these have MetS?
2. What is your experience working with patients with MetS?

**Program Outcomes Dataset Analysis**

[Show participant CLINICAL INDICATORS FIGURE]

These are the clinical indicators that we measured at baseline and at 12-months. We did paired comparisons of these to see which indicators saw a significant improvement after one year in the CHANGE program.

3. Looking at these clinical indicators measured in CHANGE, and before I reveal to you which were significantly different after CHANGE, which do you think would be significantly different after a year in CHANGE? [Mark responses on CLINICAL INDICATOR RESULTS]
    a. What is it about these indicators that make you think CHANGE would impact them?
    b. What is it about the ones that you don't think would have changed that make you think CHANGE would not impact them?
    c. Do you think that other [clinician type] would make similar choices?

[Show participant CLINICAL INDICATOR RESULTS]

According to our statistical analyses, we found statistically significant improvements in most of the clinical indicators related to MetS 12-months post-enrolment in the CHANGE program. [Point out differences from their expectations and probe the differences]

4. Looking now at the clinical indicators that you expected to show a difference compared to those that actually showed a difference, what do you think are some reasons for this discrepancy? [for each indicator that was different]
    a. Are the indicators that showed a statistically significant change surprising to you?
5. Looking at the mean difference between baseline and 12-month scores across the clinical indicators, are these clinically significant changes?
    a. How would these changes be significant for the patient?
6. There were three indicators (FBG, TC, TG) that did not show a statistically significant change. What are some possible reasons that we did not see changes in these indicators? [If answered from above, do not ask]
7. Based on your experiences, what is the typical rate of reversal for MetS?
8. How is a change in MetS status significant for a patient?

As another part of the analysis, we looked at whether or not patients met the criteria for (3 of 5 of clinical indicators…) MetS, based on their clinical indicators. We did not include pharmacotherapy status into this; this is solely based on the numeric criteria.

Including those with pharmacotherapy, we saw that 51.6% of patients had their clinical indicators under control, an increase from 26.6% from before CHANGE.

9. Would you consider this to be a success.
10. With these results, what effect would a program like CHANGE have on the wider health system?

**System Outcomes Dataset Analysis**

For this, we were interested in linking the data from CHANGE to system-utilization data to examine the effect of the program on system-level measures, such as utilization of other health services. The dataset that we were able to get access to was the emergency department visits, so we were able to look at the patients' ED utilization pre- and during- change.

11. Now that you know about the effect of CHANGE on the clinical indicators of MetS, what do you think the pattern of patients' utilization of the ED be after CHANGE?
    a. Follow up with prompts if appropriate: What is the reasoning behind that thought?

[Show CHANGE IN ED UTILIZATION RESULTS]

We found that the patients' ED utilization pre- and during- CHANGE were different, but in an unexpected way – we noticed that the rate of ED utilization increased during the CHANGE program. Although we cannot attribute the increase in visits to the CHANGE program, there appears to be a relationship there.

12. What do you think the reasons are for this increase in ED utilization after enrolment in the CHANGE program?

    [use following prompts if these ideas do not come up;]

    a. Could it be that patients are taking more care of their own health and utilizing health services more?
    b. Injuries now that they undertaking a more active lifestyle?
    c. This is an at-risk population and would have a high rate of ED utilization?
    **d.** [follow-up for issues identified] Is there any change we can make to the CHANGE program in order to mitigate these issues and decrease ED utilization?

[Show FREQUENCY OF ED VISIT]

13. Are these changes in ED visit significant for patients with MetS attending a lifestyle intervention program?

    [Show ED VISIT REASON]

14. These are the reasons for the visit to the ED. Which of these could be related to MetS?
    a. To the CHANGE program?

**Integrated Dataset Analysis**

Since the datasets were integrated, we tried to determine the relationship between the clinical indicators related to MetS and patient ED utilization to see if we would be able to predict that a patient would be at risk for visiting the ED based on these clinical indicators. This is an analysis that would not have been possible without bringing these datasets together.

15. How are high or low values of the clinical indicators related to MetS related to patient ED utilization?

For statistical reasons, not every variable was used in our analysis. The variables we used were: BP, FBG, TG, HDL and WC. We chose these based on the relationships between them, and also because they are the five indicators used as clinical criteria for MetS. Based on our analyses, there appears to be no relationship between these five MetS indicators and ED utilization ($F(8, 55) = .391$, $p=.921$; adjusted $R^2=-.084$)

16. Since we have seen an increase in ED utilization post-enrolment in CHANGE, there appears to be some relationship between CHANGE and ED utilization. What other factors should we try to measure that aren't currently measured in this population regarding ED utilization?
17. What clinical value is there in attempting to find the relationship between clinical indicators, such as the ones used here, and health system utilization?
    a. How would understanding these relationships have an impact on your clinical practice?
18. If you were a funder who was looking to fund a program, would the results of improving clinical indicators and this rate of changing MetS status be persuasive enough to continue funding the program?

a. What additional information/indicators would you be looking for to make that decision?
    i. At the patient-level?
    ii. At the program level?
    iii. At the health system-level?
b. Would you be aware of how these measurements are collected, and how they could be accessed by researchers?

19. Earlier in this interview, we discussed the possible reasons for an increase in ED utilization. How could we further explore these reasons using data?

Thanks for your time and feedback, it has been very helpful. If you think of anything else that may be helpful, please do not hesitate to contact me.

**Inclusion/Exclusion Criteria**

| Inclusion | Exclusion |
|---|---|
| 3 of 5 of…<br>**Blood Pressure** at least one of:<br>● ≥130/85 mmHg<br>● Receiving pharmacotherapy for elevated blood pressure<br>**Fasting Blood Glucose** at least one of:<br>● Fasting Blood Glucose ≥ 5.6mmol/L<br>● Receiving pharmacotherapy for elevated blood glucose levels<br>**Triglyceride** (at least one of...)<br>● Triglycerides of ≥ 1.7 mmol/L<br>● Receiving pharmacotherapy for elevated triglycerides<br>**HDL-C**<br>● <1.0mmol/L (males)<br>● <1.3mmol/L (females)<br>**Abdominal circumference** > ranges specified below:<br>● Abdominal circumference > ranges specified below<br>● Europids, Whites, sub-Saharan Africans, Mediterranean, middle east (Arab) ≥ 94 cm males, 80 cm females<br>● Asian and South Central Americans ≥ 90 cm males and 80 cm females<br>● US and Canadian Whites ≥ 102 cm males, 88 cm females<br>● Ethnicity unclear (use range for Europids) | Inability to speak, read or understanding English and/or French<br><br>Having a medical or physical condition that makes moderate intensity physical activity (like a brisk walk) difficult or unsafe.<br><br>Diagnosis of Type 1 Diabetes Mellitus<br><br>Type 2 Diabetes Mellitus only if any one of the following are present (select all that apply)<br>● Proliferative diabetic retinopathy<br>● Nephropathy (suggested parameters: serum creatinine > 160)<br>● Clinically manifest neuropathy defined as absent ankle jerks<br>● Severe fasting hyperglycemia > 11 mmol/L<br><br>Peripheral vascular disease<br><br>Significant medical co-morbidities, including uncontrolled metabolic disorders (e.g., thyroid, renal, liver), heart disease, stroke and ongoing substance abuse.<br><br>Clinically significant renal failure<br><br>BMI ≥ 35 |

**CLINICAL INDICATORS FIGURE**

Diet Interventions included:

- Individualized diet plans based on the Mediterranean diet, risk factors and needs
- Stepped approach to weight loss (i.e., lifestyle changes, weight loss, blood sugars, and then hypertension/dyslipidemia)
- Counselling weekly for the first 3 months and then follow-up with the patient's progress monthly for 9 months
- Behavioural strategies
  Exercise Intervention:

- Individualized and graded exercise plan
  **Patient Demographics:** 49.4% Female. 50.6% Male. Average Age: 60.41 (37 – 75)

| |
|---|
| **Weight** |
| **Body Mass Index (BMI)** |
| **Systolic Blood Pressure (SBP)** |
| **Diastolic Blood Pressure (DBP)** |
| **Fasting Blood Glucose (FBG)** |
| **Total Cholesterol (TC)** |
| **Low-Density Lipoprotein (LDL)** |
| **High-Density Lipoprotein (HDL)** |
| **Waist Circumference (WC)** |
| **Triglycerides (TG)** |
| **Hemoglobin A1C (HbA1C)** |

**CLINICAL INDICATOR RESULTS**

| Clinical Indicators | N | Baseline Mean | 12M Mean | DIFF (12M-pre) | p | Cohen's d |
|---|---|---|---|---|---|---|
| Weight* | 75 | 89.13kg<br>196.5lbs | 85.83kg<br>189.2lbs | -3.3kg<br>-7.3lbs | < .000 | 1.31 |
| BMI* | 75 | 31.21 | 30.05 | -1.16 | < .000 | 1.29 |
| Systolic BP* | 70 | 129.50 | 124.93 | -4.57 | .003 | .76 |
| Diastolic BP* | 75 | 78.63 | 75.32 | -3.31 | .003 | .73 |
| FBG | 73 | 6.54 | 6.57 | 0.03 | .769 | - |
| TG | 76 | 1.727 | 1.727 | 0 | .998 | - |
| LDL* | 78 | 2.57 | 2.16 | -0.41 | < .000 | 1.23 |
| HDL* | 75 | 1.11 | 1.17 | 0.06 | .001 | -.77 |
| WC* | 72 | 106.08 | 101.56 | -4.52 | < .000 | 1.51 |
| TC | 80 | 4.22 | 4.20 | -0.02 | .839 | - |
| HbA1C* | 70 | 6.34 | 6.12 | -0.22 | < .000 | 1.09 |

**CHANGE IN METS STATUS**

| Clinical Indicator | | BP (M12) | | Total |
|---|---|---|---|---|
| | | Not MetS | MetS | |
| MetS (Baseline) | Not MetS | 12 | 5 | 7 |
| | MetS | 21 | 6 | 7 |
| | Total | 33 | 1 | 4 |

| Clinical Indicator | | BP (M12) | | Total |
|---|---|---|---|---|
| | | Not MetS | MetS | |
| MetS (Baseline) | Not MetS | 18.8% | 7.8% | 26.6% |
| | MetS | 32.8% | 40.6% | 73.4% |
| | Total | 51.6% | 48.4% | 100.0% |

**CHANGE IN ED UTILIZATION RESULTS**

| | N | Baseline Mean | 12M Mean | p | Cohen's d |
|---|---|---|---|---|---|
| Emergency Department Visits* | 85 | .153 | .529 | .012 | -.563 |

## ED VISIT REASONS

| CACS DESCRIPTION | |
|---|---|
| FOLLOW-UP EXAMINATION AND OTHER NON EMERGENT CONDITION | 10 |
| DISEASE OR DISORDER DIGESTIVE SYSTEM | 6 |
| OPEN WOUND AND VASCULAR INJURY | 5 |
| DISEASE OR DISORDER MUSCULOSKELETAL AND CONNECTIVE TISSUE | 5 |
| RENAL FAILURE & OTHER DISORDERS OF THE KIDNEY/URETER | 4 |
| CLOSED FRACTURE OTHER SITE | 3 |
| OTHER DISEASE OR DISORDER CARDIAC SYSTEM | 3 |
| OTHER DISEASE OR DISORDER URINARY SYSTEM | 3 |
| DIGESTIVE SYSTEM CONDITION WITH ACUTE ADMISSION/TRANSFER | 3 |
| OTHER CONDITION WITH ACUTE ADMISSION/TRANSFER | 3 |
| DISEASE OR DISORDER SKIN & BREAST | 3 |
| DISEASE OR DISORDER BLOOD OR BLOOD FORMING ORGAN | 2 |
| EMERGENCY VISIT INTERVENTIONS | 1 |
| MIGRAINE & HEADACHE | 1 |
| BURN | 1 |
| INTERVENTIONS GENERALLY PERFORMED BY NON EMERGENCY DEPARTMENT SERVICE: GI | 1 |
| CONTUSION, DISLOCATION, NERVE & OTHER SOFT TISSUE INJURY | 1 |
| LEFT WITHOUT BEING SEEN OR TRIAGE AND NOT SEEN | 1 |
| DISEASE OR DISORDER MALE ANATOMY | 1 |
| MENTAL HEALTH & PSYCHOSOCIAL CONDITION | 1 |

## FREQUENCY OF ED VISITS

| UNIQUE IDENTIFIER ASSIGNED BY DIMR | PREVISITS | POSTVISITS | DIFF |
|---|---|---|---|
| **H** | 1 | 9 | 8 |
| **L** | 0 | 6 | 6 |
| **AR** | 0 | 3 | 3 |
| **A** | 1 | 3 | 2 |
| **O** | 0 | 2 | 2 |
| **AF** | 0 | 2 | 2 |
| **T** | 0 | 2 | 2 |
| **AS** | 0 | 2 | 2 |
| **I** | 0 | 2 | 2 |
| **BA** | 0 | 2 | 2 |
| **W** | 0 | 1 | 1 |
| **AE** | 0 | 1 | 1 |
| **Y** | 0 | 1 | 1 |
| **K** | 0 | 1 | 1 |
| **F** | 0 | 1 | 1 |
| **P** | 0 | 1 | 1 |
| **BD** | 0 | 1 | 1 |
| **AC** | 0 | 1 | 1 |
| **R** | 2 | 2 | 0 |
| **V** | 1 | 1 | 0 |
| **AA** | 2 | 1 | -1 |
| **AG** | 1 | 0 | -1 |
| **AB** | 1 | 0 | -1 |
| **G** | 4 | 0 | -4 |