8-12-2014

A Proposed Evaluation Plan for Kaiser Permanente's Diabetes Disease Management Program

Kathryn Wiedeman

A PROPOSED EVALUTION PLAN FOR KAISER PERMANENTE'S DIABETES DISEASE MANAGEMENT PROGRAM

by

KATHRYN H. WIEDEMAN, B.S.

GEORGIA STATE UNIVERSITY

A Capstone Submitted to the Graduate Faculty of Georgia State University in Partial Fulfillment of the Requirements for the Degree

MASTER OF PUBLIC HEALTH

ATLANTA, GEORGIA 30303

A Proposed Evaluation Plan for Kaiser Permanente's
Diabetes Disease Management Program

by

KATHRYN H. WIEDEMAN

Approved:

_____
Committee Chair

_____
Committee Member

_____
Date

Author's Statement Page

In presenting this thesis as a partial fulfillment of the requirements for an advanced degree from Georgia State University, I agree that the Library of the University shall make it available for inspection and circulation in accordance with its regulations governing materials of this type. I agree that permission to quote from, to copy from, or to publish this thesis may be granted by the author or, in his/her absence, by the professor under whose direction it was written, or in his/her absence, by the Associate Dean, School of Public Health. Such quoting, copying, or publishing must be solely for scholarly purposes and will not involve potential financial gain. It is understood that any copying from or publication of this dissertation which involves potential financial gain will not be allowed without written permission of the author.

_____

Signature of Author

Acknowledgements

I want to thank my capstone committee, Dr. Bruce Perry and Jessica Howell Pratt for their guidance not only throughout writing this capstone, but also for their support throughout the entire MPH program. I'm truly thankful for their advice and in helping shape my master's experience into such a positive, enlightening two years.

I appreciate the practicum opportunity Kaiser Permanente provided for me. Thank you, Miriam Bell, for this experience and your expertise.

I am extremely grateful to my family and friends, particularly my parents and in-laws. This experience would not have been possible without their support and love.

To my loving husband: Thank you for your unyielding support throughout this entire program. I am so grateful for your late-night problem solving abilities, positive attitude, and most of all, unconditional love. Your support has enabled me to achieve the first step in achieving my dreams.

Notice to Borrowers Page

All theses deposited in the Georgia State University Library must be used in accordance with the stipulations prescribed by the author in the preceding statement.

The author of this thesis is:

Student's Name:

Street Address: _____

City, State, and Zip Code: _____

The Chair of the committee for this thesis is: Professor's

Name:

Department: _____

College: _____

Georgia State University
School of Public Health
P.O. Box 3995
Atlanta, Georgia 30302-3995

Users of this thesis who not regularly enrolled as students at Georgia State University are required to attest acceptance of the preceding stipulation by signing below. Libraries borrowing this thesis for the use of their patrons are required to see that each user records here the information requested.

NAME OF USER	ADDRESS	DATE	TYPE OF USE (EXAMINATION ONLY OR COPYING)

Table of Contents

Abstract

Kathryn H. Wiedeman
A Proposed Evaluation Plan for Kaiser Permanente's Diabetes Disease Management Program
(Under the direction of Bruce Perry M.D., MPH, Faculty Member)

DM is a serious and complex public health problem in the U.S. The CDC (2013) estimated that 25.8 million people, or 8.3% of the U.S. population, were suffering from DM in 2011. DM can significantly affect patient's quality of life. Additionally, DM places a significant economic burden on the U.S. healthcare system.

Over the past two decades, DMPs have emerged as a promising intervention to improve health outcomes for patients suffering from chronic conditions, such as DM, and to bend the cost curve. DMP's aim is to improve communication and follow-up so that patients can better manage their chronic condition(s) to avoid costly hospital stays and emergency room visits (Fireman, Bartlett, & Selby, 2004).

The Georgia region of Kaiser Permanente (KPGA) is a fully integrated health system that serves 260,000 members at 28 medical offices along with two specialty offices in the metropolitan Atlanta area. The Center for Care Partnership, the population care division of KPGA, administers a chronic disease management program (DMP), *Healthy Solutions* (*HS*). *HS* exists to improve and maintain the health of chronically ill KPGA members, including patients diagnosed with diabetes mellitus (DM), by providing health coaches via telephone who counsel members on their specific chronic disease and aid members in starting or maintaining a physician approved self-care management plan.

In order to determine the impact *HS* has on KPGA members with diabetes, an evaluation plan was created to evaluate the impact *HS* has on members' glycated hemoglobin (A1C), blood pressure, and emergency department (ED) utilization. This capstone thoroughly details the proposed evaluation plan created for *HS* by using Robert Milstein and Scott Wetterhall's six-step framework for program evaluation. Additionally, further evaluation questions are suggested and discussed in order to provide a more complete picture of program performance to stakeholders.

1. Introduction

The MPH candidate completed her practicum project at Kaiser Permanente. The Georgia region of Kaiser Permanente (KPGA) is a fully integrated health system that serves 260,000 members at 28 medical offices along with two specialty offices in the metropolitan Atlanta area. The Center for Care Partnership, the population care division of KPGA, administers a chronic disease management program (DMP), *Healthy Solutions* (*HS*). *HS* exists to improve and maintain the health of chronically ill KPGA members by providing health coaches via telephone who counsel members on their specific chronic disease and aid members in starting or maintaining a physician approved self-care management plan. *HS* targets members with Chronic Obstructive Pulmonary Disease (COPD), diabetes mellitus (DM), Congestive Heart Failure (CHF), Asthma, Chronic Kidney Disease (CKD), and End Stage Renal Disease (ESRD).

The practicum opportunity involved measuring the cost effectiveness of KPGA's *HS*. Although *HS* targets a variety of chronic diseases, the scope of the practicum only focused on the DM portion of *HS*. Throughout the course of the practicum, the MPH candidate formulated an evaluation plan based on current research of DMPs that measured the effectiveness of the program. This capstone thoroughly details the evaluation plan created for the DM portion of *HS* by using Robert Milstein and Scott Wetterhall's six-step framework for program evaluation. In order to understand the application of theory to a real-world evaluation, this paper alternates between defining Milstein & Wetterhall's framework for program evaluation and the application of the framework concepts to the evaluation of *HS* at KPGA. A literature review, which is included in the third step of the framework, was conducted to understand current DMP evaluation study design methodology. The literature review focuses on how current suggestions from published literature were used to create a DMP evaluation study design for KPGA. Additionally, further evaluation

recommendations are suggested.

2. Background

2.1 Diabetes

DM is a serious and complex public health problem in the U.S. According to the Centers for Disease Control and Prevention [CDC] (2012), DM causes the pancreas to stop making insulin or to not use insulin properly, leading to high blood glucose, which can cause serious complications or even death. The prevalence of DM has been steadily increasing since 1990, which is alarming to public health officials. The CDC (2013) estimated that 25.8 million people, or 8.3% of the U.S. population, had been diagnosed with DM in 2011. It is also estimated that one in three Americans will have DM by 2050 if prevalence continues to increase at the current rate (CDC, 2012). A person who has DM is twice as likely to die when compared to the risk of a person around the same age who does not have DM (CDC, 2012). The three major types of DM are type 1 diabetes mellitus (T1DM), type 2 diabetes mellitus (T2DM), and gestational diabetes. The focus of this paper is on T1DM and T2DM.

According to the National Diabetes Information Clearinghouse (2014), T1DM prohibits the body from producing enough or any insulin because the body's immune system attacks and destroys the cells that make insulin. T1DM has often been referred to as juvenile DM because it is generally diagnosed in childhood, although it can occur at any time (CDC, 2012). According to the CDC (2012), only 5% of the U.S. population with DM has T1DM. The cause of T1DM is not fully understood, but some possible risk factors for the disease are thought to be autoimmune, genetic, or environmental (CDC, 2012). Maintaining proper insulin and healthy lifestyle management are critical to living a high quality of life with T1DM.

According to the CDC (2012), T2DM prohibits the body from using insulin properly,

which leads to high glucose levels in the bloodstream. T2DM accounts for approximately 95% of the U.S. population with DM (CDC, 2012). It is widely believed that unhealthy eating habits and decreased physical activity levels play a significant role in the onset of T2DM. Approximately 80% of people with T2DM are overweight or obese (NDIC, 2014). The CDC (2012) reports that a healthy diet and regular physical activity along with properly taking prescribed medications help control diabetes complications and prevent or delay the onset of T2DM.

According to the CDC (2012), T2DM disproportionately affects minorities in the U.S. Most minorities, such as Hispanic Americans and non-Hispanic blacks, have a higher prevalence of T2DM when compared to white, non-Hispanics. American Indians and Alaska Natives are also at particularly high risk for developing T2DM. The challenge within minority groups is that T2DM develops at a younger age in these populations, which puts minorities at a higher risk for developing complications associated with T2DM. However, age seems to be strongly correlated with T2DM because rates for the disease sharply increase with age for both sexes and for all racial and ethnic groups (CDC, 2012).

According to the American Diabetes Association [ADA] (2014), if blood sugar levels are not kept under control, DM can cause complications that can be debilitating to patients and possibly lead to premature death. Adults with DM are about 2 to 4 times more likely to have heart disease when compared to adults without diabetes. DM is the leading cause for new cases of kidney failure and blindness. From 2005 to 2008, 4.2 million, or 28.5%, of adults with DM aged 40 years or older had diabetic retinopathy, which is an eye disease that results from damage to the small blood vessels in the retina. DM is responsible for more than 60% of non-traumatic lower-limb amputations. Among adults with DM, around 60% to 70% have mild to severe forms of neuropathy, which is nerve damage caused by DM (ADA, 2014).

DM places a significant economic burden on the U.S. health care system. The article "Economic Costs of Diabetes in 2012" (2013) estimated that the national cost of DM in 2012 was $245 billion, which is a 41% increase from the estimated national costs in 2007 of $174 billion. Seventy-two percent, or $176 billion, of that national cost estimate is attributed to DM health care costs and 28%, or $69 billion, represents reduced or lost productivity. Hospital inpatient care represents the largest portion of medical expenditures at 43% of total medical costs. Additionally, health care costs for patients with diabetes are very high. On average, patients with diabetes pay $7,888 more, or 2.3 times, in excess health care costs per year when compared to their non-diabetic counterparts ("Economic costs," 2013). The U.S. healthcare system cannot sustain the growth in DM related health care costs. There is a great need for efforts to decrease health care costs related to DM in U.S. healthcare system.

2.2 Disease Management Programs

Chronic conditions are extremely costly diseases to treat and greatly impact patients' quality of life. According to the CDC (2014), eighty-four percent of health care spending was for 50% of the U.S. population who had one or more chronic condition in 2006. Additionally, chronic diseases are the leading causes of death and disability in the U.S. (CDC, 2014). The U.S. health care system desperately needs evidence-based interventions that can reduce death and disability from chronic conditions and decrease health care costs associated with chronic conditions. Over the past two decades, DMPs have emerged as a promising intervention to improve health outcomes for patients suffering from chronic conditions, such as DM, and to bend the cost curve.

The Disease Management Association of American (DMAA) defines disease management as a "system of coordinated healthcare interventions and communications for populations with conditions in which patient self-care efforts are significant" (Wilson & MacDowell, 2003, p. 146).

DMP's aim is to improve communication and follow-up so that patients can better manage their chronic condition(s) to avoid costly hospital stays and emergency room visits (Fireman, Bartlett, & Selby, 2004). Generally, DMPs are aimed at high-cost conditions that require a significant amount of patient self-care.

The unique aspect of DMPs is the emphasis placed on patient self-care in managing a chronic condition. DM is a difficult condition to manage because it requires a significant amount of lifestyle changes for the patient. Following a healthy diet, staying physically active, and managing blood glucose levels are all critical lifestyle habits for managing DM. In order for these patients to be successful in managing their condition, it is important that adequate support and education are available to patients. Often, the knowledge and tools patients need in order to successfully change lifestyle habits cannot be provided by patients' physicians in one, single appointment. Therefore, DMPs help supplement the diabetes care patients' receive from their physician. DMPs strive to provide more support to patients to manage DM.

3. Framework for Program Evaluation

In order to understand if programs are effective and implemented as intended, decision-makers must evaluate programs. Evaluation is crucial to ensure that resources are being used efficiently and that the population targeted by the program is being positively impacted. The CDC recognized the critical need for effective program evaluation of public health programs. In 1999, Robert Milstein and Scott Wetterhall, in collaboration with CDC working groups and other CDC contributors, published an article in the CDC's publication *Morbidity and Mortality Weekly Report* that detailed an evaluation framework for practitioners to effectively evaluate public health programs. According to Milstein & Wetterhall (1999), "effective program evaluation is a systematic way to improve and account for public health actions by involving procedures that are

useful, feasible, ethical, and accurate" (p. 1). A diagram of the framework is included in appendix A. This capstone project will utilize the Milstein & Wetterhall (1999) framework to define and organize crucial elements of the evaluation. The proposed evaluation is described in this paper, with an emphasis on the development of the DMP evaluation methodology in the third step of the framework.

3.1 Engage Stakeholders

According to Milstein & Wetterhall's (1999) first step in the program evaluation framework, program stakeholders must be engaged to ensure that all perspectives are taken into consideration in the creation of the evaluation. When stakeholders are not engaged, important aspects of program outcomes, goals, and operations might not be addressed. Therefore, stakeholders need to be engaged in order for the evaluation to be meaningful to them and to ensure that the results are used to either continue program success or make necessary changes. Three important groups of stakeholders for program evaluation include those involved in program operations, those served by the program, and primary users of evaluation (Milstein & Wetterhall, 1999).

Identifying stakeholders is an important step in ensuring the evaluation addresses all relevant aspects of the program. The main stakeholders who are involved in program operations for KPGA's *HS* are the program manager and the DM care manager (DCM). Both stakeholders were engaged in the evaluation. During the initial meeting with the *HS* program manager and the DCM, both expressed that they were ready to understand the impact that *HS* has on KPGA members with diabetes who participate in the program. KPGA members with diabetes are the stakeholders who are served by the program. Although understanding KPGA members with diabetes' perspectives on the program is a crucial aspect of program evaluation, this proposed

evaluation of *HS* will not be able to address this group of stakeholders due to resource constraints and patient privacy laws. The primary user of the *HS* evaluation is the Director of the Center for Care Partnership at KPGA because she can use the evaluation to make high-level program decisions. The primary user must be informed frequently in the evaluation process and her input needs to help shape the evaluation.

3.2 Describe the Program

Milstein and Wetterhall (1999) explain the importance of a detailed program description in the second step of the program evaluation framework. Creating a detailed program description is crucial in order to effectively evaluate a program. Linden & Roberts (2005) and Fitzner et al. (2004) explain that DMP evaluations must clearly describe the program so that readers can decide for themselves if evaluation outcomes are linked to the activities of the DMP. It is important to describe the program in enough detail in order to ensure all relevant aspects of the program are evaluated.

According to Milstein & Wetterhall (1999), seven critical aspects of program description are need, activities, expected effects, resources, stage of development, context, and logic model. The statement of need clearly describes a problem that the program aims to address. Defining program activities ensures that strategies and actions are displayed in a logical sequence. The expected effects, or program outcomes, are main goals that the program is trying to achieve in order to be considered successful. Program stage of development, which can be classified into the planning, implementation, or effects stage, must be assessed to understand program maturity. Understanding the program's context within the organization allows the evaluator to create an evaluation that is sensitive to the various influences, such as political or economic influences, on the program and stakeholders. The logic model is able to clearly summarize and link program

processes that affect change (Milstein & Wetterhall, 1999).

As detailed in the first section, DM is an extremely expensive disease that greatly impacts quality of life. Therefore, there is a clear need for a DM DMP at KPGA. Evidence suggests that DM prevalence and related health care costs will continue to increase in the U.S. unless major interventions are implemented. Due to the expenses related to DM, health plans, including KPGA, feel it is necessary to implement interventions that prevent and manage DM for their members.

A logic model is a helpful tool that provides a clear picture of a program by linking important program elements. The major elements of the logic model are inputs, outputs, and outcomes. A logic model is an excellent tool to ensure that evaluation is included in each component of the program. The program description elements of resources, activities, and expected effects can be explained through the context of the logic model created for *HS*. An illustration of the *HS* logic model is available in appendix B.

The *HS* program manager, DCM, and Director of the Center for Care Partnership are program staff who are major program resources. Program funding is an important resource for the success of the program. The Georgia State University MPH candidate who is creating the proposed evaluation under the direction of the Director of the Center for Care Partnership is a program resource. Technology, such as the electronic medical record (EMR) system and other KPGA clinical databases are important technology resources for *HS*.

The *HS* DCM provides health coaching telephone calls to KPGA members with diabetes. DM care management calls focus on the management of DM by covering various topics, such as providing information about DM and support so patients can follow a physician recommended DM self-management plan. The goal of DM care management calls is to have two calls the first month with a new participant, then a monthly follow-up call after the first month. Participants being

served by the program are KPGA members with a diagnosis of T1DM or T2DM with a glycated hemoglobin, or A1C, blood test result between 8% to 9%. The A1C blood test provides an idea of average blood glucose control over the past two to three months (ADA, 2013). Thus, this blood test is a commonly used to monitor the management of DM.

KPGA members with diabetes access program services through a variety of avenues. KPGA members with diabetes can contact the *HS* hotline to speak directly with the DCM. Physicians can send an internal message to the DCM recommending that she reach out to a KPGA member with diabetes. The DCM provides outreach by identifying KPGA members with diabetes who would benefit from the program through the KPGA clinical database. The DCM provides outreach to KPGA members with diabetes who have recently been discharged from a hospitalization due to DM.

Program activities aim to positively influence program expected effects, or outcomes. Generally, expected effects take place over a broad period of time. The main short-term outcome, which is expected to take place within three months, is to increase awareness of DM and healthy DM self-care management habits. *HS* aims to provide KPGA members with diabetes with the tools to start examining their own lifestyle habits, such as diet and physical activity, and to understand how these lifestyle habits can impact blood sugar. Within six months, the intermediate-term outcomes of incorporating new behaviors to improve DM self-care and increasing knowledge of DM are expected to take place. The long-term outcome, expected to take place within a year, is for KPGA members with diabetes 65 years and older to have a controlled A1C level of 8% or less and members younger than 65 years old to have an A1C level of 7% or less.

When evaluating a program, the stage of development must be taken into consideration to understand program maturity. For example, a program that has just received funding will be very

different from a well-established program. *HS* has been providing DM care management services to KPGA members with diabetes since October 2012. Enough time has passed for the expected effects of *HS* to take place; therefore, *HS* is in the effect stage of development according to Milstein and Wetterhall (1999). The *HS* evaluation goal is to understand the effect of program activities on relevant clinical and economic outcomes.

Understanding and acknowledging the context of *HS* will ensure that the evaluation is sensitive to economic or political influences. A third party vendor managed KPGA's DMP until September 30, 2012. KPGA internalized the program on October 1, 2012 because KPGA felt they could provide higher quality DMP services at a lower cost to the organization. KPGA is proud of this decision and the program. Program stakeholders are confident that the program is making a significant, positive impact on KPGA members with diabetes, but there is always a possibility that evaluation results could show little or no program impact. Therefore, special attention will need to be paid to reporting unfavorable evaluation results. It is important to emphasize that any program evaluation results provide stakeholders with valuable information on the program. Results, whether favorable or unfavorable, can be used to improve program operations and build a foundation of continuous evaluation of the program going forward.

3.3 Focus the Evaluation Design

According to Milstein & Wetterhall's (1999) third step of the framework, creating a thorough evaluation plan ensures that intended uses are made known and increases the chances that the evaluation will be useful, feasible, ethical, and accurate. The evaluation design must be clearly defined before data collection begins because changing the evaluation design after data is collected can be very difficult, if not impossible, to do. The following six components of focusing the evaluation design will be explored through the context of the proposed *HS* evaluation: purpose,

users, uses, questions, agreements, and methods. Having a clear evaluation purpose will allow all decision makers to make well-informed decisions for the evaluation plan. Users are the stakeholders who will receive the evaluation results. Uses are the specified ways the users will apply the findings. Evaluation questions need to reflect what stakeholders are ultimately trying to understand about the effectiveness of a program. Agreements are an evaluation outline that clarifies stakeholder roles and responsibilities and details the resources that will be used for the evaluation. Methods for gathering data for the evaluation are created by using opinions from scientific research, particularly from social, behavioral, or health sciences research (Milstein & Wetterhall, 1999). Although all six components are important in evaluation, the development of the evaluation methodology for *HS* will be discussed in detail in a separate, subsequent section of this paper. The component of evaluation questions will also be addressed in the separate methods section.

The purpose of the *HS* evaluation is to assess if the program is making a positive impact on health and economic outcomes of interest. Stakeholders want to understand if the investment of company financial and human resources into improving KPGA members' with diabetes health outcomes is saving KPGA money. In order to understand this relationship, the evaluation needs to examine the impact of *HS* program activities on KPGA members' with diabetes health and economic outcomes. Since *HS* is in the effects stage of development, it is appropriate to evaluate the effect of program activities on desired health and economic outcomes.

Two main users of the evaluation will be the primary user, the Director of the Center for Care Partnership, and the *HS* program manager. The primary user wants to understand the effectiveness of *HS* in order to ensure financial and human resources are being allocated appropriately to the program. The primary user is also interested in hiring additional DCM if the

evaluation suggests that the program is effective. Additionally, the primary user will use the results of the evaluation to show potential and existing KPGA clients the effectiveness of KPGA's *HS*. The *HS* program manager will use the evaluation results to improve operations and program activities, if needed.

A verbal agreement of the evaluation plan was made between the Director of the Center for Care Partnership and the GSU MPH candidate. A formal, written agreement was not executed for the *HS* evaluation plan, but the GSU MPH candidate consulted with the primary user through meetings and used the primary user's input to develop the evaluation plan. Human resources that will be needed for the evaluation are the *HS* program manager, *HS* DCM, an evaluation specialist, a KPGA data analyst, and Director of the Center for Care Partnership. The timeframe of the evaluation will need to be determined when the stakeholders are ready to execute the evaluation plan. KPGA member data will need to be available in order to conduct the *HS* evaluation data analysis.

3.3.1 Methods: Literature Review & Methodology Development

In order to understand if a DMP is effective, it must be evaluated. Although evaluation takes place in research and industry, there are differences between the two. Evaluation research aims to add to generalized knowledge often by using a randomized controlled trial (RCT) study design. In industry, health care decision-makers want to understand if a DMP is working to ensure that resources, such as staff and money, are being allocated properly. Unfortunately, it is difficult to bridge the gap between research and practice. One possible explanation for this gap is that rigorous study designs suggested in published literature are nearly impossible to use in industry due to data, resource, and time limitations. Decision-makers run the risk of misallocating scarce resources when making decisions on results produced from weak study designs. In order to

perform strong evaluations that represent the actual impact of DMPs, industry must adapt as much scientific rigor as possible when performing DMP evaluations. The study design created to evaluate the DM portion of *HS* used as much of the scientific rigor suggested in the literature as possible while working within the constraints of a real-world evaluation.

3.3.1a Outcomes of Interest

Stakeholders in the *HS* evaluation were most interested in the return on investment (ROI) of *HS*, however, determining an accurate ROI for a DMP is very challenging. A financial ROI calculation alone does not prove that the DMP is achieving its main goal: reducing or eliminating major adverse health events related to a specific condition (Lewis, 2009). It is strongly suggested that other outcomes, such as health or utilization outcomes, be taken into consideration when evaluating a DMP (Linden & Roberts, 2005). Health care costs can be greatly influenced by factors other than the DMP, such as members' financial share of the medical expense (Linden & Roberts, 2005). Therefore, it can be argued that evaluating health outcomes and health care utilization are better measurements of program success than simply a financial ROI calculation. After the outcomes of interest have been evaluated, the results can then be translated into an ROI as long as some well-defined assumptions about the population have been made (Lewis, 2009).

In order to clearly define the outcomes of most interest to stakeholders, evaluation questions must be formulated. Evaluation questions aim to answer whether or not the program is achieving desired expected effects. Evaluation questions were created with input from stakeholders regarding their opinion on the most important health and economic outcomes that can best measure program effectiveness. Additionally, research on DMP methodology helped shape the following evaluation questions:

1. Are there differences in A1C results between KPGA members with diabetes who received *HS* health coaching and KPGA members with diabetes who did not receive *HS* health coaching?
2. Are there differences in blood pressure readings between KPGA members with diabetes who received *HS* health coaching and KPGA members with diabetes who did not receive *HS* health coaching?
3. Are there differences in the number of emergency department admissions between KPGA members with diabetes who received *HS* health coaching and KPGA members with diabetes who did not receive *HS* health coaching?

The main goal in a DM DMP is to improve blood glucose management, which is measured by A1C. A1C is the most important outcome measurement for *HS* because A1C provides an accurate picture of blood glucose management in a population. Evaluating A1C control in the population provides a high-level indication of the effectiveness of the program to stakeholders. It also ensures that the end goal of successful DM self-care management is being accomplished. Therefore, it is critical to understand the relationship between *HS* and KPGA members' with diabetes A1C levels.

The ADA (2014) reports that 67% of adults age 20 or older who reported a DM diagnosis had blood pressure greater than or equal to 140/90 mmHg or used prescription blood pressure lowering medication. Since DM is a major cause of heart disease and stroke (CDC, 2013), diabetic patients with high blood pressure are at an even higher risk of an adverse cardiovascular event. DM alone can compromise quality of life and is expensive to manage; therefore, high blood pressure along with DM exacerbates health problems. It is important to understand if *HS* has an impact on KPGA members' with diabetes blood pressure.

Since two measures of program effectiveness are clinical outcomes, emergency department (ED) admissions are suggested as an economic measure of program effectiveness. DMPs can impact health care utilization more than health care costs (Serxner et al., 2006). According to the Agency for Healthcare Research and Quality [AHRQ] (2013), there were approximately 12.1 million DM-related ED visits for U.S. adults age 18 and older in 2010. Fifty-eight percent of these ED visits were treatment and release. Forty-two percent of these DM-related ED visits resulted in hospitalization while, in comparison to all ED visits, only 15.3% resulted in hospitalization. There were 771,000 DM-related ED visits; with "DM with complications" being the most frequently, first-listed condition for DM related ED visits (AHRQ, 2013). *HS* aims to provide the support and knowledge to help KPGA members manage their diabetes and prevent ED admissions. Thus, ED utilization is a good measure of understanding the self-management of DM in KPGA diabetes population.

3.3.1b Study Population

According to Wilson et al. (2004), the main goal when defining the study population is to ensure that the intervention and control population are as equivalent as possible. The more equivalent the intervention and control population are, the more likely any impact on outcomes can be attributed to the program. For example, if blood pressure is being analyzed between a control group with significantly lower cardiovascular disease than the intervention group, then it will be likely that blood pressure will be lower in the control group simply because the two groups were not equivalent at baseline. Therefore, comparing these two groups indicates nothing about program impact. It is recommended that the control group be chosen in the exact same way as the intervention group (Wilson et al., 2004). In order to ensure comparability between the two groups, identification of diabetes, age, inclusion, and exclusion criteria were clearly defined by

recommendations in the literature for this proposed study design. A graph is located in appendix C that illustrates the recommended study population for the *HS* evaluation.

Clearly defining the specifications for how DM is identified in a population is a critical first step in creating a study design to evaluate program effectiveness. Often in DMP evaluation, simply using International Classification of Diseases, 9th Revision (ICD-9) codes, identifies patients. Fetterolf, Wennberg & Devries (2004) argue that using these codes alone are not sufficient in accurately identifying the DM population. Using ICD-9 codes could lead to including patients who do not have DM or excluding patients who do have DM into the study population. Suggested practice is to identify patients by a variety of measures, including identifying patients by hospitalization for DM and certain medications for DM (Fetterolf et al., 2004).

Fortunately, it is relatively straightforward to identify patients with DM in KPGA's integrated health care system. In order to clearly identify KPGA members who have DM, it is advised to use an established, comprehensive strategy, such as the Health Plan Employer Data Information Set, or HEDIS© (Fetterolf et al., 2004). HEDIS© measures performance on 81 measures of care and service across 5 domains of care, including DM (HEDIS©, 2013). This tool is used by over 90% of health plans in the U.S. (HEDIS©, 2013). The diagnosis of DM for the *HS* evaluation should use the HEDIS Comprehensive Diabetes Care© definition. HEDIS© chooses the population with DM through measures of outpatient, ED, and inpatient encounters by appropriate ICD-9 codes along with the inclusion of certain DM medications. Since HEDIS© uses a variety of measures to choose the DM population, this comprehensive definition ensures that the appropriate members will be included in the population.

HS starts to provide DM care management telephone coaching to KPGA members' with diabetes at age 18. The oldest age used that needs to be used in this evaluation is 80 years old

because *HS* services are available to KPGA members up to age 80. Age is defined as the age of the patient at the BP or A1C reading and the age at the time of the ED admission.

The inclusion and exclusion of certain patients are necessary steps in study design, but it can be challenging when trying to maintain equivalency between the intervention and control group. To ensure comparability between the intervention and control group, the same criteria to include and exclude patients must be applied to both the intervention and control group (Wilson et al., 2004).

One inclusion criterion should be used in this study design. The criterion must be applied to both the intervention and control group. *HS* provides services to patients age 18-80, but HEDIS© uses 75 years old as the maximum age in the definition of DM. The inclusion criterion in this study is to include KPGA members with diabetes that are considered diabetic based on the HEDIS Comprehensive Diabetes Care© definition from ages 76-80 in the study population.

One exclusion criterion needs to be applied to both the intervention and control population. KPGA has two medical office buildings (MOBs) that offer members one-on-one health coaching with a certified DM health coach. Therefore, KPGA members with diabetes from these two MOBs do not need to be included in the study population. Since these two MOBs are offering intensive health coaching for DM, the diabetic populations at these MOBs could be significantly different from the remaining MOBs in Georgia. Thus, this factor could possibly skew the results of the study.

3.3.1c Study Design in Evaluation

In order to prove if a DMP is working, a causal pathway between the DMP and changes in health and economic outcomes must be clearly illustrated. However, proving this causal pathway is the most difficult part of accurately evaluating a DMP. The selection of a strong study design for

DMP evaluation is critical in order to produce "…results [that] represent an unbiased estimate of treatment effect" (Linden & Roberts, 2005, p. 113). When evaluating a DMP, it is always possible that some other factor(s) besides the DMP influenced financial and clinical outcomes (Serxner, Baker, & Gold, 2006).

One factor that can confuse the causal pathway between health outcomes and the DMP is that *HS* is not the only encounter patients with diabetes have with KPGA. KPGA members with diabetes have other providers participating in their care, such as dieticians, physicians, and nurses. Additionally, there are other possible factors, such as a DM support group, outside of KPGA health care that can influence DM self-care management. Therefore, it will be difficult to determine that *HS* alone was responsible for changes in clinical or economic outcomes.

Currently, varying DMP evaluation study designs in the literature along with a lack of methodology standardization make determining the financial and clinical impact of DMPs challenging (Wilson, MacDowell, Salber, Montrose, & Hamm, 2008). These issues have led to some skepticism when interpreting program financial impact from DMP evaluations in the literature (Lewis, 2009). In order to achieve the scientific rigor needed to accurately develop and use study designs to evaluate DMP programs, an array of knowledge is needed from disciplines such as epidemiology, health services research, evaluation research, and statistics. Often, evaluating DMPs in industry is fulfilling a business need and the individual(s) conducting the evaluation are not experts at evaluation. This could be one possible explanation for the lack of standardization and scientific rigor in DMP evaluation.

DMP methodology research overwhelmingly suggests that the strongest study design for program evaluation is a randomized controlled trial (RCT). A RCT ensures both the intervention and control group are very similar, thus increasing comparability of the two groups and the

likelihood that outcomes can be attributed to the DMP (Linden & Roberts, 2005). However, the RCT is often not possible to use when evaluating DMPs in industry, mainly because the program being evaluated has already been implemented. Additionally, there are various constraints, such as data or resource limitations, which do not allow for rigorous study design methods in practice.

Because the RCT is nearly impossible to use in industry, the most practical study design for DMP evaluation is the pre-post study design without a control group ("Standard outcome," 2003). The pre-post study design takes baseline data from a group of individuals before an intervention and assesses the same groups' outcomes of interest after the intervention (Fetterolf et al., 2004). The literature is somewhat conflicting when making suggestions of the best study design methodology. Linden & Roberts (2005) and Wilson & McDowell (2003) explain that it is absolutely necessary to have a control group to draw conclusions about the impact of the DMP.

The relative ease and feasibility of the pre-post study design appeals to DMP stakeholders; however, it does have significant limitations. According to Gordis (2009a), the difference in time periods makes it difficult to determine if any clinical or financial changes are due to the DMP itself or to factors that may have changed over time, such as improvements in health care delivery to diabetics or improvements in DM medication. The two groups lack comparability due to the time period difference, which makes it difficult to determine program impact because there are likely unknown confounders that are not controlled for in this type of study design. This type of study design can make suggestions about the impact of a program on outcomes of interest, but it cannot prove program effectiveness (Gordis, 2009a).

3.3.1d Case-Control Study Design

The goal of this proposed evaluation study design is to use as much scientific rigor as possible, but to work within the constraints of a real-world evaluation. Due to the impossibility of

using a RCT and the limitations of the pre-post study design, the case-control study design is the best option to evaluate KPGA's DM DMP. In order to understand the exposure-outcome relationship between *HS* and A1C, blood pressure, and ED utilization, a group of KPGA members with diabetes who had contact with *HS* needs to be compared to a group of KPGA members with diabetes who have not had contact with *HS*.

The timeframe and case-control definition development was adapted from a case-control study that examined the impact of a diabetes health education program on A1C and lipid levels conducted by Roblin, Ntekop, & Becker (2007). Additionally, Dr. Doug Roblin, a KPGA health services researcher and faculty member at Georgia State University, offered his expertise in developing this case-control study design. It is critical to have baseline, intervention, and post-intervention periods in order to draw comparisons of outcome measures between the cases and controls. The baseline period, which is suggested to span from 10/1/2013 to 3/31/2014, should include case and control observations with an A1C or blood pressure reading and no *HS* contact. The recommended timeframe for the intervention period is from 4/1/2014 to 9/30/2014. The intervention period needs to include cases with at least one *HS* contact along with an A1C and blood pressure reading. Controls must not have contact with *HS*, have at least one primary care encounter, and an A1C or blood pressure reading. The post-intervention period, which is suggested to span from 10/1/2014 to 3/31/2015, should include cases with an A1C or blood pressure reading. Cases can have contact with *HS* in the post-intervention timeframe. Controls need an A1C or blood pressure reading, but cannot have contact with *HS*. Members can be identified as having contact with the program by using the DM care management code in KPGA's EMR system along with the DCM's KPGA provider number. ED admission data should be collected for all cases and controls along with the first three listed ICD-9 diagnoses for the admission to ensure that the ED visit was

related to DM. A table of this study design is located in appendix D.

Unfortunately, data limitations do not allow the case-control study design to be used to evaluate *HS* before 10/1/2013. Prior to 10/1/2013, *HS* used various methods to record program data in KPGA's EMR. The variation in data collection methods makes it impossible to use data before October 2013 to make any accurate conclusions about program impact on KPGA members with diabetes. When a *HS* call was made to a KPGA member with diabetes, the call was notated in the members' medical record using a specific code for a DM care management call, but that particular code was used either if a DM care management call took place or if a voicemail regarding program information was left for the patient. Prior to October 2013, there is no way to determine if a patient actually received a high-quality DM care management call with the DCM. Starting in October 2013, the DCM began to denote if she left a message for a member or if the DM care management call was a quality call in KPGA's EMR. When evaluating program outcomes starting after October 2013, the new method of denoting quality calls in KPGA's EMR ensures that only quality calls will be used in future evaluations. This improvement in data quality will allow KPGA stakeholders to make accurate conclusion about program impact.

In addition to collecting health and economic data for the evaluation, the demographic and other relevant program data should be collected. Age, sex, and ethnicity need to be collected for each case and control observation in order to understand the demographics of the cases and controls. The number of contacts that cases have with *HS* during the *HS* and post-intervention timeframe should be collected to assess if a dose-response relationship is present between the exposure and the outcomes. It is important to understand if more contacts with *HS* result in greater improvements in A1C and blood pressure along with lower ER utilization for DM. The number of primary care visits for cases and controls should be collected in order to better understand the

health care utilization habits of the cases and controls. Visits are considered primary care if they take place in adult medicine, OBGYN, or pediatrics.

3.3.1e Challenges to Address in the Case-control Study Design

In evaluation, bias can be defined as error in the study design or the analysis that results in a mistaken interpretation of the program impact on outcomes (Gordis, 2009b). The literature overwhelmingly recommends that potential biases need to be recognized and steps to minimize bias must be taken. Bias is an issue in all types of study designs, so it must be addressed before subjects are selected for the analysis (Wilson & MacDowell, 2003). Selection and exclusion bias are the main biases that threaten the validity of the *HS* evaluation. The goal is to have an intervention and control group that are very similar. Because a RCT could not be used in the *HS* evaluation, it is likely that the intervention and control group were different. Nevertheless, steps were taken in order to minimize the effects of bias.

Selection bias enters into this evaluation because patients self-selected to participate in *HS*. For many different reasons, some people are more motivated to seek health care or perform health-related behaviors than others (Linden & Roberts, 2005). It is possible that members who decide to participate in *HS* are more motivated to improve health-related behavior than members in the control group, thus resulting in a perceived program impact. This proposed study aims to choose cases and controls that are roughly at the same level of willingness to change. In the intervention period, cases should be selected based on having contact with *HS* and controls should be selected based on having a primary care encounter. Cases and controls should be selected in this manner to ensure that both groups of members are likely to make behavior changes to improve their health. If a control attends a PCP appointment and receives the recommended A1C screenings or blood pressure readings, then it can be assumed that these patients are, at least, somewhat willing to

manage their DM.

Exclusion bias can be a concern in a case-control study design. Exclusion bias results when eligibility criteria are not applied equally in the cases and the controls (Gordis, 2009b). Wilson et al. (2004) explain that the same criteria used to identify the group who is exposed must be applied to the reference group to ensure equivalence between the two groups. This proposed study design to evaluate *HS* emphasizes that inclusion and exclusion criteria must be applied equally to both the intervention and control group so that the two groups could be as similar as possible, thus minimizing exclusion bias.

A significant challenge when trying to decide if a DMP actually has an impact on patient clinical or economic outcomes is the phenomenon of regression to the mean. Regression to the mean is defined as "a statistical property of populations... [when] individuals with extreme values one year…tend to move toward the population average the following year" ("Standard Outcome", 2003, p. 126). In theory, the sickest patients are more likely to get better with time while the healthier patients are more likely to be sicker with time (Fetterolf et al., 2004). If analysis suggests that the DMP has made improvements in the population of interest, an investigator might draw the conclusion that the DMP is making a positive impact on the population when, in reality, regression to the mean is taking place. Linden & Roberts (2005) suggest that it can take at least 6 months for the desired behavior changes to take place within the population who received the DMP intervention, so improvements in clinical or economic outcomes within one year are likely due to regression to the mean. Increasing the study timeframe can mitigate regression to the mean. The proposed study suggests a timeframe of a year and six months, which should minimize the effects of regression to the mean.

Often, confounding enters into a study design and causes error in the interpretation of data.

Confounding can confuse the exposure-outcome relationship and lead to incorrect conclusions about the impact of a program. Confounders can cause investigators "to observe a true association [and then] to derive a causal inference, when in fact, the relationship may not be causal" (Gordis, 2009b, p. 251). When evaluating a DMP, it is difficult to understand if a confounder, such as age, sex, or comorbidities, or the DMP is causing an impact on clinical or economic outcomes. Confounders are best controlled for in an RCT because "…random distribution of unmeasured factors cancels out their impact on the study results" (Wilson et al., 2004, p. 618).

A case-control study design can control for some known confounders through the process of matching. According to Gordis (2009b), matching cases to controls is an appropriate method to control for confounders in a case-control study design. Gordis (2009a) defines matching as "the process of selecting the controls so that they are similar to the cases in certain characteristics, such as age, race, sex, or socioeconomic status, and occupation" (p. 186). Matching helps to control for characteristics that can possibly confuse the true association between the DMP and its impact on patient health and economic outcomes. For example, KPGA members with diabetes who received *HS* services (cases) could be matched to KPGA members with diabetes who did not receive *HS* services (controls) based on age. After matching in this manner, if an association is observed between the cases and the controls, then it is unlikely that the association can be attributed to age. Careful attention needs to be made not to match based on too many characteristics because it can make it difficult or impossible to match a control to a case (Gordis, 2009a). Therefore, it is important to thoroughly research which factors might be possible, significant confounders in the analysis.

3.4 Gather Credible Evidence

According to Milstein and Wetterhall's (1999) fourth step of the program evaluation

framework, gathering credible evidence increases the likelihood that stakeholders will view the evaluation results as valuable and use the results to improve the program. If credible evidence is not gathered for an evaluation, DMP stakeholders could make decisions based on weak data and possibly misallocate scarce program resources. Milstein & Wetterhall (1999) suggest five aspects of data gathering that can impact evaluation credibility: indicators, sources, quality, quantity, and logistics. Indicators are meaningful measures of program attributes that provide an indication of program effectiveness. Sources in an evaluation indicate where the data originated from, such as documents or observations. It is encouraged to use multiple evaluation indicators and sources in order to enhance the quality of the evaluation. Quality refers to the reliability of the data collected for the evaluation. Quantity refers to the amount of evaluation data that is to be collected. Defining logistics outlines how the evaluation data will be collected and handled (Milstein & Wetterhall, 1999).

Indicators are considered good measures to judge whether a program is successful or not. Indicators that can be used in the *HS* evaluation need to measure the impact program activities have on program effects. As detailed in the methods section, A1C and blood pressure readings along with ED utilization are the three main indicators suggested for the *HS* evaluation. These indicators are good measures for program success because when members with diabetes successfully self-manage diabetes, A1C levels and BP will likely be controlled and ED admissions due to diabetes complications will likely be prevented.

The source for the data will be KPGA's EMR. For the purposes of this proposed evaluation, only one data source is suggested in this particular evaluation. KPGA is an integrated medical system, which allows patient care to be more coordinated and communication to be more fluid between the health plan and health care providers. KPGA uses an EMR system, which allows

member data to be easily accessible and updated in real-time. These factors along with the new method of denoting quality DM care management calls will provide high quality data for the evaluation. Since the control and intervention groups will be selected after they meet a specific set of criteria for the *HS* evaluation, it will be impossible to determine the quantity of information needed at the beginning of the evaluation. As for logistics, a data analyst, with academic and industry experience in statistics, will be responsible for pulling the control and intervention groups according to the specifications outlined by the GSU MPH candidate.

3.5 Justify Conclusions

Milstein & Wetterhall's (1999) fifth step of the program evaluation framework explains that any judgment of program performance must be justified using the evidence gathered for the evaluation. Stakeholders are not likely to use evaluation conclusions if judgments are not linked to evidence they deem as credible. Thus, it is important to get stakeholder buy-in of evaluation standards at the beginning of the evaluation to ensure that conclusions will be relevant and useful to stakeholders. Five important components for justifying conclusions include standards, analysis and synthesis, interpretation, judgment, and recommendations. Standards reflect what stakeholders believe are the best measures by which to judge a program. Analysis includes discovering important trends in the data while synthesis combines other sources of information in order to make a larger conclusion about program effectiveness. Interpretation aims to decipher what the data indicates regarding program effectiveness. Judgments are statements about program effectiveness that are made by comparing the findings and interpretations of the evaluation. Recommendations are suggestions to either improve or maintain the program based on the evaluation results (Milstein & Wetterhall, 1999).

The agreed upon standards for measuring program success in the *HS* evaluation include

A1C, blood pressure, and ED admissions. For A1C, program success is defined as KPGA members 65 years and older having an A1C of 8% or less and KPGA members 64 years and younger with an A1C of 7% or less. Program success for blood pressure is defined as a KPGA member with diabetes having a blood pressure reading of less than 140/90 mmHg. Program success for ED admissions is defined as a KPGA member with diabetes not having any ED admissions for diabetes or its related complications.

When the data for the case-control study becomes available from the KPGA data analyst, the data analyst can then run the appropriate statistical tests to determine if there are differences in A1C levels, blood pressure readings, and ED admissions between the cases and the controls using a sophisticated statistical software package, such as SAS®. Synthesis can take place in this evaluation by using the analysis along with additional sources of information to develop an ROI for the program. Some additional sources of information that will be helpful in developing an ROI will be program cost information along with research on average A1C levels, blood pressure readings, and ED admissions for diabetics. The interpretation of the data will need to be completed by the evaluation specialist and the data analyst. Program judgments will need to be made by the evaluation specialist with input from the primary user. The recommendations to either improve or maintain the program will need to be made by the primary user, the Director of the Center for Care Partnership, in collaboration with the *HS* program manager.

3.6 Ensure and Share Lessons Learned

Milstein & Wetterhall's (1999) sixth step in the framework suggests that appropriate strategic planning is necessary to ensure that evaluation findings and conclusions are disseminated and actually used. Program evaluation is a difficult and time-consuming task. Therefore, it is important to ensure that the evaluation results are used appropriately to improve program

operations. Five elements that can ensure and share lessons learned include design, preparation, feedback, follow-up, and dissemination. The design of the evaluation needs to be discussed at the onset of the evaluation and must include evaluation questions and processes along with methods. Preparation involves thinking through the use of evaluation findings, particularly how to report negative findings. Feedback is the process of continuously communicating with stakeholders during the evaluation. Follow-up is the support, both technical and emotional, that the evaluator needs to provide to the users to ensure that lessons learned from the evaluation are used to improve the program. Dissemination is reporting the evaluation procedures and findings to relevant audiences and stakeholders in an unbiased manner (Milstein & Wetterhall, 1999).

This paper can be used as a detailed outline of the design of the *HS* evaluation. Preparation will need to be taken by the evaluator in how to most effectively report positive or negative evaluation results to users. It will be important for users and the evaluator to provide feedback to each other regarding the evaluation process to increase the likelihood that trust will be developed among the team and that evaluation results will be used. The evaluator should be responsible for following-up with users to ensure that appropriate support is provided to users so that lessons learned from the evaluation are used to improve the program or maintain program success. In order to effectively disseminate the evaluation, a detailed report and presentation reporting the evaluation process and findings should be provided to users.

4. Further Evaluation Recommendations

The *HS* evaluation is considered an outcome evaluation because the impact of the program on A1C, blood pressure, and ED admissions, all major health status indicators for DM, will be analyzed. Other types of evaluation should be conducted in addition to an outcome evaluation so that a more complete picture of program performance can be provided to stakeholders. The

University of Wisconsin Cooperative Extension (2008) emphasizes that evaluation must be incorporated into every component of the logic model in order to truly understand the effectiveness of the entire program. Two additional types of evaluation that *HS* should focus on are process and impact evaluation. Additional evaluation questions related to process and impact evaluation were developed for further evaluation of *HS* based on the logic model. These questions should be used as a guide to further evaluate other aspects of *HS* in order to enhance the outcome evaluation already performed. The additional evaluation questions are included in appendix E.

Green & Kreuter (2005) explain that process evaluation focuses on how the program is being implemented. It is important to understand if the program is being implemented the way it was intended to be implemented. Process evaluation's main goal is to understand how program activities produce desired outcomes (Green & Kreuter, 2005). Therefore, evaluation questions need to be developed that address program inputs and outputs.

It is important for *HS* stakeholders to understand the amount of resources, such as time and money, that have been invested in the program. *HS* program staff should be interviewed to better understand their input on the effectiveness of the operations of the program. Process evaluation also aims to better understand program participants' experience with the program. Surveys completed by KPGA members with diabetes provide stakeholders with participants' feedback on the effectiveness and accessibility of activities of the program.

Green & Kreuter (2005) define impact evaluation as assessing "…the immediate effect the program (or some aspect of it) has on target behaviors and their predisposing, enabling, and reinforcing antecedents…" (p. 139). Some appropriate factors to take into consideration when performing an impact evaluation include changes in knowledge, attitudes, or beliefs (Green & Kreuter, 2005).

An impact evaluation can be performed on the short and intermediate-term outcomes of the *HS* logic model. KPGA members with diabetes who participated in the program can be surveyed to understand if awareness of DM and healthy lifestyle habits has increased after the first three months of starting the program. Another survey can be provided to KPGA members with diabetes six months into the program to assess if members are able to incorporate new behaviors into their daily routine and if knowledge has increased regarding DM.

5. Conclusion

DM is a debilitating disease that greatly compromises quality of life and is expensive to manage. It is placing a huge burden on the U.S. healthcare system. The significant amount of self-care required to effectively manage DM is possibly a major contributing factor for the poor management of DM. DM self-care requires the patient to make significant behavior and lifestyle changes. Without support and education to make difficult behavior and lifestyle changes, it is unreasonable to expect patients with diabetes to effectively manage their condition. This gap in DM care management in our healthcare system needs to be filled with effective programs based on public health principals of health education and prevention. More effective programs are needed to support patients in understanding and managing DM.

The health care industry has viewed DMPs as a promising solution to the management of DM. Unfortunately, many studies of the effectiveness of DMPs are not performed with the required scientific rigor needed in order to understand if these programs are effectively managing DM. Various factors, including bias and confounding, in weak study designs threaten the internal validity of DMP evaluations and do not allow decision makers truly to understand the impact DMPs have on clinical and economic outcomes. More research and evaluation is needed understand the effectiveness of DMP.

Evaluation in industry is important, but it can be complex and challenging. Milstein & Wetterhall (1999) insist that an evaluation must include procedures that are useful, feasible, ethical, and accurate. When evaluating a DMP in industry, there are strong political influences that an evaluator must work within, which can make producing an unbiased evaluation difficult. Program stakeholders might be reluctant to complete an evaluation because they may view judgments on program performance as a critique of their job performance. Therefore, it is important that organizational cultures perceive evaluation as an important component in program management. If evaluation is not being performed on industry programs, scarce resources might be used on a program that provides little to no impact on the target population. An ineffective program does a disservice to the population who is to benefit from the intervention. With the right resources, such as data, evaluation planning, and expertise, industry can effectively evaluate DMPs to ensure patients with diabetes are receiving effective interventions that improve quality of life, prevent life-threatening diabetic complications, and decrease health care costs associated with DM.

References

American Diabetes Association. (2013). *A1C and eAG.* Retrieved from http://www.diabetes.org/living-with-diabetes/treatment-and-care/blood-glucose-control/a1c/

American Diabetes Association. (2014) *Statistics about diabetes*. Retrieved from http://www.diabetes.org/diabetes-basics/statistics/

Agency for Healthcare Research and Quality. (2013). *Statistical brief 167: Emergency department visits for adults with diabetes, 2010.* Retrieved from http://www.hcup-us.ahrq.gov/reports/statbriefs/sb167.pdf

Centers for Disease Control and Prevention. (2012). *Diabetes report card 2012.* Retrieved from http://www.cdc.gov/diabetes/pubs/pdf/diabetesreportcard.pdf

Centers for Disease Control and Prevention. (2013). *2011 national diabetes fact sheet.* Retrieved from http://www.cdc.gov/diabetes/pubs/factsheet11/fastfacts.htm

Centers for Disease Control and Prevention. (2014). *Chronic diseases and health promotion.* Retrieved from http://www.cdc.gov/chronicdisease/overview/index.htm

Economic costs of diabetes in the u.s. in 2012. (2013). *Diabetes Care, 36*(4), 1033-1046. doi:10.2337/dc12-2625

Fetterolf, D., Wennberg, D., & DeVries, A. (2004). Estimating the return on investment in disease management programs using a pre-post analysis. *Disease Management*, *7*(1), 5-23.

Fireman, B., Bartlett, J., & Selby, J. (2004). Can disease management reduce health care costs by improving quality?. *Health Affairs*, *23*(6), 63-75. doi:10.1377/hlthaff.23.6.63

Fitzner, K., Sidorov, J., Fetterolf, D., Wennberg, D., Eisenberg, E., Cousins, M., & ... Duncan, I. (2004). Principles for assessing disease management outcomes. *Disease Management*, *7*(3),

191-201. doi:10.1089/dis.2004.7.191

Gordis, L. (2009a). Using epidemiology to evaluate health services. *Epidemiology* (pp. 293-309). Philadelphia: Elsevier/Saunders.

Gordis, L. (2009b). More on causal inferences: bias, confounding, and interaction. *Epidemiology* (pp. 247-264). Philadelphia: Elsevier/Saunders.

Gordis, L. (2009c). Case-control studies and other study designs. *Epidemiology* (pp. 177-200). Philadelphia: Elsevier/Saunders.

Green, L. W., & Kreuter, M. W. (2005). Epidemiological diagnosis: health, behavioral, and environmental assessments. *Health program planning: an educational and ecological approach* (pp. 78-145). New York: McGraw-Hill.

HEDIS. (2013). *HEDIS© and quality compass©.* Retrieved from http://www.ncqa.org/HEDISQualityMeasurement/WhatisHEDIS.aspx

Lewis, A. (2009). How to measure the outcomes of chronic disease management. *Population Health Management*, *12*(1), 47-54. doi:10.1089/pop.2008.0035

Linden, A., & Roberts, N. (2005). A user's guide to the disease management literature: recommendations for reporting and assessing program outcomes. *American Journal Of Managed Care*, *11*(2), 113-120.

Milstein, R. L., & Wetterhall, S. F. (1999). Framework for program evaluation in public health. *Morbidity and Mortality Weekly Report, 48*(11). Retrieved from ftp://ftp.cdc.gov/pub/Publications/mmwr/rr/rr4811.pdf

National Diabetes Information Clearinghouse. (2014) *Your guide to diabetes: type 1 and 2.* Retrieved from http://www.diabetes.niddk.nih.gov/dm/pubs/type1and2/index.aspx

Roblin, D., Ntekop, E., & Becker, E. (2007). Improved intermediate clinical outcomes from

participation in a diabetes health education program. *The Journal Of Ambulatory Care Management*, *30*(1), 64-73.

Serxner, S., Baker, K., & Gold, D. (2006). Guidelines for analysis of economic return from health management programs. *American Journal of Health Promotion*, *20*(6), 1-17.

Standard outcome metrics and evaluation methodology for disease management programs. (2003). *Disease Management*,*6*(3), 121-138.

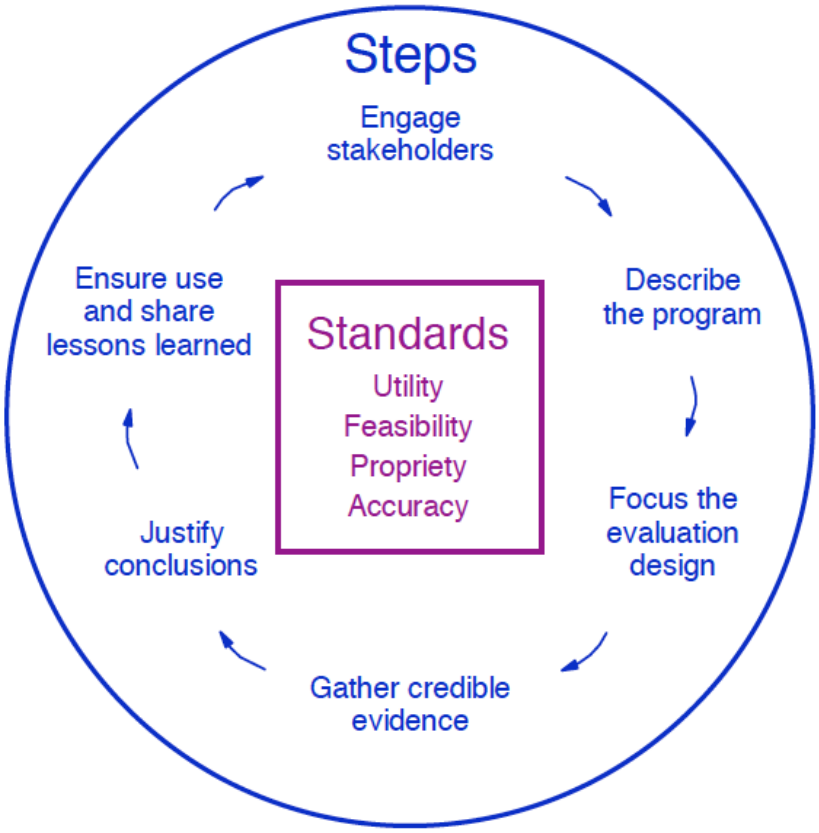
University of Wisconsin Cooperative Extension (2008). *Developing a logic model.* [PowerPoint slides]. Retrieved from http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html

Wilson, T., & MacDowell, M. (2003). Framework for assessing causality in disease management programs: principles. *Disease Management*, *6*(3), 143-158.

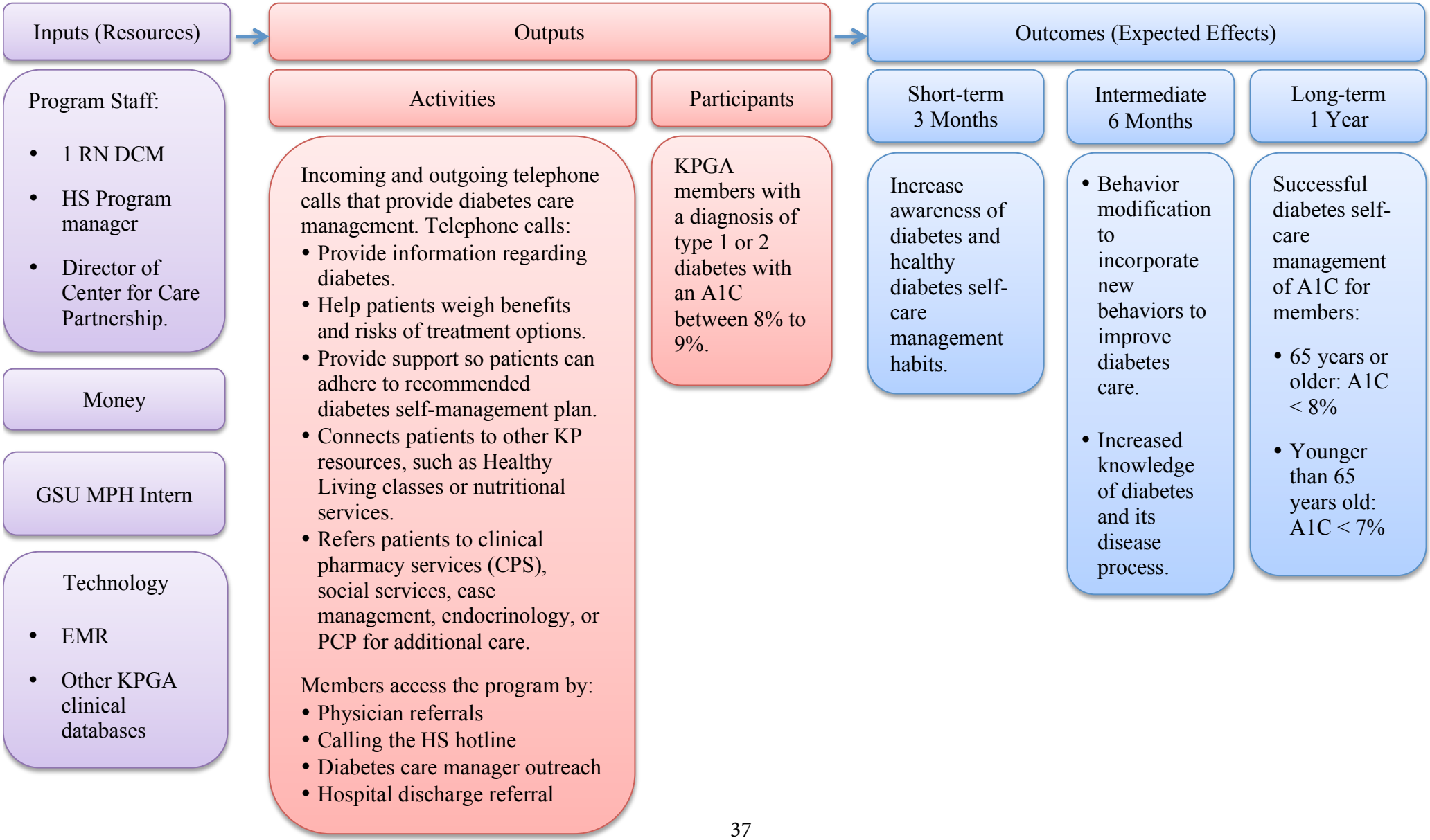
Wilson, T., Gruen, J., Thar, W., Fetterolf, D., Patel, M., Popiel, R., & ... Nash, D. (2004). Assessing return on investment of defined-population disease management interventions. *Joint Commission Journal On Quality & Safety*, *30*(11), 614-621.

Wilson, T., MacDowell, M., Salber, P., Montrose, G., & Hamm, C. (2008). Evaluation methods in disease management studies: 2004-07. *Disease Management & Health Outcomes*, *16*(5), 365-373.

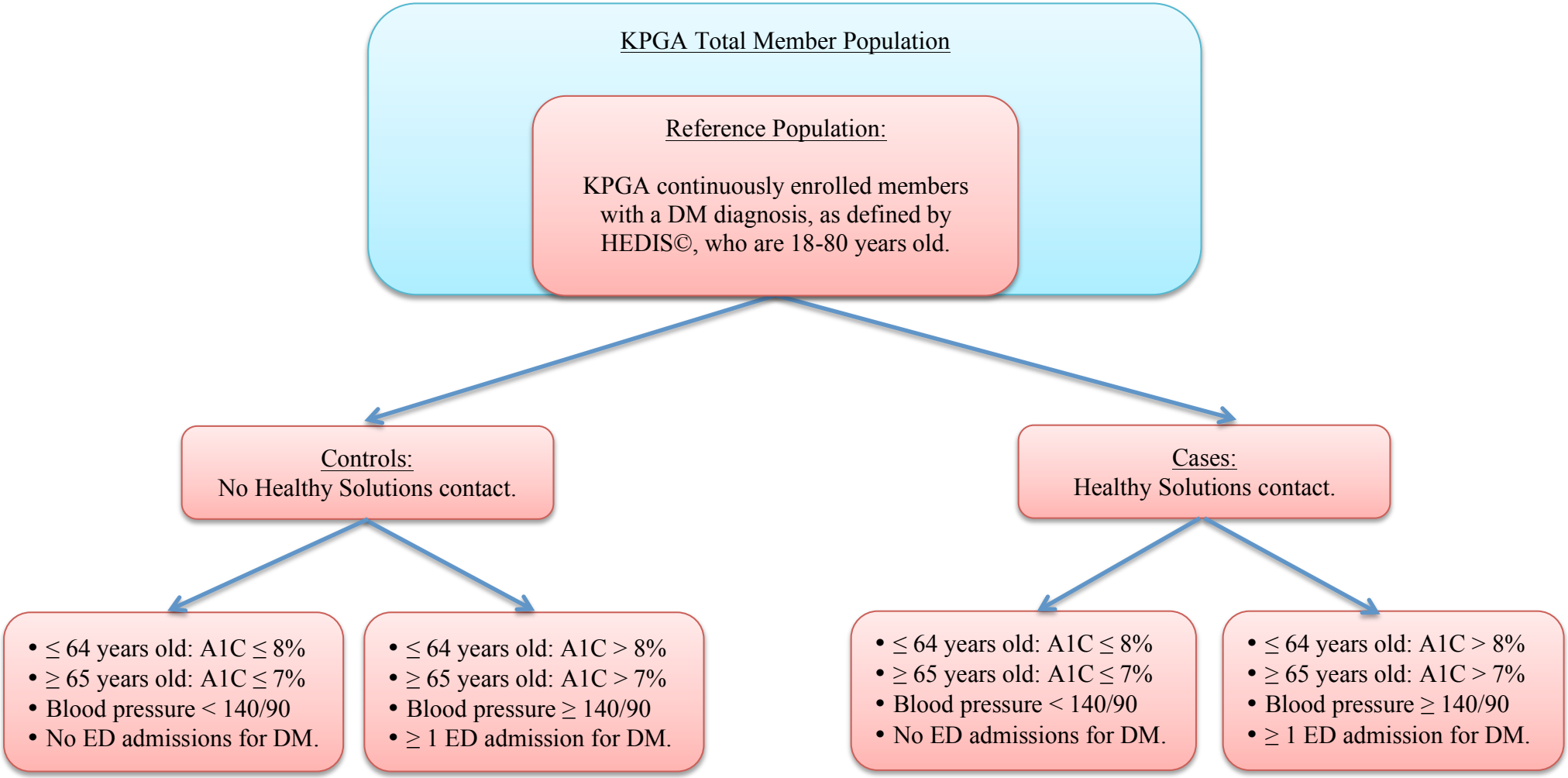
Appendix A
Program Evaluation Framework (Milstein & Wetterhall, 1999, p. 4)



Appendix B
HS Logic Model



Appendix C
Study Population[1]



1. Adapted from Gordis, L. (2009c). Case-control studies and other study designs. In L. Gordis (Ed.) *Epidemiology* (pp. 177-200). Philadelphia: Elsevier/Saunders.

Appendix D
Case-control Study Design for the Proposed *HS* Evaluation

	10/1/2013 --- 3/31/2014 **Baseline Period**	4/1/2014 --- 9/30/2014 **Intervention Period**	10/1/2014 --- 3/31/2015 **Post-Intervention Period**
Definition of Case Observation	• ***No HS contact*** • A1c result ***AND/OR*** • BP (systolic/diastolic)	• Patient receives ***at least one*** *HS* contact • A1c result ***AND/OR*** • BP (systolic/diastolic)	• A1c result ***AND/OR*** • BP (systolic/diastolic)
Definition of Control Observation	• ***No HS contact*** • A1c result ***AND/OR*** • BP (systolic/diastolic)	• ***No HS contact*** • Patient receives ***at least one*** primary care encounter. • A1c result ***AND/OR*** • BP (systolic/diastolic)	• ***No HS contact*** • A1c result ***AND/OR*** • BP (systolic/diastolic)

Appendix E
Further Evaluation Question Recommendations[1]

					Data Collection	
			Questions	Indicators	Sources	Methods
Process Evaluation	Inputs		What amount of time, money, and other resources were invested in the program?	# of staff, money used, other resources used.	Program records or existing program data.	Document review
			Do the program staff feel as though the program is effective or do changes need to be made?	Program feedback from those involved in program operations.	Program Manager, diabetes health coach, & the Center for Care Partnership Director	Interviews
	Outputs	Activities	Do KPGA members feel they are receiving the appropriate information regarding diabetes from the program?	#/% expressing information relevant	KPGA members with diabetes	Survey
			Are KPGA members with diabetes able to better weight the benefits and risks of treatment options after contact with the program?	#/% reporting being able to weight risks/benefits of treatment	KPGA members with diabetes	Survey
			Are KPGA members with diabetes receiving necessary support from the program so they can adhere to recommended diabetes management plan?	#/% reporting receiving necessary support	KPGA members with diabetes	Survey
			Are KPGA members with diabetes more knowledgeable about other KPGA resources for diabetes care management after contact with the program, such as Healthy Living classes or nutritional services?	#/% reporting being more knowledgeable about other KPGA resources available to them.	KPGA members with diabetes	Survey
		Participants	Do KPGA members with diabetes who participated in the program feel the program was helpful in the management of their diabetes?	#/% reporting the program as helpful	KPGA members with diabetes	Survey
			How do KPGA members with diabetes learn about the program?	Sources that members report that informed them of program.	KPGA members with diabetes	Survey

1. Adapted from University of Wisconsin Cooperative Extension (2008). *Developing a logic model.* [PowerPoint slides]. Retrieved from http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html

Impact Evaluation	Outcomes	Short-term (3 months)	Do KPGA members with diabetes have an increased awareness of diabetes and healthy lifestyle habits?	#/% demonstrating increased awareness of diabetes and healthy lifestyle habits.	KPGA members with diabetes	Survey
		Intermediate-term (6 months)	Are KPGA members with diabetes able to identify ways to incorporate new behaviors into their daily routine (i.e. a healthy diet or ways to increase physical activity)?	#/% reporting identification of ways to eat healthier & increase physical activity.	KPGA members with diabetes	Survey
			Do KPGA members with diabetes have an increased knowledge of diabetes and its disease process?	#/% demonstrating increased knowledge of diabetes and its disease process.	KPGA members with diabetes	Survey
Outcome Evaluation		Long-term (1 year)	Is the program successful in lowering or maintaining KPGA members with diabetes (65 years or older) A1C at 8% or less?	% of population maintaining A1C of 8% or less.	Data from KPGA EMR system.	Data analysis
			Is the program successful in lowering or maintaining KPGA members with diabetes (64 years or older) A1C at 7% or less?	% of population maintaining A1C of 7% or less.	Data from KPGA EMR system.	Data analysis

1. Adapted from University of Wisconsin Cooperative Extension (2008). *Developing a logic model.* [PowerPoint slides]. Retrieved from http://www.uwex.edu/ces/pdande/evaluation/evallogicmodel.html