

Is Positive Return on Investment (ROI) Too Good to Be True?

A Review of the Quality of Study Methodology in Behavioral Chronic Disease Interventions

Claiming Positive ROI.

by

Jiayang Hong

Department of Global Health
Duke University

Date: _____

Approved:

Eric Finkelstein, Supervisor

Elizabeth Turner

Gary Bennett

Thesis submitted in partial fulfillment of
the requirements for the degree of
Master of Science in the Department of
Global Health in the Graduate School
of Duke University

2017

ABSTRACT

Is Positive Return on Investment (ROI) Too Good to Be True?

A Review of the Quality of Study Methodology in Behavioral Chronic Disease Interventions

Claiming Positive ROI.

by

Jiayang Hong

Department of Global Health
Duke University

Date: _____

Approved:

Eric Finkelstein, Supervisor

Elizabeth Turner

Gary Bennett

An abstract of a thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in the Department of Global Health in the Graduate School of Duke University

2017

Copyright by
Jiayang Hong
2017

Abstract

Objective: To examine the quality of articles that claim cost savings resulting from chronic disease interventions. The interventions are limited to behavioral interventions, including worksite wellness programs, weight and disease management programs, and community based programs. **Methods:** Published articles were identified from a database search. Included articles were published between January 1990 and December 2016, described a positive return on investment of behavioral interventions for chronic disease. A single reviewer, following specific criteria, assessed research quality using the Quality of Health Economic Studies (QHES) instrument. **Results:** Of 1990 retrieved articles, 19 met study inclusion criteria. No study we reviewed has sufficient quality to be considered evidentiary. **Conclusions:** Economic evaluations yield positive ROI on chronic disease behavioral interventions have limited methodological quality and their results should be interpreted with caution.

Dedication

“For once you have tasted flight, you will forever walk the earth with your eyes turned skyward,
for there you have been, and there you will always long to return.”

Attributed to Leonardo da Vinci

Table of Contents

| | |
|--|-----|
| Abstract | iv |
| List of Tables | x |
| List of Figures | xi |
| Acknowledgements..... | xii |
| 1. Introduction..... | 1 |
| 2. Method | 3 |
| 2.1 Eligibility Criteria..... | 3 |
| 2.2 Information Sources & Search Strategy | 3 |
| 2.3 Study Selection..... | 5 |
| 2.4 Data Collection Process..... | 5 |
| 2.5 Methodological Quality Assessment..... | 6 |
| 3. Result | 7 |
| 3.1 Results of quality QHES scoring among 19 studies..... | 8 |
| 3.2 Detailed Analysis of Studies That Passed QHES Cutoff | 10 |
| 4. Discussion | 17 |
| 5. Conclusion | 18 |
| Appendix A. Background | 19 |
| a. Chronic disease burden..... | 19 |
| b. Chronic disease management program..... | 22 |
| c. Wellness program | 23 |
| d. Gray Literature | 26 |
| Appendix B. CHEERS Checklist..... | 29 |

| | |
|--|----|
| a. Analysis of Each Criterion | 33 |
| a.1 Baseline & Setting and location | 33 |
| a.2 Study perspective..... | 34 |
| a.3 Comparators / alternatives | 34 |
| a.4 Discount rate & Price date..... | 34 |
| a.5 Choice of outcomes & Time horizon | 34 |
| a.6 Choice of economic evaluation | 35 |
| a.7 Measurement of effectiveness | 35 |
| a.8 Confidence Interval | 35 |
| a.9 Assumptions & Analytical methods..... | 35 |
| a.10 Incremental cost and outcomes | 35 |
| a.11 Characterizing uncertainty..... | 36 |
| a.12 Limitation | 36 |
| a.13 Funding & Conflict of interest | 36 |
| Appendix C | 37 |
| 1. A Comprehensive Worksite Wellness Program in Austin, Texas: Partnership Between Steps to a Healthier Austin and Capital Metropolitan Transportation Authority (Davis et al., 2009)..... | 37 |
| 2. A Matched-Cohort Study of Health Services Utilization and Financial Outcomes for a Heart Failure Disease-Management Program in Elderly Patients (Berg et al., 2004)..... | 37 |
| 3. A Return on Investment Evaluation of the Citibank, N.A., Health Management Program (Ozminkowski et al., 1999) | 38 |
| 4. An evaluation of cost sharing to finance a diet and physical activity intervention to prevent diabetes (Ackerman et al., 2006)..... | 38 |
| 5. An Integrated Pharmacy-Based Program Improved Medication Prescription and Adherence Rates in Diabetes Patients (Brennan et al., 2012) | 39 |

| | |
|--|----|
| 6. Can Health Insurance Improve Employee Health Outcome and Reduce Cost? (Maeng et al., 2013)..... | 40 |
| 7. Diabetes Disease Management in a Community-Based Setting (Berg et al., 2002)..... | 40 |
| 8. Evaluation of an Outpatient Pharmacy Clinical Services Program on Adherence and Clinical Outcomes Among Patients with Diabetes and/or Coronary Artery Disease (Spence et al., 2014)..... | 41 |
| 9. Financial Implications of Glycemic Control: Results of An Inpatient Diabetes Management Program (Newton et al., 2006)..... | 41 |
| 10. Impact of a Patient-Centered Pharmacy Program and Intervention in a High-Risk Group (Moore et al., 2013)..... | 42 |
| 11. Impact of Worksite Wellness Intervention on Cardiac Risk Factors and One-Year Health Care Costs (Milani and Lavie, 2009)..... | 43 |
| 12. Managing Manifest Diseases, But Not Health Risks, Saved PepsiCo Money Over Seven Years (Caloyeras et al., 2014)..... | 43 |
| 13. Medical Cost Savings for Participants and Nonparticipants in Health Risk Assessments, Lifestyle Management, Disease Management, Depression Management, and Nurseline in a Large Financial Services Corporation (Serxner et al., 2012)..... | 44 |
| 14. Medication Therapy Management: 10 Years of Experience in a Large Integrated Health Care System (Ramalho de Oliveira et al., 2010)..... | 44 |
| 15. Medication therapy management and condition care services in a community-based employer setting (Johannigman et al., 2010)..... | 45 |
| 16. Quality Improvement and Cost Reduction Realized by a Purchaser Through Diabetes Disease Management (Snyder et al., 2003)..... | 45 |
| 17. Recent Experience in Health Promotion at Johnson & Johnson: Lower Health Spending, Strong Return on Investment (Henke et al., 2011)..... | 46 |
| 18. Reducing Need and Demand for Medical Services in High-risk Persons (Fries and McShane, 1998)..... | 46 |
| 19. Return on Investment for Digital Behavioral Counseling in Patients with Prediabetes and Cardiovascular Disease (Su et al., 2016)..... | 47 |
| Appendix D..... | 48 |
| a. Community Chronic Disease Intervention Program..... | 48 |

| | |
|---|----|
| b. Clinic Chronic Disease Intervention Program..... | 49 |
| d. Workplace Wellness Program..... | 50 |
| References..... | 52 |

List of Tables

| | |
|--|----|
| Table 1: Pubmed, Embase, and Google Scholar Search Strategy..... | 3 |
| Table 2: Summary of Key Characteristics of Studies Included in The Review | 13 |
| Table 3: Summary of QHES score and breakdown of Studies Included in The Review | 14 |
| Table 4: Modified Consolidated Health Economic Evaluation Reporting Standards (CHEERS) Checklist | 29 |
| Table 5: Modified Consolidated Health Economic Evaluation Reporting Standards (CHEERS) Checklist | 31 |
| Table 6: Frequency of Each Criterion Met by ROI Studies in Chronic Disease Intervention Program (N= 19)..... | 31 |

List of Figures

| | |
|---|----|
| Figure 1: Flow Diagram of Study Selection. | 7 |
| Figure 2: Percentage of Studies Included in The Review Addressing Each Question of The QHES Instrument (N=19). | 9 |
| Figure 3: QHES Instrument Score Summary for All Studies (N=19). | 10 |

Acknowledgements

I would like to express the deepest appreciation to my committee chair, Dr. Eric Finkelstein, who has the attitude and the substance of a great mentor; he truly understands what hard work and dedication can bring to one's life. Without his guidance and persistent help this thesis would not have been possible.

I would also like to express gratitude toward committee members, Dr. Elizabeth Turner and Dr. Gary Bennett, who devote their time to guide me along the path.

Finally, I would like to express appreciation to every faculty and staff in DGHI who helped me along the journey. Without the support and motivation provided by the DGHI family, life in graduate school would have been mundane.

This thesis is formatted according to Health Affair systematic review guideline for publishing. Any part of the study that do not fit manuscript is in the Appendix section.

1. Introduction

Chronic diseases are responsible for a sizeable and growing burden of largely preventable morbidity and mortality. Results from the 2010 National Health Interview Survey reveal that over half of adults in the USA have at least one chronic condition and 26% have two or more conditions (Schiller et al., 2012). In 2014, seven of the top ten leading causes of death in the USA were chronic conditions. Since 2011, death due to chronic disease has steadily increased both in total and in the percentage of all deaths (CDC, 2015).

The provision of care for patients with chronic disease is costly. The Agency for Healthcare Research and Quality reports that care for people with chronic conditions accounts for over 70% of total healthcare spending. Among Medicare fee-for-service beneficiaries, patients with two or more chronic conditions account for 93% of Medicare spending (Gerteis et al., 2014). Loss of productivity causes additional costs for people remaining in the workforce, for those who prematurely leave due to disabilities caused by chronic diseases, and for insurers and employers who subsidize these costs (CMMS, 2012).

As a result of the high disease burden and costs, chronic disease management and wellness programs have grown into a \$6 billion industry per year (Mattke et al., 2013). Employers spend roughly \$600 per employee on wellness programs per year (NBGH, 2016) with 83% of US large firms (200 or more employees) and 46% of small firms (less than 200 employees) offering some type of wellness program (RAND, 2016). Insurers have also invested heavily in chronic disease management programs (Mattke et al., 2015).

The theoretical underpinning behind chronic disease management and employee wellness programs is well documented (Weingarten et al., 2002; Carnethon et al., 2009). The goal is to address the risk factors that cause chronic diseases via prevention and tailored interventions. The

empirical literature also shows that many of these programs show effectiveness and cost effectiveness (Pelletier, 2009; Barlow et al., 2002), meaning the cost per quality adjusted life year (QALY) gained, or related measure, is within accepted bounds of what third parties may be willing to pay (Neumann et al., 2014). Yet, most private payers (and increasingly government payers) are looking for more than cost effectiveness. Their goal is to reduce health care costs (Mattke et al., 2015). As a result, there is a high demand for programs that can make the case of cost savings. Yet, this is a tall order given that the nature of these programs is to reduce risk factors for conditions that sometimes do not materialize for years or even decades into the future. This makes the chances of showing a positive return on investment (ROI) unlikely. However, there is an increasing number of studies claiming to show positive ROI, with some estimating returns as high as \$5.82 for each dollar spent (Baicker et al., 2010).

Given the nature of chronic disease prevention, we find these claims suspect. Therefore, the nature of this review is to explore whether or not the quality of these studies is sufficient such that the claims of cost savings may be considered credible.

2. Method

This systematic review was conducted following the Preferred Reporting Items for Systematic Review and Meta-Analyses (PRISMA) statement.

2.1 Eligibility Criteria

Eligible studies included all published articles that show a positive return on investment of behavioral interventions for chronic disease, including worksite wellness programs, weight management programs, community based programs, and disease management programs.

2.2 Information Sources & Search Strategy

A comprehensive systematic search of the literature was performed. Three databases—PubMed, Embase, and Google Scholar—were searched to identify relevant studies published from 1990 through December 2016. In addition, a manual search of citations from selected systematic reviews was undertaken. Medical subject headings (MeSH) terms were used. Only English language articles are included.

Table 1 shows the exact search strategy employed in PubMed, Embase and Google Scholar search.

Table 1: Pubmed, Embase, and Google Scholar Search Strategy

| PubMed Search Strategy | |
|---|---|
| Long Search | #1 And #2 And #3 |
| #1 Chronic Disease | “Chronic Disease” [Mesh] OR “Chronic Illness” [tiab] OR “Heart Disease”[Mesh] OR “Cardiac Diseases”[Mesh] OR “Stroke” [Mesh] OR “cancer”[tiab] OR ” Neoplasms” [Mesh] OR “type 2 diabetes” [tiab] OR “Diabetes Mellitus, Type 2” [Mesh] OR “obesity” [Mesh] OR “arthritis” [Mesh] |
| #2 Intervention | “health promotion”[Mesh] OR “Promotional Items”[tw] OR “Wellness Programs”[tw] OR “Health Campaigns”[tw] OR “prevention”[tw] OR “health” [tw] OR “wellbeing” [tw] OR “weight management”[tw] OR “disease management”[tw] OR “behavioral” [tw] OR “counseling” [tw] OR “education” [tw], OR “self management” [tw] |
| #3 Economic filters | “net cost”[tw] OR “cost saving”[Mesh] OR “Costs and Cost Analyses”[tw] OR “Cost Comparison”[tw] OR “Cost”[Mesh] OR “Return on Investment”[tw] OR “Investment”[Mesh] |
| Embase and Google Scholar Search Strategy | |
| Long Search | #1 And # 2 And # 3 |
| #1 Chronic Disease | (“Chronic Disease” OR “Heart Disease” OR “Cardiac Diseases” OR “Stroke” OR “cancer” OR “Neoplasms” OR “type 2 diabetes” OR “Diabetes” OR “obesity” OR “arthritis”) |
| #2 Intervention | (“health promotion” OR “Wellness Programs” OR “Health Campaigns” OR “prevention” OR “health” OR “wellbeing” OR |

| | |
|--------------------------------------|---|
| | <p>“weight management” OR “disease management” OR “behavioral” OR “counseling” OR “education”, OR “self management”)</p> |
| <p>#3 Economic filters</p> | <p>(“net cost” OR “cost saving” OR “Costs and Cost Analyses” OR “Cost Comparison” OR “Cost” OR “Return on Investment” OR “Investment”</p> |

* Long search performed on January 23, 2017, yielded Pubmed (525), Embase (255), and Google Scholar (1210) studies.

2.3 Study Selection

From the initial yield, each article’s title and abstract were reviewed for relevance, defined as containing at least one behavioral intervention, targeting at least one chronic disease, and a claim of cost savings or positive ROI. The reference list of each article was scanned to identify any citations that were not initially captured. Final decision of inclusion was made after review of the entire manuscript. Literature review, systematic review, and editorials were excluded.

2.4 Data Collection Process

Data extraction was performed independently by one reviewer. Information was extracted for each including the affiliation of authors, program provider, study design, duration, the number of participants and control in these studies, participant eligibility, the type of the intervention, the

duration of the intervention, outcome measure, cost of the program, benefit of the program, ROI, and the ROI calculation method.

We developed a data extraction sheet (based on the PRISMA template), pilot-tested it on three randomly-selected included studies. The initial result was checked with mentor to ensure accuracy. Then extracted the above information from included studies.

2.5 Methodological Quality Assessment

The methodological quality of studies was assessed using the Quality of Health Economic Studies (QHES) instrument (Ofman et al., 2003). It consists of 16 criteria in the form of “yes or no” questions that were selected by a panel of 8 experts in health economics (Ofman et al., 2003). Each question has a weighted point value ranging from 1-9, which are used to generate a summary score from 0-100 (Ofman et al., 2003). QHES scores below 50 are considered to be of poor quality (Ofman et al., 2003; Spiegel et al., 2004).

We applied the QHES questionnaire to assess the methodological rigor of all identified studies. Results from the studies falling below the cutoff line are considered poor quality. However, studies that exceed the threshold may still contain biases and/or other shortcomings that undermine the credibility of the conclusions. Therefore, we conducted in-depth methodological reviews on studies exceeding the threshold to identify potential biases and other concerns. For these studies, we focus mainly on three key aspects of the study: 1) Does it include a clinical endpoint as a mediating factor between the intervention and cost savings, 2) does the design

adequately control for potential biases, and 3) does the study incorporate sensitivity analyses around key variables.

3. Result

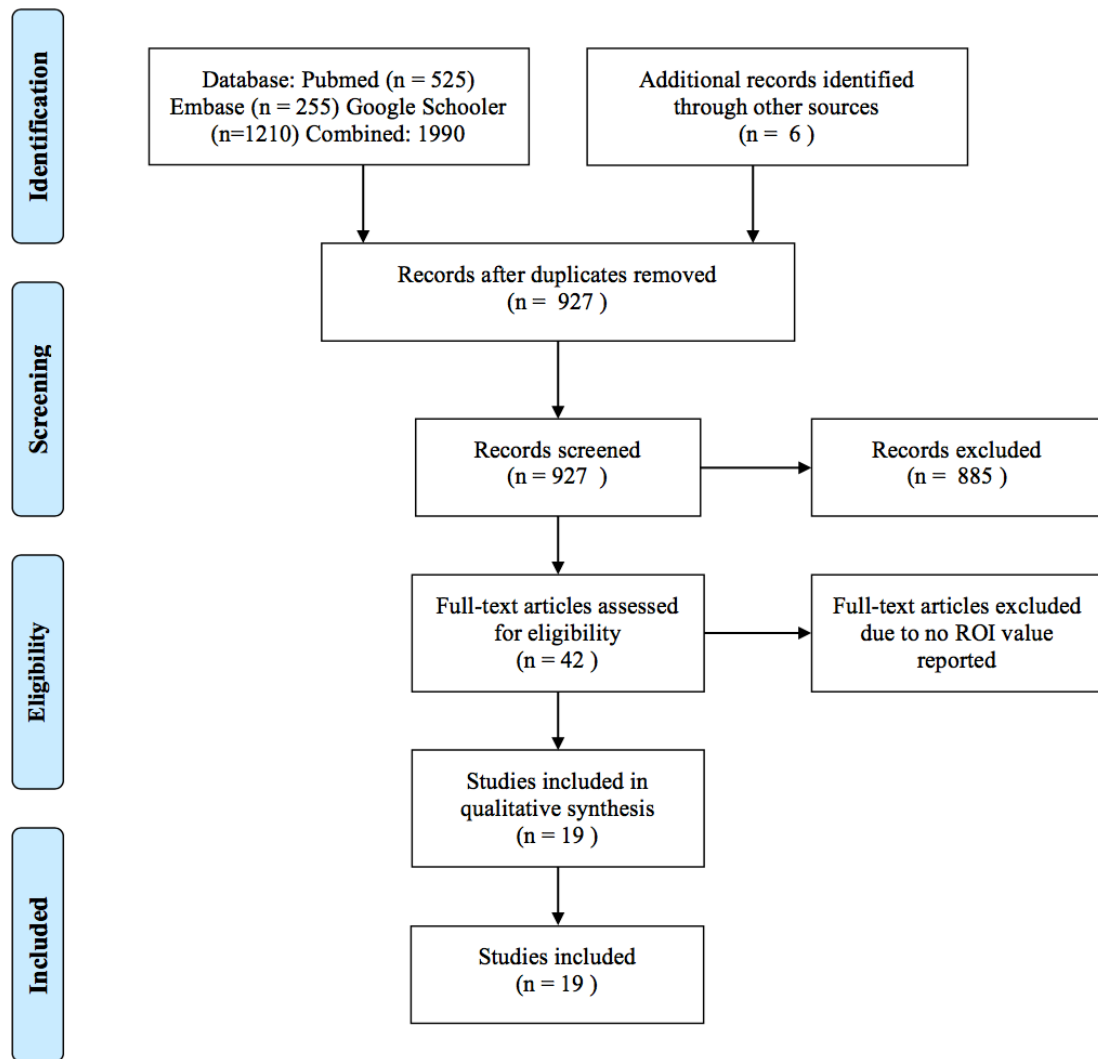


Figure 1: Flow Diagram of Study Selection.

The search process identified a total of 927 articles after excluding duplications, including 525 from PubMed, 255 from EMBASE database, 1210 from Google Scholar and 6 from the reference of articles. 19 articles fulfilled our inclusion criteria. Fig. 1 shows the flow diagram for the study selection from the three databases.

Table 2 summarizes the key features of the studies. Three types of chronic disease intervention programs were identified based on their principal emphasis and target population: (1) Community Chronic Disease Intervention Programs (n = 6); (2) Worksite Wellness Programs (n = 9); and (3) Clinic Chronic Disease Management Programs (n = 4). The majority of the studies used non-randomized designs, including ten case control studies, five pre-post studies, and three model simulation studies. Compared to randomized designs, these studies are more prone to selection bias and omitted variable bias (Sterne et al., 2016). Only one study used a randomized design (Table 2).

3.1 Results of quality QHES scoring among 19 studies

For all studies combined, the questions of the Quality of Health Economic Studies (QHES) instrument that were most frequently addressed in the studies related to the specification of clear, measurable objectives (Q1) (19 of 19), study perspective (Q2) (12 of 19), pre-specification of subgroups for subgroup analyses (Q4) (11 of 19), and the inclusion of appropriate reliable and valid economic outcomes (Q10) (14 of 19) (Fig. 2). By comparison, the areas of the QHES questions that were not well addressed related to use the best study methodology possible

(Q3) (6 of 19), the handling of uncertainty (Q5) (5 of 19), whether the duration of the study allowed for all important or relevant outcomes to be captured (Q8) (2 of 19), discussion of potential biases (Q14) (9 of 19), and other methodological items such as the inclusion of appropriate reliable and valid health outcomes (Q11) (6 of 19), data abstraction methods (Q7) (6 of 19) appropriate economic modeling (Q6, Q12, Q13) (1 of 19, 4 of 19, 5 of 19), justified conclusions based on study results (Q15) (9 of 19), and disclosure of the study funding source (Q16) (5 of 19).

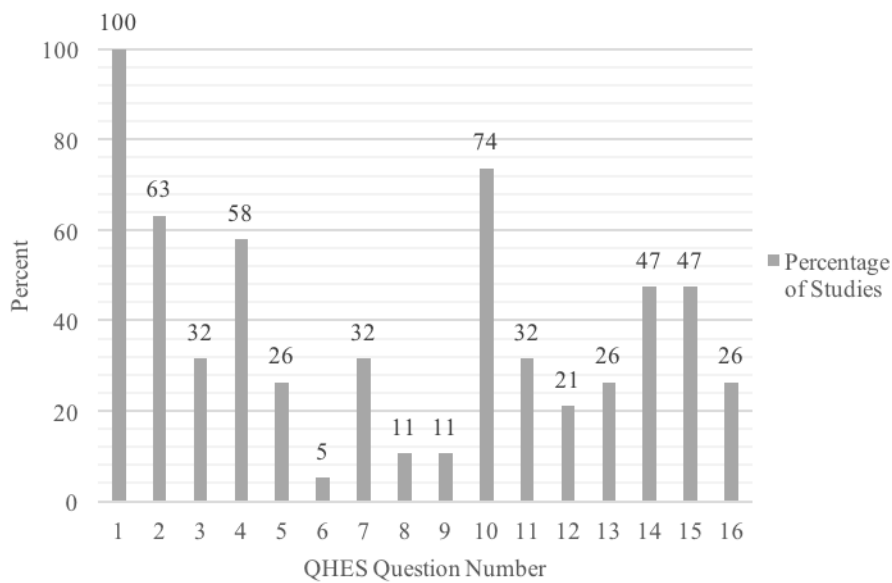


Figure 2: Percentage of Studies Included in The Review Addressing Each Question of The QHES Instrument (N=19).

Component scores for each study are shown in Table 4, with summary scores included in Figure 3. The figure shows that only 3 of the 19 studies had a score above the 50% cutoff for

what is considered a high quality study. These are Ozminkowski et al. (1999), Brennan et al. (2012), and Maeng et al. (2013) with respective ROIs of \$4.56-4.73, \$3, and \$1.52. However, confidence intervals reveal that Maeng’s estimate is not statistically significant, therefore, the authors cannot reject the null hypothesis of no ROI. This leaves only two studies that may provide convincing evidence of a potential ROI.

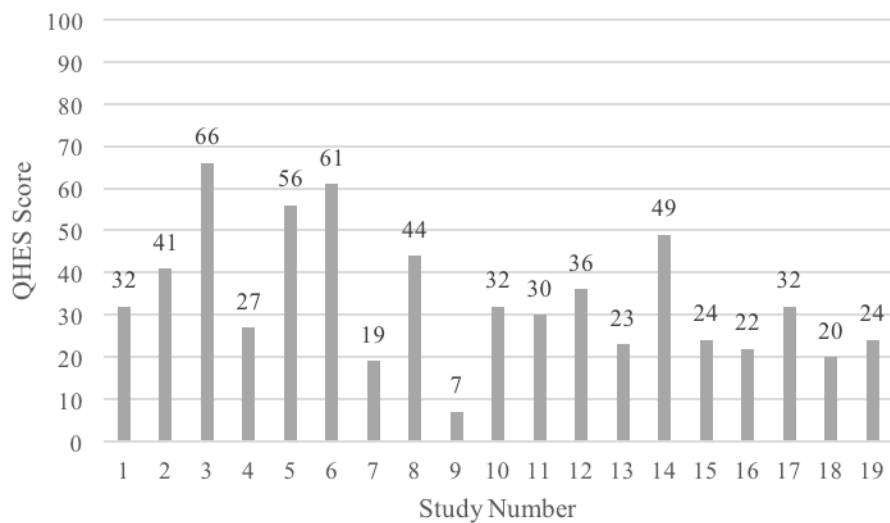


Figure 3: QHES Instrument Score Summary for All Studies (N=19).

3.2 Detailed Analysis of Studies That Passed QHES Cutoff

Ozminkowski’s (1999) case-control study examined the net medical expenditures of 22,838 self-selected participants and 11,644 non-participants. The study did not include any evidence of improvements in clinical endpoints or variables that would act as indicators for the cost savings. The study confounded by the classical issue of self-selection and non-comparable comparison groups. Characteristics for intervention and control group participants differed at

baseline for several key indicators, including coverage type, job type, salary, the department they were working, and gender.

Pre- and post-period logistic and ordinary least-squares regressions were carried out to assess the difference in medical expenditure, and examine the odds ratio for different variables between active and control group over the time horizon. However, there is a predicted direction of bias due to the selection of participants into the study. Non-participant in the study are more likely to be hourly wage earners, earn less income, and be on a single health coverage (as opposed to family coverage); all of which affect health-seeking behavior as a result of less job security, less money, more stress, and more expensive health insurance. Also, no confidence interval was found in the ROI evaluation.

Brennan's (2012) case-control study consisted 5,123 self-selected participants and 24,124 matched non-employees control group based on sex, age, Ingenix pharmacy risk group score, preferred pharmacy, and medication possession ratio. However, the authors did not describe how the matching was done. Key characteristics like pharmacy risk group score, age, and median household income were significant different in baseline between intervention and control group, which made the study vulnerable to selection bias and confounding.

A six-month time horizon was used in this study. The authors did not quantify ROI as part of their study. The ROI was only speculated in the discussion section using models from other studies that look at the present value of future cost savings from improved medication

adherence. It makes the positive ROI claim dubious in this study. Despite the limitations in the economic analysis, the authors claimed in the abstract that the study showed a positive ROI as a main finding of the study.

The study also did not report a valuation method for the cost offset. The authors did not describe how adherence was defined. Segment linear regression was performed to estimate the change in adherence, however, the authors did not explain their segmentation methods. Also, no confidence interval or probability distribution was found out throughout the study.

Table 2: Summary of Key Characteristics of Studies Included in The Review

| Authors | Type | Study Design | Sample Size | | Time Horizon (Years) | Intervention Program Cost | | Intervention Program Benefits | | ROI | |
|---|-----------|--------------|--------------|---------|----------------------|---------------------------|------|-------------------------------|------|-------------------------|-----------|
| | | | Intervention | Control | | Cost (\$) | Unit | Benefit (\$) | Unit | Perspective | ROI (\$) |
| Ackerman et al., 2006 | Community | M | 3234 | 0 | 3 | NA | - | 2,136 | PP | Private Payer | 1 |
| Berg et al., 2002 | Community | PP | 127 | 0 | 1 | 49,429 | T | 214,486 | T | NA | 4.34 |
| Ramalho de Oliveira et al., 2010 | Community | PP | 9068 | 0 | 10 | 2,258,302 | T | 2,913,850 | T | Health service provider | 1.29 |
| Snyder et al., 2003 | Community | PP | 422 | 0 | 3 | 415,533 | T | 1,402,071 | T | Program Purchaser | 3.37 |
| Fries and McShane, 1998 | Community | CC | 2586 | 89652 | 0.5 | 50 | PPP6 | 484 | PPP6 | NA | 6 |
| Su et al., 2016 | Community | M | 2371 | 0 | 3 | NA | - | NA | - | NA | 1.07 |
| Berg et al., 2004 | Clinic | CC | 533 | 236 | 1 | 619,902 | T | 1,430,281 | T | Employer | 2.31 |
| Spence et al., 2014 | Clinic | CC | 1480 | 1477 | 1 | 1,713,468 | T | 11,640,296 | T | Health plan provider | 5.79 |
| Newton et al., 2006 | Clinic | M | NA | NA | 1 | NA | - | 2,224,029 | T | NA | 4.67 |
| Moore et al., 2013 | Clinic | CC | 2250 | 2250 | 1 | 478 | PPPY | 977 | PPPY | NA | 2 |
| Davis et al., 2009 | Worksite | PP | 300 | 0 | 4 | 750,000 | T | NA | - | Employer | 2.43 |
| Ozminkowski et al., 1999 | Worksite | CC | 22838 | 11644 | 1.5 | 1,879,522 | T | 8,888,730 | T | Employer | 4.56-4.73 |
| Brennan et al., 2012 | Worksite | CC | 5123 | 24124 | 0.5 | 200,000 | T | 830,000 | T | Health care provider | 3 |
| Maeng et al., 2013 | Worksite | CC | 4895 | 12077 | 5 | NA | - | 70 | PPPQ | Employer | 1.52 |

| | | | | | | | | | | | | | | | | | | | | | |
|---|---|-----|----|----|----|----|----|----|----|----|---|----|----|----|----|----|----|----|----|----|----|
| study justified and based on the study results? | | | | | | | | | | | | | | | | | | | | | |
| 16 | Was there a statement disclosing the source of funding for the study? | 3 | * | * | * | * | * | | | | | | | | | | | | | | |
| Total Points | | 100 | 32 | 41 | 66 | 27 | 56 | 61 | 19 | 44 | 7 | 32 | 30 | 36 | 23 | 49 | 24 | 22 | 32 | 20 | 24 |

*Study number was assigned to studies according to the alphabetical order of their article titles.

1 - Davis et al., 2009; 2 - Berg et al., 2004; 3 - Ozminkowski et al., 1999; 4 - Ackerman et al., 2006; 5 - Brennan et al., 2012; 6 - Maeng et al., 2013; 7 - Berg et al., 2002; 8 - Spence et al., 2014; 9 - Newton and Young., 2006; 10 - Moore et al., 2007; 11 - Milani and Lavie, 2009; 12 - Caloyeras et al., 2014; 13 - Serxner et al., 2012; 14 - Ramalho de Oliveira et al., 2010; 15 - Johannigman et al., 2010; 16 - Snyder et al., 2003; 17 - Henke et al., 2011; 18 - Fries et al., 1998; 19 - Su et al., 2016

4. Discussion

Based on this systematic review, we do not find sufficient evidence to support the claims of a positive ROI for chronic disease behavioral intervention programs in community, clinic, or workplace setting. There are no methodologically robust studies for any single type of program. The majority of the studies (84%) we reviewed were considered poor quality based on QHES Instrument. For the three studies passed cutoff score, weak study design and other shortcomings also called their results into question. The studies reviewed did not adhere to current standards for reporting economic evaluations, and due to their vulnerability to bias, they provide a very limited evidence base to guide decision-makers in other settings. These findings are concerning and reflect the need for future economic evaluations of the ROI for chronic disease behavioral interventions that better address the criteria outlined in the QHES.

Transparency in the economic analysis and cost components is another important aspect in economic evaluations we found lacking across studies reviewed. Unclear reporting of estimation of resource use and costs may make it hard to judge applicability in a different healthcare setting, which leads to low external validity. Lack of transparency also likely to cause reporting bias, which amplified by the lack of funding source and potential conflict of interest from the authors.

This review had its weaknesses. Although we made an effort to be as comprehensive as possible in our search, articles may have been missed by the single reviewer. We did not reach out to the authors of studies to obtain information that may have been available but not in the articles themselves. Length restrictions on manuscripts may have discouraged the inclusion of important study details. This review's strengths included the use of a structured, multidisciplinary

search and review process. This review also benefited from its consideration of both study design and execution.

A persistent barrier to developing a critical mass of high-quality research is the absence of standardization. Even within program categories, we found a lack of consistency in program elements, employee eligibility criteria, outcome, and the performance metrics used to assess impact. The consistency problem has been identified previously, but little effort has been made to rectify these issues. With the demand for chronic disease behavioral intervention increase, the need for high-quality evaluation may generate an incentive for producing high methodological research. However, for now, the evidence stakeholders need to make sound choices remains a critical missing part of the puzzle.

5. Conclusion

This paper demonstrates economic evaluations yield positive ROI on chronic disease behavioral interventions have limited methodological quality and their results should be interpreted with caution. The rationale behind chronic disease behavioral interventions is the improvement in risk factors and its associated conditions. However, the health outcome from such programs usually do not materialize for years or even decades into the future. Therefore, the evidence coming from methodologically poor studies may force decision-makers to make uninformed decisions on employ such programs and expect quick returns. It is important to improve the methodology and reporting of economic evaluations so that decisions based on ROI are informed by robust evidence of improvements in chronic conditions and in a reasonable time horizon. Future economic evaluations of chronic disease behavioral intervention should adhere to standard guidelines on economic evaluations.

Appendix A. Background

a. Chronic disease burden

Chronic diseases have emerged as a major challenge to health systems around the globe in the past decade. Over two-third of global deaths in 2010 can trace back to one of four chronic diseases: chorionic respiratory diseases, cardiovascular disease, cancer, and diabetes (Bauer et al., 2014). Global burden of disease is continually showing the shift from communicable disease to non-communicable disease since indicated the emergence of the problem and identified key risk factors that liable for the trends, including tobacco smoking, alcohol use, high blood pressure, high BMI, and unhealthy diets.

Risk factors, social and environmental factors, along with increased life expectancy are the three major components responsible for chronic disease epidemiology in the USA (NCHS, 2013). The Health system is burdened by the increasing number and rates of chronic disease that demands allocate a significant proportion of budget and resources to tackle the issue.

In the USA, key modifiable behavior risk factors including tobacco smoking, alcohol consumption, unhealthy diets, and physical inactivity primarily responsible for the burden of chronic disease (Woolf & Aron, 2013). Tobacco smoking is the most avoidable cause of chronic disease in the USA according to the Surgeon General's reports since 1964. The CDC (2009) estimate smoking or secondhand smoke exposure accounts for more than 440,000 premature death annually. Although tobacco control program can significantly reduce tobacco use epidemic and tobacco-related disease, smoking rates remained unimproved from 2003 to 2007 with more than 1/5 of American still smoke (CDC, 2009). Research shows that despite 70% of smokers in

the USA have the intention to quit smoke, yet most of the attempts are unaided and unsuccessful (Fiore et al., 2000).

Alcohol consumption is the third lifestyle-related cause of death, chronic disease and social problems in the USA. Although current Dietary Guidelines for Americans specifically outlines the recommended alcoholic beverage consumption amount per day and specified population should avoid drinking the alcoholic beverage, the survey shows about 30% of adult alcoholic beverage consumer report over drinking up to 4 times more than recommended amount in the past 30 days (Blazer & Wu, 2009). Despite the National Minimum Legal Drinking Act that enforce a minimum legal drinking age of 21 nationwide, about 45% of high school students report consuming alcohol in the past 30days, among which 60% reported binge drinking over 5 times more than recommended amount in the past 30 days (Naimi et al., 2007; Miller et al., 2007). Intervention for preventing heavy episodic drinking are often overlooked and associated with the stigma of addiction.

The risk of many chronic diseases, like cardiovascular disease and diabetes, can be reduced by the assistance of a balanced diet. Dietary Guidelines for Americans suggested 5 or more servings of fruits and vegetables per day per adults maintain a healthy diet. However, less than 1/4 of adults consume the recommended amount of fruits and vegetables per day (CDC, 2009). Another major dietary issue lies in saturated fat consumption, as more than 60% of children and teenage eat more than recommended saturated fat in their diet in the USA (CDC, 2009). Thus, increase their risk of diabetes, hyperlipidemia, high blood pressure, and eventually chronic illness.

Lastly, physical inactivity greatly increases the potential exposure to overweight, type 2 diabetes, cardiovascular disease, and metabolic syndrome (CDC, 2009). CDC Physical Activity

Guidelines for Americans recommends a minimum of 150 minutes of moderate-intensity aerobic activity each week, including running, brisk walking, play sports, and swimming. Regular physical activity can a low-cost measure a person can take to stay healthy. However, more than 60% of adults and 25% of teenager failed to meet the minimum recommendations according to a 2007 national survey. Also, a reduction in the participation rate in high school physical education classes and a spike of student spend more than 3 hours per day on computer or television has added to the problem of physical inactivity (CDC, 2009).

In the USA, the distribution of the burden of chronic disease follows a certain demographic and geographical pattern. Factors such as education level, incomes, races or ethnic backgrounds, geographical locations can cause disadvantages and vulnerability when affected by chronic diseases (Braveman et al., 2011). For example, studies reveal that Southeastern states have highest death rate caused by stroke, and low education or low-income population have the highest obesity rates (CDC, 2013). Although many important determining factors of health, like poverty or education, are not directly addressed by the health sector, the consequences of those social and economic factors can be mitigated by target interventions through reaching the population suffers the most from burden of disease (Mokdad et al., 2004).

Life expectancy at birth in the USA sees a steady increase since 1993 at 76.5 years for males, and 81.3 for the female in 2014 (WHO, 2014). Increased longevity directly links to the increased prevalence of many age-related chronic diseases, along with disability, and end-of-life issues. The great challenge on macroeconomic burden as well as healthcare resource allocation came with an aging society lead by Baby Boom generation (Knickman & Snell, 2002).

Despite the substantial improvement in some chronic conditions including smoking and uncontrolled hypertension. Many preventable chronic diseases with effective intervention

methods still dominant top-ranked risk factors and preventable causes of death in the USA (Bauer et al., 2014).

b. Chronic disease management program

Research suggests that risk factors for chronic disease can be adequately addressed by modifying health behaviors (Long et al., 2010). Early detection and better management within the health-care system can improve outcomes, slow disease progression, mitigate complications, and avert adverse outcomes. For people at high risk for chronic conditions, community-deployed and evidence-based disease management programs can help to improve the quality of life, reduce the number of health-care visits needed to maintain good health, and ultimately reduce health cost (Bauer et al., 2014).

The chronic disease management programs are active prevention and control interventions focused on the needs of certain participants or community with or at high risk of chronic disease. Chronic disease management program usually includes multidisciplinary collaboration including healthcare professionals, patient education, and patient self-management to improve health outcome and QALYs while reducing healthcare burden and cost (Bodenheimer et al., 2002). Interventions employed by chronic disease management program usually aim to assist patient's behavioral change in achieving desired health outcome. Chronic disease management can deploy in different patient groups with different chronic conditions and various health care setting.

Community, workplace and clinic based management programs can effectively address high burden chronic diseases including obesity, hypertension, and diabetes (Bauer et al., 2014). These management programs create an environment that supports and ensures that participants

have what is needed to manage chronic conditions outside clinical setting. Through interventions like risk assessment, physical activities, nutrition consulting, and health education to reduce risk factors, and ultimately prevent chronic diseases.

Economic evaluations of chronic disease management programs are a standard method undertaken to assess benefits, either potential or realized, for a given cost of program implementation. Cost-effectiveness analysis, a significant economic evaluation tool in the health sector, has been widely deployed to aid policy-makers for the decision-making process. The literature on the cost-effectiveness of chronic disease management program is extensive. Although many chronic disease management program shows a good value (defined as costing less than \$100,000 per Quality Adjusted Life Year), only two intervention - childhood immunizations and aspirin low dosage counseling for adults are widely regarded as cost-saving (Neumann & Cohen, 2009). Furthermore, cost-effectiveness does not necessarily translate to cost-saving. Cost-effectiveness analysis focuses on reducing risk factor that leads to increasing QALYs, while cost saving study focuses on the reduction of direct health cost and indirect health benefit. While a program is cost effective, costs to reduce risk factors, screening costs, and the cost of treatment when diseases are found may not be cost saving against the investment and maintenance cost for chronic disease management program.

c. Wellness program

Firms have become increasingly interested in preventing chronic illness and reducing associated costs over the last decade (Goetzel & Ozminkowski, 2008). The workplace is a unique environment where the employee spends a considerable amount of time. According to Gallup Consulting report (Gallup, 2016), an average adult employee working 34.4 hours a week with a

full-time adult employee working an average of 47 hours per week in the USA. Potential cost-saving and financial benefits for employers, employees as well as healthcare providers are the primary driver of this growing interest, which eventually prompts a whole industry around worksite healthcare cost reduction (Mattke et al., 2013).

The workplace wellness program is a chronic disease management program aim to promote healthy behavior and improve health outcomes in a workplace setting (Goetzel & Ozminkowski, 2008). Wellness program often employed health education, activities, weight management programs, health counseling, or nutrition advising to achieve behavioral change. Similar to the three level of disease prevention, wellness program interventions can be categorized into the primary, secondary and tertiary intervention based on their primary goal. Primary prevention aims to prevent chronic disease before it occurs. Primary intervention, such as exercise or healthy diet, helps encourage relatively healthy employees to maintain their healthy status and reduce the potential risk of chronic disease. Secondary prevention aims to reduce the impact of chronic disease at its early stage. Secondary intervention, such as tobacco use management or health risk assessment, eliminate risk factor and behavior to reduce the possibility that chronic disease that may occur. Tertiary prevention aims to soften the impact of existing chronic disease. Tertiary intervention, such as self-management, helps slow the progression of chronic disease, or to reduce symptoms by increasing medical adherence.

Firms often subsidize the cost for wellness program or provide a cash incentive for completion in the hope that such program will save money in the long term by improving employee health, productivity, and overall working spirit (Mattke et al., 2013). The Kaiser Family Foundation 2016 annual survey on employer-sponsored health plans shows a steady increase of interests in workplace chronic diseases management programs (also known as wellness program)

by employers since 2013 (Kaiser, 2016). The report shows that 83% of all organizations with 200 or more employees offer at least one wellness program to their employees, among which 32% have incentives in place to encourage employees to participate in or complete wellness program. The report also shows an increase in health risk assessment program provided in organizations with 200 or more employees, with a record high 59% implementation rate among large organizations surveyed in survey history. Similar trends can be seen in small organizations with less than 200 employees as well, a record high 46% of small organizations offered at least one wellness program (Kaiser, 2016). Compare to 2013 Kaiser survey, almost all indicators show a positive increase in program adoption in 2016 Kaiser survey, suggest an increasing adoption of workplace wellness program. It is also worth noting that average annual total premium for single employee and employer increased 2.2% from 2015, compare to the national average of 4% increase from 2015 before tax credit and 1.5% after (Kaiser, 2013 & Kaiser, 2016).

However, low participation in workplace wellness program may significantly limit the potential benefits and offset savings. The previous study on workplace wellness program participation indicates an average of 33% participation rate among 23 included wellness programs, with lowest at 10% and highest at 64% (Robroek et al., 2009). Robroek et al. (2009) suggest married employees show higher interest in participating workplace wellness program, and women show a higher participation rate in wellness program than men. Other studies identified full-time employees, secured contract employees, older in age, and small company employees display relatively high participation rate in workplace wellness program (Person et al., 2010).

Most organization prefer wellness programs over other approaches such as on-site health clinics, narrow network, or private exchanges. The reason for such preference is due to over 71 percent of all firms considered wellness program saves money per 2013 Kaiser survey (Kaiser,

2013). Firms believe in wellness program as a cost-saving chronic disease intervention is largely because of the heavily promoted concept of Return on Investment (ROI). Prior studies published in peer-reviewed journal suggested a high ROI from \$3.27 to \$5.82 across implemented wellness programs (Baicker et al., 2010). Similar high ROI can be found in disease management program as RAND Cooperation survey in 2013, which suggested an overall ROI of \$3.80. These results have been overwhelmingly adopted by gray literature such as promotional reports, news release, and policy brief regarding chronic disease management program to market their own products and programs.

d. Gray Literature

The term Gray literature is officially defined in 1997 as “produced at all levels of governmental, academics, business and industry in print and electronic formats, but which is not controlled by commercial publishers” at the Third International Conference on Grey Literature (Auger, 1998). Gray literature often not formally peer-reviewed and often can be found on the internet with open access. An example of gray literature includes articles from policy brief, government publications, promotion material, news article, unpublished data, research reports, and student thesis.

Gray literature often used by organizations to store and share information for quicker and wider distribution compared to the peer-reviewed journal and book publishing (Hopewell et al., 2007). An example of gray literature includes government publications like early findings for sponsored study; briefings for a status update or troubleshooting; white papers for policy evaluation and advocacy; and many other forms and reasons. Organizations can share it with relevant parties without delays, complications, and restrictions of academic publishing. In most

cases, these literatures are openly accessible without a charge, and with little to none incentive or reasons to publish in academic journals or books.

Gray literature is often considered as a less reliable, significant, and recognized format than articles in a peer-reviewed journal by the academic community (Hopewell et al., 2007). The quality and validity of gray literature varies by the creator, and highly dependent on the organizational reputation. The reason for such inconsistency is mainly caused by lack of an extensive and rigorous methodology and bibliographic control (Lawrence et al., 2014). Gray literature often lack basic information like name of the authors, affiliations, and publish dates. Also, lack of a professional layout and formats and non-conventional distribution make searching, citing and organizing gray literature complicate compared to peer-reviewed journal and books (Auger, 1998). Furthermore, gray literature often has a low print count and lack of a permanent Uniform Resource Locator (URL) or Digital Object Identifier (DOI), make it hard to access and hard to cite (Lawrence et al., 2014).

Gray literature starts to gain recognition in some academic fields as a good recourse to reduce public bias in research. Social science researchers show that unpublished studies often provide smaller experimental effects than published articles (Glass, 1981). A Recent review in health research also suggested that published trials provided a higher overall treatment effect than gray literature trails (Hopewell et al., 2007).

Gray literature that is promoting wellness program usually produced by parties who have interests in chronic disease management program, such as program provider, health insurance company, and government. Most gray literature cites cost saving results from published journals as tools to claim their program can yield similar cost-saving results. Most ROI number adapted by the gray literature advertising chronic disease management program can chase back to one of the

cooperation in-house evaluation such as Citibank, PepsiCo, or Johnson & Johnson Company (Lerner et al., 2013). Almost none of the gray literature conduct either an experimental or quasi-experimental research design to examine their claim of positive ROI.

Rigorous study design and financial impact evaluation can eliminate potential bias and error in economic evaluation. However, quality of economic evaluations of such programs varies. Some of the studies are conducted by firms providing such service and thus cast doubts on their credibility (Lerner et al., 2013). Furthermore, most of the studies consider mainly the short-term effects of the intervention like weight loss, which will not be sustained in many cases. Lerner et al. (2013) addressed in their most recent review of methodology quality in wellness program study, that low methodological quality of economic evaluations hinders the credibility of estimation about cost saving. This finding echoes with Warner et al. (1988) on study design and methodological quality inconsistencies of economic evaluation (Warner et al., 1988). Recent evidence from Dongen et al. (2011) further identified that study design affects the accuracy of reported financial outcomes, concluding that randomized control trials (RCTs) were more likely to render a negative financial return.

Appendix B. CHEERS Checklist

Studies were also scored against a modified version of Consolidated Health Economic Evaluation Reporting Standards (CHEERS), a guideline of methodological and essential elements to improve robustness of economic evaluations (Husereau et al., 2013). We modified original CHEERS checklist to include indicators that are most relevant for our research question (Table 4). The CHEERS checklist was mainly used to check the consistency of scoring each article due to the limitation of only one reviewer was employed during the process.

We employed the following scoring algorithm to provide a methodology quality score for each study. 21 indicators in modified CHEERS checklist were given equal weight. Each study was assessed against all 21 indicators. If the methodology of this study passes the criteria, the value of this indicator for this study will be given a 1, otherwise 0. If the study uses a biased design then the result was discounted no matter the score. The total methodology score of each individual study was calculated following the equation below:

$$MS_a = \sum_{i=1}^{21} Score_{a,i} / 21 \times 100$$

Where:

MS_a is the methodological score of study a according to CHEERS checklist,

$Score_{a,i}$ is the score of study a on indicator i (can only take the value of 0 or 1),

and 21 is the total number of indicators.

Table 4: Modified Consolidated Health Economic Evaluation Reporting Standards (CHEERS) Checklist

| | | |
|----------------------|---|--|
| Baseline | 1 | Describe characteristics of the base case population and subgroups analysed, including why they were chosen. |
| Setting and location | 2 | State relevant aspects of the system(s) in which the decision(s) need(s) to be made. |

| | | |
|-------------------------------|--------|---|
| Study perspective | 3 | Describe the perspective of the study and relate this to the costs being evaluated. |
| Comparators / alternatives | 4 | Describe the interventions or strategies being compared and state why they were chosen. |
| Time horizon | 5 | State the time horizon(s) over which costs and consequences are being evaluated and say why appropriate. |
| Discount rate | 6 | Report the choice of discount rate(s) used for costs and outcomes and say why appropriate. |
| Choice of outcomes | 7 | Describe what outcomes were used as the measure(s) of benefit in the evaluation and their relevance for the type of analysis performed. |
| Choice of Economic Evaluation | 8 | Was the choice of form of economic evaluation stated and justified in relation to the questions addressed? |
| Effectiveness | 9 | Describe fully the design features of the single effectiveness study and why the single study was a sufficient source of clinical effectiveness data. |
| Confidence intervals | 11 | Describe the details of statistical test(s) and confidence intervals given for ROI. |
| Price date | 12 | Report the dates of the estimated resource quantities and unit costs. Describe methods for adjusting estimated unit costs to the year of reported costs if necessary. |
| Explain Choice of model | 13 | Describe and give reasons for the specific type of decision-analytical model used. Providing a figure to show model structure is strongly recommended. |
| Assumptions | 14 | Describe all structural or other assumptions underpinning the decision-analytical model. |
| Analytical methods | 15 | Describe all analytical methods supporting the evaluation. This could include methods for dealing with skewed, missing, or censored data; extrapolation methods; methods for pooling data; approaches to validate or make adjustments (such as half cycle corrections) to a model; and methods for handling population heterogeneity and uncertainty. |
| Costs and outcomes | 16 | For each intervention, report mean values for the main categories of estimated costs and outcomes of interest, as well as mean differences between the comparator groups. If applicable, report incremental cost-effectiveness ratios. |
| Characterizing uncertainty | 17 (a) | <i>Single study-based economic evaluation:</i> Describe the effects of sampling uncertainty for the estimated incremental cost and incremental effectiveness parameters, together with the impact of methodological assumptions. |
| | 17 (b) | <i>Model-based economic evaluation:</i> Describe the effects on the results of uncertainty for all input parameters, and uncertainty related to the structure of the model and assumptions. |

| | | |
|------------------------------|----|--|
| Characterizing heterogeneity | 18 | If applicable, report differences in costs, outcomes, or cost-effectiveness that can be explained by variations between subgroups of patients with different baseline characteristics or other observed variability in effects that are not reducible by more information. |
| Limitation | 19 | Discuss limitations and the generalizability of the findings and how the findings fit with current knowledge. |
| Source of funding | 20 | Describe how the study was funded and the role of the funder in the identification, design, conduct, and reporting of the analysis. Describe other non-monetary sources of support. |
| Conflicts of interest | 21 | Describe any potential for conflict of interest of study contributors in accordance with journal policy. |

Rating the studies with the modified CHEERS checklist produced scores ranging from 13 to 72 with a mean of 34.5 (SD=14.6). Approximately 90% of the studies rated had scores less than 50 (n= 17), while only 10% had scores above or equal to 50 (n = 2) (Table 5).

Table 5: Modified Consolidated Health Economic Evaluation Reporting Standards (CHEERS) Checklist

| Score | Number of Studies | % |
|--------|-------------------|-----|
| 0-24 | 4 | 21 |
| 25-49 | 12 | 63 |
| 50-74 | 3 | 16 |
| 75-100 | 0 | 0 |
| Total | 19 | 100 |

Average score: 34.5; standard deviation: 14.6

Table 6 presents information regarding how frequently each CHEERS criterion was met by the 19 studies.

Table 6: Frequency of Each Criterion Met by ROI Studies in Chronic Disease Intervention Program (N= 19)

| Item # | Indicators | Frequency | % Studies |
|--------|------------|-----------|-----------|
| 1 | Baseline | 11 | 58 |

| | | | |
|----|-------------------------------|----|-----|
| 2 | Setting and location | 15 | 79 |
| 3 | Study perspective | 13 | 68 |
| 4 | Comparators / alternatives | 3 | 16 |
| 5 | Time horizon | 19 | 100 |
| 6 | Discount rate | 2 | 11 |
| 7 | Choice of outcomes | 19 | 100 |
| 8 | Choice of economic evaluation | 0 | 0 |
| 9 | Effectiveness | 9 | 47 |
| 10 | Confidence intervals | 1 | 5 |
| 11 | Price date | 0 | 0 |
| 12 | Explain Choice of model | 5 | 26 |
| 13 | Assumptions | 0 | 0 |
| 14 | Analytical methods | 4 | 21 |
| 15 | Study parameters | 1 | 5 |
| 16 | Incremental cost and outcomes | 1 | 5 |
| 17 | Characterizing uncertainty | 5 | 26 |
| 18 | Characterizing heterogeneity | 5 | 26 |
| 19 | Limitation | 9 | 47 |
| 20 | Source of funding | 5 | 26 |
| 21 | Conflicts of interest | 5 | 26 |

Linear regression was performed between CHEERS checklist score and QHES instrument score (Fig. 4). The results show a strong linear correlation ($r^2 = 0.756$, $p < 0.0001$), which suggests relative strong consistent over the two evaluation methods on methodology.

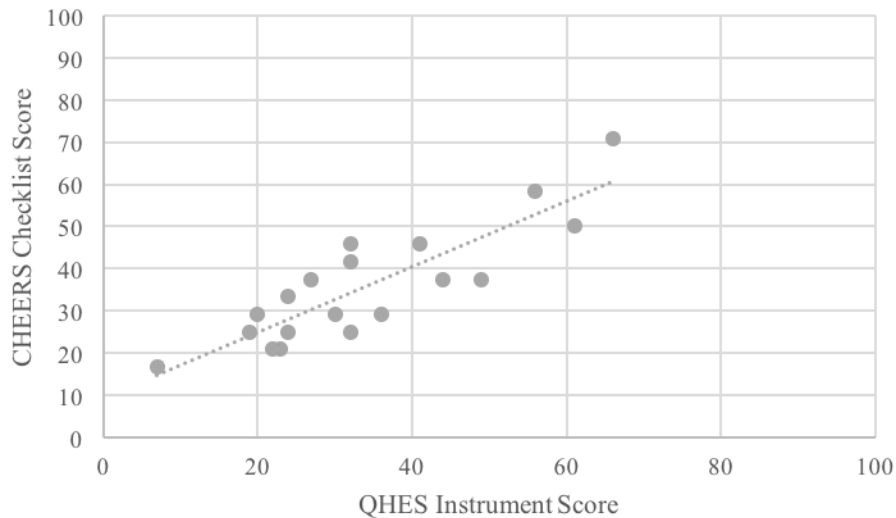


Figure 4: Linear Regression Plot of QHES Instrument Score and CHEERS Checklist Score.

a. Analysis of Each Criterion

a.1 Baseline & Setting and location

Lack of baseline lead to selection bias caused by imbalance between active and control groups. However, only half of the studies (58%) described characteristics of the base case population and subgroups analyzed, including why they were chosen. Although most of the studies (79%) described the study setting and location for study groups, selection bias and performance bias remains a threat since difference in setting and location may introduce systematic differences between groups in the intervention that was provided.

a.2 Study perspective

Most of the studies (68%) clearly stated their study perspective. Among these, seven reported the use of employer perspective, two reported health care perspective, one reported hospital perspective, one reported program purchaser perspective, and one reported private payer perspective. Lack of a clear study perspective cast doubt on the credibility of the study because cost-saving may vary with the perspective and an intervention which is cost-saving from one viewpoint and may not be from another.

a.3 Comparators / alternatives

Only three studies (16%) described any interventions or strategies (include status quo) being compared and stated why they were chosen. This is important for case-control and RCTs to control unintended exposure to interventions in the control group.

a.4 Discount rate & Price date

Only two studies (11%) reported discount rates used for costs and outcomes. Consider ten studies had time horizon that longer than 1 year, lack of discount rate raise concern over external validity and credibility of the economic evaluation. Furthermore, none of the study stated price date or adjusted cost to the year of report.

a.5 Choice of outcomes & Time horizon

All of the studies describe what outcomes were used as the measure of benefit in the evaluation. The commonly used outcome measures in studies were health care cost, hospitalize rate, absentness, and length of hospital stay. All study stated time horizon of cost and benefits, which is another fundamental aspect in study design and economic evaluation.

a.6 Choice of economic evaluation

None of the studies describe their choice of economic evaluation. This raises concern over reporting bias, that unfavorable findings might be unreported.

a.7 Measurement of effectiveness

Less than half (47%) of the study describe why the single study was a sufficient source for program effectiveness. This indicator is important for external validity as it explains why the results of the study can be generalized to other situations and to other people.

a.8 Confidence Interval

Only one study (5%) provides confidence intervals for ROI. However, in that study, the 95% confidence interval for ROI and outcome includes the null value, which means that there is no statistically meaningful difference between intervention and control groups (Maeng et al., 2013). Therefore, the lack of confidence intervals in studies cause concern over reporting bias, and the credibility of the positive ROI figure.

a.9 Assumptions & Analytical methods

None of the model-based studies described or explain any assumptions underpinning the decision-analytical model. Reporting analytic methods can help to assess the appropriateness of the methods and the corresponding results in economic evaluations (Husereau et al., 2013).

a.10 Incremental cost and outcomes

One study (5%) included an analysis of the full costs and benefits of intervention. Although most studies provided incremental outcomes, we found that many programs lack comprehensive program cost.

a.11 Characterizing uncertainty

Effects of uncertainties in parameters arising out of methodological assumptions, sampling variation or structure of analyses should be described in economic evaluations (Husereau et al., 2013). However, only five studies (26%) performed sensitivity analysis: four used one-way sensitivity analysis and one used two-way sensitivity analysis. All of studies in this review were single study-based economic evaluations and among these only four studies (21%) reported the regression methods and other statistical tests used in their analysis. Reporting of statistical analysis in those four studies was very limited.

a.12 Limitation

Less than half of the study (47%) discuss limitations and the generalizability of the findings. Four studies mentioned limitation in study design and protentional selection bias. One study acknowledged the low external validity and requested the result to be interpreted with caution (Maeng et al., 2013).

a.13 Funding & Conflict of interest

Funding and conflict of interest is critical to bias results that in favorite of sponsor or organization (Young, 2009). Five studies (26%) describe how the study was funded, and five studies (26%) described any potential for conflict of interest, or declare no conflict of interest. Out of all nineteen studies, eight studies were conducted by authors that affiliated with the program providers, two of which declared no conflicts of interest with no explanation.

Appendix C. Characteristics Summary of the Studies

1. A Comprehensive Worksite Wellness Program in Austin, Texas: Partnership Between Steps to a Healthier Austin and Capital Metropolitan Transportation Authority (Davis et al., 2009)

In this article, the authors examined the Capital Metro wellness program that carried out by Health & Lifestyles Corporate Wellness, Inc and CDC Steps to a Healthier Austin program. The program provides health information, one-on-one consulting, education, on-site fitness, healthier food options, workshops, smoking cessation, and cash incentives to all employees that willing to participate. Most of the services are free of charge; healthier food options and on-site fitness offered at a discount price. The pre-post study examined the annual health care cost for 300 participants over the four years' implementation period. The total cost of the four-year program was \$750,000, and savings from absenteeism avoidance was \$450,000. However, the authors did not show the amount of health care saving in this study. The authors calculated ROI as follow: health care and absenteeism savings divided by the cost of the program during the four years. The total ROI from 2004 to 2007 is \$2.43 according to the authors.

2. A Matched-Cohort Study of Health Services Utilization and Financial Outcomes for a Heart Failure Disease-Management Program in Elderly Patients (Berg et al., 2004)

In this article, the authors examined the McKesson Health Solutions clinic chronic disease intervention program that targets Medicare + Choice recipients residing in Ohio, Kentucky, and Indiana. The program provides risk stratification, education sessions, 24-hour telephone nurse counseling, action plans, individualized risk assessment, and medication compliance reminders. All the services are free of charge to the patients. The case-control study

examined the annual medical and pharmacy cost different between 533 self-selected participants and 236 matched non-participants. Total program cost for the one-year implementation was \$619,902, and savings from the program was calculated at \$1,430,281. The total ROI for the one-year implementation is \$2.31. The authors did not show the ROI calculation formula.

3. A Return on Investment Evaluation of the Citibank, N.A., Health Management Program (Ozminkowski et al., 1999)

In this article, the authors examined the Citibank N.A. wellness program that provided by Healthrac, Inc. The program provides education, health risk assessment and telephone counseling to self-selected participants completed a health risk appraisal survey. All the services are free of charge to the participant, with cash incentive provided to encourage program participation rate. The case-control study examined the net medical expenditures of 22,838 self-selected participants (54.3% participation rate) and 11,644 non-participants. Total program benefit and cost for the 3-year timeline was \$8,888,730 and \$1,879,522 at 0% discount rate; \$8,653,535 and \$1,864,657 at 4% discount rate; \$8,439,372 and \$1,850,893 at 8% discount rate. The total ROI for the 3-year implementation is \$4.73, \$4.64, and \$4.56 at 0%, 4%, and 8% discount rate. ROI was calculated as follow: benefits from the program divided by the cost of the program.

4. An evaluation of cost sharing to finance a diet and physical activity intervention to prevent diabetes (Ackerman et al., 2006)

In the article, the authors examined the ROI of cost sharing economic model for diet and physical activity intervention to prevent diabetes. The program, originally carried out by The Diabetes Prevention Program (DPP) Research Group was used to build a Markov simulation model to estimate the cost, QALYs, and cost per QALY gained. The DPP interventions include diet adjustment, and moderate intensity physical activities, group education session and one-on-

one case management to participants; placebo group received one time 30-min standard lifestyle recommendations. The economic model did not explain the number of intervention participants or placebo group, nor explore the baseline characteristic of each group. The author defined the direct medical cost of the hospital, emergency room, urgent care, and outpatient services, prescription medications cost, and the cost for telephone calls to healthcare providers. The model yielded a \$2,136 per person saving in direct medical cost vs. placebo intervention group. The authors did not disclose the program cost. The authors suggested that in single private healthcare payer situation results in a 0.58 ROI. However, the author also discussed the situation that in the real world, 1.00 ROI in three years could be achieved by letting Medicare co-pay 10% of the intervention costs for participants. No ROI calculation method was included in this study.

5. An Integrated Pharmacy-Based Program Improved Medication Prescription and Adherence Rates in Diabetes Patients (Brennan et al., 2012)

In this article, the authors examined the wellness program implanted to CVS retail pharmacies by CVS Caremark. The program provides health education, individual evaluation, adherence enforcement, and first-fill counseling to CVS employees and covered dependents. The case-control study consisted 5,123 self-selected participants and 24,124 matched non-employees control group based on sex, age, Ingenix pharmacy risk group score, preferred pharmacy, and medication possession ratio. Health care cost was examined as the outcome. The total labor and support costs of delivering the program were \$200,000 with \$830,000 total benefit (\$630,000 in adherence increase and \$200,000 in pharmacy costs reduction) in health care cost avoidance for the study population. The ROI for the 6-month implementation period was calculated at approximately \$3.00.

6. Can Health Insurance Improve Employee Health Outcome and Reduce Cost? (Maeng et al., 2013)

In this article, the authors examined the Geisinger Health System's employee health and wellness program provided by Geisinger Health Plan. The program's intervention includes health risk assessment, subsidized medications, self-management training, nutrition and physical activity, medication management, and acute exacerbation management. The company offers \$200 enrollment incentive payment to each participant, additional \$200 was offered to a participant for achieving 6-month and 1-year goal. In this case-control study, the authors examined the total medical cost of care, defined as the sum of all payments made by GHP to all providers plus member out-of-pocket expenses, between 4,895 self-selected participants and 12,077 matched non-participants and non-employees. The difference in total medical cost between intervention and control group for the five years are -\$70 per-member-per-quarter (1,634 and 1,703 respectively). No implementation or program cost was disclosed in this article. The author did perform sensitivity analysis on ROI, but, four out of five years were not significant at 5% level, and the total all five-year ROI was not significant at 5% level as well.

7. Diabetes Disease Management in a Community-Based Setting (Berg et al., 2002)

In this article, the authors examined the McKesson Health Solutions' community chronic disease intervention program for HMO and PPO health plan member. The program provides individualized monitoring, education, and counseling services to improve patient's self-management practices. In the pre-post study, change in medical service utilization rate, including inpatient visits, ER visits, MD visits, and facility visits, were examined for the 127 self-selected participants. The total cost for the one-year implementation was reported at \$49,429, with total

savings of \$214,486 from utilization changes. The authors reported ROI at \$4.34 with no information of ROI calculation.

8. Evaluation of an Outpatient Pharmacy Clinical Services Program on Adherence and Clinical Outcomes Among Patients with Diabetes and/or Coronary Artery Disease (Spence et al., 2014)

In this article, the authors examined the clinic chronic disease intervention program carried out by Kaiser Permanente Southern California for patients in KPSC Diabetes and/or CAD registry. The interventions include a face-to-face B-SMART consult, which consisted identify nonadherence, personalized solutions, motivated, adherence tool, and triaged the patient. The outcomes of the study included rates of disease-related non-elective inpatient admissions and ED visits. The case-controlled study consisted 1,480 participants in the intervention group and 1,477 matched participants in the usual care group. The cost of program training and implementation was reported at \$1,713,465 for the one year intervention period. The total cost saving through reduced hospitalizations and ED visits was reported at \$11,640,296 for the one year period, with majority saving came from reduced hospitalizations. The ROI was calculated as $(\text{total cost saving} - \text{total cost of the program}) / (\text{total cost of the program})$, yield a result of \$5.79 saving per dollar spending.

9. Financial Implications of Glycemic Control: Results of An Inpatient Diabetes Management Program (Newton et al., 2006)

In this article, the authors examined the Pitt County Memorial Hospital's clinic chronic disease management program provided by American Healthways, Inc and Pitt County Memorial Hospital. The program, targeting nurse, and physicians in the hospital involve education, guidelines, patient screening, individual patient consults, and daily interaction to achieve the

behavioral change of diabetes patients. Lengths-of-stay of diabetes patients was evaluated as the outcome. The economic model evaluation study did not disclose sample size, baseline characteristic of participants, or study design. The LOS aversion result averaged 5.11 days per patient, and transfer to \$2,224,029 savings for the year. Program implementation cost was not disclosed, but the authors mentioned that cost was calculated based on salaries and product services from American Healthways. The ROI was estimated at \$4.67 with no detail calculation shown.

10. Impact of a Patient-Centered Pharmacy Program and Intervention in a High-Risk Group (Moore et al., 2013)

In this article, the authors examined the clinic chronic disease management program carried out by a major pharmacy benefit plan to their national patient with at least 14 pharmacy claims in a 120-day period. The intervention, medication therapy management (MTM) program comprised risk assessment, telephone counseling, care plan, and follow-up telephone call. The case-control study consisted 2,250 self-selected participants (out of 8,723 eligible) and 2,250 matched non-participants from those eligible but decline participation. The study examined health care cost (total plan-paid medical and pharmacy costs), health care utilization (inpatient visits, physician visits, ER visits, and days supply), and condition-specific MPR (hypertension MPR, dyslipidemia MPR, diabetes MPR, depression MPR, and asthma MPR). ROI evaluation only considered the health care savings, which is \$977 per patient per year. The cost of the program was reported at \$478 per patient per year. The ROI was reported at \$2.0, no detail evaluation was reported.

11. Impact of Worksite Wellness Intervention on Cardiac Risk Factors and One-Year Health Care Costs (Milani and Lavie, 2009)

In this article, the authors examined a wellness program at a single employer in New Orleans area with two geographically disparate work locations carried out by Ochsner Health System to employees and covered dependents. The intervention including onsite health education, group smoking cessation, stress management, lipid clinic, hypertension and diabetes management, drug and alcohol addiction treatment, and fitness centers. The randomized control trial consisted 308 eligible employees, randomized by family units into the active intervention group (n = 185) and usual care group (n = 154). Total medical claim cost was examined as the outcome. Program cost was not provided in the article, the total medical cost difference between intervention and control are \$983 per subject per year (\$1,539 for the intervention group and \$ 2,522 for control group respectively). The ROI was reported at \$6.00 with no detail calculation and explanation.

12. Managing Manifest Diseases, But Not Health Risks, Saved PepsiCo Money Over Seven Years (Caloyeras et al., 2014)

In this article, the authors examined PepsiCo's wellness program provided by American Specialty Health, Inc to PepsiCo employees and their dependents. The program includes health risk assessment, on-site wellness event, lifestyle management, disease management, complex care management, 24/7 telephone counseling, individual counseling, and maternity management. The case-control study consisted 22,204 self-selected participants and 22,204 matched eligible non-participants. The study assessed health care cost as the outcome. Program cost, including vendors per participant per year fees, and the health risk assessment fee per completed survey, was not disclosed by authors. The saving in health care cost over the 7-years implementation period was reported at \$360 per person per year. The ROI was reported to be \$1.46.

13. Medical Cost Savings for Participants and Nonparticipants in Health Risk Assessments, Lifestyle Management, Disease Management, Depression Management, and Nurseline in a Large Financial Services Corporation (Serxner et al., 2012)

In this article, the authors examined a wellness program in a large financial services corporation to employees and Consolidated Omnibus Budget Reconciliation Act (COBRA) participants. The authors did not disclose the information of provider. The program includes health risk assessment, lifestyle management, and specific disease management. The case-control study consisted total 65,401 self-selected participants and 64,828 non-participants over the three-year implementation period. Gross medical cost savings was measured as the outcome. The total three-year HPM program fees were reported at \$12,207,000, and three-year total gross medical cost savings was reported at \$29,944,000. The three-year total ROI was reported at \$2.45 (0.59 for the first year, 3.33 for the second year, and 2.59 for the third year). The ROI was calculated as follow: gross medical cost savings/HPM program fees.

14. Medication Therapy Management: 10 Years of Experience in a Large Integrated Health Care System (Ramalho de Oliveira et al., 2010)

In this article, the authors examined a community chronic disease intervention by Fairview Pharmacy Service to Medicaid beneficiaries and their covered sponsors and private-pay patients. The program interventions are highly personalized ranging from drug utilization reviews, face-to-face pharmaceutical care services, health risk assessment, and chronic disease specific medical therapy. The pre-post study consisted total 9,068 participants from retrospective data over the ten-year timeline. The economic outcome was focused on total health care cost, including medical services avoided from the office visit, emergency room visits, urgent care visits, long-term care stays, and hospitalizations, avoidance of lost work time was also included.

Total program cost estimated at \$2,258,302, with a total saving estimated at \$2,913,850. The ROI was reported at \$1.29 with no detail formula disclosed.

15. Medication therapy management and condition care services in a community-based employer setting (Johannigman et al., 2010)

In this article, the authors examined a community chronic disease intervention by Blanchard Valley Health System to local company employees in the community setting. The interventions consist health risk assessment, education, personal counseling, risk factor screening, and condition care program. The pre-post study consisted 216 self-selected participants. The study only reported startup cost for wellness and condition care program, at less than \$10,000. No information on program management cost or total cost was disclosed. The total cost saving, including medication-related changes, lipid level change, HbA level change, weight loss, was estimated at \$1,011 per patient per year over the one year period. The return on investment was reported at \$2.21 with no formula or calculation included.

16. Quality Improvement and Cost Reduction Realized by a Purchaser Through Diabetes Disease Management (Snyder et al., 2003)

In this article, the authors examined the Clark County community disease intervention program carried out by Teachers Health Trust targeting active and retired educators. The programs included disease management, telephone counseling, group section, and health education. The pre-post study consisted 422 self-selected participants over the three-year period. The individual medical cost was evaluated as the outcome. The study reported a gross saving of \$1,402,071, with total program cost at \$415,533. The return of ROI was reported at \$3.37. No detail on program cost or ROI calculation was reported.

17. Recent Experience in Health Promotion at Johnson & Johnson: Lower Health Spending, Strong Return on Investment (Henke et al., 2011)

In this article, the authors examined the Johnson & Johnson Family of Companies' wellness program carried out by Johnson & Johnson Health and Wellness to US employees. The program consisted health assessment, physical activity, exercise expenditure reimbursement, seasonal fitness challenges, nutrition choices, weight watch program, online weight management, lifestyle management, computerized coaching and education programs, and chronic disease management. The medical and drug cost was evaluated as the outcome from 2002-08 health care cost data. The case-control study consisted 31,823 participants identified through a health assessment. The matched control group comprised 31,823 employees from a comparison company. Program savings were calculated by the estimated average adjusted annual costs per employee minus expected costs from comparison company. The program cost was estimated at \$144 per person per year, a more conservative cost estimate of \$300 per person per year was also employed for economic evaluation. Adjusted savings of \$565 according to 2009 dollars was reported, which yield an ROI of 3.92 based on \$144 PPPY cost, and 1.88 based on \$300 PPPY cost.

18. Reducing Need and Demand for Medical Services in High-risk Persons (Fries and McShane, 1998)

In this article, the author examined the community disease intervention program carried out by Healthrac, Inc targeting behavioral change in high-risk population. The intervention program includes routine health evaluation, report, and education based on individual health risk problems. Two intervention groups were examined in this case control study; one consisted all high-risk population identified by health assessment questionnaire (n = 2,586), the other consisted

comparison employees and seniors (n = 89,652). Both groups received same standard intervention; the program repeats every six months for standard program and three months for the high-risk population. The cost per participant is about \$30 per year for the standard program and about \$100 per year for the high-risk program. Total cost saving for the high-risk program was \$484 per person per 6 months; standard program was \$103 per person per 6 months. The ROI was reported at \$6.00 for the high-risk group and \$4:00 for standard groups without detail explaining.

19. Return on Investment for Digital Behavioral Counseling in Patients with Prediabetes and Cardiovascular Disease (Su et al., 2016)

In this article, the author examined the Prevent community disease management program carried out by Omada Health. The intervention mainly involved a digital behavioral counseling program. The economic model study examined 2,371 people enrolled in Prevent between 2012 and 2014. Markov-based microsimulation model was used to predict annual medical expenditures. The article reported break-even point at third year for both the population with prediabetes and the population with cardiovascular disease risk factors. The ROI was calculated as follow: $(\text{total return at or before current year}) / \text{total investment at or before current year}$. The authors reported ROI of 0.48 for year 1, 0.75 for year 2, and 1.07 for year 3. No detail on cost was disclosed.

Appendix D. Quantitative Methodology Analysis Based on Program Target Group

a. Community Chronic Disease Intervention Program

Six community disease management programs were examined in this review. One was a case-control studies with comparison group (Fries and McShane, 1998), three were pre-post studies (Berg et al., 2002; Ramalho de Oliveira et al., 2010; Snyder et al., 2003), and two were model based with no economic data collected as part of the intervention (Ackerman et al., 2006; Su et al., 2016).

Intervention participants in most studies were self-selected, which may introduce bias into the study sample. This bias is related to differences in the characteristics of participants who agree to participate in programs and those who decline. Only one study includes a non-randomized comparison group (Fries and McShane, 1998). However, the comparison group was not matched in characteristics to reduce the effects of selection bias.

Across the six studies, three studies examine medical service utilization avoided and associated cost reduction as outcome (Ackerman et al., 2006; Berg et al., 2002; Ramalho de Oliveira et al., 2010); three studies look in to medical care cost as outcome (Snyder et al., 2003; Fries and McShane, 1998; Su et al., 2016). All six studies did not report specific breakdown for the cost of the interventions.

The benefits of intervention program are well explained in most of the articles. Three studies consider savings in medical care cost as benefits, and all three studies measure the benefit based on pre-post difference (Snyder et al., 2003; Fries and McShane, 1998; Su et al., 2016);

however, Fries and McShane (1998) study has a non-matched comparison group. Three studies consider benefit as service utilization avoidance, including inpatient visits, ER visits, MD visits, facility visits reduction, and hospitalization (Ackerman et al., 2006; Berg et al., 2002; Ramalho de Oliveira et al., 2010). However, each of the three studies include different services with little to no standardized measurement of benefits.

Four of the six articles calculated ROI using the benefit (gross savings) divided by the cost; two articles do not disclose their ROI calculation method. The ROIs for the three diabetic studies average \$2.90 in savings per dollar spent on the program for the implementation period (\$1, \$3.37 and \$4.34, respectively). Three diabetic studies have different outcome variables and time horizons. Ackerman et al., (2006) examined life time direct medical cost as outcome and expected 1.00 ROI at the end of three-year period; Berg et al., (2002) examined medical service utilization reduction as outcome and calculated 4.34 ROI at the end of one-year timeline; and Snyder et al., (2003) examined medical cost as outcome and calculated 3.37 ROI for the 3-year implementation.

One cardiovascular studies reported \$1.07 return per dollar spent at year 3. Two general health risk studies reported average \$3.65 in savings per spent (\$1.29 and \$6 respectively). None of these six studies reported confidence interval on their ROI estimation analysis.

b. Clinic Chronic Disease Intervention Program

Four clinic chronic disease intervention programs were examined. Three were case-control studies with comparison group (Berg et al., 2004; Spence et al., 2014; Moore et al., 2013), and one was economic model analysis study (Newton et al., 2006).

Three case-control studies have self-selected participants in the intervention group with matched non-participants in comparison group. However, the economic model study did not disclose study design, sample size, and participant characteristics. Saving was generated using (Newton et al., 2006).

Each of four studies have different outcome metrics: medical and pharmacy costs (Berg et al., 2004); hospitalization and ED visits rate (Spence et al., 2014); length of hospital stay (Newton et al., 2006); and plan-paid health cost (Moore et al., 2013). The inconsistency in program outcome metric echoes the problem in standardized measures of outcomes. Three out of four studies lack specific breakdown of the cost of the interventions. Spence et al. (2014) studies provide information on salary and additional drug cost. Newton et al. (2016) studies do not mention cost at all.

Newton et al. (2016) does not disclose their ROI calculation method in their study; two articles calculated ROI using the benefit (gross savings) divided by the cost; one article calculated ROI using benefit subtract cost (net saving) then divided by the cost. The ROIs for the clinic chronic disease intervention program average at \$3.69 in savings per dollar spent on the program based on their highest estimate. Two studies (Berg et al., 2004; Moore et al., 2013) shown saving in health care cost in intervention group compared to control group, but neglected to mention that the saving is none significant at 5% level in economic evaluation.

d. Workplace Wellness Program

Nine workplace wellness programs were examined in this review. Six were case-control studies with the comparison group (Ozminkowski et al., 1999; Brennan et al., 2012; Maeng et al., 2013; Caloyeras et al., 2014; Serxner et al., 2012; Henke et al., 2011), two were pre-post studies

(Davis et al., 2009; Johannigman et al., 2010), and one was RCTs randomized at worksite level (Milani and Lavie, 2009). However, in the RCTs study, the intervention provider also performed the outcome assessment.

All programs offered at least one behavioral intervention to the employee. Two studies provide financial incentive for participating and completing wellness program (Davis et al., 2009; Maeng et al., 2013). Also, one program provided participants with specific medications at no cost (Milani and Lavie, 2009).

As we observed previously, there was little standardization in program outcome metric. All studies include one form of medical and pharmacy claims costs as outcome, but each define and measure program benefit in a different way. Two studies also include absenteeism in their outcome evaluation. One study includes potential cost saving from medication-related changes, including lipid level reduction, HbA level reduction, and weight loss into their outcome measurement. Three out of nine studies lack any information on program cost. None of the studies provided a detail breakdown of each component in the program cost.

Four articles do not disclose their ROI calculation method; five of the nine articles calculated ROI using the benefit (gross savings) divided by the cost. The ROIs for the nine wellness programs average \$3.96 in savings per dollar spent on the program based on their highest estimate. Range from lowest at \$1.96 (Caloyeras et al., 2014) to highest at \$6.00 (Milani and Lavie, 2009) in savings per spent. Maeng et al. (2013) performed sensitivity analysis on ROI, but, four out of five years were not significant at 5% level, and the total all five-year ROI was not significant at 5% level as well.

References

- Ackermann, R. T., Marrero, D. G., Hicks, K. A., Hoerger, T. J., Sorensen, S., Zhang, P., ... & Herman, W. H. (2006). An evaluation of cost sharing to finance a diet and physical activity intervention to prevent diabetes. *Diabetes care*, 29(6), 1237-1241.
- Auger, C. P. (1994). Information sources in grey literature. *Bowker-Saur*.
- Baicker, K., Cutler, D., & Song, Z. (2010). Workplace wellness programs can generate savings. *Health affairs*, 29(2), 304-311.
- Barlow, J., Wright, C., Sheasby, J., Turner, A., & Hainsworth, J. (2002). Self-management approaches for people with chronic conditions: a review. *Patient education and counseling*, 48(2), 177-187.
- Basu, R., Ory, M. G., Towne Jr, S. D., Smith, M. L., Hochhalter, A. K., & Ahn, S. (2015). Cost-effectiveness of the chronic disease self-management program: implications for community-based organizations. *Frontiers in public health*, 3.
- Bauer, U. E., Briss, P. A., Goodman, R. A., & Bowman, B. A. (2014). Prevention of chronic disease in the 21st century: elimination of the leading preventable causes of premature death and disability in the USA. *The Lancet*; London, 384(9937), 45–52.
[https://doi.org/http://dx.doi.org.proxy.lib.duke.edu/10.1016/S0140-6736\(14\)60648-6](https://doi.org/http://dx.doi.org.proxy.lib.duke.edu/10.1016/S0140-6736(14)60648-6)
- Baxter, Siyan, Kristy Sanderson, Alison J. Venn, C. Leigh Blizzard, and Andrew J. Palmer. "The relationship between return on investment and quality of study methodology in workplace health promotion programs." *American Journal of Health Promotion* 28, no. 6 (2014): 347-363.
- Berg, G. D., & Wadhwa, S. (2002). Diabetes disease management in a community-based setting. *Managed Care* (Langhorne, Pa.), 11(6), 42-45.
- Berg, G. D., Wadhwa, S., & Johnson, A. E. (2004). A Matched-Cohort Study of Health Services Utilization and Financial Outcomes for a Heart Failure Disease-Management Program in Elderly Patients. *Journal of the American Geriatrics Society*, 52(10), 1655-1661.

- Blazer, D. G., & Wu, L. T. (2009). The epidemiology of at-risk and binge drinking among middle-aged and elderly community adults: National Survey on Drug Use and Health. *American Journal of Psychiatry*, 166(10), 1162-1169.
- Bodenheimer, T., Lorig, K., Holman, H., & Grumbach, K. (2002). Patient self-management of chronic disease in primary care. *Jama*, 288(19), 2469-2475.
- Braveman, P. A., Kumanyika, S., Fielding, J., LaVeist, T., Borrell, L. N., Manderscheid, R., & Troutman, A. (2011). Health disparities and health equity: the issue is justice. *American Journal of Public Health*, 101(S1), S149-S155.
- Brennan, T. A., Dollear, T. J., Hu, M., Matlin, O. S., Shrank, W. H., Choudhry, N. K., & Grambley, W. (2012). An integrated pharmacy-based program improved medication prescription and adherence rates in diabetes patients. *Health Affairs*, 31(1), 120-129.
- Caloyeras, J. P., Liu, H., Exum, E., Broderick, M., & Mattke, S. (2014). Managing manifest diseases, but not health risks, saved PepsiCo money over seven years. *Health Affairs*, 33(1), 124-131.
- Cancelliere, C., Cassidy, J. D., Ammendolia, C., & Côté, P. (2011). Are workplace health promotion programs effective at improving presenteeism in workers? A systematic review and best evidence synthesis of the literature. *BMC public health*, 11(1), 395.
- Carnethon, M., Whitsel, L. P., Franklin, B. A., Kris-Etherton, P., Milani, R., Pratt, C. A., & Wagner, G. R. (2009). Worksite wellness programs for cardiovascular disease prevention. *Circulation*, 120(17), 1725-1741.
- Centers for Disease Control and Prevention, & Centers for Disease Control and Prevention. (2013). CDC health disparities and inequalities report – United States 2013. *MMWR* 2013;62.
- Centers for Disease Control and Prevention. (2009). The power of prevention: Chronic disease... the public health challenge of the 21st century. *US Department of Health and Human Services*.
- Centers for Medicare and Medicaid Services. (2012). Chronic conditions among Medicare beneficiaries, chartbook, 2012 edition. Baltimore, MD, 2012.

- Davis, L., Loyo, K., Glowka, A., Schwertfeger, R., Danielson, L., & Brea, C. (2009). A comprehensive worksite wellness program in Austin, Texas: Steps to a Healthier Austin and Capital Metropolitan Transportation Authority. *Prev Chronic Dis*, 6(2).
- Fiore, M. C., Bailey, W. C., Cohen, S. J., Dorfman, S. F., Goldstein, M. G., Gritz, E. R., ... & Mecklenburg, R. E. (2000). Treating tobacco use and dependence: clinical practice guideline. Rockville, MD: *US Department of Health and Human Services*, 00-0032.
- Fries, J. F., & McShane, D. (1998). Reducing need and demand for medical services in high-risk persons. A health education approach. *Western Journal of Medicine*, 169(4), 201.
- Gallup. (2016). Work and Workplace. Retrieved March 24, 2017, from <http://www.gallup.com/poll/1720/Work-Work-Place.aspx>
- Gerteis, J., Izrael, D., Deitz, D., LeRoy, L., Ricciardi, R., Miller, T., & Basu, J. (2014). Multiple chronic conditions chartbook. Rockville, MD: *Agency for Healthcare Research and Quality*.
- Goetzel, R. Z., & Ozminkowski, R. J. (2008). The health and cost benefits of work site health-promotion programs. *Annu. Rev. Public Health*, 29, 303-323.
- Henke, R. M., Goetzel, R. Z., McHugh, J., & Isaac, F. (2011). Recent experience in health promotion at Johnson & Johnson: lower health spending, strong return on investment. *Health Affairs*, 30(3), 490-499.
- Hopewell, S., McDonald, S., Clarke, M., & Egger, M. (2007). Grey literature in meta-analyses of randomized trials of health care interventions. *Cochrane Database Syst Rev*, 2(2).
- Husereau, D., Drummond, M., Petrou, S., Carswell, C., Moher, D., Greenberg, D., ... & Loder, E. (2013). Consolidated health economic evaluation reporting standards (CHEERS) statement. *Cost Effectiveness and Resource Allocation*, 11(1), 6.
- Johannigman, M. J., Leifheit, M., Bellman, N., Pierce, T., Marriott, A., & Bishop, C. (2010). PRACTICE REPORT. Medication therapy management and condition care services in a community-based employer setting. *American Journal of Health-System Pharmacy*, 67(16).

- Kaiser Family Foundation (2013). 2013 Employer Health Benefits Survey. *Kaiser Family Foundation Annual Survey*.
- Kaiser Family Foundation (2016). 2013 Employer Health Benefits Survey. *Kaiser Family Foundation Annual Survey*.
- Knickman, J. R., & Snell, E. K. (2002). The 2030 problem: caring for aging baby boomers. *Health services research, 37*(4), 849-884.
- Lawrence, A., Houghton, J., Thomas, J., & Weldon, P. R. (2014). Where is the evidence? Realising the value of grey literature for public policy & practice: a discussion paper.
- Lerner, D., Rodday, A. M., Cohen, J. T., & Rogers, W. H. (2013). A systematic review of the evidence concerning the economic impact of employee-focused health promotion and wellness programs. *Journal of Occupational and Environmental Medicine, 55*(2), 209-222.
- Lorig, K. R., Sobel, D. S., Ritter, P. L., Laurent, D., & Hobbs, M. (2000). Effect of a self-management program on patients with chronic disease. *Effective clinical practice: ECP, 4*(6), 256-262.
- Maeng, D. D., Pitcavage, J. M., Tomcavage, J., & Steinhubl, S. R. (2013). Can health insurance improve employee health outcome and reduce cost? An evaluation of Geisinger's employee health and wellness program. *Journal of occupational and environmental medicine, 55*(11), 1271-1275.
- Mattke, S., Liu, H., Caloyeras, J., Huang, C. Y., Van Busum, K. R., Khodyakov, D., & Shier, V. (2013). Workplace wellness programs study. *Rand health quarterly, 3*(2).
- Mattke, S., Mengistu, T., Klautzer, L., Sloss, E. M., & Brook, R. H. (2015). Improving care for chronic conditions: current practices and future trends in health plan programs. *Rand Health Quarterly, 5*(2).
- Milani, R. V., & Lavie, C. J. (2009). Impact of worksite wellness intervention on cardiac risk factors and one-year health care costs. *The American journal of cardiology, 104*(10), 1389-1392.

- Miller, J. W., Naimi, T. S., Brewer, R. D., & Jones, S. E. (2007). Binge drinking and associated health risk behaviors among high school students. *Pediatrics*, 119(1), 76-85.
- Mokdad, A. H., Marks, J. S., Stroup, D. F., & Gerberding, J. L. (2004). Actual causes of death in the United States, 2000. *Jama*, 291(10), 1238-1245.
- Moore, J. M., Shartle, D., Faudskar, L., Matlin, O. S., & Brennan, T. A. (2013). Impact of a patient-centered pharmacy program and intervention in a high-risk group. *Journal of Managed Care Pharmacy*, 19(3), 228-236.
- Murray, C. J., Vos, T., Lozano, R., Naghavi, M., Flaxman, A. D., Michaud, C., ... & Aboyans, V. (2013). Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. *The lancet*, 380(9859), 2197-2223.
- Naimi, T. S., Brewer, R. D., Miller, J. W., Okoro, C., & Mehrotra, C. (2007). What do binge drinkers drink?: Implications for alcohol control policy. *American journal of preventive medicine*, 33(3), 188-193.
- National Center for Health Statistics (US. (2013). Health, United States, 2012: With special feature on emergency care.
- National Center for Health Statistics, Centers for Disease Control, & Preventi (Eds.). (2015). Health, United States, 2013, with special feature on prescription drugs. Government Printing Office.
- NBGH. Fidelity Benefits Consulting & National Business Group on Health. Moving from Wellness to Wellbeing: Seventh Annual Employer-Sponsored Health and Wellbeing Survey. National Business Group on Health. 2016. Available at: <http://www.businessgrouphealth.org/pub/a0d30d2b-782b-cb6e-2763-7380d7691b8a>
- Neumann, P. J., & Cohen, J. T. (2009). Cost savings and cost-effectiveness of clinical preventive care. The Synthesis project. Research synthesis report, (18), 189-189.
- Neumann, P. J., Cohen, J. T., & Weinstein, M. C. (2014). Updating cost-effectiveness—the curious resilience of the \$50,000-per-QALY threshold. *New England Journal of Medicine*, 371(9), 796-797.

Newton, C. A., & Young, S. (2006). Financial implications of glycemic control: results of an inpatient diabetes management program. *Endocrine Practice*, 12(Supplement 3), 43-48.

Ofman, J. J., Sullivan, S. D., Neumann, P. J., Chiou, C. F., Henning, J. M., Wade, S. W., & Hay, J. W. (2003). Examining the value and quality of health economic analyses: implications of utilizing the QHES. *Journal of Managed Care Pharmacy*, 9(1), 53-61.

Online survey includes responses from 129 organizations. Survey was fielded from November 2015 through December 2015 among National Business Group on Health members and clients of Fidelity Investments.

Doi:<https://www.businessgrouphealth.org/pressroom/pressRelease.cfm?ID=276>

Ozminkowski, R. J., Dunn, R. L., Goetzel, R. Z., Cantor, R. I., Murnane, J., & Harrison, M. (1999). A return on investment evaluation of the Citibank, NA, Health Management Program. *American Journal of Health Promotion*, 14(1), 31-43.

Page, P. (2014). Beyond statistical significance: clinical interpretation of rehabilitation research literature. *International journal of sports physical therapy*, 9(5), 726.

Pelletier, K. R. (2009). A review and analysis of the clinical and cost-effectiveness studies of comprehensive health promotion and disease management programs at the worksite: update VII 2004–2008. *Journal of Occupational and Environmental Medicine*, 51(7), 822-837.

Person, A. L., Colby, S. E., Bulova, J. A., & Eubanks, J. W. (2010). Barriers to participation in a worksite wellness program. *Nutrition research and practice*, 4(2), 149-154.

Ramalho de Oliveira, D., Brummel, A. R., & Miller, D. B. (2010). Medication therapy management: 10 years of experience in a large integrated health care system. *Journal of Managed Care Pharmacy*, 16(3), 185-195.

Robroek, S. J., Van Lenthe, F. J., Van Empelen, P., & Burdorf, A. (2009). Determinants of participation in worksite health promotion programmes: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 6(1), 26.

Schiller, J. S., Lucas, J. W., Ward, B. W., & Peregoy, J. A. (2012). Summary health statistics for US Adults: National health interview survey, 2010. *Vital and Health Statistics. Series 10, Data from The National Health Survey*, (252), 1-207.

- Serxner, S., Alberti, A., & Weinberger, S. (2012). Medical cost savings for participants and nonparticipants in health risk assessments, lifestyle management, disease management, depression management, and nurseline in a large financial services corporation. *American Journal of Health Promotion*, 26(4), 245-252.
- Snyder, J. W., Malaskovitz, J., Griego, J., Persson, J., & Flatt, K. (2003). Quality improvement and cost reduction realized by a purchaser through diabetes disease management. *Disease Management*, 6(4), 233-241.
- Spence, M. M., Makarem, A. F., Reyes, S. L., Rosa, L. L., Nguyen, C., Oyekan, E. A., & Kiyohara, A. T. (2014). Evaluation of an outpatient pharmacy clinical services program on adherence and clinical outcomes among patients with diabetes and/or coronary artery disease. *Journal of Managed Care Pharmacy*, 20(10), 1036-1045.
- Spiegel, B. M., Targownik, L. E., Kanwal, F., DeRosa, V., Dulai, G. S., Gralnek, I. M., & Chiou, C. F. (2004). The quality of published health economic analyses in digestive diseases: a systematic review and quantitative appraisal. *Gastroenterology*, 127(2), 403-411.
- Su, W., Chen, F., Dall, T. M., Iacobucci, W., & Perreault, L. (2016). Peer Reviewed: Return on Investment for Digital Behavioral Counseling in Patients With Prediabetes and Cardiovascular Disease. *Preventing chronic disease*, 13.
- Su, W., Chen, F., Dall, T. M., Iacobucci, W., & Perreault, L. (2016). Peer Reviewed: Return on Investment for Digital Behavioral Counseling in Patients With Prediabetes and Cardiovascular Disease. *Preventing chronic disease*, 13.
- Thabane, L., Mbuagbaw, L., Zhang, S., Samaan, Z., Marcucci, M., Ye, C., ... & Debono, V. B. (2013). A tutorial on sensitivity analyses in clinical trials: the what, why, when and how. *BMC medical research methodology*, 13(1), 92.
- van Dongen JM, Proper KI, van Wier MF, et al. Systematic review on the financial return of worksite health promotion programmes aimed at improving nutrition and/or increasing physical activity. *Obes Rev*. 2011;12:1031–1049.
- Vickery, C. G., Nyberg, G., & Whiteaker, D. (Eds.). (2015). *Modern Clinic Design: Strategies for an Era of Change*. John Wiley & Sons.

- Wamer, K. E., Wickizer, T. M., Wolfe, R. A., Schildroth, J. E., & Samuelson, M. H. (1988). Economic implications of workplace health promotion programs: review of the literature. *Journal of Occupational and Environmental Medicine*, 30(2), 106-112.
- Weingarten, S. R., Henning, J. M., Badamgarav, E., Knight, K., Hasselblad, V., Gano Jr, A., & Ofman, J. J. (2002). Interventions used in disease management programmes for patients with chronic illness which ones work? Meta-analysis of published reports. *Bmj*, 325(7370), 925.
- Weingarten, S. R., Henning, J. M., Badamgarav, E., Knight, K., Hasselblad, V., Gano Jr, A., & Ofman, J. J. (2002). Interventions used in disease management programmes for patients with chronic illness which ones work? Meta-analysis of published reports. *Bmj*, 325(7370), 925.
- WHO, Life expectancy at birth, total (years) | Data. (2014.). Retrieved March 23, 2017, from <http://data.worldbank.org/indicator/SP.DYN.LE00.IN?locations=US&view=chart>
- Woolf, S. H., & Aron, L. Y. (2013). The US health disadvantage relative to other high-income countries: findings from a National Research Council/Institute of Medicine report. *Jama*, 309(8), 771-772.
- Young, S. N. (2009). Bias in the research literature and conflict of interest: an issue for publishers, editors, reviewers and authors, and it is not just about the money. *J Psychiatry Neurosci*, 34(6), 412-7.