ACTEX Professional Series

Managing & Evaluating Healthcare Intervention Programs

Second Edition

Ian Duncan, FSA, FIA, FCIA, MAAA

Requests for permission should be addressed to

ACTEX Learning
PO Box 715
New Hartford CT 06057

Manufactured in the United States of America

Cover design by Jeff Melaragno

ISBN: 978-1-62542-816-5

CONTENTS

PART II
ACTUARIAL ISSUES IN CARE MANAGEMENT INTERVENTIONS

PART III
PRACTICAL APPLICATIONS

PART IV
WELLNESS, INTEGRATED PROGRAMS AND OTHER NEW DEVELOPMENTS IN CARE MANAGEMENT PROGRAMS

INTRODUCTION TO THE SECOND EDITION
BY
PROF. JONATHAN GRUBER,
MASSACHUSETTS INSTITUTE OF TECHNOLOGY

The year 2014 will usher the U.S. into a new era of health insurance markets. The Patient Protection and Affordable Care Act (ACA) will impose the most significant national changes in the regulation and pricing of health insurance in our nation's history. This is an incredible moment for research on the health insurance market. Yet the two major groups that undertake such research, economists and actuaries, continue to conduct their research in largely separate realms. This is an unfortunate state of affairs that misses critical opportunities for advancing our understanding of health insurance markets.

Economic research on health insurance markets tends to focus on estimating the primitives of economic and firm behavior. Economists are focused on questions such as the price sensitivity of employer insurance offering or the price sensitivity of individual enrollment in insurance and the impact of benefits design on health care spending. The centerpiece of such research efforts is the search for useful treatment/control groups that we can use to identify such behavioral responses. This has led in recent years to the focus on "natural experiments" whereby insurance prices or benefits vary across individuals in quasi-random ways that allow us to form treatment/control groups that are similar except for these prices or benefits. At the same time economists largely ignore underlying structural factors that determine insurance pricing, as well as the risk composition of population movements.

Actuaries, on the other hand (and if I may be so bold!), tend to focus on developing models that allow them to predict how changes in insurance market structure will affect pricing in insurance markets. This analysis takes underlying behavioral parameters as given and largely determined by others, or infers them from data in a less rigorous statistical framework than do economists. Instead, the major innovation in research is in developing sophisticated predictive models of how changes in regulation, for example, will impact the mix of risks in a pool and how that translates to changes in underlying costs, and thereby pricing.

The potential complementarities of these research approaches are clear. Economists can provide the primitives on population movement that focus on both totals and on risk composition, within a rigorous empirical framework. Actuaries can then incorporate those estimates into predictive models that can be used to most effectively forecast the impact of changes in insurance markets on pricing. But such a coordinated effort will require professionals in both disciplines to communicate to and not past each other.

That is where Ian Duncan's fantastic book comes in. This book is an impressive tour of substantive actuarial research on a wide variety of critical topics. Ian raises and answers a wide variety of questions about how to think actuarially about healthcare interventions. The book positions

itself nicely in the space between economics and actuarial science, using cutting edge empirical methodologies to answer the underlying research questions that matter quite a lot for actuaries. As such, it can hopefully serve as a first step towards bridging this important gap between the disciplines. I hope that both economists and actuaries can use this book as a starting point for thinking about healthcare invention programs, and that in doing so we can bring the fields closer together.

INTRODUCTION TO THE FIRST EDITION
BY
Howard Bolnick, FSA, MAAA, Hon. FIA

Health actuarial practice has been a growing and dynamic part of the profession for many decades. When I began working in the area in the 1970s, indemnity insurance was the only game in town: data were quite limited, and actuarial tools were basic. All this changed with the advent and growth of managed care. There were new types of insurance arrangements that required more robust data, and, new actuarial tools were needed to successfully manage a growing variety of managed care plans, each with its own distinct characteristics and its own distinct health actuarial needs. The profession successfully responded to these changes. Health actuaries broadened their scope of practice and developed the tools and analyses needed to successfully support this new environment. Health actuaries remained leaders in a changing environment by using our background, unique skills, and creativity to become a key resource in the success of managed care plans.

Throughout its existence, managed care has continued to evolve. Older prescriptive approaches to care management are being replaced with newer supportive approaches to population health management such as disease management programs and wellness programs. As these new types of care management evolve, they clearly pose a new challenge to health actuaries. To sustain our leadership role, health actuarial practice will need to continue to broaden its scope and create new tools to support the changing environment.

Ian Duncan's new book, *Managing and Evaluating Healthcare Intervention Programs,* continues this tradition of actuaries responding to a changing environment. This book is a thoughtful, well written, and well-researched study that provides actuaries, senior managers, financial managers, and others interested in the topic with a wealth of information, careful analyses, and a strong intellectual basis for expanding actuarial and financial leadership to population health management.

ACKNOWLEDGMENTS

SECOND EDITION

The second edition has benefited significantly from review and comments by Henry Dove, PhD, Yale University, Division of Health Policy and Administration; Iver Juster, MD, ActiveHealth Management; and Steve Siegel, ASA, MAAA, Society of Actuaries.

For this second edition I have added a number of new chapters and completely revised others.

- I have added an early chapter on the clinical foundations of care management. This chapter covers some of the common chronic conditions that are the subject of care management programs, as well as their treatments. This chapter was developed in response to the increased clinical focus of care management programs since the publication of the first edition.

- Chapter 3 (Care Management Programs and Interventions) has been significantly revised to take into account the new intervention models developed since the first edition.

- Chapter 4 of the prior edition covered an exhaustive discussion of the (then) evidence in peer-reviewed research of program impact on utilization and savings. I have updated this chapter (now Chapter 5) with reviews of the outcomes for the new programs that are covered in Chapter 3, and selective outcomes from the prior programs in the first edition.

- The chapter on care management program design (Formerly Chapter 5) had an appendix about care management quality measurement. Recent developments (such as pay-for-performance, Meaningful Use and Accountable Care Organizations) have placed more focus on the management and measurement of quality, and this expanded discussion is now part of Chapter 7.

- Understanding the economics of care management programs (formerly Chapter 6) was at the heart of the message of the prior edition. This remains true, but to reflect our recent work in what we call "opportunity analysis," or the identification of members for management who are not merely high risk but who also represent an opportunity for reduction in utilization, we have added Chapter 9 ("Opportunity Analysis").

- The prior edition of this book discussed the use of risk adjustment to create equivalent comparison groups. We have expanded the range of discussion with the addition of propensity score matching (a technique that is prevalent in health systems research) in a new Chapter 11.

- As in the prior edition, I cover a number of the issues involved in comparing populations: risk adjustment, exclusions, and the size of population necessary to draw conclusions about an intervention or program.
- Finally, the previous edition covered wellness and other population-focused programs. The literature about these programs is vast, and I have selectively updated the text to accommodate some of the recent literature.

Contributors and Research Assistants

I am grateful to the following individuals for their contributions to specific chapters:

Chapter 2: Tricia Nguyen, MD, MBA, President, Texas Health Population Health, Education & Innovation Center, Dallas TX.
Marie Hatam, MD, MBA, Bryan University, Phoenix AZ
Mark Matusik, PharmD. Walgreens Co. Deerfield IL

Chapter 3: Jeff Davis MBA, Walgreens Co. Franklin TN
Michael Einodshofer, RPh, MBA, Walgreens Specialty Pharmacy Services, Pittsburgh PA
Avinash Patwardhan, MD, Walgreens Co. Deerfield IL.

Chapter 5: Susan Furlong, MA, Health Educator, Phoenix AZ.

Chapter 6: Carl F. Husa Jr., MA, Bryan University, Phoenix AZ.

Chapter 9: Geraint Lewis, MB, BChir, FRCP, FFPH, Chief Data Officer, NHS Commissioning Board, UK.

Chapter 11: Quijuan Li, MPH. SCIO Health Analytics, West Hartford, CT.

FIRST EDITION

The first edition of this book grew out of a series of papers that was developed with grants from the Society of Actuaries Health Section and the Society of Actuaries Council for Knowledge Extension and Research, together with additions from other actuaries and researchers in the field of measurement of chronic care and other care programs. I am grateful for the peer review and valuable input on the original papers by the project oversight group. The project oversight group was:

Bryan Miller, FSA, MAAA, Blue Cross Blue Shield of Kansas City (Chairman)
John Cookson, FSA, MAAA, Milliman USA
Stacey Lampkin, FSA, MAAA, Anthem, Inc.
John Stark, FSA, MAAA, Anthem Inc.
Margie Rosenberg, PhD, FSA, MAAA, University of Wisconsin-Madison
Ronora Stryker, ASA, MAAA, Society of Actuaries
Stephen Siegel, ASA, MAAA, Society of Actuaries

The valuable assistance, support, and input from the Society of Actuaries Health Section (Karl Volkmar, FSA, MAAA, Chairman) and the Committee on Knowledge Extension Research (Curtis Huntington, JD, FSA, MAAA, Chairman) is also acknowledged.

Readers familiar with the Society of Actuaries papers will recognize that they form the basis for Chapters 2 through 12. A number of chapters are, however, entirely new material. We have added new material on a growing, related area, namely, wellness and employer worksite health programs in Chapters 17 and 18.

We acknowledge also the contributions of Munich Re Health Care, Highmark Inc., McKesson Corporation and Nationwide Better Health Inc., without whose support some of the analytical work would not have been completed.

Gail Hall, FSA, MAAA of ACTEX Publications should be commended for taking a risk on an unknown author and a topic that is still outside the actuarial mainstream. I am grateful to Marilyn Baleshiski for her help in editing and laying out the manuscript.

Contributors and Research Assistants

Chapters 2-4: Henry Dove, PhD, Yale University, Division of Health Policy and Administration

Chapter 4: Koren Odierna, Yale University, Division of Health Policy Research and Research Administration
Assistant: Elizabeth Heckinger, UCLA

Chapter 9: Mita Lodh, MBA, PhD, FSA, MAAA Schrammraleigh Health Strategy
Gregory Berg, PhD, and ..David Mattingly, BA, McKesson Health Solutions.

Chapter 10: Thomas Messer PhD, ASA, MAAA, Solucia Inc.
J. Ramsay Farah, MD, MPH, FAAP, FACPM, FASAM, CPE, CMRO; Kyahn Kamali, MHS, CPHIMS
Jeffrey Harner, MS, Nationwide Better Health Inc.

Chapter 11: Robert Bachler, FSA, FCAS, MAAA, Munich Re Healthcare
Iver Juster, MD, ActiveHealth Management Inc.

Chapter 12: Rebecca Owen, FSA, MAAA, Solucia Inc.
Henry Dove, PhD Yale University, Division of Health Policy and Administration

Chapters 13-14: Sheryl Coughlin, PhD MHA, B. App. Sci., Solucia Inc.

Chapter 15: Christian Birkmeyer, MS, Solucia Inc.
Manjula Singh, MA, PhD, Solucia Inc.
Appendix: Vincent Kane, FSA, MAAA, DxCG Inc.
Anju Joglekar, PhD, DxCG Inc.

This book is dedicated to

my wife Janet Duncan, FCAS, MAAA

who shared me with the first edition for 5 years,

and now for another two of the second edition.

PART I

AN INTRODUCTION TO CARE MANAGEMENT INTERVENTIONS

1 INTRODUCTION

This book covers the developing topic of care management interventions: design, management, and evaluation. Originally written for actuaries (the financial engineers of the healthcare payment system) I hope that it will be of value to anyone interested in these aspects of the management of healthcare intervention programs. The first edition addressed selected operational topics (such as the organizational structure and management of a disease management program) but its focus (appropriately for actuaries) was generally on cost, outcomes, and other financial issues. The topic of care management programs has expanded considerably since the publication of the first edition in 2008. Focus has shifted from single disease or multiple chronic disease management to a Population Health Management (PHM) approach as payers and policymakers recognize the importance of person- and population-centricity. This shift impacts not only the structure and operations of care management but also payment, metrics and reporting, topics that we address later in the book.

In the 1990s and 2000s, private, commercial initiatives drove the industry, with government largely uninvolved. Government became more involved as the 2000s progressed with CMS sponsoring a large-scale test of disease management interventions (the Medicare Health Support initiative). Increasing disappointment with the results of some of the widely-implemented programs in the 2000s coincided with work of, among others, Don Berwick MD at the Institute for Health Improvement in Boston, which resulted in the coining of the well-known term "Triple Aim" of healthcare. Dr. Berwick was, however, not alone in developing innovative care management models. Many integrated health systems with access to complete medical records, such as the Mayo Clinic in Minnesota, the Geisinger Clinic in Danville, Pennsylvania and the Kaiser Permanente Medical Group became recognized as models for the movement toward co-ordinated care based on integration of data and systems. Many of the developments in both programs and financial incentives in the recent past have been introduced in an attempt to replicate within a non-integrated system the methods and infrastructure available within these integrated systems.

This book is for the most part analytical, objective and based on research. When the research began in 2003 that led to the first edition of this book, funded by the Society of Actuaries (SOA) and overseen by the SOA's Project Oversight Group, we had little idea of the scope and duration of the work that would ensue. Ten years is a long time in which to be engaged in a single project, albeit part-time and with the assistance of volunteers, co-authors, reviewers, and others and the publication of another book (about risk adjustment and predictive modeling) along the way. In total the SOA-sponsored study generated eight research papers. These papers, together with the addition of a number of topics that were not part of the original study formed part of the first edition of this book. Many practitioners, both actuaries and non-actuaries, have

downloaded one or more of the original papers from the Society of Actuaries website, and have used some of the principles we developed in their own work. Some of the terms that we have coined in the course of the study (**migration bias** for example) have found their way into day-to-day discussion of disease management (DM) outcomes. The popularity of the papers vindicates the Society of Actuaries Health Section's and the Committee on Knowledge Extension Research's decisions to support the research and confirms the increasing role that actuaries are playing in this new and exciting area of managed care.

Since the publication of the first edition, this area has seen an explosion in research, innovations in interventions, techniques, and programs. Five years since the publication of the first edition, therefore, we have updated the original study with discussion of a number of new intervention programs that have been developed since the first edition, selected literature on programs previously discussed, new techniques (propensity matching and opportunity analysis, for example) and a more detailed discussion of other important topics, such as quality measurement and reporting.

A note about literature and research: the field of medical, clinical, managed care program, health policy and health economics is a vast one and generates considerable peer-reviewed research each year. A glance at the bibliography of this book, which lists over 300 references, illustrates the tip of the research iceberg. An actuary practicing in this field will be surrounded by professionals from many health-related disciplines all of whom will be familiar, to a greater or lesser extent, with the relevant literature. While the actuary will not need to be familiar with all the research on a particular topic, it is important to be aware of the literature, how to search for relevant articles and how to evaluate them (we address this in Chapter 10).

While the subject area continues to develop ever-more rapidly, the focus of this edition (like the first) is to arm the reader with fundamental principles that can be applied within any type of population or intervention.

1.1 A GUIDE TO THE CONTENTS OF THIS BOOK

This first Chapter is an overview and background reading for those with little prior exposure to the topic. Chapter 2 is new in this edition and provides a survey of the clinical background for diagnosis and treatment of some of the conditions frequently encountered in care management intervention programs: Respiratory, Endocrine, Circulatory, Immune, Musculoskeletal, and Cancer. We also provide some broad data on the prevalence and cost of different conditions and populations, so that the reader may judge the importance and scope for savings due to interventions aimed at managing these populations. Chapter 3 discusses different care management intervention programs. This chapter has been expanded with the addition of a number of new programs and providers not previously discussed (such as pharmacists and nurse practitioners) In chapter 4 we discuss specific actuarial topics that arise in program evaluation, including study design and the importance of understanding and accounting for risk.

Chapter 5 is a review of the published literature about the financial outcomes of different types of programs. This chapter includes a summary of literature previously discussed in the first edition, as well as newly-published literature since that edition. Detailed summaries of papers included in the chapter are available on-line at www.ActexMadRiver.com. We have not re-

printed summaries of the literature reviewed in the first edition; instead, the complete text of the corresponding chapter from the first edition (Chapter 4) is also available at the ACTEX website.

Chapter 7 is a new chapter in this edition. There is increased focus on reporting, benchmarking and reimbursement of clinical quality, partly as a result of initiatives under the Affordable Care Act. Chapter 7 surveys many of the different clinical quality measures required by different agencies.

Chapters 6 through 9 represent more of an operational perspective on programs. If the reader is to be able to assess financial and other results, it is important that he or she understand the underlying mechanics and economics of programs (and if necessary, plan for them). We begin with a discussion of Value Chain analysis in Chapter 6 – it is worth considering for its potential contribution to understanding the components of a program. Chapter 8 discusses the economics of Care Management, while Chapter 9 addresses a specific application of the economic model, which we call "Opportunity Analysis."

Chapters 10 through 14 discuss principles for designing a study to assess financial results of a program, and the practical implementation of those principles in a particular study design, the adjusted historical control methodology. Chapter 11 is new in this edition and covers propensity score matching, a technique that will be new to many actuaries but is commonly used in performing program evaluations. Chapter 12 discusses an actuarial method for evaluating programs, using trend adjustment. In Chapter 13 we examine the issue of changes in the risk profile of the intervention population, its potential effect on the measured outcomes of a program, and methods for mitigating this effect. Chapter 14 contains an extended discussion of the topic of sample size and the difficulties inherent in measuring outcomes in small populations.

Chapter 15 examines what is arguably the most important single factor in program evaluations – healthcare cost trend. This chapter, a re-print of an article that appeared in the *North American Actuarial Journal* in 2006, examines both theoretically and practically the potential measurement bias that can arise in measuring trend under different assumptions. Chapter 16 is a practical test of the effect that varying different assumptions has on the measured outcomes of a program.

Chapter 17 through 19 cover the topic of Wellness programs, a broader set of healthcare interventions that have attracted considerable attention. Chapter 17 reviews the published literature on outcomes and savings, and Chapter 18 looks at the possible savings from a program in a commercial population, using one measurement methodology. Finally, Chapter 19 presents the results of a study of the relationship between self-reported (Health Risk Assessment) risk factors, underlying health conditions, and claims.

1.2 THE STATE OF THE UNION

In Chapter 5 we provide a detailed review of the literature on financial outcomes of different care management programs. Over the years that we have been engaged in this study, however, the world of Care Management interventions has not stood still. The history of evaluation of disease management outcomes is an example: it is interesting to consider what has been achieved and what has *not* been achieved in the last ten years. In 2004, the Disease

Management Association of America[1] (DMAA) published "Principles for Assessing Disease Management Outcomes." [75] Far from establishing once and for all methodology and principles to be followed by practitioners, it is widely-agreed, including I believe by DMAA, that the guide fell short of the needs of the industry in this area. Accordingly, DMAA convened another work group in 2006 to tackle the subject again. The findings of this workgroup, entitled "DMAA Outcome Guidelines Report" [227] were published in December 2006. Because it is an industry consensus document, the DMAA workgroup report made a number of recommendations with which readers of this book may be familiar. In addition, the guidelines identified a number of potentially controversial issues, many of which were deferred for future consideration. Accordingly, DMAA convened a third series of work groups in 2007, which led to the publication of a second edition of Outcomes Guidelines [228] in the same year. The second edition addressed some of the gaps left by the first – for example, DMAA now recommends a particular method of selecting members for inclusion in a study population (which we discuss in greater detail in Chapter 16, and refer to as a re-qualification standard) when applying the adjusted historical control methodology, to overcome one of the more glaring areas of potential difference between comparison populations. DMAA did not make specific recommendations for sample sizes although the workgroup performed an analysis of the consequences of various sample sizes on the confidence interval of a savings calculation. We have provided guidance regarding credibility of outcomes for different population sizes in Chapter 14.

While its guidelines may help practitioners and purchasers, DMAA, as the industry trade association, was perceived by purchasers as representing an industry viewpoint, and thus at least somewhat suspect. The professional North American actuarial associations[2], (Society of Actuaries, Casualty Actuarial Society, Canadian Institute of Actuaries, and American Academy of Actuaries), on the other hand, have a reputation for being objective. Recommendations from these professional actuarial bodies, therefore, will carry more weight, particularly given the increasing involvement of actuaries in the performance and review of studies. The American Academy of Actuaries released its paper "Disease Management Programs: What's the Cost?" [75] in 2005, and released a Practice Note for actuaries practicing in the field in early 2008. It is the nature of actuarial practice notes to be descriptive, rather than prescriptive, providing a compendium of acceptable approaches taken by actuaries in tackling a particular problem, rather than choosing a particular approach as the "best practice." Actuarial best practices in DM may eventually be published in a Standard of Practice for DM, but, given the shift away from DM towards population health management and wellness, the profession has not focused on practice standards for DM. Since the publication of the Academy's practice note, the passage of the Affordable Care Act has diverted actuarial attention away from care management and evaluation issues. The introduction of new programs and reimbursement methodologies (some of which are discussed in Chapter 3) will, however require renewed focus on design and evaluation in the future[3].

[1] Now re-named Care Continuum Alliance, CCA.

[2] The Society of Actuaries mission is to provide education and research for North American Life, Health, Pensions and Investment Actuaries. The American Academy of Actuaries is the U.S. profession's interface with regulators, and is responsible for professional standards and accreditation. The Canadian Institute of Actuaries combines both educational and regulatory roles in Canada.

[3] In 2011 the CCA and HERO launched a collaborative effort to recommend guidance and an effectiveness metric scorecard for employers. This is slated for publication in January 2014. Financial metrics are included though not to the level of prescriptive 'recipes' near the order of HEDIS measures.

Some peer-reviewed papers and other notable studies have been published since we began this study. Ariel Linden, a well-known researcher in this field, published a paper in 2006 that attracted considerable attention [175]. This paper addresses what the author calls "number needed to treat," and which may also be called (as we do in Chapter 8 of this book) the "Economics of Care Management." In addition, this paper draws attention to the need for identification of a causal relationship between any savings estimated or measured, and the underlying inpatient admission experience of the population (where the major portion of savings are to be found). Soeren Mattke, MD, and others from RAND published a paper with a provocative title: "Evidence for the Effect of Disease Management: Is $1 Billion a Year a Good Investment?" [190]. The Congressional Budget Office and others published research into the results of the Medicare Health Support initiative between 2008 and 2011. While these papers are covered in more detail in Chapter 5, the authors' conclusions will not come as a surprise to anyone who has read any of the literature, namely that there is some evidence that DM improves quality of care but that there is little reliable evidence of financial improvement. What remains puzzling is the absence of practical papers that examine the biases in measurement and the impact that these have on outcomes, as, for example, we have done in Chapter 16 of this book, and, for studies that fail to show successful outcomes, deeper analysis of what elements worked, did not work, and what could be changed.

Chapter 5 updates Chapter 4 of the previous edition (which is not reproduced here but is available in its entirety online at www.ActexMadRiver.com for any interested reader). Chapter 4 was written and published early in the life of the SOA project, and the revision takes account (selectively) of newly-published articles as well as outcomes from studies of the newly-added programs discussed in Chapter 3.

We have seen increased actuarial involvement in care management outcomes studies and audits since the publication of the original Society of Actuaries studies, including the inclusion of care management as a topic on the Society's Fellowship health track syllabus. The fundamental building blocks of studies – rigorous reconciliation of data and understanding of Per Member Per Month (PMPM) costs and trends for example – lend themselves to analysis by actuaries. We also suggest in Chapter 15 that a relatively new technique in the actuarial arsenal, but one gaining wide acceptance – risk adjustment – also has a role to play in ensuring equivalence between populations. This new edition contains a discussion of a related subject (Propensity Score Matching) that is used in health services research to generate comparable populations and which will be important in the future for actuaries working in this field.

To the extent that the prior edition of this book has helped educate actuaries and others about intervention program design, management and evaluation and equipped them to work with health services professionals and clinicians, the Society of Actuaries study will have made a contribution.

1.3 WHAT HAS CHANGED?

1.3.1 MEDICARE HEALTH SUPPORT (MHS)

At the time of the publication of the first edition, the Medicare Health Support (MHS) program was in full swing. We looked forward to this program "finally provide(ing) the industry with the answers to two questions:

1. Does care management "work" (that is, produce a statistically-significant difference in financial and clinical results in the managed population)?

2. Potentially more important, how do the financial results measured by the randomized control methodology differ from results measured by a standard industry methodology (such as the actuarially-adjusted methodology described in Chapter 8)? While this comparative analysis is not part of the program, many researchers are anxiously awaiting the opportunity to perform just such a comparative analysis.[4]

The MHS program was introduced in 2005 under Section 721 of the Medicare Modernization Act of 2003 (MMA) (the same act that passed Medicare Part D coverage for prescription drugs and expanded accessibility to health savings accounts). The act authorized development and testing of voluntary chronic care improvement programs, (later re-named Medicare Health Support) to improve the quality of care and life for people living with multiple chronic illnesses. This program applied to Medicare fee for service members with diabetes and/or heart failure. The Centers for Medicare and Medicaid Services (CMS) awarded eight different programs to disease managers in different regions. Three vendors subsequently withdrew from the program, and CMS reduced the savings target from Fees Plus 5% of Total Chronic Claims] to just fees (break-even). These developments implied that even during the program, enrollment and savings targets were not being met, an impression later confirmed by outcomes evaluation.

As we discuss in more detail in Chapter 3, the MHS programs demonstrated some benefit in terms of improved quality of care but little financial savings. The conclusions were challenged by the DM industry based on uneven assignment and lack of comparability of the patient populations. The lack of demonstrated success of the largely nurse call-center-based MHS model coincided with the increased interest in provider-based models, as we discuss later.

One may conclude that the MHS did provide a clear answer, at least to the first of our two questions. Perhaps less noticed is the fact that there is also an implicit answer to the second question: an intervention program that the industry accepted as being financially successful, based on a number of different evaluation methods, was demonstrated by a large randomized controlled trial as failing to provide a positive financial outcome. To our knowledge no study of the MHS program has ever compared the calculated outcomes based on more traditional evaluation methods and the randomized control method used by MHS. Such a study, particularly if it were able to identify the sources of deviation between the randomized outcomes and those from other methods, would be valuable for identifying ways to improve our population studies (which will rarely be able to use randomization).

1.3.2 PLAUSIBILITY ANALYSIS

Practitioners have also contributed to advances in outcomes measurement, although the techniques have not been published in the peer-reviewed literature. Al Lewis, president of the Disease Management Purchasing Consortium International, recommends the use of what he calls "Plausibility Factors." These factors are not a method for calculating savings, but rather, a method for evaluating the reasonability of the published outcomes, (whether the calculated savings are "plausible," based on the underlying utilization of the population and what we know

[4] From Managing and Evaluating Healthcare Intervention Programs, 1st edition.

of the success of similar programs). Plausibility analysis requires the calculation of the following statistic (the plausibility factor) for the entire health plan's entire condition-specific population:

$$\frac{\text{Disease-Specific Admissions/1000 (Program Year)}}{\text{Disease-Specific Admissions/1000 (Baseline Year)}}$$

USE OF PLAUSIBILITY FACTORS

The theory of plausibility factors is that they independently validate the measured financial results of a care management savings calculation, by demonstrating that actual utilization is reduced by the intervention, consistent with the financial measurement. Plausibility factors are generally utilization rates per 1,000 of the overall population for hospital admissions and emergency room visits for certain primary diagnoses. The primary diagnoses are: diabetes, coronary artery disease, chronic obstructive pulmonary disorder, heart failure and asthma. The proposed interpretation of the Plausibility measures is that if the savings calculation results in positive savings but the utilization-based measure do not, the savings are not validated. Rather than reconciling the two contradictory results, the Plausibility factors are so dispositive that their results always trump any other outcomes calculation.

HOW VALID ARE UTILIZATION-BASED CALCULATIONS?

In order to be a valid test of the outcomes of a savings calculation, utilization-based measures must be calculated on the same basis as the savings. With plausibility factors this is not always the case. The plausibility factors may be a poor validator because:

1. In a population evaluation, the measurement population is carefully constructed to consist of members with sufficient eligibility to be enrolled and managed by the program and to exclude members and conditions that may confound the calculation. As calculated the plausibility factors bear only a tenuous relationship to the population being managed and measured. Their use implicitly assumes comparability between populations, but this comparability must be demonstrated and cannot be assumed.

2. Plausibility factors, because they apply to admissions and ER visits for primary diagnoses only, represent a very small percentage of all admissions and costs for chronic patients. Within a commercial population, for example, these admissions and ER visits only account for 3% of the total claims costs for members with diabetes, and the admissions only account for approximately 7 % of inpatient expenditures. Even a very successful program that avoided 25% of diabetes admissions could never demonstrate enough savings to warrant program costs under this methodology. By definition, therefore purchasers must be assuming that the program beneficially affects other utilization measures of the population, and indeed, programs aim to do precisely that. So failure to demonstrate reduction in the direct utilization measures does not necessarily imply lack of success with other types of utilization[5].

[5] A more valid indicator might be inclusion of certain admissions with the condition in the secondary diagnostic position and specified principal diagnoses in the principal position. Most admissions of diabetics, for example, do not have a principal diagnosis of diabetes but rather a complication or comorbidity. Including ischemic peripheral vascular disease (diagnosis or procedures for, including cellulitis, lower extremity ulcers, gangrene, and interventions) would yield a more complete picture of the program impact. Similarly for COPD, consider including claims with a principal diagnosis of pneumonia and influenza.

3. Plausibility factors do not take account of changes in population. Because the denominator is the entire population, a change in population size and composition will change the measured rate of chronic admissions per 1000, independent of any impact of a DM program (positive or negative).

4. Plausibility factors do not take account of the risk profile of a population. It is entirely possible, for example, that a new group of relatively high-risk members may replace a relatively low-risk group, increasing the measured chronic admission rate per 1000, independent of any program effect.

5. The plausibility factors take no account of volatility in admission rates. As Table 1.1 illustrates, the standard deviation of admission rates per 1000 in even a population as large as Medicare can be fairly large. Thus one cannot accept a hypothesis that the program effect is positive unless the deviation from the prior year's admission rate is outside a confidence interval based on the standard deviation; conversely, one cannot reject the hypothesis that the program effect is positive simply because the difference between two rates is small (or even positive!).

6. The plausibility factors, unlike the underlying adjusted historical control methodology, take no account of existing trends in the population. As Table 1.1 illustrates, admission trends are frequently low (lower than overall trend). The results shown in Table 1.1 are significant in that they illustrate that (for Medicare members) discharge trends have generally been slightly negative for many chronic conditions, in an environment in which chronic prevalence has been increasing in the Medicare population.

To illustrate this last point, consider Table 1.1 which illustrates the actual trends in discharges per 1000 for certain chronic conditions for Medicare patients for selected years between 1984 and 2000, and annually between 2000 and 2011. We can ignore the early years (because of program changes and changes in definitions of Diagnosis-related Groups (DRGs) and focus on the period 2000 to 2011. In 2008, DRGs were re-defined and data prior to 2008 will not always be comparable to those after 2007, although the trends remain obvious. Bronchitis and Asthma discharges are no longer reported every year because they do not meet the minimum threshold for reporting.

Over the 11-year period, the average trends in admissions for major chronic conditions (diabetes, heart conditions and COPD) while low, were negative. Even the overall trend in Medicare Admissions is negative as well. Admissions for Renal Failure, a condition that is generally omitted from studies of program outcomes, however, have increased significantly during this period. Table 1.1 reports the trend (calculated as the coefficient of the admission rate in a simple linear regression fitted to the years 2000-2011 (2007 in the case of diabetes and bronchitis/asthma). The standard deviation is also shown. Figure 1.1 illustrates the same results in graphical form.

Trends in Medicare Discharges/1000 for Major Chronic Conditions

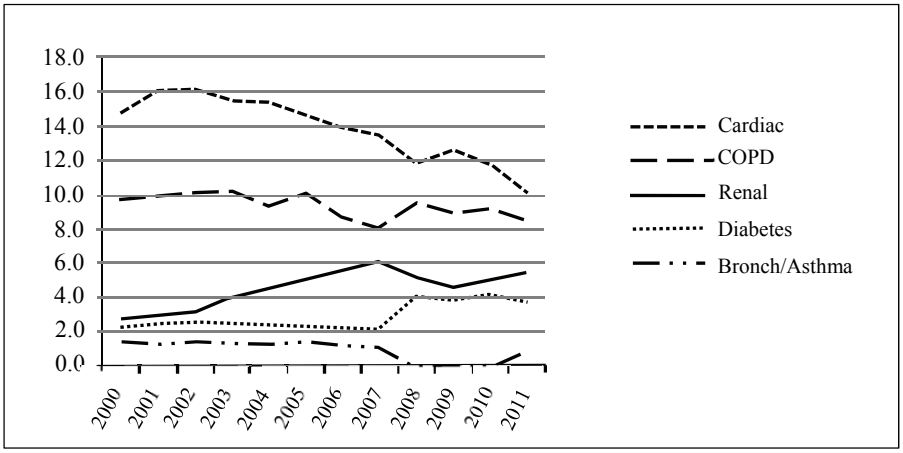

FIGURE 1.1

TABLE 1.1

Medicare Discharges/1000 for selected Conditions

Year	Member -ship	Diabetes	Renal Failure	Bronchitis & Asthma	COPD	Heart	All Discharges
DRG		294*	316**	096***	088+	138, 139, 143-4++	
1984	29,996	4.717	1.547	5.937	7.084	11.906	363.213
1990	33,731	2.743	1.443	5.624	4.294	12.493	311.936
2000	39,211	2.280	2.768	1.470	9.925	14.914	298.895
2001	39,625	2.458	3.001	1.352	10.047	16.130	308.660
2002	40,079	2.516	3.174	1.428	10.275	16.229	314.563
2003	40,696	2.450	3.985	1.385	10.335	15.575	315.941
2004	41,391	2.425	4.498	1.276	9.564	15.459	312.100
2005	42,129	2.368	5.055	1.442	10.191	14.772	306.294
2006	42,975	2.267	5.632	1.217	8.878	14.050	288.170
2007	43,910	2.172	6.105	1.128	8.295	13.608	274.112
2008	45,067	4.008	5.175	n/a	9.616	11.910	262.294
2009	46,195	3.858	4.686	n/a	9.104	12.633	250.205
2010	47,316	4.096	5.002	n/a	9.295	11.888	260.817
2011	48,511	3.750	5.486	0.847	8.684	10.239	236.908
Mean (2000-)		2.367	4.547	1.337	9.517	13.951	285.747
Std. Deviation		0.110	1.046	0.113	0.644	1.838	26.642
Trend 2000-2011^		−2.6%	8.8%	−2.5%	−1.3%	−3.3%	−2.3%

^ For diabetes and COPD, trends are calculated for 2000-7 only because of the re-definition in DRGs that occurred in 2008. Data for 2007 are reported for only 9 months and have been annualized.

* Data are no longer reported for the Diabetes DRG and reported data from 2008 onward are for ICD-9 250.x. Prior to 2008, limited to age 35+.

** Now DRG 682/3

*** Now DRG 202; not reported in all years because of relative insignificance.

+ Now DRG 190/1/2

++ Arrhythmia and Conduction disorders, chest pain and other circulatory disorders. Now DRG 308-10 and 313-4.

As with any healthcare statistics, definitions matter and should be considered before any data are used in any comparison.

Some caution should be exercised when reading this table: underlying definitions may have changed during the illustrated periods and data are for Medicare only (comparable commercial data are not yet available, although we expect the new Health Care Cost Institute data to become available in the near future, which will make comparable Commercial reporting possible). Some conclusions may be drawn from the data, and should be kept in mind when reading the remainder of this book:

- At least in Medicare, while there are exceptions, there has been a downward trend in admissions for major chronic diseases for some years. The external trend in any population or measure must be considered in any study that considers a longitudinal population.
- Admission rates are subject to wide variability. Simple analyses that compare pre- and post- admission rates should consider this underlying variability in outcomes.
- The Medicare population experience illustrated here is likely to be more stable than commercial experience because the Medicare population itself is more stable; a commercial insurer or employer is subject to "churn" of both employees and groups, and we therefore cannot simply assume that the *underlying* population whose admissions we are measuring is comparable from year to year.

TOWARDS A UNIFIED THEORY OF UTILIZATION BASED MEASUREMENT

Tom Wilson's article published some years ago highlighted the importance of demonstrating a causal pathway for any savings calculation. When measured appropriately to take account of underlying trends in the population, the reduction in utilization is a powerful demonstration of causality. In the future, with large volumes of both Medicare and Commercial data (both utilization and cost) becoming available, it is possible that we will be able to model expected utilization and cost for a specified condition-population, and compare this with that population's actual utilization and cost. A statistically-significantly lower utilization rate in a managed population would provide a powerful demonstration that an intervention or program had worked.

1.4 WHERE TO FROM HERE?

More work needs to be done to understand some of the areas we analyze in this book, and those discussed in this chapter. Other areas for future research include:

1.4.1 CONDITION IDENTIFICATION

In Chapter 16 we consider the effect on the measured results of changes in the way chronic members are identified. In Chapter 15 we also demonstrate that *when* the member was identified as having a chronic condition can have a significant effect on trend, and thus, on the estimated savings from a program. Understanding the impact of these issues on a study is not just an actuarial task and will require involvement of clinical and actuarial researchers.

1.4.2 TRANSITION STATES

We have discussed some of the implications of a transition state model earlier. If we understood chronic members' propensity to change states (particularly as their disease condition matures over time) we could perhaps do a better job of analyzing how and whether an intervention has changed that trajectory.

1.4.3 WHAT "WORKS" IN CARE MANAGEMENT?

Those of us who are practitioners in this area have been focused, because of the needs of our employers and clients, on assessing the impact of a program, particularly on financial outcomes. This focus has often been on program results at the expense of attempting to discern the impact of different types of intervention within sub-populations. A typical disease management program, for example, may include different types of interventions delivered to many different member sub-populations (with different conditions; co-morbidities; level of severity and risk). Programs often co-exist within a health plan, with case management interventions that apply yet more intensive management to a member's problems. My prediction for care management in the future is that we will see fewer, more-intensive interventions targeted at smaller chronic populations, within integrated programs that include both intensive case management and broader population management (or wellness), often delivered through a more cost-effective medium such as the internet and social media. This trend will increase our need to know what works, with whom. It will also increase the need for more accurate predictive models to be able to identify those members who match the "target" profiles (those candidates who have been identified through an algorithm, predictive model or other means for intervention). The Value Chain approach, outlined briefly in Chapter 7, may provide a basis for understanding program components. But it will require the care management companies to be willing to share much more detailed data if we are to answer questions like "what works?"

1.4.4 A "STANDARD" METHODOLOGY

The DM industry has struggled and failed for a number of years to agree on a standard measurement methodology. By default, most evaluations tend to be performed using a variant of the actuarially-adjusted historical control methodology. Given that a large percentage of industry evaluations are performed using a similar methodology, with variation being in the details (chronic definitions; timing; exclusions and inclusions), I have earlier suggested that a more potentially useful expenditure of the industry's resources would be in understanding the impact on the measured results of these definitions, as a pre-cursor to developing a common set of definitions. The industry has for too long struggled to respond to the demand for an absolute result (how much was saved), a problem that was answered by the Medicare Health Support program, rendering industry efforts to develop outcomes standards somewhat redundant. Instead, the industry should borrow a leaf from the National Committee for Quality Assessment (NCQA) book and develop a set of measures *together with standard definitions* that health plans and those performing interventions could produce that would allow comparisons to be performed. I do not think that any user of NCQA's Health Effectiveness Data and Information Set (HEDIS) measures would necessarily believe that these are an *absolute* measure of health plan quality or, for that matter, that they are the only measures of health plan quality. But the measures, imperfect as they are, have the advantage of being standardized, produced by all health plans, and therefore comparable. The care management industry could perhaps learn from the experience of NCQA and develop similar measures (and definitions) that would allow valid comparisons between programs and vendors.

1.5 WHAT HAVE WE LEARNED FROM OUR RESEARCH?

The key conclusions from the research can be summarized as follows:

1. The most important objective in any care management outcomes study is to ensure comparability between the intervention and comparison populations. The existing care management evaluation literature tends to encourage a belief that there are two "threats to validity" in studies: **selection bias**, which will be observed when participants are compared with non-participants, and **regression to the mean**. But as we show in Chapter 3, regression to the mean is an *individual,* not population concept (except in the default case of a population comprising similar individuals). As discussion throughout our research suggests, the identification and correction of regression to the mean is a much larger and more complicated issue than some of the literature suggests, particularly if the definition of who is included in a population is not clear.

2. Understanding the Economics of Care Management (as discussed in Chapter 8) is probably even more important than it was when the first edition was published. This is due to the explosion in the number of programs and program sponsors, particularly among agents who often are unsophisticated financially, the shift towards provider-based programs and the increasing focus on quality measures rather than financial outcomes. An important question to ask about any program is whether the claimed (or projected, in the case of a proposed program) savings outcomes are plausible. Application of a simple economic model to the underlying population data allows users to estimate a range of likely outcomes, as well as test the sensitivity of those outcomes to different program components. More importantly, understanding the key variables of the financial model and their contribution to the overall financial outcome will allow analysis of individual proxy variables that can be directly measured (the enrollment rate, for example). Actuaries and other financial professionals have a critical role to play in understanding the economics of programs, being able to evaluate for clients those interventions that will help to meet their (increasingly stringent) financial objectives and to recommend program changes that will enable those financial goals to be achieved.

3. Population studies, a common study design in care management evaluation, may achieve comparability if the populations being studied do not change much from period to period. A major challenge for actuaries is to demonstrate this stability. Fortunately, actuaries understand the issues involved in ensuring comparability over time and the implication for PMPM costs when comparability is not achieved. Actuarial tools such as risk-adjustment make assessment of risk profiles over time and demonstration of equivalence simpler.

4. As discussed in Chapter 12, the actuarially-adjusted historic (pre-post) design, which is the most prevalent in the industry, offers a reasonable compromise between validity and practicality. Many would wish to use more scientifically-pure methods, but, as we discuss, these are seldom achievable. Instead, the popularity of the actuarially-adjusted historical control method in the industry is testament to the fact that a well-executed study is viewed as being reasonably reliable. The work that we and other researchers have done attempts to address some of the areas of sensitivity in outcomes, for example the identification of patients for different populations.

5. While the fundamental evaluation methodology does not vary much between practitioners, the assumptions and methods used to deal with data issues do vary considerably. Definitions matter. We cover in Chapter 12 many of the issues that are usually considered in a study – exclusions, inclusions, timing, and so on; principles that are widely applicable within many program evaluations.

6. Chapter 15, published in stand-alone form in the *North American Actuarial Journal* in October 2006, highlights the issue of chronic identification and its impact on chronic prevalence and trends. In a population study, the issues of *what* claims codes identify a chronic population, *when* those codes have to be observed, how frequently and over what time period, are crucial. As an industry we have only begun to scratch the surface of these issues, but I believe that it is probably the single most important issue for the industry to focus on in the future.

7. It is important to understand the impact or "value" of different assumptions on the final results of a study. It is surprising that much of the discussion in the literature remains at a theoretical level when many practitioners have access to data sets and could simply test out some of the issues that they debate. The industry would greatly benefit from it. It would make the current methodology more robust and would reduce the need for the industry to search for alternative methodologies. In Chapter 16 we examine some of the sensitivities of the results calculated using one such methodology for one client, under different assumptions. Much more of this type of analysis needs to be published, to gain knowledge about the methodology.

The care management industry continues to expand. One area of growth is Wellness and Worksite Health. Recognizing this, DMAA (formerly the Disease Management Association of America) was re-named "DMAA-The Care Continuum Alliance" in 2007. DMAA now covers the new, broader spectrum of interventions. The chapters that address some of the issues of Wellness and Worksite Health programs have been selectively updated to reflect changes in this area and some of the recent published literature.

1.6 A CHANGE IN EMPHASIS FOR ACTUARIES?

Traditionally, actuaries have focused on financial analysis and worked with aggregate data, often at a category of service level. An example of the change in emphasis in recent years is the need to deal with chronic populations, which requires that the actuary have a more detailed knowledge of the medical conditions, underlying services and treatments (and the claims that they generate) that a particular member requires. This, in turn, requires actuaries to have more clinical knowledge than was the case in the past. The Society of Actuaries has enthusiastically supported actuaries' involvement in broader healthcare topics. One indicator of this change in emphasis is the popular "Medical School for Actuaries" seminar that the Society of Actuaries hosts twice each year. Actuaries also need to become increasingly fluent in topics such as risk adjustment that require knowledge of claims data and the medical conditions and procedures that generate them.

This change in emphasis turns the old "financial" analytical paradigm on its head, with the condition-population becoming the unit of interest and analysis. We still have much to learn

about the behavior of traditional actuarial measures (for example cost PMPM and trend) when applied to sub-populations with common characteristics, such as a health condition. And because our clinical colleagues have barely begun to scratch the surface of what constitutes "clinical best practice" for members with specific conditions, we have a long way to go before we can begin to benchmark utilization and cost for these populations. Nevertheless, for the actuary interested in pursuing this area of practice, the techniques and tools described in this book (and in the companion volume, *Healthcare Risk Adjustment and Predictive Modeling*) are a good place to start.