The Dissertation Committee for Ryan Christopher Leslie certifies that this is the approved version of the following dissertation:

Development of a Data-Driven Method for Selecting Candidates for Case Management Intervention in a Community's Medically Indigent Population

Committee:

_____

Marvin D. Shepherd, Supervisor

_____

Karen Rascati

_____

Kenneth A. Lawson

_____

David C. Warner

_____

Thomas Erlinger

DEVELOPMENT OF A DATA-DRIVEN METHOD FOR SELECTING CANDIDATES FOR CASE MANAGEMENT INTERVENTION IN A COMMUNITY'S MEDICALLY INDIGENT POPULATION

by

Ryan Christopher Leslie, B.B.A., M.S.

Dissertation

Presented to the Faculty of the Graduate School of

The University of Texas at Austin

in Partial Fulfillment

of the Requirements

for the Degree of

Doctor of Philosophy

The University of Texas at Austin

May 2008

Dedication

This dissertation is dedicated to my parents, Steven and Denese Leslie, who have, throughout this process and throughout my entire life, given me their unyielding support, confidence, and love.

Acknowledgements

I cannot imagine a better graduate education experience than that created by the Pharmacy Practice and Administration faculty and staff at the University of Texas at Austin. Dr. Marvin Shepherd, my supervising professor for both my thesis and dissertation, has been a wonderful mentor.

Dr. Kenneth Lawson and Dr. Karen Rascati provided invaluable guidance and support. I am very fortunate to have had them as committee members. Without Dr. David Warner I would never have been introduced to the ICC and its staff. Furthermore, Dr. Warner has helped me look at my research from a perspective that I otherwise would have missed.

In addition to being a great committee member, Dr. Thomas Erlinger, with each conversation we have, seems to make me more and more excited about health care research that focuses on the underserved. His enthusiasm is contagious, and it is a wonderful thing.

Once again, the ICC and its staff made this research project possible. Sandy Coe Simmons helped me access and understand the data and was a great sounding board for my research ideas. Ann Kitchen helped me identify the salient details among my results and helped me view the analysis from the perspective of the partner organizations. I also had the pleasure of working with Cindy Batcher who provided the invaluable perspective of a nurse case manager.

Finally, I want to thank my beautiful wife, Lindsay, who has been patient and supportive throughout this long, extended process. You are a treasure.

Abstract

Development of a Data-Driven Method for Selecting Candidates for Case Management Intervention in a Community's Medically Indigent Population

Ryan Christopher Leslie, Ph.D.

The University of Texas at Austin, 2008

SUPERVISOR: Marvin D. Shepherd

The Indigent Care Collaboration (ICC), a partnership of Austin, Texas, safety net providers, gathers encounter data and manages initiatives for the community's medically indigent patients. One such initiative is the establishment of a care management program designed to reduce avoidable hospitalizations.

This study developed predictive models designed to take year-one encounter data and predict inpatient utilization in the following two years. The models were calibrated using 2003 through 2005 data for the 41,260 patients with encounters with ICC partner providers in all three years. Predictor variables included prior inpatient admissions, age, sex, and a summary measure of overall health status: the relative risk score produced by the Diagnostic Cost Groups prospective Medicaid risk-adjustment model.

Using the 44,738 patients with encounter data in each of years 2004 through 2006 data, the performance of the predictive models was cross-validated

and compared against the performance of the "common sense" method of choosing candidate patients based on prior year chronic disease diagnoses and high utilization, referred to herein as the Utilization Method (UM). The 620 patients with three or more 2005 through 2006 inpatient admissions were considered the actual high use patient subset. Each model's highest-risk 620 patients comprised its high-risk subset. Only 344 high-risk patients met the UM's criteria. Prediction accuracy was described in terms of positive predictive value (PPV), i.e., the proportion of identified high-risk patients who were high-use patients.

Three of the predictive models had a PPV of near 25% or greater, with the highest, the linear model using the DCG relative risk score, at 26.8%. The PPV of the UM was 17.1%, lower than that of all predictive models. When all high-risk subsets were limited to 344 patients (the number identified by the UM), the performance of the UM and the predictive models was similar.

This study demonstrated that "common sense" targets for case management can be identified via simple filter as effectively as through empirically-based predictive models. However, once the supply of easily identifiable targets is exhausted, predictive models using a measure of health status identify high-risk patients who could not be easily identified by other means.

Table of Contents

List of Tables

List of Figures

Chapter 1: Background

Overview

The Indigent Care Collaboration (ICC) is a partnership between central Texas hospitals and safety net clinics that gathers data on hospital and clinic visits for uninsured patients in the greater Austin, Texas, metropolitan area. The data are shared electronically by ICC member clinics and hospitals through Health Insurance Portability and Accountability Act (HIPAA) compliant Business Associate Agreements. At the individual patient level, the data provide a history that can aid providers in diagnoses and therapeutic decision-making. At an aggregate level, the data can be analyzed to monitor diseases in the population and measure the effects of policy changes.

In addition to collecting and maintaining this patient data, the ICC also manages various health initiatives focused on the area's indigent, uninsured population. One such proposed initiative is the establishment of a care management program designed to improve safety net efficiency by reducing inpatient admissions and emergency department (ED) use by high-cost patients. This program would assign case managers to patients with expected high utilization in an attempt to improve chronic disease management, increase coordination of care, and educate patients on correct use of resources.

The success of such a program depends on the ability to correctly identify future high-cost patients who have a high opportunity for improved health outcomes and utilization behavior. In Travis County an estimated 171,752 people, 19.3 percent of the population, lacked health insurance in 2005. In the neighboring counties of Hays and

Williamson, an estimated 23,325 (18.7 percent) and 70,782 (21.2 percent) lacked health insurance in 2005.[1] Since the overall uninsured population in the Austin area is quite large, it is unreasonable to conduct a full evaluation on each individual patient (including interviews and other time-intensive processes) to determine enrollees for the care management program. Therefore, a list of candidates must first be generated.

To create such a list, the ICC is querying their aggregated dataset for patients who have high inpatient and/or ED utilization and a hospital diagnosis of one or more of six pre-chosen chronic diseases (diabetes, hypertension, congestive heart failure, asthma, obesity, and tobacco abuse) in the previous year. This method of identifying candidates, which will be referred to herein as the Utilization Method, may prove to be effective; however, limitations do exist.

The first limitation is that historical utilization of health care resources has been shown to be a poor predictor of future utilization.[2,3] Second, to make valid inferences about health care outcomes using administrative data, one must account for disease severity and comorbid conditions.[4] High-risk patients with multiple chronic conditions may avoid consideration for intervention if their prior year utilization is beneath an arbitrary threshold. Third, outpatient diagnoses which may yield better insight into patients' total health status are not considered. Fourth, the focus on number of ED visits

[1] Murdoch and Armstrong. County Estimates of the Uninsured for Texas, 2005. *Texas State Data Center*. Available at: http://txsdc.utsa.edu/download/pdf/presentations/2006_12_01_County_Estimates_of_the_Uninsured_for_Texas_Tables.pdf.
[2] Ash, Zhao, Ellis and Schlein Kramer. Finding future high-cost cases: comparing prior cost versus diagnosis-based methods. *Health Serv Res.* Dec 2001;36(6 Pt 2):194-206.
[3] Meenan, O'Keeffe-Rosetti, Hornbrook, Bachman, Goodman, Fishman and Hurtado. The sensitivity and specificity of forecasting high-cost users of medical care. *Med Care.* Aug 1999;37(8):815-823.
[4] Deyo, Cherkin and Ciol. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology.* 1992;45(6):613-619.

may identify patients who wrongly use the ED for primary care but, due to triaging, have relatively low avoidable costs compared to patients with serious chronic illness. While intervening to discourage unnecessary use of the ED may be appropriate, this falls outside of the scope of a care management program when utilization is driven by improper use and not by poor health. Finally, the Utilization Method has not been back-tested on historical data to gauge expected performance relative to other methods of selecting program candidates.

The current study sought to identify a data-driven method of selecting candidate patients through an analysis of the historical data facilitated by the use of a diagnosis-based risk adjustment model. This project tested the combined predictive capability of utilization and demographic variables from the ICC's database along with diagnosis-based risk scores. The resulting models, which will be referred to herein as the Risk Models, were then compared with the Utilization Method using historical data to identify which method better predicted utilization. Data from 2004 was used to generate candidates, while data from 2005 and 2006 was used to assess candidate outcomes.

Literature Review

Background

This section describes the various methods used to control the rising costs of health care and explains why care management initiatives focusing on high-risk patients are applicable to uninsured populations. Risk is defined both conceptually and operationally for this project. Finally, this chapter discusses methods of assigning risk to individual patients through the use of predictive models.

The Rising Costs of Health Care

In the United States, national health expenditures have grown from 12.3% of gross domestic product (GDP) in 1990 to 13.8% of GDP in 2000 and 16.0% in 2005.[5] National expenditures related to hospital care and retail sales of prescription drugs have increased from $417.0 and $120.8 billion in 2000 to $611.6 billion and $200.7 billion in 2005, respectively.

The rising costs have been attributed to a wide variety of factors, including overuse of services by non-cost-conscious consumers,[6,7] the increasing use of innovative prescription drugs and treatments,[8] the aging of the population,[9] and a lack of true

[5] U.S. Department of Health and Human Services. National Health Expenditures Aggregate, Per Capita Amounts, Percent Distribution, and Average Annual Percent Growth, by Source of Funds: Selected Calendar Years 1960-2005. Available at: http://www.cms.hhs.gov/NationalHealthExpendData/downloads/tables.pdf. Accessed 1/22/2007.
[6] Thorpe. The rise in health care spending and what to do about it. *Health Aff*. Nov-Dec 2005;24(6):1436-1445.
[7] Liebowitz. Why health care costs too much. *Cato Institute*. Available at: http://www.cato.org/pubs/pas/pa211.html. Accessed 1/21/2007.
[8] Bodenheimer. High and rising health care costs. Part 2: technologic innovation. *Ann Intern Med.* Jun 7 2005;142(11):932-937.

competition in the health care marketplace.[10] Borger et al. predict that the growth in national health spending will outpace the growth of the gross domestic product (GDP) in each year through 2015, causing national health spending as a percentage of GDP to grow to 20 percent in 2015.[11] Borger et al.'s analysis shows that between 2006 and 2015 growth in hospital care spending, the largest component (37 percent) of personal health spending, will exceed growth in GDP by over two percent per year. Prescription drug spending is expected to grow between 8.0 and 8.4 percent annually from 2007 to 2015, as factors pushing spending up and those pushing spending down largely offset each other.[12] The proportion of the general population taking prescription drugs will continue to expand due to evolving physician practice patterns, and high-cost drugs will increasingly be prescribed for rare conditions.[13] Meanwhile, use of generic drugs and consumer cost-sharing are both expected to increase, causing downward pressure on spending growth.[14,15]

Aaron describes two general strategies for reducing the growth of health care spending: (1) reducing pure waste (i.e., high-cost care which produces no benefit) and (2)

---

[9] Reinhardt argues against this popular notion in: Reinhardt. Does the aging of the population really drive the demand for health care? *Health Aff.* Nov-Dec 2003;22(6):27-39.
[10] Enthoven. Employment-based health insurance is failing: now what? *Health Aff.* Jan-Jun 2003;Suppl Web Exclusives:W3-237-249.
[11] Borger, Smith, Truffer, Keehan, Sisko, Poisal and Clemens. Health Spending Projections Through 2015: Changes On The Horizon. *Health Aff.* March 1, 2006 2006;25(2):w61-73.
[12] Ibid.
[13] Mullins, DeVries, Hsu, Meng and Palumbo. Variability And Growth In Spending For Outpatient Specialty Pharmaceuticals. *Health Aff.* July 1, 2005 2005;24(4):1117-1127.
[14] Medco Health Solutions. *Drug Trend Report 2005.* Franklin Lakes, N.J: Medco; 2005.
[15] Watson Wyatt. *Managing Health Care Costs in a New Era: Tenth Annual National Business Group on Health/Watson Wyatt Survey Report 2005.* Washington: Watson Wyatt; 2005.

reducing high-cost care whose benefits are not justified by the costs.[16] Both strategies focus on a single concept: avoiding care that is not cost-beneficial, i.e., where costs in dollars exceed the estimated dollar value of the resulting health outcomes. The merits of sweeping, national-level initiatives to identify and reduce non-cost-beneficial care (e.g., imposing rationing, price controls, or regulatory barriers to new, expensive treatments) are hotly debated. Drastic, national changes to health care systems tend to occur only as a result of economic catastrophe or war,[17] so participants in the American health care system, including patients, providers, insurers, and employers, can expect the same drivers of increasing health care costs to be present into the foreseeable future.

To combat rising health care costs, managed care organizations (MCOs) have implemented a variety of strategies that generally fall into one of two categories: (1) managing utilization, and (2) disease and case management programs.[18] The following sections discuss MCOs' cost reduction initiatives in both categories and the respective applicability of each initiative in the care of the uninsured by an organization like the ICC.

[16] Aaron. Should public policy seek to control the growth of health care spending? *Health Aff.* Jan-Jun 2003;Suppl Web Exclusives:W3-28-36.
[17] Ibid.
[18] Felt-Lisk and Mays. Back to the drawing board: new directions in health plans' care management strategies. *Health Aff.* Sep-Oct 2002;21(5):210-217.

Utilization Management

The Institute of Medicine defines utilization management as:

> [A] set of techniques used by or on behalf of purchasers of health care benefits to manage health care costs by influencing patient care decision-making through case-by-case assessments of the appropriateness of care prior to its provision.[19]

Although some consider disease and case management to be subsets of utilization management, this manuscript treats them as separate and distinct forms of cost control, indirectly impacting costs through patient outcomes. Herein, utilization management is viewed as a more direct attempt to reduce demand for health care services from an MCO's perspective. Hospital or provider initiatives, like referral management, are therefore not included in this discussion. The types of utilization management discussed include: prior-authorization, provider profiling (quality management), and consumer cost sharing.

Prior Authorization

Prior authorization is a form of prospective utilization management where high-cost drugs and services are available only to those patients who have been screened and found needing the drugs or services. The screening process helps avoid unnecessary treatments when less expensive alternatives exist. Smalley et al. demonstrated the cost

[19] Institute of Medicine Committe on Utilization Management by Third Parties. *Controlling Costs and Changing Patient Care? The Role of Utilization Management*. Washington, D.C.: National Academy Press; 1989. p.17

savings that prior authorization programs can achieve.[20] However, the process itself, which can be time consuming and frustrating for patients and providers alike, has been shown by Guirguis et al. to delay treatments and increase acute complications.[21] For this reason, prior authorization programs have met with fierce opposition from patient advocacy groups. MCOs have responded by reducing prior authorization requirements and streamlining the process.[22]

In the uninsured population, an MCO is not paying providers to perform services. The safety net provider organizations must maintain their own internal policies regarding high-cost treatments and the uninsured, while keeping in mind the legal obligation to provide necessary emergency care to patients established in the Emergency Medical Treatment and Active Labor Act.[23]

While a safety net partnership organization like the ICC has no prospective role to play in prior-authorization-type decisions made by providers, the data can be reviewed retrospectively to determine whether treatment patterns for the uninsured match established guidelines.

---

[20] Smalley, Griffin, Fought, Sullivan and Ray. Effect of a prior-authorization requirement on the use of nonsteroidal antiinflammatory drugs by Medicaid patients. *N Engl J Med.* Jun 15 1995;332(24):1612-1617.
[21] Guirguis and Taylor. The complications of cholelithiasis caused by state authorization delays. *Surg Endosc.* Sep 1995;9(9):974-976.
[22] Felt-Lisk and Mays. Back to the drawing board: new directions in health plans' care management strategies. *Health Aff* . Sep-Oct 2002;21(5):210-217.
[23] Fields, Asplin, Larkin, Marco, Johnson, Yeh, Ghezzi and Rapp. The Emergency Medical Treatment and Labor Act as a Federal health care safety net program. *Acad Emerg Med.* November 1, 2001 2001;8(11):1064-1069.

Case Reviews and Provider Profiling

For particularly expensive and complex inpatient cases, MCOs may perform concurrent reviews of care, assessing case status and ascertaining whether the current treatment plan is appropriate.[24] By reviewing the appropriateness of care as care is delivered, MCOs attempt to reduce length of inpatient stays, diagnostic tests, and procedures. MCOs may also evaluate the appropriateness of care delivered to patients retrospectively, investigating in more detail cases which appear irregular.[25]

The history of individual providers may also be profiled retrospectively to determine whether guidelines and protocols are being followed.[26] DeLong et al. describe how provider quality is statistically estimated in coronary artery bypass grafting through the use of risk adjustment provider profiling. The observed occurrences of events (e.g., mortality) are compared with expected occurrences derived from large samples.[27]

A safety net partnership with aggregated patient utilization data could perform retrospective provider profiling analysis to ensure that treatment of uninsured patients follows guidelines and protocols. However, such an organization may have difficulty conducting concurrent case review analysis due to time delays in receiving patient data, the incompleteness of the data in comparison to full chart data maintained by the providers, and the administrative and practical difficulties in critiquing ongoing care across organizational boundaries.

---

[24] Kongstvedt PR. *Managed Care: What It Is and How It Works.* 2nd Edition ed. Gaithersburg, Maryland: Aspen Publishers; 2002. p.137
[25] Ibid. p.138
[26] Ibid. p.139
[27] DeLong, Peterson, DeLong, Muhlbaier, Hackett and Mark. Comparing risk-adjustment methods for provider profiling. *Statistics in Medicine.* 1997;16(23):2645-2664.

Consumer Cost Sharing

Arrow and Debreu argued that free market systems using marginal pricing can produce Pareto efficient outcomes, provided no externalities exist (among other idealized assumptions).[28] In health care, however, the presence of insurance isolates the consumer from some of the costs related to services they may seek. In the words of Arrow: "Insurance removes the incentive on the part of individuals, patients, and physicians to shop around for better prices for hospitalization and surgical care."[29] This externality leads to Pareto inefficiencies via moral hazard. Moral hazard exists when an insured consumer's behavior differs from how that consumer would have behaved without insurance. For example, Blustein found that women covered by Medicare were much more likely to get mammograms when they had supplemental insurance that paid the out-of-pocket costs.[30]

To increase the price sensitivity of patients, MCOs have implemented various cost-sharing strategies, primarily through premiums, copayments, and deductibles.

Premiums, Copayments, and Deductibles

The most straightforward method of increasing the share of costs borne by consumers is increasing the premium, the base charge for insurance. Premiums are

---

[28] Arrow and Debreu. Existence of an equilibrium for a competitive economy. *Econometrica.* 1954/07 1954;22(3):265-290.
[29] Arrow. Uncertainty and the welfare economics of medical care. *The American Economic Review.* 1963;53(5):941-973. p.962
[30] Blustein. Medicare coverage, supplemental insurance, and the use of mammography by older women. *N Engl J Med.* Apr 27 1995;332(17):1138-1143.

generally paid monthly (or annually), and when health insurance is employer-sponsored, premiums are typically deducted from employees' paychecks. While increasing premiums helps increase the total amount of costs paid by consumers, premiums do little to increase the price sensitivity of patients.[31] A high premium may disincline a patient to choose a particular health plan,[32] but once that plan has been chosen and premium payments are made on regular intervals, the premiums have become a sunk cost to consumers.[33] Dowd demonstrates the game-theoretic framework in which premiums exist, showing that premiums may inhibit conservation (via moral hazard) simply because rational consumers want to feel like value has been obtained for paying the premium.[34]

Copayments (copays) are requirements that patients pay out of pocket for a portion of their health care services at the time the services are rendered. The amount paid out of pocket may be either a fixed amount or a percentage of the total expenses. The latter option is sometimes called coinsurance. The presence of a copay requirement is intended to increase a patient's sensitivity to costs, resulting in reduced utilization. The RAND Health Insurance Experiment (HIE), operating between 1974 and 1982, tested patient sensitivity to varying copayment rates including no cost sharing, or 25, 50, or 95% copay, with a maximum yearly out-of-pocket expense of $1,000. The RAND HIE

[31] Dowd. The logic of moral hazard: a game theoretic illustration. *The Journal of Risk and Insurance.* 1982;49(3):443-447.
[32] Buchmueller and Feldstein. Consumers' sensitivity to health plan premiums: evidence from a natural experiment in California. *Health Aff* . Spring 1996;15(1):143-151.
[33] Dowd. The logic of moral hazard: a game theoretic illustration. *The Journal of Risk and Insurance.* 1982;49(3):443-447.
[34] Ibid.

found that spending was consistently reduced among the cost-sharing groups, as patients sought care less frequently.[35]

Copayments are frequently used with prescription drug benefit programs. As of March 2004, 39 state Medicaid programs had in place cost-sharing requirements for enrollees.[36] These requirements generally consisted of a small copay of $3 or less, and 14 states limited the number of prescriptions an enrollee could fill in a single month (e.g., 3 monthly in Texas).[37] The 1982 Tax Equity and Fiscal Responsibility Act (TERFA) requires that Medicaid copays be nominal: the lesser of 10% of the charge or three dollars. The copays must be waived for Medicaid HMO enrollees, pregnant women, children, and nursing home residents. However, even small copays may discourage filling prescriptions, and eligible patients may not realize that copays may be waived.[38]

Nelson et al. found that the implementation of a copay requirement on outpatient prescriptions drugs in South Carolina's Medicaid program resulted in reduced utilization rates and expenditures.[39] Expanding on this, Stuart and Zacker found that Medicaid patients in states requiring prescription drug copays filled an average of five fewer prescriptions than patients in states with no copay requirement.[40] Soumerai et al. showed

---

[35] Keeler. Effects of cost sharing on use of medical services and health. *Journal of Medical Practice Management.* 1994;8:317-321.
[36] United States General Accounting Office. *Medicaid and SCHIP: States' Premium and Cost Sharing Requirements for Beneficiaries* March 2004.
[37] Kaiser Commission on Medicaid and the Uninsured. Medicaid Outpatient Prescription Drug Benefits: Findings from a National Survey, 2003. Available at: http://www.kff.org/medicaid/upload/Medicaid-Outpatient-Prescription-Drug-Benefits-Findings-from-a-National-Survey-2003.pdf. Accessed 7/20/2007
[38] Stuart and Zacker. Who bears the burden of Medicaid drug copayment policies? *Health Aff.* March 1, 1999 1999;18(2):201-212.
[39] Nelson, Reeder and Dickson. The effect of a Medicaid drug copayment program on the utilization and cost of prescription services. *Med Care.* Aug 1984;22(8):724-736.
[40] Stuart and Zacker. Who bears the burden of Medicaid drug copayment policies? *Health Aff.* March 1, 1999 1999;18(2):201-212.

that when New Hampshire's Medicaid program implemented a limit of three prescriptions per month, the decrease in filled prescriptions included essential medications like insulin.[41] A follow-up study, also by Soumerai et al., showed that among New Hampshire's frail elderly patients, the reduction in use of prescription drugs due to the newly imposed monthly limit was associated with an increase in rates of admission to nursing homes.[42]

Lurk et al. demonstrated the sensitivity of under- and uninsured patients to increases in prescription drug copays in a safety-net provider setting. Following a 50 percent increase in copays ($5.00 to $7.50 for generics, $10.00 to $15.00 for brand name medications) clinic prescription costs per patient per month dropped $26.07. However, the impact on patient health outcomes, total health care utilization, and total costs were not studied.[43] In 2005, Gibson et al. conducted a review of the published literature on the effects of prescription drug cost-sharing initiatives. Based on the results of the 30 articles included in their analysis, Gibson et al. came to the following conclusion.

> It is also becoming clear that cost sharing is not always a benign instrument, and at times it may come at a price. Although not consistently reported in the literature, the most troublesome effects associated with higher levels of cost sharing are reports of treatment disruption for chronically ill patients who depend on a regular regimen of prescription drugs. In addition, higher levels of cost sharing can have significant effects on the use of essential or maintenance medications, the outcomes of care, and the process of care.[44]

---

[41] Soumerai, Avorn, Ross-Degnan and Gortmaker. Payment restrictions for prescription drugs under Medicaid. Effects on therapy, cost, and equity. *N Engl J Med.* August 27, 1987 1987;317(9):550-556.
[42] Soumerai, Ross-Degnan, Avorn, McLaughlin and Choodnovskiy. Effects of Medicaid drug-payment limits on admission to hospitals and nursing homes. *N Engl J Med.* October 10, 1991 1991;325(15):1072-1077.
[43] Lurk, Dejong, Woods, Knell and Carroll. Effects of changes in patient cost sharing and drug sample policies on prescription drug costs and utilization in a safety-net-provider setting. *American Journal of Health-System Pharmacy.* 2004/02/01/ 2004;61(3):267-272.
[44] Gibson, Ozminkowski and Goetzel. The effects of prescription drug cost sharing: a review of the evidence. *Am J Manag Care.* Nov 2005;11(11):730-740. p.739

With health services other than prescription drugs, copayments may create other dilemmas for indigent patients and their providers. For example, enacting copayment requirements for emergency treatment of indigent patients is problematic. The Emergency Medical Treatment and Labor Act (EMTALA) requires performance of a medical screening exam on anyone who comes to an emergency department and, if necessary, treatment to stabilize the patient, including inpatient admission, without requiring any advance payment or financial means testing.[45] For non-emergency treatment of indigent patients, direct patient charges, whether in the form of copays, reduced rates, or full-price charges, may exceed a patient's ability to pay, leaving the clinician with an uncomfortable choice: proceed despite the potential economic repercussions on the patient, reduce the level of care to meet the patient's budget, or provide no care at all.[46] If a copayment amount is hindering needed care of an indigent patient, a clinician may feel pressured to undercode (e.g., bill for a minor, less expensive visit than actually occurred). However, undercoding jeopardizes the integrity of health data that may be the basis for research.[47]

A deductible is "an exclusion of a certain amount of expense from coverage."[48] For example, a patient with an annual $100 deductible would pay the total expense of any service with a cost of $100 or less and would pay only $100 for any service costing over

[45] Lawsuits imply EMTALA requires EDs to admit all uninsured patients. *ED Manag.* Aug 2004;16(8):85-87.
[46] Weiner. "I can't afford that!". Dilemmas in the care of the uninsured and underinsured. *Journal of General Internal Medicine.* 2001;16(6):412-418.
[47] Ibid.
[48] Pauly. The Economics of Moral Hazard: Comment. *American Economic Review.* 1968;58(3):531-537. p.535

$100. Excluding income effects, the presence of a deductible creates a binary outcome on a patient's demand for an individual service. When the deductible amount is less than or equal to the patient's perceived value of the service, the presence of the deductible has no effect on demand. If the deductible amount is greater than the patient's perceived value of the service, the patient's demand for the service will be the same as if the patient had no insurance.[49]

The presence of high deductibles may allow for more affordable premiums; however, low income patients may still be unable to afford such plans if the premiums account for over five percent of income.[50] Even if the premiums do not cause a barrier to coverage, high deductibles may lead to barriers in care. The results of the Commonwealth Fund Biennial Health Insurance Survey (2003) suggest that high deductible plans cause disproportionately high access problems among low income Americans.[51]

The Medicare Part D prescription drug plan uses a two-part deductible: a base deductible and a gap or "doughnut hole." In 2006, beneficiaries with income above 150 percent of poverty paid a $250 deductible and had expenses covered to $2,250. Any expenses between $2,250 and $5,100 were paid out of pocket by the beneficiary. The plan covered expenses beyond $5,100.[52] Stuart et al.'s analysis shows that many chronically ill patients will not reach the top of the coverage gap, making the gap

[49] Ibid.
[50] Ku and Coughlin. Sliding-scale premium health insurance programs: four states' experiences. *Inquiry.* Winter 1999;36(4):471-480.
[51] Davis, Doty and Ho. *How High is Too High? Implications of High-Deductible Health Plans*: The Commonwealth Fund; April, 2005.
[52] Stuart, Simoni-Wastila and Chauncey. Assessing The Impact Of Coverage Gaps In The Medicare Part D Drug Benefit. *Health Aff.* April 19, 2005 2005:167.

function more like a benefit cap.[53] Such a cap, according to Hsu et al., leads to poorer adherence in drug therapy and poorer clinical outcomes.[54]

Disease and Case Management

Disease and case management programs try to reduce utilization by improving the health status of patients with high-cost chronic disease states, such as diabetes, asthma, and depression. Unlike utilization management initiatives that attempt to curb spending either immediately (in the case of prior authorization and case reviews) or in the very near future (as with provider profiling and patient cost sharing initiatives), case and disease management programs are more long-term focused, relying on the future improvement (or at least stabilization) of potentially costly patients to offset the costs of enacting such programs. For example, a program aimed at improving adherence to antihypertensive drug therapy would hope to reduce or delay future high-cost cardiovascular problems.

Disease Management – Overview

Chronic diseases that are undiagnosed, poorly managed, and/or poorly treated may lead to high-cost complications, increased health care utilization, and increased costs. Chronically ill patients frequently must monitor their own diseases, often without the education and knowledge necessary to properly do so. Furthermore, these patients

[53] Ibid.
[54] Hsu, Price, Huang, Brand, Fung, Hui, Fireman, Newhouse and Selby. Unintended consequences of caps on Medicare drug benefits. *N Engl J Med.* June 1, 2006 2006;354(22):2349-2359.

may see a variety of disparate providers with little to no care coordination, leading to care that varies from established evidence-based treatment guidelines.[55]

Disease management programs focus on specific, costly chronic diseases and coordinate the efforts of a multidisciplinary team (e.g., physicians, pharmacists, nurses). Patient education is usually a key feature of disease management programs with the aim of changing patient behavior. Additionally, patient progress and health are monitored centrally by experts in the specific targeted disease, reducing some of the burden of responsibility on the patient and allowing for standardization of care and compliance with established guidelines.[56]

Three main methods exist for selecting patients within a managed care population to enroll in disease management programs. The most general method is a broad screening effort considering all population members. A more specific method would utilize predictive modeling to identify specific individuals most likely to benefit from disease management. The most specific method would be to wait until a warning event, such as hospitalization, before enrolling the patient. General methods of patient selection may identify most high-risk patients but waste time and money on patients who do not need intervention. Highly specific methods may properly identify the correct patients but only after their conditions have worsened, making interventions less effective.[57]

[55] Congressional Budget Office. An analysis of the literature on disease management programs. Available at: http://www.cbo.gov/ftpdocs/59xx/doc5909/10-13-DiseaseMngmnt.pdf. Accessed 2/9/2007.
[56] Ibid.
[57] Ibid.

Disease Management – Measuring Impact

Measuring the impact of a disease management program can be difficult, since the main health outcome measures may only occur several years in the future. For example, the costly events associated with diabetes, such as circulatory damage, nerve damage, blindness, kidney failure, amputation, and heart attack, typically occur years after the onset of diabetes. Short-term attempts to measure the impact of a diabetes disease management program on these specific events are, therefore, difficult if not impossible.

Despite these limitations, the impact of a disease management program may be measured in changes in intermediate outcomes. For diabetes, these may include indicators of glycemic control (e.g., glycated hemoglobin and fasting blood glucose), physiology (e.g., blood pressure, weight, microalbuminuria), and lifestyle (e.g., diet and activity level).[58] For coronary heart disease, these may include blood pressure, serum lipid parameters, and body-mass index.[59]

However, translating impact on intermediate outcomes into the overall cost-effectiveness of the disease management program is no simple feat. Currently, the lack of a standardized framework used to measure the economic impact of disease management programs complicates comparative analysis of reported results.[60] In 2004, the Congressional Budget Office (CBO) conducted an analysis of the published disease management literature to help the U.S. Senate determine whether such programs would

[58] Norris, Nichols, Caspersen, Glasgow, Engelgau, Jack, Isham, Snyder, Carande-Kulis, Garfield, Briss and McCulloch. The effectiveness of disease and case management for people with diabetes: A systematic review. *American Journal of Preventive Medicine.* 2002/5 2002;22(4, Supplement 1):15-38.
[59] Anderson and Konz. Obesity and disease management: effects of weight loss on comorbid conditions. *Obesity Res.* November 1, 2001 2001;9(90004):326S-334.
[60] Villagra. Strategies to control costs and quality: a focus on outcomes research for disease management. *Med Care.* Apr 2004;42(4 Suppl):III24-30.

prove beneficial to Medicare beneficiaries and to the Medicare budget. The CBO reported that too many uncertainties currently exist to reliably translate improvements in intermediate outcomes into actual cost savings. Furthermore, the CBO found that studies often did not adequately measure the costs associated with disease management programs or failed to directly address costs at all. These findings led the CBO to suggest that general implementation of disease management programs in broad populations may not generate cost savings, although when focused on a select group of patients, such programs could be designed to reduce costs.[61]

In 2005, Goetzel et al. conducted a review of the disease management literature to identify how expected return on investment (ROI) varied by chronic disease. Programs focused on coronary heart disease and multiple comorbid conditions showed the most evidence for positive ROI. Diabetes-focused programs showed more short-term promise when tightly focused on diagnosed diabetic patients and less when geared toward prevention, looking for pre-diabetic patients. Asthma programs showed mixed results, while depression programs consistently showed negative ROI, although impact on productivity was not assessed.[62]

Case Management

While disease management programs start with a disease of interest and then screen for high-risk patients within the disease, case management programs determine

[61] Congressional Budget Office. An analysis of the literature on disease management programs. Available at: http://www.cbo.gov/ftpdocs/59xx/doc5909/10-13-DiseaseMngmnt.pdf. Accessed 2/9/2007.
[62] Goetzel, Ozminkowski, Villagra and Duffy. Return on investment in disease management: a review. *Health Care Financing Review.* Summer 2005;26(4):1-19.

enrollees in nearly the exact opposite manner, starting by identifying all high-risk patients and then determining which ones would potentially benefit from increased focus and coordination of care.

Holloway et al. suggests that the case management literature broadly agrees on the following components of case management programs:

1) assessment of client need;
2) development of a comprehensive service plan;
3) arrangement of service delivery;
4) monitoring and assessment of services; and
5) evaluation and follow-up.[63]

Since case management programs are highly individualized and dependent upon the resources and capabilities of the program sponsor(s), case management programs may vary greatly between organizations and patient populations. Researchers have attempted to categorize case management programs into archetypes or models (e.g., assertive community therapy, intensive case management, broker service model), debating the relative merits of each type of model.[64,65] While these models are instructive in detailing the necessary "assumptions, methods, structures, and tools that dictate the intervention," real-world implementations of case management programs rarely replicate these models exactly.[66] Rapp dissects the various case management models to identify the common requirements for effective case management. Table 1.1 lists Rapp's guidelines.

---

[63] Holloway F, Oliver N, Collins E and Carson J. Case management: a critical review of the outcome literature. *European Psychiatry.* 1995;10:113-128. p.113

[64] Solomon. The efficacy of case management services for severely mentally disabled clients. *Community Ment Health J.* Jun 1992;28(3):163-180.

[65] Mueser, Bond, Drake and Resnick. Models of community care for severe mental illness: a review of research on case management. *Schizophr Bull.* 1998;24(1):37-74.

[66] Rapp. The active ingredients of effective case management: a research synthesis. *Community Ment Health J.* Aug 1998;34(4):363-380. p.364

Table 1.1 Requirements of Effective Case Management

Structure	*Service*
1. Team structure for the purpose of creative case planning, problem- solving, sharing knowledge of resources and support to team members.	1. Case management contact with clients should be in-vivo (limit office-based contacts.)
2. Team leaders/supervisors should be experienced, professionally trained mental health professionals.	2. Frequency of in-vivo client contact will vary based on the client but should average across clients a minimum of six in person contacts per month. This should be supplemented by telephone and collateral contacts.
3. Case managers can be paraprofessional (e.g., B.A. level) but need access to specialists; involvement of nurses seems particularly important.	3. Case managers should deliver as much of the "help" (e.g., modeling, resource acquisition, skill training, advice) directly as possible.
4. Case load sizes can vary based on client severity, geography, etc. but should never exceed 20:1. The average across program clients should probably be 12:1 to 15:1.	4. Referrals to traditional mental health programs (e.g., partial hospitalization, day treatment, in-office counseling, sheltered and many transitional employment programs, congregate housing, etc.) should be avoided.
5. Efforts should be made to enhance the continuity of relationship between the client and case manager.	5. The use of naturally occurring community resources (landlords, employers, coaches, neighbors, churches, friends, clubs, junior colleges, etc.) should be encouraged.
6. Clients need 24 hour, 7 days a week access to crisis and emergency services. That service should require access to staff who have familiarity and a relationship with the client (can be and perhaps should be the case manager).	6. Case managers should have ultimate responsibility for client services (with the exception of medication). They retain authority even in referral situations.
7. Pre-service, in-service, and technical assistance should be available.	7. Clients should be given equal or greater authority than case managers or other professionals in treatment and life decisions with the exception of hospitalization decisions.
8. Length of case management service should be indeterminate and expected to be on-going (although intensity at any point in time would vary).	

Source: Rapp CA. The active ingredients of effective case management: a research synthesis. *Community Ment Health J.* Aug 1998;34(4):363-380. pp.377-378

Although Rapp's requirements are geared toward patients with mental illness, most address fundamental aspects of care quality: e.g., qualification of team members, ease of access, quantity and quality of patient contact, patient empowerment. Although his specific requirements for case load sizes and frequency of in-vivo contact may not be

ideal (or feasible) for all disease states and patient populations, the domains themselves nevertheless deserve attention, and applicable standards should be set that are appropriate to the unique circumstances.

Case Management – Measuring Impact

General conclusions about the effectiveness of case management programs (both in terms of health improvement and cost reduction) are difficult to make due to the uniqueness of individual programs. Most published reports focus on patients with specific chronic diseases[67,68,69] or on specific subsets of patients (e.g., those recently discharged from a hospital,[70,71] elderly in community settings[72]).

As found in the disease management literature, published analyses of case management programs often lack detailed information about program costs, which is needed to accurately assess economic impact. In 1995, Holloway et al. reviewed 23 studies detailing the outcomes of patients enrolled in case management programs focused

[67] Latour, Bosmans, van Tulder, de Vos, Huyse, de Jonge, van Gemert and Stalman. Cost-effectiveness of a nurse-led case management intervention in general medical outpatients compared with usual care: An economic evaluation alongside a randomized controlled trial. *Journal of Psychosomatic Research.* 2007/3 2007;62(3):363-370.
[68] Harris, Luft, Rudy, Kesterson and Tierney. Effects of multidisciplinary case management in patients with chronic renal insufficiency. *The American Journal of Medicine.* 1998/12 1998;105(6):464-471.
[69] Kushel, Colfax, Ragland, Heineman, Palacio and Bangsberg. Case management is associated with improved antiretroviral adherence and CD4+ cell counts in homeless and marginally housed individuals with HIV infection. *Clinical Infectious Diseases: An Official Publication Of The Infectious Diseases Society Of America.* 2006/07/15/ 2006;43(2):234-242.
[70] Stewart, Marley and Horowitz. Effects of a multidisciplinary, home-based intervention on planned readmissions and survival among patients with chronic congestive heart failure: a randomised controlled study. *The Lancet.* 1999/9/25 1999;354(9184):1077-1083.
[71] Spillane, Lumb, Cobaugh, Wilcox, Clark and Schneider. Frequent users of the emergency department: can we intervene? *Acad Emerg Med.* June 1, 1997 1997;4(6):574-580.
[72] Lu, Lin, Tzeng, Huang and Chang. Effectiveness of case management for community elderly with hypertension, diabetes mellitus, and hypercholesterolemia in Taiwan: a record review. *International Journal Of Nursing Studies.* 2006/11/27/ 2006;43(8):1001-1010.

specifically on mentally ill patients. All but three studies failed to supply adequate cost detail.[73] Still, by looking at reported changes in utilization, Holloway et al. were able to conclude that "[t]here is impressive evidence to suggest that case management can reduce the number and length of hospital admissions for psychiatric patients."[74]

Studies focusing on case management intervention on patients with high hospitalization utilization tend to have small sample sizes and have given inconclusive results.[75,76] In a 2000 pilot study comparing the hospital utilization of 53 frequent emergency department users before and after case management intervention, Okin et al. found that each dollar spent on the program yielded $1.44 in cost savings.[77]

The Synthesis of Disease and Case Management

A review of the disease and case management literature suggests that both can be effective if programs are well-designed and patients are well-targeted. For disease management programs, this means increasingly focusing on comorbidities and high-risk patients—elements of case management. For case management, this means focusing on patients with chronic diseases that have been shown to be responsive to care interventions—a key focus of disease management. Increasingly the boundaries between

---

[73] Holloway F, Oliver N, Collins E and Carson J. Case management: a critical review of the outcome literature. *European Psychiatry.* 1995;10:113-128.
[74] Ibid. p.123
[75] Lee and Davenport. Can case management interventions reduce the number of emergency department visits by frequent users? *Health Care Manag (Frederick).* Apr-Jun 2006;25(2):155-159.
[76] Phillips, Brophy, Weiland, Chenhall and Dent. The effect of multidisciplinary case management on selected outcomes for frequent attenders at an emergency department. *The Medical Journal Of Australia.* 2006/06/19/ 2006;184(12):602-606.
[77] Okin, Boccellari, Azocar, Shumway, O'Brien, Gelb, Kohn, Harding and Wachsmuth. The effects of clinical case management on hospital service use among ED frequent users. *The American Journal of Emergency Medicine.* 2000/9 2000;18(5):603-608.

what have historically been considered disease and case management have become blurred, with both being referred to singularly as efforts of care management,[78] coordinated care,[79] and integrated care coordination.[80]

As the following section demonstrates, care management intervention is particularly well-suited for combating the rising costs of caring for the uninsured.

[78] Bernstein. New arrows in the quiver for targeting care management: high-risk versus high-opportunity case identification. *J Ambul Care Manage.* Jan-Mar 2007;30(1):39-51.
[79] Chen, Brown, Archibald, Aliotta and Fox. Best practices in coordinated care. Available at: http://www.mathematica-mpr.com/PDFs/bestpractices.pdf. Accessed 3/16/2007.
[80] National Chronic Care Consortium. Case Management: Methods and Issues. Available at: http://www.nccconline.org/pdf/CaseManagement.pdf. Accessed 3/16/2007.

Managing Health Care Costs of Community Uninsured Patients

Increasing health care costs have led to a decline in the number of Americans with insurance coverage. Chernew, Cutler, and Keenan demonstrated via probit regression and instrumental variable techniques that rising premiums explained 64% of the decline in coverage between 1990 and 1998.[81]

In 2004, the uninsured, including those with coverage only part of the year, accounted for nearly $125 billion in health care costs. Almost a third of this figure, $40.7 billion, was uncompensated, i.e., provided by safety net hospitals, clinics, and private practices without reimbursement.[82]

The bulk of uncompensated care dollars, nearly 60% in 2001, are accounted for by hospitals, where care is the most expensive.[83] Berk and Monheit found in a review of 1996 Medical Expenditure Panel Survey (MEPS) data that the costliest 5 percent of patients among the uninsured accounted for 60 percent of total expenditures.[84] Leslie similarly found a skewed distribution in the cost of the Austin, Texas, metropolitan area uninsured, where the costliest 10 percent of patients accounted for approximately 35.4 percent of total area expenditures.[85]

---

[81] Chernew, Cutler and Keenan. Increasing health insurance costs and the decline in insurance coverage. *Health Serv Res.* Aug 2005;40(4):1021-1039.
[82] Hadley J and Holahan J. The Cost of Care for the Uninsured: What Do We Spend, Who Pays, and What Would Full Coverage Add to Medical Spending? *Kaiser Commission on Medicaid and the Uninsured*. Available at: http://www.kff.org/uninsured/7084.cfm.
[83] Ibid.
[84] Berk and Monheit. The concentration of health care expenditures, revisited. *Health Aff*. Mar-Apr 2001;20(2):9-18.
[85] Leslie R. Cost Analysis of the Uninsured, Indigent Population of Central Texas. Master's Thesis. The University of Texas at Austin. 2006.

Although this skewed distribution of expenditures is found in both the uninsured and insured populations, the costliest of the uninsured receive far less care annually than their insured counterparts, with the top 5% of uninsured and insured patients consuming $6,651 and $17,871 respectively.[86] Outcomes comparisons between uninsured and insured patients suggest that insured patients benefit from the additional care. Bradbury et al. found in a cross-sectional study of 100 hospitals that among admitted patients, the uninsured tended to be sicker and experience worse outcomes.[87] This finding is consistent with the bulk of the literature comparing the outcomes of insured and uninsured patients. In 2003, Hadley reviewed the past 25 years of research, finding that 43 out of 54 identified studies showed that uninsured patients have reduced access leading to worse health.[88]

In 2000, Ayanian et al. surveyed over 100,000 patients in all 50 US states and the District of Columbia to identify differences in unmet needs between insured and uninsured patients. The authors found that uninsured patients in poor health tended to have fewer routine checkups than their insured counterparts. Furthermore, long-term uninsured patients received significantly less preventative services in the areas of cancer,

[86] Berk and Monheit. The concentration of health care expenditures, revisited. *Health Aff* . Mar-Apr 2001;20(2):9-18.
[87] Bradbury, Golec and Steen. Comparing uninsured and privately insured hospital patients: admission severity, health outcomes and resource use. *Health Serv Manage Res.* Aug 2001;14(3):203-210.
[88] Hadley. Sicker and poorer--the consequences of being uninsured: a review of the research on the relationship between health insurance, medical care use, health, work, and income. *Med Care Res Rev.* Jun 2003;60(2 Suppl):3S-75S; discussion 76S-112S.

cardiovascular disease (including hypertension screening, cholesterol screening, weight loss advice, and smoking cessation advice) and diabetes.[89]

The various methods of controlling costs described above have differing levels of applicability to a community uninsured population like that tracked by the ICC. Attempts to reduce consumption through patient cost-sharing have little to no utility with the uninsured, who largely rely on charity care. Herring's analysis of 1996-2000 Medical Expenditure Panel Survey (MEPS) data showed that the uninsured paid $336 out-of-pocket annually in year 2000 dollars for health, representing only 36.4% of the health care they consumed.[90]

Utilization management strategies focusing on provider behavior (e.g., prior authorization, case reviews, provider profiling) may be more difficult to implement community-wide in the uninsured than in a managed care population since boundaries exist between disparate safety net provider organizations. Since reimbursement payments cannot be used as leverage, it may be difficult for an organization like the ICC to change the behavior of providers even when potential savings (or outright waste) can be identified.

Disease and case management initiatives, on the other hand, could be implemented and managed by a third-party organization like the ICC with much less buy-in by provider organizations than would be required by utilization initiatives like case reviews.

---

[89] Ayanian, Weissman, Schneider, Ginsburg and Zaslavsky. Unmet health needs of uninsured adults in the United States. *JAMA.* Oct 25 2000;284(16):2061-2069.

[90] Herring. The effect of the availability of charity care to the uninsured on the demand for private health insurance. *Journal of Health Economics.* 2005/3 2005;24(2):225-252.

The high-cost subset of patients within the ICC population represents a valuable target for a case management initiative. In the literature, expert opinion and empirical evidence support the notion that a small percentage of the population consumes a disproportionately large portion of health care.[91,92,93] However, from year to year, the specific individuals who make up the high-cost subgroup within a population should not be expected to remain constant.[94]

The viability of a case management program relies on the ability to correctly identify the members of the ever-changing subset of high-cost patients who have a high likelihood of benefiting from increased intensity of care – i.e., the high-opportunity patients among the high-risk patients.[95]

Choosing Care Management Focus Based on Potential for Positive Return-on-Investment

Although reviews of the care management literature provide some insight as to which conditions have been shown to be the focus of cost-beneficial care management programs,[96,97] a number-needed-to-decrease (NND) analysis can show whether, for a

[91] Eggert and Friedman. The need for special interventions for multiple hospital admission patients. *Health Care Financ Rev.* Dec 1988;Spec No:57-67.
[92] Zook and Moore. High-cost users of medical care. *N Engl J Med.* May 1 1980;302(18):996-1002.
[93] Harris. Disease management: new wine in new bottles? *Ann Intern Med.* May 1 1996;124(9):838-842.
[94] Ash, Zhao, Ellis and Schlein Kramer. Finding future high-cost cases: comparing prior cost versus diagnosis-based methods. *Health Serv Res.* Dec 2001;36(6 Pt 2):194-206.
[95] Bernstein. New arrows in the quiver for targeting care management: high-risk versus high-opportunity case identification. *J Ambul Care Manage.* Jan-Mar 2007;30(1):39-51.
[96] Goetzel, Ozminkowski, Villagra and Duffy. Return on Investment in Disease Management: A Review. *Health Care Financing Review.* Summer 2005;26(4):1-19.
[97] Congressional Budget Office. An analysis of the literature on disease management programs. Available at: http://www.cbo.gov/ftpdocs/59xx/doc5909/10-13-DiseaseMngmnt.pdf. Accessed 2/9/2007.

specific population, a care management program has the potential to show a positive return-on-investment (ROI).

Linden, who introduced the concept of NND analysis in 2006, suggests that the economic feasibility of a care management program can be evaluated *a priori* by determining the numeric reduction in inpatient days or emergency department visits that must be attributable to a proposed care management program in order to offset the program's costs.[98] Researchers would create a table showing the NND with various assumptions on costs and level of ROI. The range of NND values could be put in context by comparing historical target population statistics with national statistics. Of specific interest would be inpatient and emergency department utilization per 10,000 patients and mean length-of-stay. For example, a program targeting heart disease would compare the national and population-of-interest trends on heart disease related hospitalization. If the population of interest compares favorably with national statistics (i.e., has equal or lower utilization), sufficient opportunity for reduction may not exist. However, if admissions and/or length-of-stay are above national averages, reductions may be possible.

Table 1.2 is an example of NND analysis. In this example, a hypothetical program costing $0.90 per-member-per-month must reduce annual inpatient days by 225, assuming the cost of each inpatient day is approximately $1,000. At a cost per day of $2,000, the break-even NND would be 113. The proposed care management program would only be considered further if these reduction levels appear feasible after comparing these numbers with the difference in national and population-of-interest statistics.

[98] Linden AL. What will it take for disease management to demonstrate a return on investment? New perspectives on an old theme. Am J Manag Care. 2006;12(4):217-22.

The ICC, with their aggregated encounter data, can easily determine the utilization statistics for the area's medically indigent patients with specified chronic diseases and compare them with their national counterparts. With estimates of program and utilization costs, a NND analysis could be performed for any potentially targeted subgroup.

Table 1.2 Number-Needed-to-Decrease (NND) Analysis of the Number of Inpatient Days That Must Be Reduced to Meet Varying Levels of Return on Investment (ROI), Assuming a $0.90 Per-Member-Per-Month Program Fee

	Cost per Day*			
	$1,000		$2,000	
ROI	NND	% Decrease**	NND	% Decrease**
1.0***	225	21.5	113	10.8
1.5	338	32.3	170	16.3
2.0	450	43.1	226	21.6
2.5	563	53.9	283	27.1
3.0	675	64.6	339	32.4

*The mean cost per hospital day ($1000 per Medicare Provider Analysis and Review data and $2000 per Maryland Health Services Cost Review Commission data)

**From the current discharge rate

***1.0 indicates that program fees equal current hospital costs

Source: Linden AL. What will it take for disease management to demonstrate a return on investment? New perspectives on an old theme. Am J Manag Care. 2006;12(4):217-22.

Defining Risk

Before risk can be discussed in the context of patients and health care expenditures, a conceptual understanding of risk is necessary. Holton defines risk as "exposure to a proposition of which one is uncertain."[99] Holton's definition centers on two components of risk: uncertainty and exposure.

The term uncertainty connotes a conscious awareness of a lack of information with which to judge the likelihood of an event. In practical application, however, uncertainty may exist even when the possibility of the event is unknown.[100] For example, a man with a family history of coronary heart disease (CHD) who understands the

[99] Holton. Defining risk. *Financial Analysts Journal.* 2004;60(6):19-25. p.22
[100] Ibid.

hereditary propensity of CHD knows that he may eventually suffer from the disease. Another man with no understanding of the causes of CHD may never consider his likelihood of experiencing the disease. Since neither man can accurately state whether or not he will experience CHD at some point within his lifetime, uncertainty exists for both men, even if only the former perceives the uncertainty.

Exposure is the impact of an event on an individual. Unlike uncertainty, exposure is by nature subjective. For an individual to be exposed to the consequences of an event, the individual must care about the event's consequences.[101] For example, a pregnant woman may fear contracting chicken pox (infection with varicella zoster virus) due to its adverse effects on her unborn child. However, a woman who plans on becoming pregnant in the future may not mind contracting chicken pox now, so that the likelihood of contracting it again during pregnancy is dramatically reduced. In this example, greater exposure exists for the pregnant woman than for the woman planning on becoming pregnant, since the former would consider the consequences of contracting chicken pox far more severe.

One drawback to defining risk solely in terms of uncertainty and exposure is the difficulty in operationalizing the definition. Holton argues that neither uncertainty nor exposure can be defined operationally. Only *perception* of exposure and *perception* of uncertainty can be defined operationally. Consequently, "[a]t best, we can operationally define our perception of risk. There is no true risk."[102]

[101] Ibid.
[102] Ibid. p.24

Cameron and Peloso define risk (or, as Holton would likely argue, *perceived* risk) in a more straightforward fashion, saying that risk is simply the product of the impact of adverse events and the likelihood of those events occurring.[103] Based on this definition, risk in a given situation may be estimated in the form of a risk metric with estimated values for impact and likelihood. Holton argues that such risk metrics should not be evaluated based on how accurately they represent the underlying, actual risk, since true risk is an unmeasurable construct. Instead, risk metrics should be evaluated based on practical utility.[104] The implications for risk metrics used by organizations are straightforward: a risk metric is valuable if actions (or interventions) based on the metric result in a net improvement to the bottom line, whether from reduced costs or increased profit.

Risk in the Context of Patient Care

The term *high-risk patient* may have different meanings depending on the environment and patient population. For example, in hospital emergency departments, patients may be categorized as high risk (i.e., needing immediate care) or non-emergency (i.e., able to be triaged out of the emergency department) based on screenings by triage nurses using criteria like those presented by Derlet et al.[105] Within disease-specific patient subgroups, high-risk patients may be identified by the presence of known risk

[103] Cameron and Peloso. Risk management and the precautionary principle: a fuzzy logic model. *Risk Anal.* Aug 2005;25(4):901-911.
[104] Holton. Defining Risk. *Financial Analysts Journal.* 2004;60(6):19-25.
[105] Derlet, Kinser, Ray, Hamilton and McKenzie. Prospective identification and triage of nonemergency patients out of an emergency department: a 5-year study. *Ann Emerg Med.* Feb 1995;25(2):215-223.

factors, like high blood pressure in patients at risk for cardiovascular disease.[106] In both above examples, the value of utilized risk metrics (triage and disease severity scores) can be measured in terms of impact on patient health status and health care service utilization.

In a managed care setting with a diverse population of patients, high-risk patients may be defined as the small subset of patients who are expected to consume a disproportionately large share of resources.[107,108] A description of risk in terms of expected costliness allows for comparative analysis of patients who, like those in the present study's population, vary in age, gender, race, socio-economic status, and medical conditions.[109] Risk metrics based on future cost used in such populations range in complexity from a simple sum of an individual patient's prior annual health care costs to relative risk scores calculated through predictive modeling.

Actual historical cost data, when available, can be used to identify the high-cost patient subgroup within a population retrospectively. Berk and Monheit's review of the findings of the 1996 Medical Expenditure Panel Survey found that the costliest 1 percent of the population accounted for 27% of total health care expenditures, and the top 5% accounted for 55% of total health care expenditures.[110] However, from year to year, the specific individuals who make up the high-cost subgroup within a population should not

---

[106] Haskell, Berra, Arias, Christopherson, Clark, George, Hyde, Klieman and Myll. Multifactor cardiovascular disease risk reduction in medically underserved, high-risk patients. *Am J Cardiol.* Dec 1 2006;98(11):1472-1479.
[107] Zook and Moore. High-cost users of medical care. *N Engl J Med.* May 1 1980;302(18):996-1002.
[108] Harris. Disease management: new wine in new bottles? *Ann Intern Med.* May 1 1996;124(9):838-842.
[109] ICC population document – find specific source
[110] Berk and Monheit. The concentration of health care expenditures, revisited. *Health Aff* . Mar-Apr 2001;20(2):9-18.

be expected to remain constant.[111] Acute medical conditions or injuries may have high current year medical costs but low future cost implications, while chronic diseases may have low current year medical costs but high future cost implications.[112] Furthermore, the knowledge that a particular patient will have high future costs is of little practical value if the costs themselves are not preventable. Stated more conceptually, without uncertainty there is no risk to mitigate.

Exposure: Beyond Direct Medical Costs

So far, risk in the context of patient care has been discussed purely in terms of expected future direct medical costs. However, adverse health events impact more than just the pocketbooks of those paying for care. To accurately quantify the exposure to an uncertain event, one must first define the perspective of the analysis—i.e., answer the question: who or what is exposed?

From the perspective of an organization, defining exposure purely in terms of direct medical costs is appropriate. A hospital admitting a charity care patient is exposed primarily to the unreimbursed cost of the services rendered to the patient. Likewise, an insurer whose covered patient needs emergency care is exposed to the negotiated reimbursement that must be paid to the care provider.

From the patient's perspective, exposure includes out-of-pocket direct medical expenses, like copayments, deductibles, or, if the patient has no insurance, the charges

[111] Ash, Zhao, Ellis and Schlein Kramer. Finding future high-cost cases: comparing prior cost versus diagnosis-based methods. *Health Serv Res.* Dec 2001;36(6 Pt 2):194-206.
[112] Ibid.

associated with the care services. The patient is also exposed to the intangible costs of the pain and suffering associated with the adverse event, which are frequently quantified in terms of reductions to the patient's quality of life. Furthermore, a patient is also exposed to direct non-medical costs, which are costs necessary to enable patients to receive medical care. Patients also may be exposed to some indirect costs, including lost wages due to time off work.

From the perspective of the community or society, exposure includes all direct costs related to the event, plus indirect costs, which in this perspective can be generally defined as "the value of production lost to society as a result of the illness."[113]

The ICC has data that typically would be found only within an MCO. However, since the ICC is not making payments on behalf of patients and since its mission is to help improve care for the area's medically uninsured, its perspective is similar to the community perspective described above.

A valuable risk metric, in this context, would identify patients with future high utilization of the community's costly hospital services. How such a metric will be calculated is the subject of the following chapter.

[113] Koopmanschap and van Ineveld. Towards a new approach for estimating indirect costs of disease. *Soc Sci Med.* May 1992;34(9):1005-1010. p.1005

Chapter 2: Methods

This chapter discusses the applicability of various concepts, methods, and techniques described in the literature that have been applied to research similar in nature to the current project. This chapter concludes with a description of the current study's objectives, hypotheses, and data.

Methods and Techniques in the Literature

Construction of a Patient Care Risk Metric

To be consistent with Holton's recommendations, a risk metric useful in population-level patient care would be a measure of *preventable* expected future health care costs. Such a metric would identify specific patients who could be targeted for case management intervention. Bernstein elucidates this point within case management initiatives by distinguishing between high-risk patients and high-opportunity patients (i.e., those with a high likelihood of benefiting from more intensive intervention).[114]

Therefore calculating such a risk metric would be a two step process: 1) assessing patient expected future costs; and 2) assessing potential intervention benefits (i.e., patient opportunity). To accomplish step one, the researcher must identify the primary cost drivers of patient care (i.e., what makes a patient high-cost) and a method by which to estimate future costs.

[114] Bernstein. New arrows in the quiver for targeting care management: high-risk versus high-opportunity case identification. *J Ambul Care Manage.* Jan-Mar 2007;30(1):39-51.

To accomplish step two, one must find a way to sub-group the high-risk patients by their relative *intervenability*, a term coined by Knabel and Louwers that means "the potential for improving the health of an at-risk population [or individual]."[115] Knabel and Louwers recommend comparing the current care of high-risk patients with evidence-based medicine (EBM) guidelines.[116] For example, two patients with congestive heart failure may have similar predictive model-based risk scores. However, EBM analysis may show that one has recently visited a cardiologist and is taking a recommended therapeutic regimen while the other has not seen a physician in 11 months and is only taking a diuretic. In this case, the latter patient would be a more effectible target of care management since the identified gaps in care serve as a clear target for intervention.[117] This analysis can use an EBM analytic software engine or be part of the manual screening process.

The current analysis focuses on step one, above, seeking to provide a method by which to identify a subset of patients with expected high cost. Furthermore, the method should be flexible enough to work with all patients (as opposed to those with a particular disease state). In this context, the specific methods of step two would depend on the type of subgroup sought. Intervenability depends on policy, resources, and goals, both initiative-specific and long-term. Step one is an approach that can be applied population wide, while step two requires more in depth consideration of each individual patient's situation.

---

[115] Knabel and Louwers. Intervenability: another measure of health risk: by coupling predictive modeling with evidence-based medicine, health plans can identify patients who will benefit the most from care management intervention. *Health Management Technology.* 2004;25(7):36(33). p.36
[116] Ibid.
[117] Ibid.

What is Being Predicted?

Conceptually, equating patient risk to expected patient future cost is straightforward. However, operationally this definition is difficult with the uninsured who receive uncompensated care, since actual patient cost, both historical and current, is not known with the certainty that exists among insured patients whose care can be described in terms of payer reimbursements.

This operational difficulty presents two dilemmas: 1) a predictive model using historical data cannot have cost as its dependent variable if historical cost is unknown; and 2) the real-world performance of such a model cannot be tested against future cost if future cost is unknown.

Fortunately, a review of the literature provides a solution. Models used to support initiatives targeting high-cost patients generally focus on specific subsets of patients and attempt to predict hospital inpatient and emergency department utilization. Adams et al. utilized multiple logistic regression to determine the odds ratio of baseline factors associated with inpatient and emergency department hospital use in adults with asthma.[118] "Costs due to asthma are substantial, and this relates both to treatment costs and the use of health services consequent on poor disease control. A small proportion of patients with asthma who have poor disease control account for a disproportionate amount of these costs."[119] These statements, supported by the research of Weiss, Gergen, and Hodgson[120]

[118] Adams, Smith and Ruffin. Factors associated with hospital admissions and repeat emergency department visits for adults with asthma. *Thorax.* July 1, 2000 2000;55(7):566-573.
[119] Ibid. p.566
[120] Weiss, Gergen and Hodgson. An economic evaluation of asthma in the United States. *N Engl J Med.* Mar 26 1992;326(13):862-866.

and by Weiss and Sullivan,[121] explain Adams et al.'s focus on hospital utilization. The authors created three separate dichotomous dependent variables describing utilization for three separate analyses: 1) one or more inpatient admission; 2) two or more inpatient admissions; and 3) two or more emergency department visits.[122] The use of three equations allowed the researchers to identify the associations of the risk factors with the various outcomes of interest. For example, a factor that is highly predictive of emergency department use may be less predictive of inpatient admissions.

Fan et al. demonstrated how quality of life (QoL) measures can serve as predictors of hospital utilization for patients with chronic obstructive pulmonary disease (COPD).[123] Fan states his rationale for using hospitalization as a dependent variable thusly, "Because hospitalizations account for such a major portion of the cost of caring for patients with COPD, it might be helpful to identify those who are at the highest risk of being hospitalized in order to target preventive interventions."[124] Fan et al. used two inpatient-admission-based, dichotomous dependent variables describing utilization: hospitalization for any reason and hospitalization due to COPD.

Engel, Korff, and Katon attempted to identify predictor variables that would be useful in identifying high-cost back pain patients. The authors calculated the association of predictor variables (including chronic pain grade, days in pain, SCL-90 depression score) with two dependent variables: cost and utilization. The strongest predictors of cost

---

[121] Weiss and Sullivan. The economic costs of asthma: a review and conceptual model. *Pharmacoeconomics.* Jul 1993;4(1):14-30.
[122] Adams, Smith and Ruffin. Factors associated with hospital admissions and repeat emergency department visits for adults with asthma. *Thorax.* July 1, 2000 2000;55(7):566-573.
[123] Fan, Curtis, Tu, McDonell and Fihn. Using Quality of Life to Predict Hospitalization and Mortality in Patients With Obstructive Lung Diseases. *Chest.* August 1, 2002 2002;122(2):429-436.
[124] Ibid. p.430

tended to be the strongest predictors of utilization: increasing chronic pain grade, disc disorder, increasing depression, and increasing pain persistence.[125] The dependent variables describing utilization included: two or more back pain primary care visits, one or more back pain specialty visits, and one or more back pain inpatient admissions.

Beddhu et al. used a modified Charlson Comorbidity Index to predict high-cost dialysis patients.[126] The authors' previous research showed that among dialysis patients "costs related to hospitalization were responsible for 37% to 40% of total costs, and that 25% of patients accounted for 50% of total costs and 42% of deaths."[127, 128] Their comorbidity index showed strong association with both utilization and costs.[129] For their utilization dependent variable, the researchers used the number of inpatient days.

The wide usage of hospital utilization as a dependent variable in research attempting to predict high-cost patients suggests that some measure of utilization would be a suitable dependent variable for analysis in the uninsured when cost data are unavailable or non-existent. The specific dependent variable used in such analysis would depend on the underlying data available and the distribution of the various utilization measures.

Diehr et al. suggest that while ordinary least squares (OLS) regression is often used to analyze utilization data, the data fail to satisfy the method's assumptions of

---

[125] Engel, von Korff and Katon. Back pain in primary care: predictors of high health-care costs. *Pain.* May-Jun 1996;65(2-3):197-204.
[126] Beddhu, Bruns, Saul, Seddon and Zeidel. A simple comorbidity scale predicts clinical outcomes and costs in dialysis patients. *Am J Med.* Jun 1 2000;108(8):609-613.
[127] Ibid. p.609
[128] Bruns, Seddon, Saul and Zeidel. The cost of caring for end-stage kidney disease patients: an analysis based on hospital financial transaction records. *J Am Soc Nephrol.* May 1998;9(5):884-890.
[129] Beddhu, Bruns, Saul, Seddon and Zeidel. A simple comorbidity scale predicts clinical outcomes and costs in dialysis patients. *Am J Med.* Jun 1 2000;108(8):609-613.

normality, homoscedasticity, and independence. Counts of visits or admissions could be analyzed with Poisson models, but the variance will usually be underestimated. Diehr et al. state: "Health care utilization variables are usually not normally distributed, as they tend to have a mode at zero and distribution with a long, heavy right tail. The distribution often looks more like a lognormal than a Poisson distribution, even when the data are counts." For this reason, "utilization data are often transformed to the log scale."[130]

However, using a transformation of utilization has its drawbacks. In linear regression, non-normal dependent variables may be transformed to conform to the normality and constancy of variance requirements of the general linear model. In addition to the confounding element of discussing relationships between the predictor variables and a transformed response variable, a single transformation may not exist that "yields simultaneously additive effects in the systematic part of the model and constancy of variance for the random part."[131] The generalized linear model allows the use of a non-normal dependent variable and allows using non-constant variance if the variance function is known or can be estimated.[132] The section *Generalized Linear Model*, later in this chapter, discusses in detail a common method of performing regression with non-normally distributed outcome variables.

[130] Diehr, Yanez, Ash, Hornbrook and Lin. Methods for analyzing health care utilization and costs. *Annu Rev Public Health.* 1999;20:125-144. p.129
[131] McCulloch and Nelder. *Generalized Linear Models*. 2nd ed. Cambridge: Chapman & Hall; 1991. p.378
[132] Nelder and Wedderburn. Generalized linear models. *J. R. Statist. Soc.* 1972;A135:370-384.

Predictor Variables

With hospital utilization selected as a reasonable dependent variable for identifying future high-cost uninsured patients, independent (i.e., predictor) variables need to be identified for potential inclusion in the model.

The study in the literature that most closely resembles the current study was performed by Radcliff et al. The researchers attempted to use administrative data at a safety net hospital to look for key identifying characteristics among historically high-cost uninsured patients to enable *a priori* identification of future high-cost cases.[133] The researchers attempted to use multivariate regression to predict membership in the high-cost patient subgroup with little success. While the researchers had encounter-level data on over 30,000 patients, they had no method of summarizing the diagnostic data (in the form of ICD-9 codes) into a format useable in a multivariate regression model. Consequently, their models, including only location, age, and sex data, had little predictive capability and practical utility. The researchers attempted to pick specific diagnoses in the high-cost patient group and add these variables to their models; however, little additional predicted value was added.

Although Radcliff et al. had access to uncompensated charge data for the patients in their study which could be used as a dependent variable, their available independent variables are similar to those available in the current project. Diehr et al. provide an extensive review of methods of analyzing utilization which cover age and sex as

[133] Radcliff, Cote and Duncan. The identification of high-cost patients. *Hosp Top.* Summer 2005;83(3):17-24.

independent variables.[134] Diehr et al. use data from Washington State's Basic Health Plan to demonstrate the complex relationship between cost and age. Mean cost by age is nonlinear and differs by sex. "Very young children have high utilization. Males and females have similar utilization until puberty, at which time women increase their utilization because of childbearing. Men's utilization is low until about age 40."[135] Various methods of adjusting for age are shown in Table 2.1.

Table 2.1 Cost Prediction Model Types that Adjust for Age

Type	***Model***	***Relationship between cost and age***
Linear	cost = a + b(sex) + c(age)	linear, with parallel regression lines for males and females
Linear Interactive	cost = a + b(sex) + c(age) + d(age)(sex)	linear, with different slopes for males and females
Quadratic	cost = a + b(sex) + c(age) + d(age)(sex) + e(age^2) + f(age^2)(sex)	quadratic, different for males and females
Source: adapted from Diehr et al.		

Bernstein suggests that identifying patients with outlier utilization history can be "an important complement to other approaches to risk stratify a population."[136] However, models that overly rely on historical outlier utilization or costs may yield unstable predictions due to regression toward the mean.[137]

Diagnosis data, in the form of ICD-9 codes, provides valuable clinical information about the health status of patients and, therefore, would be valuable as

[134] Diehr, Yanez, Ash, Hornbrook and Lin. Methods for analyzing health care utilization and costs. *Annu Rev Public Health.* 1999;20:125-144.
[135] Ibid. p.134
[136] Bernstein. New arrows in the quiver for targeting care management: high-risk versus high-opportunity case identification. *J Ambul Care Manage.* Jan-Mar 2007;30(1):39-51. p.41
[137] Meenan, O'Keeffe-Rosetti, Hornbrook, Bachman, Goodman, Fishman and Hurtado. The sensitivity and specificity of forecasting high-cost users of medical care. *Med Care.* Aug 1999;37(8):815-823.

independent variables in a model predicting utilization or cost. However, as was seen in the Radcliff et al. study, ICD-9 codes are awkward to work with due to their large numbers. The over 14,000 codes cannot be used as binary independent variables, because of the implications on sample size. Using the general recommendation of at least 20 observations (patients in this case) per independent variable in linear regression, the minimum sample size would be 280,000 patients, or almost ten times the sample size used in the Radcliff et al. study.

Fortunately, several popular diagnosis-based risk adjustment models can supply a means of summarizing these data into a single measure of health status (including age and sex). One such model, Diagnostic Cost Groups, has been evaluated in the current study population.[138] Its output (or similar output from an alternative model) could have been used by Radcliff et al. as a means of incorporating health status in their prediction model. Although the Diagnostic Cost Groups model was initially developed to help calculate capitation for Medicare recipients, its risk-adjustment technique is currently used to help identify future high-cost patients for intervention, as discussed in the following section.

Creating a Single Measure of Patient Health Status with a Diagnosis-Based Risk Adjustment Model

The 1982 Tax Equity and Fiscal Responsibility Act allowed for risk contracts with a prospective payment system for Medicare patients to be utilized by health

[138] Leslie R. Cost Analysis of the Uninsured, Indigent Population of Central Texas. Master's Thesis. The University of Texas at Austin. 2006.

maintenance organizations (HMOs) in place of the previous fee for service (FFS) payment system.[139] At the time, the adjusted average per capita cost (AAPCC) method, which adjusted for age, sex, welfare status, and institutional status, was used to estimate the yearly cost of each patient on Medicare in each county.[140] Using this method, the risk contracts were intended to provide HMOs with 95 percent of what the enrollees would have cost under FFS in that county. However, several researchers found that the AAPCC failed to adjust capitation levels based on systematic differences in health status between enrollee groups.[141] This shortcoming led researchers to begin proposing alternatives to the AAPCC that incorporate prior use data. During the following two decades of use and refinement, early models became increasingly sophisticated and new models were proposed. Therefore, researchers needing a diagnosis-based risk adjustment model now have several from which to choose.

Previous research by Leslie has demonstrated the utility of the Diagnostic Cost Groups (DCG) models for profiling and describing the ICC patient population.[142] Leslie also demonstrated how the DCG models could be used to identify high-risk patients for case management intervention.[143] Therefore, the DCG models will be used again in this study.

---

[139] Ash, Porell, Gruenberg, Sawitz and Beiser. Adjusting Medicare capitation payments using prior hospitalization data. *Health Care Financ Rev.* Summer 1989;10(4):17-29.
[140] Ibid.
[141] Ibid.
[142] Leslie R. Cost Analysis of the Uninsured, Indigent Population of Central Texas. Master's Thesis. The University of Texas at Austin. 2006.
[143] Ibid.

DCG Models: Background

The DCG models use linear regression to predict health care expenditures based on demographic and diagnostic data. Ash et al. aggregated the diagnosis codes from the International Classification of Diseases System Ninth Revision (ICD-9) into subgroups designed to pool diagnoses judged clinically similar. Some subgroups were merged when too few episodes (less than 100) existed in the researchers' dataset. The resulting 78 subgroups used in the original models were then aggregated into nine DCGs. The DCGs were designed to group different types of diagnoses with similar costs. Ash et al.'s analysis demonstrated that the early DCG models showed promise, with higher predictive ratios than AAPCC. However, prior use models with administrative data from the Health Care Financing Administration's Health Insurance Master Accretions file showed higher predictive ratios than the DCG models due to their improved ability to predict costs on patients with no hospital utilization data.[144]

The initial DCG models focused only on the patient's principal diagnosis, which Iezzoni says is insufficient.[145] Secondary and tertiary diagnoses are often not only necessary to determine the severity of the patient's condition, but are critical in identifying co-morbid conditions that influence future patient encounters. The initial DCG models also focused only on inpatient encounters.

The DCG models have since gone through several refinements, including the addition of minimum stay periods, a more complex and sophisticated method of grouping

[144] Ibid.
[145] Iezzoni L. *Risk Adjustment for Measuring Health Care Outcomes*. 3rd ed; 2003. p.48

diagnoses, and the ability to distinguish "discretionary" hospitalizations,[146] but the most notable refinements of the models came in 1996 when Ellis et al. added the ability to include diagnoses from additional types of encounters (outpatient, ambulatory, etc.) and hierarchical categorization of patients based on secondary and tertiary diagnoses. The revised models are now referred to as Diagnostic Cost Group / Hierarchical Coexisting Conditions (DCG/HCC) models.[147] In 1998, Ash and Ellis adapted the DCG/HCC models for use with both privately-insured populations and Medicaid recipients.[148]

In 2000, the Centers for Medicare and Medicaid Services (CMS) began using the principal inpatient DCG (PIP-DCG) models for Medicare health plan reimbursement based on predicted health-based medical needs of beneficiaries.[149] In 2004, CMS began using the all encounter DCG model.[150]

Along with prior-use prescription drug models (RxGroups), the DCG models are now part of a software package called RiskSmart, which is owned by DxCG, Inc., of Boston, Massachusetts.[151]

Zhao et al. proposed using the DCG methodology to understand the prevalence of diseases in populations and the associated planning necessary to minimize costs.[152] This

---

[146] Ellis and Ash. Refinements to the Diagnostic Cost Group (DCG) model. *Inquiry.* Winter 1995;32(4):418-429.
[147] Ellis, Pope, Iezzoni, Ayanian, Bates, Burstin and Ash. Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financ Rev.* Spring 1996;17(3):101-128.
[148] Ash, Ellis, Yu and et al. *Risk Adjustment for the Non-Elderly.* Boston University. Boston, MA: Final report submitted to the Health Care Financing Organization under cooperative agreement No. 18-C-90462/1-02; Jun 1998.
[149] Pope, Ellis, Ash, Liu, Ayanian, Bates, Burstin, Iezzoni and Ingber. Principal inpatient diagnostic cost group model for Medicare risk adjustment. *Health Care Financ Rev.* Spring 2000;21(3):93-118. p.94
[150] DxCG Inc. *DxCG RiskSmart Models and Methodologies Guide*; Nov 2002.
[151] Zhao, Ash, Ellis, Ayanian, Pope, Bowen and Weyuker. Predicting pharmacy costs and other medical costs using diagnoses and drug claims. *Med Care.* Jan 2005;43(1):34-43.

approach differs from traditional disease management, because high risk individuals are identified based on their inclusion in expensive Cost Categories, which consider combinations of co-morbidities rather than simply the severity of specific diseases. Using year-one data, the researchers compared the DCG-based categorization with disease-specific severity (DSS) and prior cost categorization for predicting high year-two costs for Medicaid and commercially insured diabetic patients with administrative data. The DCG-based method showed a higher R^2 than those of the DSS and prior cost methods (12.19% vs. 4.74% and 10.49%, respectively).[153]

Zhao et al.'s method is especially appealing for use with the ICC data because both the information necessary to categorize patients based on DSS (e.g., HbA1C level) and prior-cost data are unavailable. Leslie has demonstrated how this method can be used with the ICC population.[154]

---

[152] Zhao, Ash, Ellis and Slaughter. Disease burden profiles: an emerging tool for managing managed care. *Health Care Manag Sci.* Aug 2002;5(3):211-219.
[153] Ibid.
[154] Leslie R. Cost Analysis of the Uninsured, Indigent Population of Central Texas. Master's Thesis. The University of Texas at Austin. 2006.

DCG/HCC Model Methodology[155]

The set of ICD-9 diagnosis codes provides a detail that is necessary for administrative recordkeeping yet is too granular for conducting analysis. The DCG model maps each ICD-9 code to one of 781 clinically-related DxGroups except in cases where the cost profile for a diagnosis differs dramatically for pediatric and adult patients. In these cases, the ICD-9 code will map to one DxGroup for pediatric patients and a different DxGroup for adult patients. Patients with multiple diagnosed conditions will fall into multiple DxGroups.

Aggregating over 14,000 ICD-9 codes into 781 DxGroups lets patients with related diagnoses be grouped together so that the population is more easily described. Each of the 781 DxGroups is then mapped to one of 184 clinical summary Condition Categories (CCs). Patients with multiple diagnosed conditions will generally have multiple CCs.

The CCs all fall into Disease Hierarchies, which are used to avoid over-categorization of patients and coding idiosyncrasies. For example, a schizophrenic patient (CC 54) cannot be additionally categorized as clinically depressed (CC 58). Similarly, a patient with cystic fibrosis (CC 107) cannot be additionally categorized as asthmatic (CC 110) but can be additionally categorized by viral pneumonia (CC 113).

The CCs have associated cost-weights which are used in the calculation of both relative risk (RR) scores and estimated costs for patients. The models have sets of cost-

[155] DxCG Inc. *DxCG Risk Smart Stand Alone Version 2.2 Help*; 2007.

weights for three patient population types: Medicare, Medicaid, and Commercial. The cost-weights were generated using multivariate linear regression from the following large databases:

- Medicare: 5% national Medicare database sample, years 2000-2002, with over 2 million beneficiaries;
- Medicaid: Data on 2 million Massachusetts Medicaid patients for the fiscal years 2003 through 2005, with results validated against a million enrollees from another state's program; and
- Commercial: Fee-for-service and managed care data from the 2000-2003 MEDSTAT Group, Inc., MarketScan databases, consisting of 1.8 million lives for the prospective model and 6 million lives for the concurrent model.

As an alternative to using the supplied cost-weights for the CCs, new weights can be calculated using data from the population of interest.

The DCG models perform both concurrent and prospective analysis using ordinary least squares (OLS) regression, with CCs and demographic data (e.g., age, sex) as independent variables. The concurrent models are additive, summing the effects of each clinical and demographic classification to explain current year costs. Episodic conditions such as appendicitis receive cost weights in concurrent models. However, while episodic conditions reflect current year costs, they are not indicative of future conditions or costs. Therefore, the prospective models do not assign cost-weights to episodic conditions. The prospective models also include risk adjustments based on interactions between clinical groupings. The DxCG Models and Methodologies guide provides the example of the interaction of Diabetes with Renal Manifestations, Congestive Heart Failure, and Renal Failure CCs in a Medicare patient. In the

prospective model, this patient would have a higher RR score to capture the elevated risk associated with the combination of these CCs.

The DCG models calculate a relative risk (RR) for each patient that is a multiple of the population's benchmark per capita cost. A patient with a RR of 21 would exhibit (or be expected to exhibit, in the prospective model) approximately 21 times the cost of the benchmark patient. The benchmark per capita costs are listed in Table 2.2. The costs in Table 2.2 can be adjusted to fit the population of interest.

Table 2.2 Predictive and Concurrent Benchmark Per Capita Costs for Medicare, Medicaid, and Commercial DCG Models

	Model Type	Benchmark Per Capita Cost
Concurrent*	Medicare	$6,477
	Medicaid Fee For Service	$12,066
	Medicaid Managed Care	$4,688
	Commercial	$2,687
Prospective**	Medicare	$6,485
	Medicaid Fee For Service	$13,143
	Medicaid Managed Care	$4,614
	Commercial	$2,710

*The concurrent model estimates risk and costs in the current year.

**The prospective model estimates risk and costs in the following year.

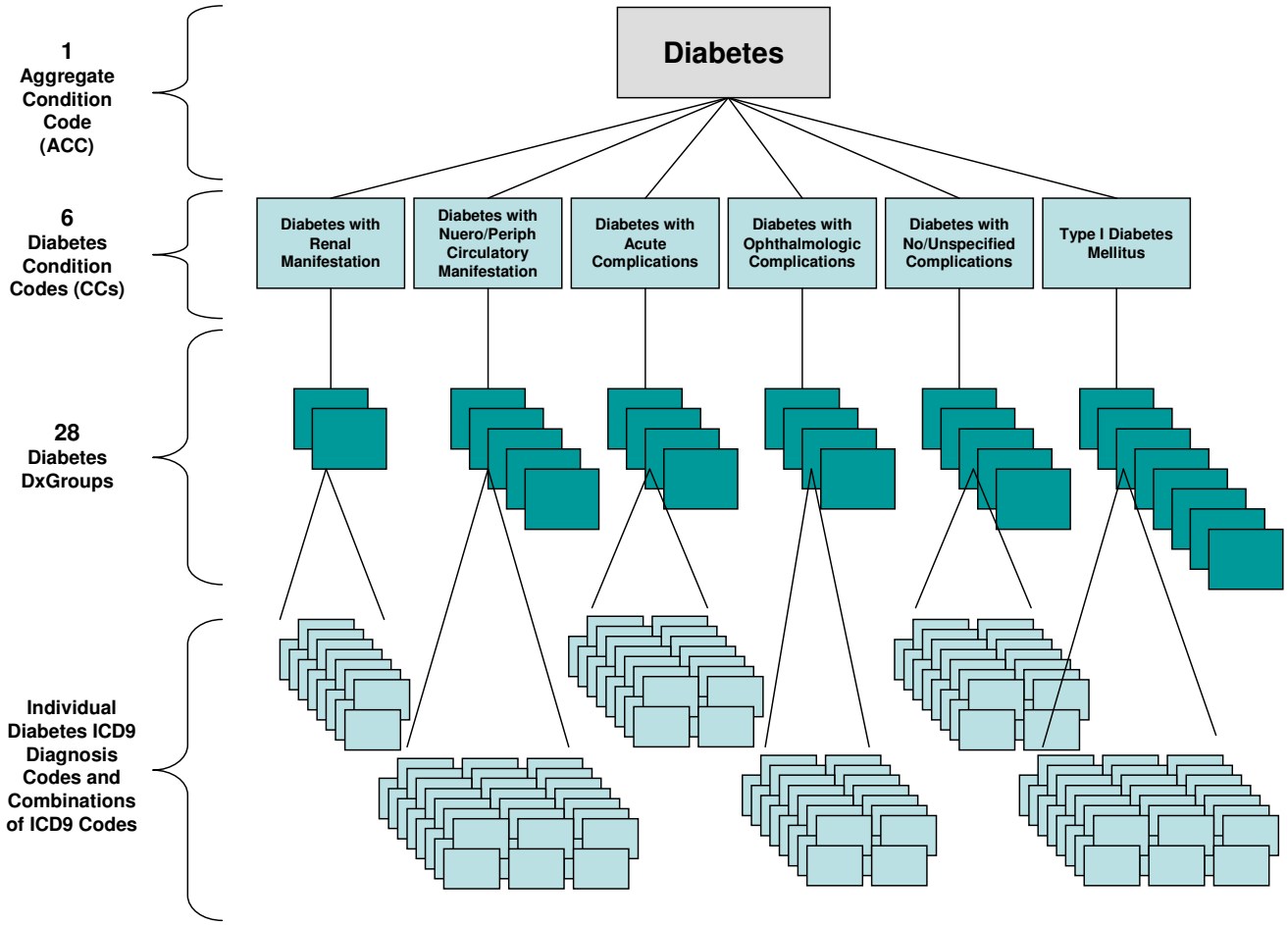
Source: DxCG Inc. DxCG RiskSmart Stand Alone Help; 2007.

The benchmark costs represent the mean cost of the population of patients used to calibrate each model. Therefore, the RR scores produced by each model can be multiplied by the benchmark per capita cost to produce a patient-specific estimated cost.

Once the models have produced RR scores for each patient, further adjustments can be made. If default model calibrations are used, the resulting RR scores are unlikely to have a mean of 1.0. Dividing all patient RR scores by the mean RR score of the population will result in population-specific RR scores. Although not used in calculating RR scores, a higher level categorization of 30 Aggregate Category Codes (ACCs) can also be used for general patient profiling or presentation of summarized data.

Figure 2.1 provides an example of the DCG model using diabetes as the disease. It shows the categorization hierarchy of ICD9s, DxGroups, CCs, and ACCs.

Figure 2.1 Depiction of DCG Model Patient Categorization - Diabetes Aggregate Condition Code



The DCG Relative Risk Score as a Predictor Variable

The relative risk score produced by the DCG model could be used as a predictor variable as a single measure of patient health status. This is a parsimonious alternative to creating binary variables for specific conditions based on ICD-9 codes, and its hierarchical structure incorporates diagnosis interaction effects that would be otherwise difficult to model.

Subjective Prediction: Physician Referrals

Some patients may be specifically singled out by physicians as candidates for case management. Although recommended patients deserve consideration, a physician recommendation alone should not mean a guaranteed inclusion in a case management program, since the physician likely lacks the information necessary to determine if an individual patient is more worthy than other identified patients in a population when scarce resources limit program enrollment. Physician referrals, therefore, should supplement other case identification methods.

Physician referrals also present a dilemma when designing a method of testing the impact of an intervention program. If historical controls are used, no historical patients will have had physician recommendations. If a prospective, randomized trial is employed with patients in both groups matched on key attributes, too few referred patients may exist to ensure equal group distribution. Referred patients may therefore need to be excluded from studies measuring program effectiveness. If sufficient referred patients are

available within the study data, a binary “referred” variable could be included in the set of predictor variables.

Model Type: One-Part Versus Two-Part

The following table shows the dependent and independent variables identified for use in predicting high-utilization uninsured patients.

Table 2.3 Variables Used to Predict High-Cost Uninsured

Variable	Variable Type	Format(s)
Hospital Utilization	Dependent (year 2) and Independent (year 1)	Count of admissions, count of emergency department visits, count of inpatient days, binary coding of patient grouping (high/low utilizer)
Age*	Independent	Numeric age
Sex*	Independent	Binary
Diagnosis data*	Independent	Must be summarized into a smaller number of high-level binary groups (e.g., Diabetes)
*By using a diagnosis-based risk adjustment model that incorporates age and sex, these variables can be replaced with a single measure of health status, generally in the form of a relative risk score.		

Once the dependent and independent variables have been identified, the actual model type must be determined. The simplest option is a one-part model, which is fit for all patients, irrespective of whether they have zero utilization. To accommodate non-normal distributions, modeling may be performed with Poisson, gamma, or negative-binomial distributions.[156]

A two-part model uses two equations. The first predicts whether a patient will have any utilization, often using logistic regression. The second equation takes those

[156] Diehr, Yanez, Ash, Hornbrook and Lin. Methods for analyzing health care utilization and costs. *Annu Rev Public Health.* 1999;20:125-144.

patients with predicted utilization and predicts the magnitude of the utilization. Two-part models can be especially useful for estimating costs, because the relationship between the independent variables and the choice to seek care may be different from the relationship between the independent variables and the magnitude of the cost of care. Patient viewpoints may dictate the former and health care infrastructure may dictate the latter.[157]

Blough et al. describe specifically how a two-part model could utilize patient demographic data and a case-mix risk adjustment model variable (provided by Ambulatory Care Groups[158], an alternative to the DCG models) to predict patient risk. The authors use logistic regression to predict whether a patient will incur medical expenses followed by another form of the generalized linear model to predict the expenses of all utilizers identified in step 1.[159]

Generalized Linear Model

The general linear model (GLM) is the basis for ordinary least squares regression and analysis of variance (ANOVA). The GLM suggests that a linear relationship exists between the dependent variable (y) and one or more independent variables (x_i) as shown in the following equation:

$$y_i = x_i'\beta + \varepsilon_i$$

[157] Ibid.
[158] Weiner, Starfield and Lieberman. Johns Hopkins Ambulatory Care Groups (ACGs). A case-mix system for UR, QA and capitation adjustment. *HMO Pract.* Mar 1992;6(1):13-19.
[159] Blough, Madden and Hornbrook. Modeling risk using generalized linear models. *Journal of Health Economics.* 1999/4 1999;18(2):153-171.

where y_i is the response (dependent) variable for the i-th observation, x_i' is the vector of independent variable values, β is the vector of coefficients, and ε_i represents independent, normally distributed error.

The GLM has seen wide usage in social and behavioral science research since the publication of Cohen's 1968 article discussing its applications.[160] However, in cases where the dependent variable is non-normally distributed or the relationship between the dependent and independent variables is non-linear, the GLM may fail to adequately explain the relationships between variables.[161]

The general*ized* linear model extends the GLM framework to exponential distribution functions via a link function *g(…)* as shown:

$$y_i = g(x_i'\beta) + \varepsilon_i$$

The link function manipulates the linear predictor to fit the expected value of *y*. By using the identify function, f(x) = x, as the link function, the generalized linear model functions as the GLM, following the normal distribution. For this reason, the GLM can be viewed as a subset of the umbrella concept of the generalized linear model, with the normal distribution simply one of the exponential distributions available via an identity link function. Although the link function can be any function that fits the linear predictor to the dependent variable, certain link functions are used most frequently because of the common occurrence of their corresponding univariate exponential distributions. These

[160] Cohen. Multiple regression as a general data-analytic system. *Psychological Bulletin.* 1968;70:426-443.
[161] McCulloch and Nelder. *Generalized Linear Models*. 2nd ed. Cambridge: Chapman & Hall; 1991.

'canonical' links exist for the normal, binomial, Poisson, and gamma distributions, among others.[162]

For binomial dependent variables, the logit is used as the generalized linear model's link function. This form of the generalized linear model, called logistic regression, is quite common in outcomes research when the dependent variable is an odds ratio.[163,164,165,166]

By using the log as the link function, the generalized linear model can conform to the Poisson distribution, which may be useful when the dependent variable represents count data or as an alternative to logistic regression when the preferred parameter of interest is relative risk[167,168] or incident rate ratio.[169]

The gamma distribution may be modeled by using the reciprocal as the link function. Using the reciprocal imposes the constraint that predicted values must be non-negative. Gamma regression is common with dependent variables like hospital costs, the

---

[162] Ibid.
[163] Pradhan, Manson, Rossouw, Siscovick, Mouton, Rifai, Wallace, Jackson, Pettinger and Ridker. Inflammatory Biomarkers, Hormone Replacement Therapy, and Incident Coronary Heart Disease: Prospective Analysis From the Women's Health Initiative Observational Study. *JAMA.* August 28, 2002 2002;288(8):980-987.
[164] Paterson, Swindells, Mohr, Brester, Vergis, Squier, Wagener and Singh. Adherence to Protease Inhibitor Therapy and Outcomes in Patients with HIV Infection. *Ann Intern Med.* July 4, 2000 2000;133(1):21-30.
[165] Crystal, Sambamoorthi, Walkup and Akincigil. Diagnosis and treatment of depression in the elderly Medicare population: predictors, disparities, and trends. *Journal of the American Geriatrics Society.* Dec. 2003;51(12):1718.
[166] Frasure-Smith, Lesperance and Talajic. Depression and 18-month prognosis after myocardial infarction. *Circulation.* Feb 15 1995;91(4):999-1005.
[167] McNutt, Wu, Xue and Hafner. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. *Am. J. Epidemiol.* May 15, 2003 2003;157(10):940-943.
[168] Zou. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol.* Apr 1 2004;159(7):702-706.
[169] Fihn, Callahan, Martin, McDonell, Henikoff and White. The Risk for and Severity of Bleeding Complications in Elderly Patients Treated with Warfarin. *Ann Intern Med.* June 1, 1996 1996;124(11):970-979.

distribution of which are positive with a long right tail, indicating a small number of cases with extreme scores.[170,171]

Unlike with the GLM, generalized linear models do not require constancy of variance and instead use a variance function. The natural link functions and variance functions of commonly used exponential distributions are shown in the following table.

Table 2.4 Generalized Linear Model Link and Variance Functions of Commonly Used Exponential Distributions

Distribution	**Natural Link Function**	**Variance Function**
Gaussian	M	1
Binomial	$\text{Log}\{\mu/(1-\mu)\}$	$\mu\,(1-\mu)\,/\,n$
Poisson	$\text{Log}(\mu)$	M
Gamma	$1/\mu$	μ^2

Source: Adapted from Blough et al.

Measuring the Goodness of Fit of a Model

The choice between a one-part and two-part model may be made by comparing each model's goodness of fit. The smaller the deviance, the better the model fit. McCullagh and Nelder suggest using residuals (i.e., the difference between the data values and the predicted values) to determine how adequately a model fits the data. "Residuals can be used to explore the adequacy of fit of a model, in respect of choice of

[170] Gilmer, Dolder, Lacro, Folsom, Lindamer, Garcia and Jeste. Adherence to Treatment With Antipsychotic Medication and Health Care Costs Among Medicaid Beneficiaries With Schizophrenia. *Am J Psychiatry.* April 1, 2004 2004;161(4):692-699.
[171] Simon, Manning, Katzelnick, Pearson, Henk and Helstad. Cost-effectiveness of Systematic Depression Treatment for High Utilizers of General Medical Care 10.1001/archpsyc.58.2.181. *Arch Gen Psychiatry.* February 1, 2001 2001;58(2):181-187.

variance function, link function and terms in the linear predictor."[172] The residual sum of squares can be used for least squares models. [173] Various similar statistics exist for other distributions.[174] Blough et al. provide the following deviance functions for use with generalized linear models using common distribution functions.[175]

Table 2.5 Deviance Functions for Exponential Distributions Commonly Used in Generalized Linear Models

Distribution	**Deviance Function**
Gaussian	$2\Sigma_i w_i(y_i - \mu_i)^2$
Binomial	$2\Sigma_i w_i m_i\{y_i \log(y_i / \mu_i) + (1 - y_i)\log((1 - y_i)/(1 - \mu_i))\}$
Poisson	$2\Sigma_i w_i\{y_i \log(y_i / \mu_i) - (y - \mu_i)\}$
Gamma	$2\Sigma_i w_i\{-\log(y_i / \mu_i) - (y - \mu_i) / \mu_i\}$
Source: Adapted from Blough et al.	

[172] McCulloch and Nelder. *Generalized Linear Models*. 2nd ed. Cambridge: Chapman & Hall; 1991. p.37
[173] Diehr, Yanez, Ash, Hornbrook and Lin. Methods for analyzing health care utilization and costs. *Annu Rev Public Health.* 1999;20:125-144.
[174] Davison and Gigli. Deviance residuals and normal scores plots. *Biometrika.* June 1, 1989 1989;76(2):211-221.
[175] Blough, Madden and Hornbrook. Modeling risk using generalized linear models. *Journal of Health Economics.* 1999/4 1999;18(2):153-171.

Cumulative Accuracy Profiles and Accuracy Ratios

Another method of evaluating the fit and performance of predictive models can be borrowed from the credit industry, where the risk of loan default is analogous to the risk of high-cost hospitalization. In 1999, Keenan and Sobehart proposed a method of assessing the utility of credit risk models.[176] This method involves the construction of cumulative accuracy profiles (CAPs), a type of chart that displays how well the risk scores calculated by predictive models correspond in a usable fashion with actual outcomes.

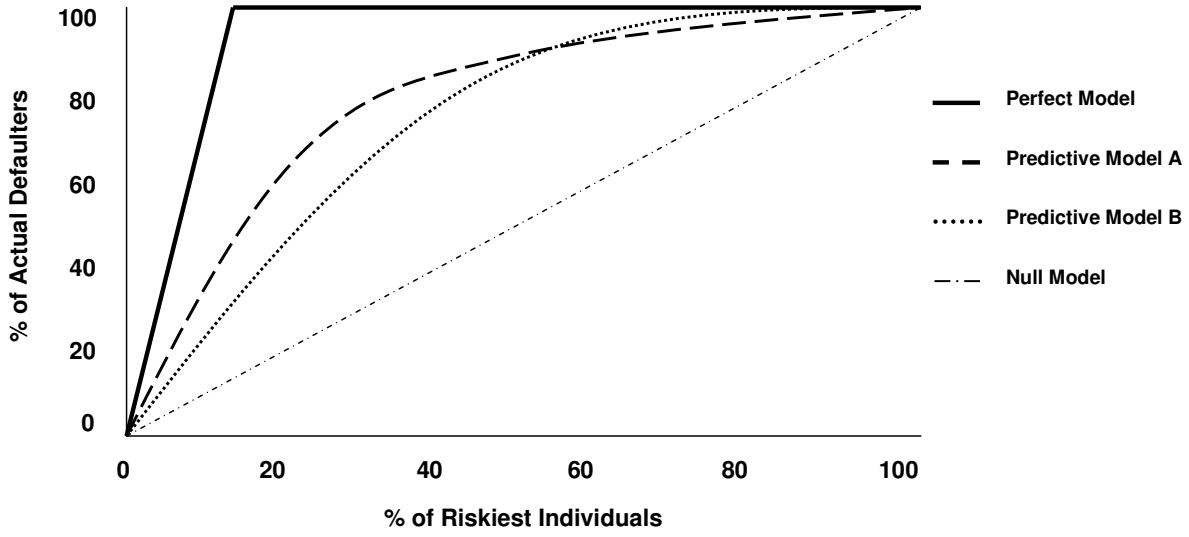
To construct a CAP, risk scores must first be assigned to each individual, and the individuals are then ranked in descending order of risk score. This ranking process would occur for each predictive model being evaluated. The CAP answers the following question: How many actual defaulters are contained in the subset of individuals with the *x*% highest risk scores? For example, if the top 50% riskiest individuals contained only 50% of actual defaulters, the performance of the risk model, in this context, is no better than assigning risk scores randomly. However, if the top 10% riskiest individual contained 80% of actual defaulters, the risk model performs much better than random risk assignment and may be useful.

Figure 2.2 shows an example CAP as it would be used in the credit industry, with two ersatz predictive models. The CAP curves answer the question described above for each value of *x* between 0 and 100. To aid in interpreting the curves, two additional

[176] Keenan SC, Sobehart JR. Performance measures for credit risk models, Research report# 1-10-10-99. *Risk Management Services, Moody's Investors Service*. 1999.

curves are added for context: a null model curve, which shows the performance of random risk assignment, and a perfect model curve, which shows the performance of a model that perfectly predicts actual defaults.

Figure 2.2 Example Cumulative Accuracy Profile (CAP) Chart



The area between the null model and the curve of a particular predictive model shows the performance of that model above and beyond random chance. By dividing this area by the area below the perfect model but above the null model, a model's CAP score can be described on a one-point scale, with a score of one representing perfect performance. Keenan and Sobehart label this score the accuracy ratio (AR) of the model.

To use this concept in the context of the current analysis, "defaulters" may be substituted for actual high-use patients. Therefore, a CAP would show how many actual

high-use patients are contained in the subset of individuals with the $x\%$ highest risk scores.

The main advantage of the CAP curve is that it provides a visual way to evaluate the performance of predictive models. However, no method currently exists to determine if the differences between CAP curves or AR scores are significant. Therefore, these tools may complement other methods of model evaluation but do not provide a complete evaluation on their own.

The Utilization Method

The ICC's proposed method of selecting candidate patients for care management focuses on hospital utilization data. Patients are considered candidates for care management if, during the previous twelve months, they satisfy at least one of the two utilization requirements and at least one of the chronic disease requirements.

The following are the single-year utilization requirements:

- the patient has at least one inpatient admission, and
- the patient has 12 or more emergency department (ED) encounters.

The chronic disease requirements include inpatient or ED diagnoses in ICD9 codes that roll up into any of the following summary categories defined by the Agency for Healthcare Research and Quality (AHRQ):

- diabetes,
- hypertension,
- congestive heart failure,
- asthma,
- obesity,
- and tobacco abuse.

Furthermore, any patient recommended by a clinician for inclusion in the care management program will automatically become a candidate. However, referral patients will not be considered as part of this retrospective analysis since no referrals have yet occurred.

Specific Methods of this Study

Objectives and Hypotheses

The purpose of this study was to develop and test predictive models used to identify candidate indigent, uninsured patients for care management intervention within the Austin, Texas, metropolitan area. This study is separated into two parts: the first part is the construction of the predictive models, and the second is the comparison of the predictive models and the ICC's current proposed method of selecting patients, the Utilization Method.

Objective 1

The first objective was to create and compare models M1 – M5 that used the various methods outlined below to predict year-two and -three utilization with year-one data.

M1: ordinary least squares multivariate regression model with 2003 utilization and DCG RR score as independent variables and the sum of 2004 and 2005 inpatient admissions as the dependent variable

M2: ordinary least squares multivariate regression model with 2003 utilization, age, and sex data as independent variables and the sum of 2004 and 2005 inpatient admissions as the dependent variable

M3: ordinary least squares multivariate regression model with 2003 utilization as the independent variable and the sum of 2004 and 2005 inpatient admissions as the dependent variable

M4: generalized linear multivariate regression model with 2003 utilization and DCG RR score as independent variables and the sum of 2004 and 2005 inpatient admissions as the dependent variable

M5: two-step model utilizing logistic regression to predict utilization and generalized linear regression to predict magnitude of utilization with 2003 utilization and DCG RR score as independent variables and the sum of 2004 and 2005 inpatient admissions as the dependent variable

Table 2.6 Description of Models Used in Study

Model	**Type**	**DV**	**IVs**
M1	Ordinary least squares	sum of 2004 and 2005 inpatient admissions	2003 inpatient admissions, DCG RR score
M2	Ordinary least squares	sum of 2004 and 2005 inpatient admissions	2003 inpatient admissions, age, sex
M3	Ordinary least squares	sum of 2004 and 2005 inpatient admissions	2003 inpatient admissions
M4	Generalized linear model (Poisson)	sum of 2004 and 2005 inpatient admissions	2003 inpatient admissions, DCG RR score
M5	2 step generalized linear model (logistic + Gamma/Poisson)	sum of 2004 and 2005 inpatient admissions	2003 inpatient admissions, DCG RR score

Hypotheses for Objective 1:

H1: Model fit will vary between the five models.

H2: Models M1, M4, and M5 will have lower deviance residuals than models M2 and M3.

H3: Model M1 will have a higher R^2 than the square of the Pearson product moment correlation between the DCG RR scores and the sum of 2004 and 2005 inpatient admissions.

H4: Model M4 will have lower deviance residuals than model M1.

H5: Model M5 will have the lowest deviance residuals of all models.

Objective 2:

The second objective was to test whether the predictive models (M1-M5) identified patients with a greater number of 2005 and 2006 inpatient admissions than those patients identified by the Utilization Method. The Utilization Method was used on 2004 data to identify the highest risk patients. Models M1-M5 also used 2004 data to identify the highest risk patients. Each model's respective highest risk patients were the patients with the top approximately 2 percent of predicted utilization scores. The Utilization Method identified a smaller subset of high-risk patients due to its query-like approach.

Hypotheses for Objective 2:

H6: The Utilization Method and Models M1-M5 will identify different patients as highest-risk.

H7: The number of 2005 and 2006 inpatient admissions of the *j* patients identified by Models M1, M4, and M5 will be higher than the number of 2005 and 2006 inpatient admissions of the *j* patients identified by models M2 and M3.

H8: The data-driven models (M1 – M5) will identify case management candidates with a higher mean number of comorbid chronic conditions than those candidates identified by the Utilization Method.

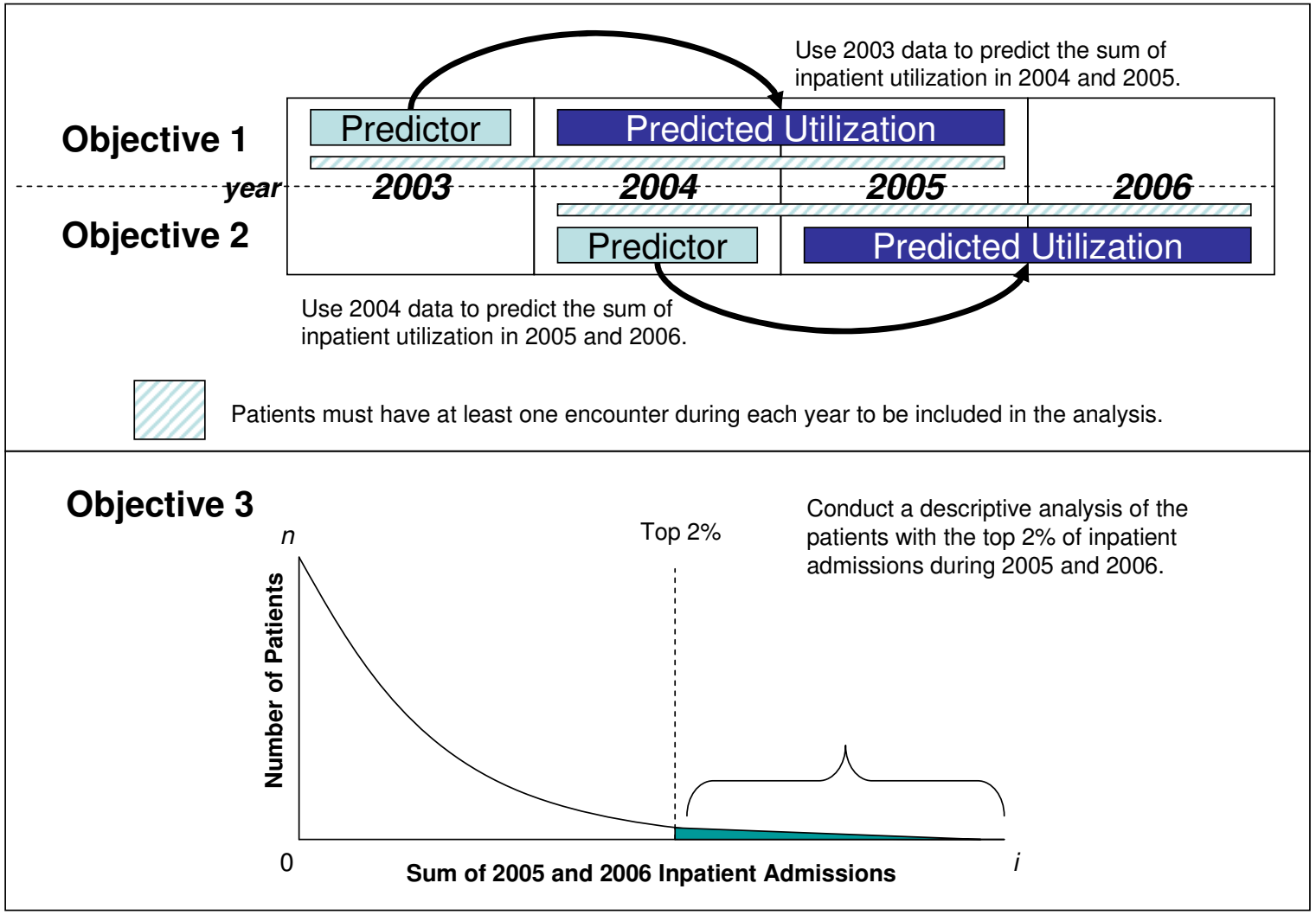
Objective 3

The third objective was to describe the top 2 percent of patients with the greatest number of 2005 and 2006 inpatient admissions. This descriptive analysis should facilitate an understanding of this subgroup of patients, which can be used to identify specific groups of similar patients that can be targeted for intervention.

Hypothesis for Objective 3

H9: The subset of patients with the top 2 percent of actual 2005 and 2006 inpatient utilization will contain a subgroup of patients with 1 or fewer 2004 inpatient encounters and at least 1 chronic condition HCC in 2004.

Figure 2.3 Visual Representation of the Study Objectives



Institutional Review Board Approval

The University of Texas at Austin (UT) Institutional Review Board (IRB) has conducted an ethics review on this project's proposed details and design. The project received IRB approval on October 18, 2007.

Data Source

The study used encounter data captured by the ICC and stored in a MS SQL Server database managed and secured by Ascension Health Systems. Ascension set up both virtual private network (VPN) encryption and database-level security to ensure data were accessed only by the lead researcher. Two separate passwords were required for access to the data. Each record included a unique patient identifier with no personal, patient-specific meaning. The data fields available to the lead researcher included:

- Unique Patient Identification Number;
- Patient Zip Code;
- Patient Age;
- Patient Gender;
- Encounter Type (e.g., emergency department, in-patient, clinic);
- Location ID (i.e., facility at which patient encounter occurred);
- Admit Date / Time;
- Discharge Date / Time;
- Diagnosis Code(s) (ICD9); and
- Funding Source (Insurance Provider).

Accuracy of Data

With each passing year, the ICC's individual partner member organizations have become increasingly proficient with managing and supplying the data aggregated in the ICC's patient database. While past data omissions have been addressed procedurally (to

improve current and future data collection), not all omitted past data has been obtained and retroactively populated in the ICC database.

Specifically, the discharge date of inpatient encounters (and therefore the length of stay) is almost completely missing in 2003 data and progressively becomes more complete through 2006, when it is populated in every encounter record. For this reason, the number of inpatient days cannot be used as a dependent variable for predictive models created using 2003 and 2004 data. The number of inpatient admissions was used instead, although this data element lacks the severity of illness information that would be present in length-of-stay data.

Study Time Frame

This study used patient encounters that occurred between January 1, 2003, and December 31, 2006.

Study Population

The study population included 494,551 medically indigent individuals who received uncompensated care from at least one of the clinics or hospitals in the Austin area during the study time frame. A list of participating ICC member locations can be found in Appendix 1. As a matter of policy, some ICC partner members have withheld certain types of patients and/or encounters from the data files submitted to the ICC. In these cases, patient history or entire patients may be missing from the ICC database and subsequently from this analysis. HIV diagnoses were omitted from St. David's hospitals

data and from the Austin Travis County HIV specialty clinic. The People's Community Clinics omitted all encounters with minors.

Inclusion Criteria

To be included in the study population, a patient must have been medically indigent with at least one diagnostic encounter at one of the clinics or hospitals in the Austin area during the study time frame.

Any patient whose first encounter occurred at or after age 65 was excluded from the study population. Additionally, any patients with illogical data (e.g., future birthdates, impossible gender/diagnosis combinations) were excluded from the study population.

Objectives 1 and 2 both utilized three-year windows of data: year one for prediction and years two and three for outcomes. A patient who has a year-one encounter but no encounters in years two or three may: legitimately have no year-two and three utilization, have moved out of the area after year one, or have become eligible for a health care plan after year one and thereafter become considered medically indigent. Similarly, a patient with no year-one encounters but encounters in either years two or three may: legitimately have no year-one utilization, have moved into the area after year one, or have lost coverage after year one and thereafter become medically indigent. By restricting the analyses in Objectives 1 and 2 to patients with a least one diagnostic encounter in years one, two, and three, the models' performances can be evaluated with less interference from unknown factors that cause patients to move in and out of the

ICC's patient population. A sensitivity analysis tested whether relaxing this constraint changed the results of Objective 2.

Study Design

This study was retrospective, using patient age, sex, and encounter data to predict future costs. Five regression-based models and one policy-based process used 2004 patient data to predict hospital utilization in 2005 and 2006. The relative predictive accuracy of each model and the policy process were assessed, and the population of identified high-risk patients were described.

Description of Data Element: Count of Inpatient Admissions

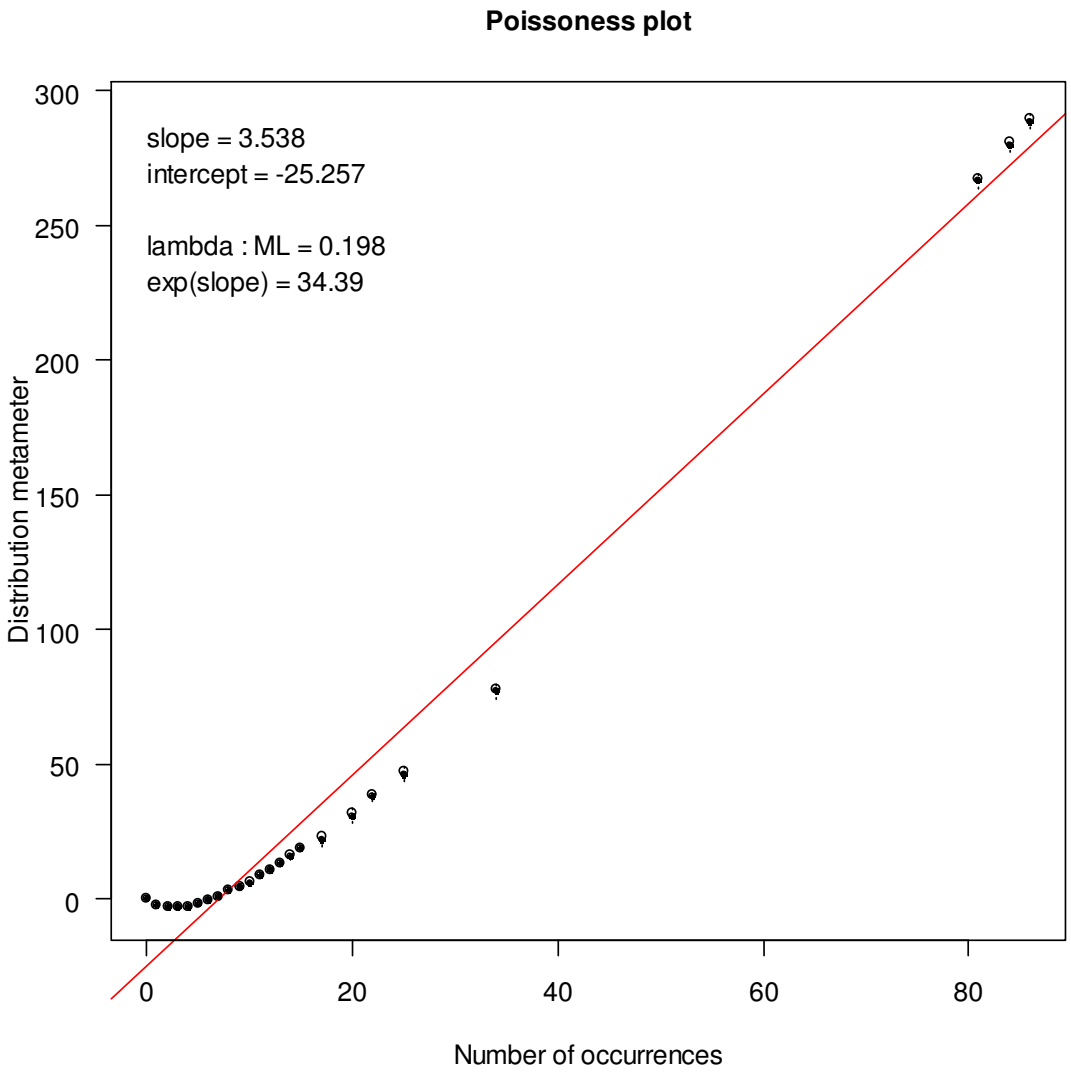
The count of inpatient admissions were used both as an independent variable (year 1) and a dependent variable (sum of year 2 and year 3). To choose the appropriate generalized linear model link functions (used in M4 and M5), the distribution of the variance of the count of inpatient admissions must be approximated. The Kolmogorov-Smirnov (K-S) test compares the observed values of the data with values of a known distribution function.

K-S tests were run on the sum of 2004 and 2005 inpatient admissions to test the fit of the normal and Poisson distributions. The lower the K-S D statistic, the more likely that the empirical data conform to the test distribution. Although the Poisson K-S test yielded a lower D statistic than the normal K-S test (0.024 vs. 0.466), the p-values associated with both D statistics fell below the critical limit at alpha = .05. While the K-S

test can indicate whether a substantial discrepancy exists between the empirical and test distributions, it cannot indicate if the discrepancy exists only in a single cell while the remainder of the data fit the test distribution well. A graphical goodness-of-fit method can provide this insight. Similar to a probability plot, like a quantile-quantile (Q-Q) plot, Hoaglin's "Poissonness Plot" is designed for observed frequency distributions.[177] In R, this plot is available through the *distplot* function. Figure 2.4 depicts the Poissonness Plot of the sum of 2004 and 2005 inpatient admissions.

[177] Hoaglin. A Poissonness plot. *The American Statistician.* 1980;34(3):146-149.

Figure 2.4 Poissonness Plot of the Sum of 2004 and 2005 Inpatient Admissions



As with a Q-Q plot, the closer the data fall to a straight line, the better the fit of the test distribution. As seen in the above figure, the data follow the line close enough to suggest that modeling the data as Poisson distributed appears tenable.

Specific Methods Related to Study Objectives

Objective 1:

H1: Model fit will vary between the five models.

While models using more relevant independent variables are expected to predict future utilization better, this expectation has not been tested in this population. Deviance residual statistics were presented for each model in a table. Model fit for linear models were described in terms of deviance residuals and R^2. For non-linear models, model fit was described in terms of deviances residuals statistics solely.

H2: Models M1, M4, and M5 will have lower median deviance residuals than models M2 and M3.

The benefit of using the DCG model must be assessed to determine if the incremental costs and efforts of implementing this model have an expected benefit on the final model's accuracy. In other words, this hypothesis tests whether utilization, age, and sex data predict future inpatient days as well as models that use the DCG RR score as an independent variable.

H3: Model M1 will have a higher R^2 than the square of the Pearson product moment correlation between the 2003 DCG RR scores and the number of inpatient admissions between 2004 and 2005.

A model that uses past utilization as well as the DCG RR score to predict 2004 through 2005 inpatient admissions must perform as well (i.e., have at least as high an R^2) as a

model using only the DCG RR score, since adding independent variables to a regression model cannot lower R^2. However, whether R^2 will increase and the extent to which it will increase was unknown. Some utilization information is contained in the DCG RR score, since the RR score is partly dependent upon inpatient diagnoses. This hypothesis suggests that some valuable utilization information is not captured in the DCG RR score.

H4: Model M4 will have lower deviance residuals than model M1.

The literature suggests that cost and utilization data are non-normally distributed and more closely follow highly skewed distributions like the gamma distribution. This hypothesis tests whether a generalized linear model using a non-normal distribution results in lower deviance scores.

H5: Model M5 will have the lowest deviance residuals of all models.

Model M5 is two-part and uses the generalized linear model. The literature suggests that this type of model should be best at identifying patient risk. This hypothesis tests that assumption.

Objective 2:

A descriptive analysis compared the highest-risk patients as identified by the Utilization Method and each model. For each model and the Utilization Method, the following were calculated as part of this descriptive analysis:

- *Positive Predictive Value (PPV)*: the percentage of high-risk patients who were actually high use; please see Figure 2.4 for details;

- *Mean absolute prediction error (MAPE)*: the mean difference between each patient's predicted number of 2005 through 2006 inpatient admissions and the actual number of Outcome Period inpatient admissions of the subset of patients identified as high-risk;

- *Total 2005 and 2006 inpatient admissions*: the total number of 2005 through 2006 inpatient admissions for the subset of patients identified as high-risk;

- *Mean number of 2005 through 2006 inpatient admissions per correctly identified patient*: the mean number of 2005 through 2006 inpatient admissions for patients who are in the predicted high-risk subset and the actual high-use patient subset;

- *Mean rescaled relative risk scores for 2004, 2005, and 2006*: the mean rescaled relative risk scores for patients in the predicted high-risk subset will be presented for each individual year;

- *Mean number of assigned 2004 Hierarchical Condition Categories (HCCs)*: the mean number of known comorbid conditions of patients identified as high-risk;

- *Mean number of assigned 2005 through 2006 HCCs*: the mean number of 2005 through 2006 comorbid conditions of the patients identified as high-risk; and

- *Number of patients in predicted subset with pregnancy-related diagnoses*: this is included to compare how each method handles patients with pregnancies.

H6: The Utilization Method and Models M1-M5 will identify different patients as highest-risk.

The Utilization Method identified *j* patients who meet its criteria. This hypothesis tests whether the *j* patients with the highest predicted utilization from each model and the Utilization Method will be the same subset of patients. No matter how accurate a data-driven statistical model is that predicts high risk patients, if the same subset of patients can be identified in a simpler fashion, utilization of the model has no unique benefit.

The diagnoses of each group were described with tables showing the occurrence of selected high-cost chronic diseases and conditions in each model subset. Additionally, Venn diagrams were used to show how many patients were shared between the Utilization Method patient subset and the subsets of the predictive models.

H7: The number of 2005 and 2006 inpatient admissions of the *j* patients identified by Models M1, M4, and M5 will be higher than the 2005 and 2006 inpatient admissions of the *j* patients identified by models M2 and M3.

Loading the models with 2004 data and testing their predicted outputs against actual 2005 and 2006 utilization not only cross-validated the models, it simulated how the models would be used in a real setting. A real-world test of the models would follow the patients identified as high-risk for two years. A patient with low 2005 utilization but high 2006 utilization may have been a good target for intervention. Furthermore, the ability of the models to identify the top 2 percent of patients with the highest actual 2005 and 2006 utilization was tested using positive predictive value (PPV).

H8: The data-driven models will identify case management candidates with a higher mean number of 2005 and 2006 comorbid chronic conditions than those candidates identified by the Utilization Method.

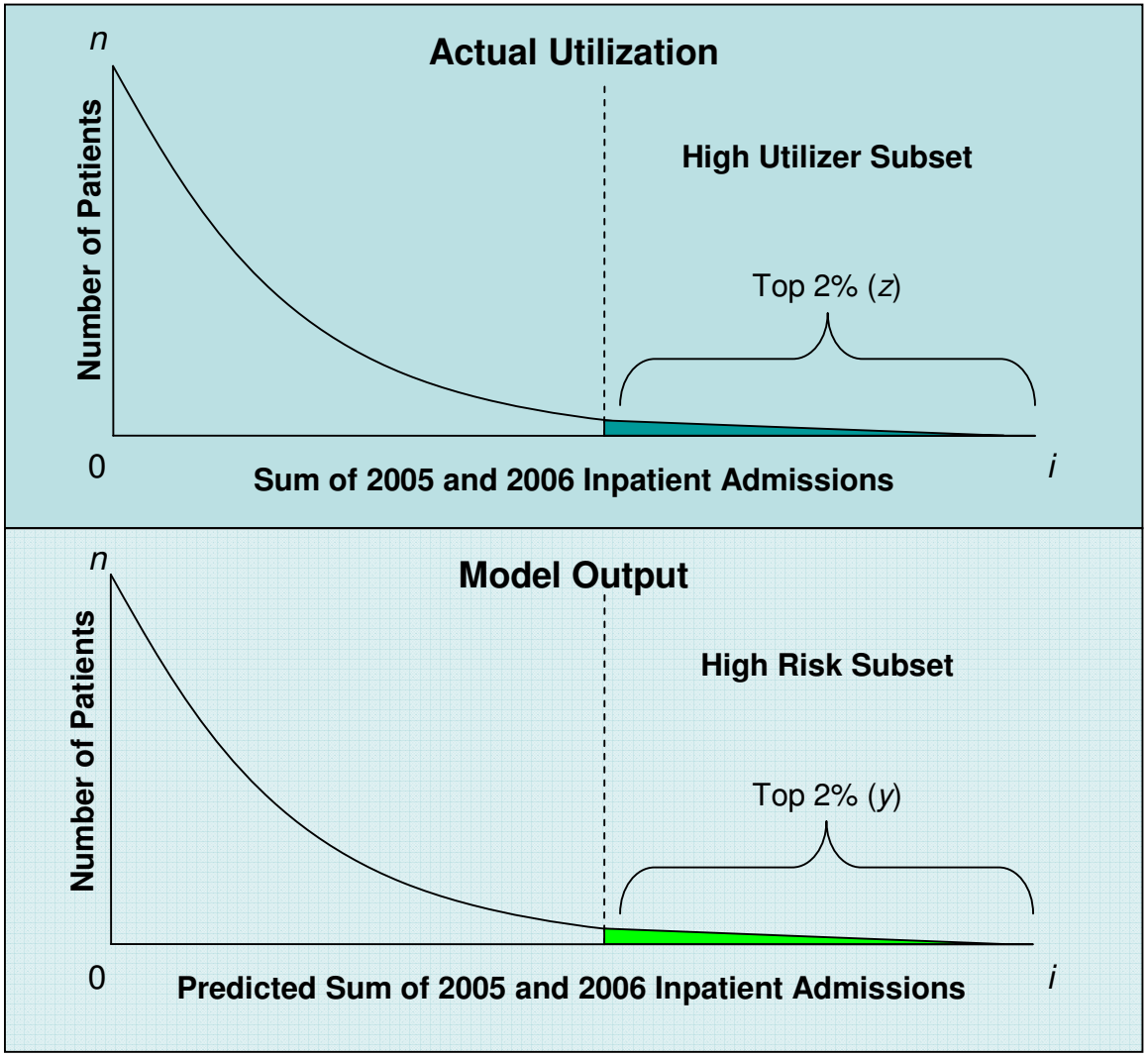
Hospitalizations caused by comorbid chronic conditions may be more preventable than those by acute conditions. Chronic conditions will be identified in patients via the occurrence of the respective HCCs. The specific chronic conditions that were counted

were the six specific preventable chronic conditions identified by the World Health Organization: heart disease, stroke, cancer, asthma, chronic obstructive pulmonary disease, and diabetes.[178]

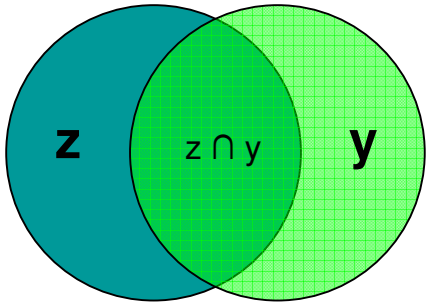
Any condition that has multiple HCCs will be counted only once, even if the patient has several HCCs within the condition. The mean number of chronic conditions was calculated for the *j* patients in each subset identified by Models 1-5 and the Utilization Method. Furthermore, the PPV of identifying the top 2 percent of patients with the highest actual 2005 and 2006 utilization was compared. Figure 2.5 explains how PPV will be calculated and used in Objective 2.

[178] World Health Organization. *Preventing Chronic Diseases: A Vital Investment* 2005.

Figure 2.5 Calculation of Positive Predictive Value for Objective 2



A model's positive predictive value (PPV) is the percentage of high-utilizers correctly identified as such. Stated another way, it is the percentage of patients in the high-risk subset (y) who are also in the high-utilizer subset (z).



$$PPV = \frac{z \cap y}{y}$$

Objective 3:

H9: The subset of patients with the top 2 percent actual 2005 and 2006 inpatient utilization will contain a subgroup of patients with 1 or fewer 2004 inpatient encounters and at least 1 chronic condition HCC in 2004.

An extensive descriptive analysis will be conducted on the patients with the top 2 percent actual 2005 and 2006 inpatient utilization. Of particular interest is the subgroup of those patients who have 1 or fewer 2004 inpatient encounters and at least 1 chronic condition HCC in 2004. In evaluating this hypothesis, the chronic conditions that are considered included diabetes, hypertension, asthma, chronic obstructive pulmonary disease, substance abuse, congestive heart failure, and depression. This list of conditions differs slightly from that used to evaluate hypothesis H8. This is because H8 focuses on patient outcomes during the 2005 through 2006 period, while H9 focuses on 2004, the prediction period. For the purposes of this study, outcomes such as stroke and cancer (both of which were used to evaluate H8) were of less interest in the evaluation of H9 than outcomes such as hypertension and depression, because patients with outcomes such as stroke and cancer in 2004 would already be costly and the focus of responsive medicine rather than the preventative medicine this project is attempting to facilitate.

Patients without a 2004 inpatient admission would be impossible to identify using only past utilization as a predictor for future hospital stays. However models incorporating a measure of health status (e.g., the DCG RR score) may identify these patients. The number of such patients identified by Models 1-5 and the Utilization Method will be presented.

Chapter 3: Results

Overview

This chapter presents a description of the study population and the results for Objectives 1, 2, and 3. Objective 1 is the creation of statistical models designed to take year one patient data and predict the number of inpatient admissions over the period including years two and three. Objective 1 uses 2003 as year one and 2004 and 2005 as years two and three, respectively.

Objective 2 tests the models created in Objective 1 on a different subset of patients, with 2004 as year one and 2005 and 2006 as years two and three, respectively. To make the results of Objective 1 and Objective 2 more comprehensible, the following standard nomenclature will be used.

The ***Calibration Subgroup*** is the subset of patients used in Objective 1 to create the statistical models. For a patient to be part of the Calibration Subgroup, the patient must have had at least one diagnostic encounter in years 2003, 2004, and 2005.

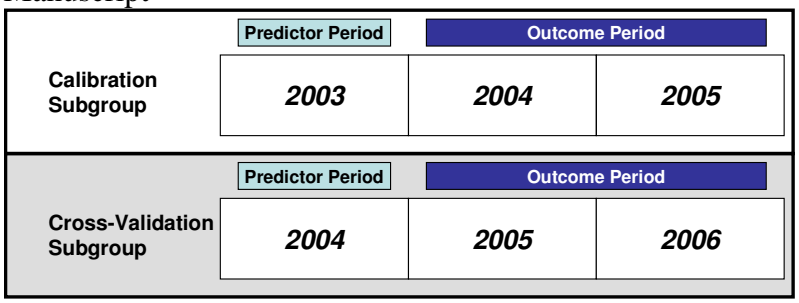
The ***Cross-Validation Subgroup*** is the subset of patients used in Objective 2 to test the performance of the statistical models created in Objective 1. For a patient to be part of the Cross-Validation Subgroup, the patient must have had at least one diagnostic encounter in years 2004, 2005, and 2006.

The descriptions of both the Calibration Subgroup and the Cross-Validation Subgroup will focus on two distinct time periods: year one and the combined period of

years two and three. The term ***Predictor Period*** refers to year one and is so named because data from this period is the basis for model predictions. The Predictor Period for the Calibration Subgroup is 2003. The Predictor Period for the Cross-Validation Subgroup is 2004. The term ***Outcome Period*** refers to the combined period of years two and three. The Outcome Period for the Calibration Subgroup is 2004 through 2005. The Outcome Period for the Cross-Validation Subgroup is 2005 through 2006.

Figure 3.1 is a simplified version of Figure 2.2 modified to illustrate this nomenclature. This figure (or a portion thereof) will be included in various sections throughout the Results chapter to assist the reader in following the analyses.

Figure 3.1 Illustration of the Nomenclature Used in the Results Chapter of this Manuscript



Objective 3 is a detailed description of the patients in the Cross-Validation Subgroup with the highest number of inpatient admissions during the Outcome Period.

Study Population

Between January 1, 2003, and December 31, 2006, a total of 494,551 patients had 3,556,034 recorded encounters at ICC partner provider locations. In 2003, a total of 156,796 unique patients had encounters in ICC partner provider locations. In 2004, a total

of 184,362 unique patients had encounters in ICC partner provider locations. In 2005, a total of 190,425 unique patients had encounters in ICC partner provider locations. In 2006, a total of 234,769 unique patients had encounters in ICC partner provider locations.

For the Calibration Subgroup, only patients with an encounter in each of years 2003, 2004, and 2005 were included in the analysis. For the Cross-Validation Subgroup, only patients with an encounter in each of years 2004, 2005, and 2006 were included in the analysis. Furthermore, for both the Calibration and Cross-Validation Subgroups, the encounters must have been one of the following types: Inpatient, Outpatient, Clinic Visit, Emergency Room, or Office Visit.

Table 3.1 displays how the requirements of specific encounter types and of encounters in years 2003, 2004, and 2005 resulted in the 41,260 patients in the Calibration Subgroup.

Table 3.1 Number of Patients who Meet Encounter Type and Year Criteria for Inclusion in the Calibration Subgroup of Patients

Calibration Subgroup

Period	Year	Number of Unique Patients Total	w/ Diagnoses*	Both 2003 and 2004	Both 2004 and 2005	All 2003, 2004, and 2005
Predictor	2003	156,796	151,788	71,148	77,759	**41,260**
Outcome	2004	179,208	175,677			
	2005	190,425	185,998			

*Encounter of the year must have been of type "Inpatient", "Outpatient", "Emergency Room", "Clinic Visit", or "Office Visit"

As shown in Table 3.1, adding the requirement of a diagnostic encounter in 2003, 2004, and 2005 for the Calibration Subgroup greatly reduced the number of patients. In years 2003 and 2004, a total of 156,796 and 179,208 patients, respectively, had one or more encounter records in the ICC database. However, for 5,008 patients in 2003 and

3,531 patients in 2004 these encounters were not of the types listed above. Only 71,148 patients had one or more valid encounters in both 2003 and 2004. Similarly, only 77,759 patients had one or more valid encounters in both 2004 and 2005. Eliminating patients who did not have a valid encounter in each of 2003, 2004, and 2005 further reduces the number of patients to 41,260, which is the Calibration Subgroup.

Table 3.2 Number of Patients who Meet Encounter Type and Year Criteria for Inclusion in the Cross-Validation Subgroup of Patients

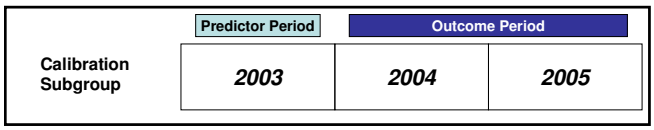
Cross-Validation Subgroup

Period	Year	Number of Unique Patients Total	Number of Unique Patients w/ Diagnoses*	Both 2004 and 2005	Both 2005 and 2006	All 2004, 2005, and 2006
Predictor	2004	179,208	175,677	77,759		
Outcome	2005	190,425	185,998		80,231	**44,738**
	2006	234,769	229,757			

*Encounter of the year must have been of type "Inpatient", "Outpatient", "Emergency Room", "Clinic Visit", or "Office Visit"

Table 3.2 shows how a similar process of exclusions was performed to arrive at the 44,738 patients in the Cross-Validation Subgroup. A total of 29,663 patients had one or more valid encounters in all the years from 2003 through 2006. These patients were included in both the Calibration Subgroup and the Cross-Validation Subgroup.

Age and Sex Characteristics for Calibration Subgroup Patients



As described above, all Calibration Subgroup patients had at least one valid encounter in years 2003, 2004, and 2005. In the ICC data set, patient birthdates were

removed to increase the anonymity of patient records. However, each encounter record listed the patient's age at the time of the encounter. Since only ages at encounters were known, a mean age for calendar year 2003 (the Predictor Period) was calculated for patients with multiple 2003 encounters and rounded to the nearest whole number. For example, a patient with 2 encounters at age 23 and 4 encounters at age 24 would have a mean age of 23.66 for 2003. This was rounded up to 24 years.

Age was used as an independent variable in the calibration of model M2. Age is also used by the DxCG model in assigning relative risk scores, so age indirectly was used in the calibration of the models that used relative risk score: M1, M4, and M5. The mean age for Calibration Subgroup patients was 25.07 years (SD = 17.60). The mean age for female patients was higher than for males (26.60, SD = 15.98 vs. 21.30, SD = 20.57) and was found to be significantly different using a t-test ($t = -24.14$, $p < 0.01$).

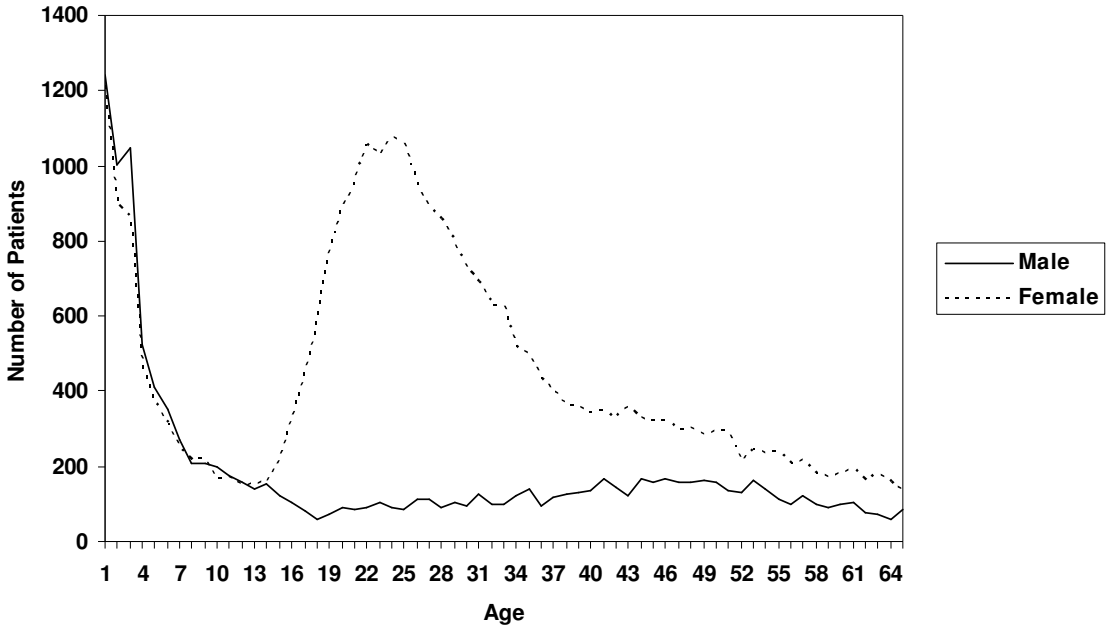
Table 3.3 shows a breakdown of Calibration Subgroup patients by age categories and gender in the Predictor Period.

Table 3.3 Distribution of Calibration Subgroup Patients by Predictor Period (2003) Age Category and Gender

Age Category	Males			Females			Total		
	N	(Row %)	[Col %]	N	(Row %)	[Col %]	N	(Row %)	[Col %]
Under 10	5,468	(52.30)	[45.79]	4,988	(47.70)	[17.01]	10,456	(100.00)	[25.34]
10-19	1,152	(23.05)	[9.65]	3,846	(76.95)	[13.12]	4,998	(100.00)	[12.11]
20-29	975	(9.40)	[8.17]	9,400	(90.60)	[32.06]	10,375	(100.00)	[25.15]
30-39	1,196	(19.70)	[10.02]	4,874	(80.30)	[16.62]	6,070	(100.00)	[14.71]
40-49	1,562	(32.91)	[13.08]	3,184	(67.09)	[10.86]	4,746	(100.00)	[11.50]
50-59	1,192	(35.27)	[9.98]	2,188	(64.73)	[7.46]	3,380	(100.00)	[8.19]
60-64	396	(32.06)	[3.32]	839	(67.94)	[2.86]	1,235	(100.00)	[2.99]
Total	11,941	(28.94)	[100.00]	29,319	(71.06)	[100.00]	41,260	(100.00)	[100.00]

The Under 10 age group accounted for 45.79% of all male patients but only 17.01% of all female patients. For females, the 20-29 age group was the largest, accounting for 32.06% of female patients. Figure 3.2 shows the change in gender distribution controlling for age for Calibration Subgroup patients.

Figure 3.2 Distribution of Calibration Subgroup Patients by Predictor Period (2003) Age and Gender



The numbers of male and female patients are relatively similar until the ages of 13 or 14 years. It is during the child-bearing years between 14 and the late 30s that the number of females increases dramatically. In the early 20s, the number of females is approximately ten times higher than the number of males (9,400 vs. 975, respectively). In the 40s, the number of females drops while the number of males increases.

Age and Sex Characteristics for Cross-Validation Subgroup Patients

	Predictor Period	Outcome Period	
Cross-Validation Subgroup	*2004*	*2005*	*2006*

All Cross-Validation Subgroup patients had at least one valid encounter in years 2004, 2005, and 2006. The mean age for calendar year 2004 (the Predictor Period) was

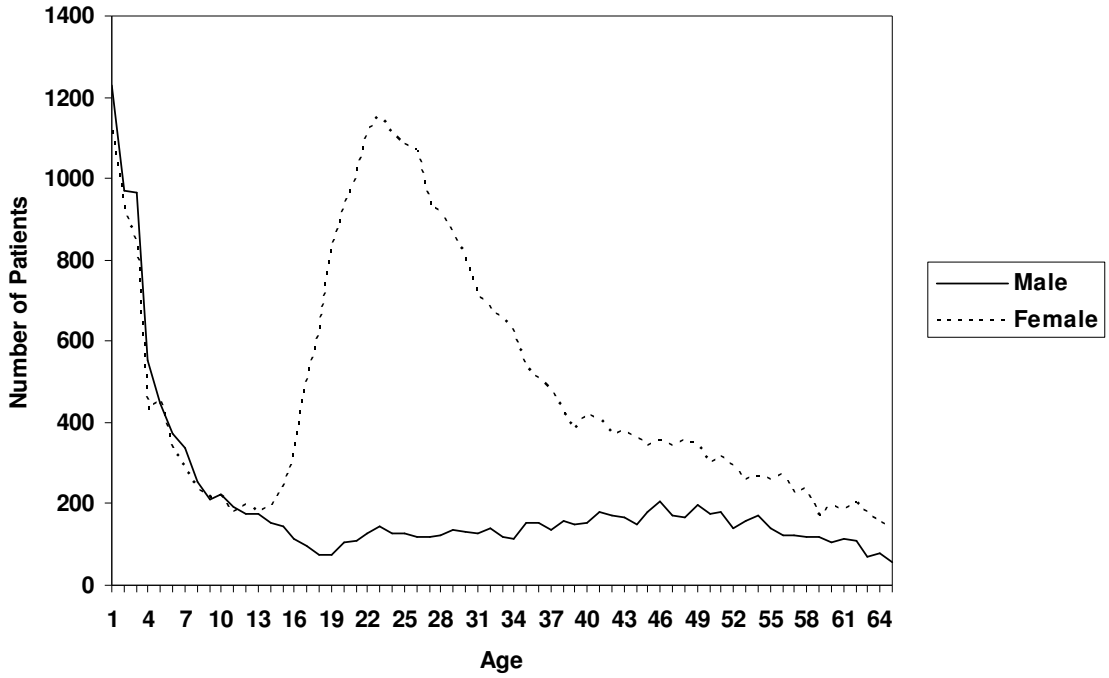
calculated for patients with multiple 2004 encounters and rounded to the nearest whole number. The mean age for Cross-Validation patients was 25.68 years (SD = 17.49). The mean age for males was 22.35 years (SD = 20.40) and for females was 27.06 years (SD = 15.86). The mean age of female patients was found to be significantly higher than for male patients ($t = -26.19$, $p < 0.01$). Table 3.4 shows a breakdown of Cross-Validation Subgroup patients by age category and gender.

Table 3.4 Distribution of Cross-Validation Subgroup Patients by Predictor Period (2004) Age Category and Gender

Patients by	Males			Females			Total		
Age	N	(Row %)	[Col %]	N	(Row %)	[Col %]	N	(Row %)	[Col %]
Under 10	5,567	(52.42)	[42.42]	5,053	(47.58)	[15.98]	10,620	(100.00)	[23.74]
10-19	1,307	(23.88)	[9.96]	4,167	(76.12)	[13.18]	5,474	(100.00)	[12.24]
20-29	1,262	(11.15)	[9.62]	10,053	(88.85)	[31.80]	11,315	(100.00)	[25.29]
30-39	1,411	(20.63)	[10.75]	5,427	(79.37)	[17.17]	6,838	(100.00)	[15.28]
40-49	1,769	(33.24)	[13.48]	3,553	(66.76)	[11.24]	5,322	(100.00)	[11.90]
50-59	1,380	(35.59)	[10.52]	2,498	(64.41)	[7.90]	3,878	(100.00)	[8.67]
60-64	428	(33.15)	[3.26]	863	(66.85)	[2.73]	1,291	(100.00)	[2.89]
Total	13,124	(29.34)	[100.00]	31,614	(70.66)	[100.00]	44,738	(100.00)	[100.00]

The Under 10 age group accounted for 42.42% of all male patients, compared to 15.98% for female patients. For females, the 20-29 age group was the largest, accounting for 31.80% of all female patients. Figure 3.3 shows the change in gender distribution controlling for age for Cross-Validation Subgroup patients.

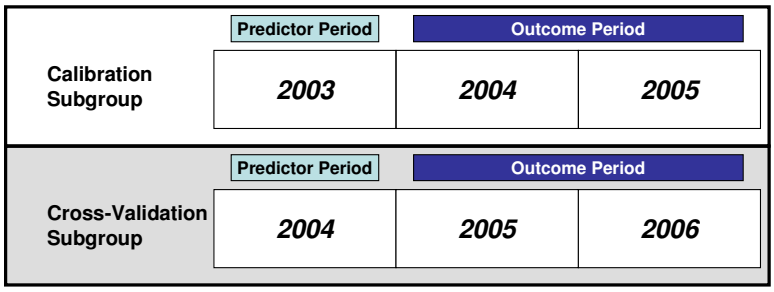
Figure 3.3 Distribution of Cross-Validation Subgroup Patients by Predictor Period (2004) Age and Gender



Among patients included in the Cross-Validation Subgroup analyses, the distribution of age by sex follows the same pattern as the age-sex distribution of Calibration Subgroup patients. The largest disparity is between the numbers of females and males again occurring between the ages of 14 to 40.

Calibration Subgroup patients had a lower mean age than their Cross-Validation Subgroup counterparts (25.07 vs. 25.68, $t = -5.14$, $p < 0.01$). The percentage of patients who were female was similar between Calibration Subgroup patients and Cross-Validation Subgroup patients (52.2 vs. 51.9, chi-squared = 0.50, $p = 0.48$).

Predictor Period Relative Risk Scores and Clinical Conditions



Three of the statistical models (Models M1, M4, and M5) use the relative risk (RR) score calculated by the DxCG RiskSmart software as an independent variable representing a measure of patient health status. The RR score is largely determined by the diagnosed conditions of the patients. To calculate the RR score, the DxCG RiskSmart software first categorizes patients into condition categories (CCs) based on the specific ICD-9 codes associated with their encounters. This method of categorization can also be used to quickly identify patients with specific disease states or conditions. In this section, the Predictor Period diagnoses of the Calibration Subgroup and Cross-Validation Subgroup patients are compared using CCs to get a sense of whether the prevalence of various conditions was consistent. The Predictor Period for the Calibration Subgroup was calendar year 2003, while the Prediction Period for the Cross-Validation subgroup was calendar year 2004.

Table 3.4 compares the number of patients in the Calibration Subgroup and the Cross-Validation Subgroup with CCs indicating the presence of one or more of the eight specific conditions found to be among the costliest among the ICC's population.[179]

[179] Leslie. Cost Analysis of the Uninsured, Indigent Population of Central Texas: Master's Thesis. College of Pharmacy, The University of Texas at Austin; 2006.

Assignment of CCs is non-exclusive, meaning individual patients can be categorized into multiple CCs. In Table 3.5 (and all other tables describing frequency of CC assignment) patients with multiple conditions are counted in each condition. For example, a patient with both asthma and diabetes would be included in the counts of both conditions.

Table 3.5 Number of Calibration Subgroup and Cross-Validation Subgroup Patients with the Costliest* Chronic Disease States and Conditions

Condition	Calibration Subgroup		Cross-Validation Subgroup	
	N	%	N	%
Pregnancy	4,097	(9.93%)	4,719	(10.55%)
Hypertension	3,520	(8.53%)	5,129	(11.46%)
Diabetes	3,105	(7.53%)	3,821	(8.54%)
Asthma	1,607	(3.89%)	2,177	(4.87%)
Substance Abuse	1,211	(2.94%)	1,520	(3.40%)
Non-major Depression	1,044	(2.53%)	2,009	(4.49%)
Congestive Heart Failure	354	(0.86%)	428	(0.96%)
COPD	329	(0.80%)	510	(1.14%)

Note: The numbers in parentheses represent the percentage of the subgroup with the condition during the Predictor Period. Patients may be counted in multiple conditions if diagnoses for each condition is present. For the Calibration Subgroup, the Predictor Period is calendar year 2003. For the Cross Validation Subgroup the Predictor Period is calendar year 2004.

*As identified in: Leslie RC. Cost Analysis of the Uninsured, Indigent Population of Central Texas: Master's Thesis. College of Pharmacy, The University of Texas at Austin; 2006.

Although most of the above conditions showed similar frequencies in years 2003 and 2004, the number of hypertension patients increased from 8.53% of the Objective 1 study population in 2003 to 11.46% of the Objective 2 study population in 2004.

Table 3.6 describes the frequencies of all 183 HCCs. The Change column displays the increase or decrease in frequency between the Calibration Subgroup and the Cross-Validation Subgroup. Patients with multiple diagnosed conditions are counted in each condition.

Table 3.6 Number and Percent of Patients with Assigned Hierarchical Condition Codes for Calibration and Cross-Validation Subgroups in the Respective Predictor Periods

	Number of Patients with HCC in Predictor Period				
	Calibration		Cross-Validation		
Hierarchical Condition Category	**N**	**Percent**	**N**	**Percent**	**Change**
Screening/Observation/Special Exams	20,965	50.81%	23,313	52.11%	1.30%
Other Ear, Nose, Throat, and Mouth Disorders	8,245	19.98%	9,746	21.78%	1.80%
Minor Symptoms, Signs, Findings	7,032	17.04%	9,107	20.36%	3.31%
Major Symptoms, Abnormalities	6,620	16.04%	8,117	18.14%	2.10%
Other Musculoskeletal and Connective Tissue Disorders	4,826	11.70%	6,262	14.00%	2.30%
Other Gastrointestinal Disorders	4,078	9.88%	5,249	11.73%	1.85%
Uncompleted Pregnancy With No or Minor Complications	3,631	8.80%	4,196	9.38%	0.58%
Other Female Genital Disorders	3,528	8.55%	4,697	10.50%	1.95%
Other Injuries	3,520	8.53%	5,000	11.18%	2.64%
Hypertension	3,520	8.53%	5,129	11.46%	2.93%
Other Infectious Diseases	3,162	7.66%	4,010	8.96%	1.30%
Diabetes with No or Unspecified Complications	2,943	7.13%	3,697	8.26%	1.13%
Other Endocrine/Metabolic/Nutritional Disorders	2,694	6.53%	4,267	9.54%	3.01%
Other Dermatological Disorders	2,450	5.94%	3,234	7.23%	1.29%
Other Lung Disorders	2,084	5.05%	2,226	4.98%	-0.08%
Urinary Tract Infection	1,999	4.84%	2,625	5.87%	1.02%
Asthma	1,607	3.89%	2,177	4.87%	0.97%
Completed Pregnancy Without Complications (Normal Delivery)	1,429	3.46%	1,525	3.41%	-0.05%
Cellulitis, Local Skin Infection	1,267	3.07%	1,892	4.23%	1.16%
Completed Pregnancy With Complications	1,239	3.00%	1,231	2.75%	-0.25%
Iron Deficiency and Other/Unspecified Anemias and Blood Disease	1,162	2.82%	1,531	3.42%	0.61%
Other Eye Disorders	1,158	2.81%	1,524	3.41%	0.60%
Drug/Alcohol Abuse, Without Dependence	1,083	2.62%	1,374	3.07%	0.45%
Depression	1,044	2.53%	2,009	4.49%	1.96%
Normal, Single Birth	1,027	2.49%	857	1.92%	-0.57%
Benign Neoplasms of Skin, Breast, Eye	1,011	2.45%	1,376	3.08%	0.63%
Post-Surgical States/Aftercare/Elective	969	2.35%	1,379	3.08%	0.73%
Other Psychiatric Disorders	878	2.13%	1,356	3.03%	0.90%
Disorders of Fluid/Electrolyte/Acid-Base Balance	786	1.90%	1,127	2.52%	0.61%
Other Perinatal Problems Affecting Newborn	779	1.89%	808	1.81%	-0.08%
History of Disease	775	1.88%	1,195	2.67%	0.79%
Viral and Unspecified Pneumonia, Pleurisy	774	1.88%	739	1.65%	-0.22%
Seizure Disorders and Convulsions	766	1.86%	923	2.06%	0.21%
Uncompleted Pregnancy With Complications	730	1.77%	853	1.91%	0.14%
Mononeuropathy, Other Neurological Conditions/Injuries	720	1.75%	1,010	2.26%	0.51%
Pelvic Inflammatory Disease and Other Specified Female Genital Disorders	607	1.47%	731	1.63%	0.16%
Major Depressive, Bipolar, and Paranoid Disorders	582	1.41%	1,024	2.29%	0.88%
Other Urinary Tract Disorders	579	1.40%	771	1.72%	0.32%
Other Neoplasms	576	1.40%	632	1.41%	0.02%
Chronic Hepatitis	538	1.30%	648	1.45%	0.14%
Diabetes with Renal Manifestation	523	1.27%	486	1.09%	-0.18%
Poisonings and Allegic Reactions	486	1.18%	719	1.61%	0.43%
Male Genital Disorders	462	1.12%	468	1.05%	-0.07%
Type I Diabetes Mellitus	454	1.10%	525	1.17%	0.07%
Disorders of the Vertebrae and Spinal Discs	453	1.10%	620	1.39%	0.29%
Other Hepatitis and Liver Disease	402	0.97%	499	1.12%	0.14%
Peptic Ulcer, Hemorrhage, Other Specified Gastrointestinal Disorders	399	0.97%	396	0.89%	-0.08%
Urinary Obstruction and Retention	384	0.93%	456	1.02%	0.09%
Congestive Heart Failure	354	0.86%	428	0.96%	0.10%
Chronic Obstructive Pulmonary Disease	329	0.80%	510	1.14%	0.34%
Coronary Atherosclerosis/Other Chronic Ischemic Heart Disease	314	0.76%	439	0.98%	0.22%
Diabetes with Acute Complications	304	0.74%	225	0.50%	-0.23%
Other Developmental Disability	302	0.73%	260	0.58%	-0.15%
Breast, Prostate, Colorectal and Other Cancers and Tumors	266	0.64%	299	0.67%	0.02%
Rheumatoid Arthritis and Inflammatory Connective Tissue Disease	263	0.64%	320	0.72%	0.08%
Miscarriage/Abortion	259	0.63%	304	0.68%	0.05%
Osteoporosis and Other Bone/Cartilage Disorders	253	0.61%	322	0.72%	0.11%
Completed Pregnancy With Major Complications	247	0.60%	255	0.57%	-0.03%
Serious Perinatal Problem Affecting Newborn	244	0.59%	229	0.51%	-0.08%
Other Circulatory Disease	241	0.58%	322	0.72%	0.14%
Diplegia (Upper), Monoplegia, and Other Paralytic Syndromes	203	0.49%	196	0.44%	-0.05%

Table 3.6 (continued)

	Number of Patients with HCC in Predictor Period				
	Calibration		Cross-Validation		
Hierarchical Condition Category	**N**	**Percent**	**N**	**Percent**	**Change**
Drug/Alcohol Dependence	188	0.46%	240	0.54%	0.08%
Coagulation Defects and Other Specified Hematological Disorders	187	0.45%	237	0.53%	0.08%
Other Complications of Medical Care	187	0.45%	192	0.43%	-0.02%
Concussion or Unspecified Head Injury	175	0.42%	247	0.55%	0.13%
Hearing Loss	175	0.42%	243	0.54%	0.12%
Spinal Cord Disorders/Injuries	173	0.42%	176	0.39%	-0.03%
Anxiety Disorders	170	0.41%	230	0.51%	0.10%
Diabetes with Ophthalmologic Manifestation	167	0.40%	102	0.23%	-0.18%
Cirrhosis of Liver	152	0.37%	183	0.41%	0.04%
Diabetic and Other Vascular Retinopathies	152	0.37%	85	0.19%	-0.18%
Osteoarthritis of Hip or Knee	151	0.37%	206	0.46%	0.09%
Polyneuropathy	150	0.36%	187	0.42%	0.05%
Attention Deficit Disorder	148	0.36%	170	0.38%	0.02%
Cardio-Respiratory Failure and Shock	147	0.36%	162	0.36%	0.01%
Pancreatic Disease	147	0.36%	175	0.39%	0.03%
Valvular and Rheumatic Heart Disease	144	0.35%	222	0.50%	0.15%
Major Complications of Medical Care and Trauma	140	0.34%	155	0.35%	0.01%
Renal Failure	139	0.34%	187	0.42%	0.08%
Incontinence	136	0.33%	191	0.43%	0.10%
Vascular Disease	130	0.32%	179	0.40%	0.09%
Other Significant Endocrine and Metabolic Disorders	128	0.31%	188	0.42%	0.11%
Diabetes with Neurologic or Peripheral Circulatory Manifestation	115	0.28%	137	0.31%	0.03%
Mild/Unspecified Mental Retardation/Developmental Disability	113	0.27%	109	0.24%	-0.03%
Artificial Openings for Feeding or Elimination	113	0.27%	113	0.25%	-0.02%
Drug/Alcohol Psychosis	111	0.27%	112	0.25%	-0.02%
Chronic Ulcer of Skin, Except Decubitus	111	0.27%	122	0.27%	0.00%
Specified Heart Arrhythmias	109	0.26%	128	0.29%	0.02%
Severe Hematological Disorders	109	0.26%	122	0.27%	0.01%
Other Heart Rhythm and Conduction Disorders	108	0.26%	140	0.31%	0.05%
Cataract	104	0.25%	90	0.20%	-0.05%
Other Digestive and Urinary Neoplasms	101	0.24%	178	0.40%	0.15%
Angina Pectoris/Old Myocardial Infarction	98	0.24%	166	0.37%	0.13%
Lymphatic, Head and Neck, Brain, and Other Major Cancers	90	0.22%	133	0.30%	0.08%
HIV/AIDS	88	0.21%	82	0.18%	-0.03%
Congenital/Developmental Skeletal and Connective Tissue Disorders	88	0.21%	80	0.18%	-0.03%
Paraplegia	87	0.21%	106	0.24%	0.03%
Central Nervous System Infection	82	0.20%	89	0.20%	0.00%
Nephritis	82	0.20%	82	0.18%	-0.02%
Other Congenital Heart/Circulatory Disease	79	0.19%	88	0.20%	0.01%
Schizophrenia	78	0.19%	110	0.25%	0.06%
Delirium and Encephalopathy	78	0.19%	79	0.18%	-0.01%
Disorders of Immunity	77	0.19%	92	0.21%	0.02%
Ischemic or Unspecified Stroke	77	0.19%	113	0.25%	0.07%
Appendicitis	73	0.18%	76	0.17%	-0.01%
Bone/Joint/Muscle Infections/Necrosis	73	0.18%	104	0.23%	0.06%
Fibrosis of Lung and Other Chronic Lung Disorders	71	0.17%	80	0.18%	0.01%
Glaucoma	70	0.17%	101	0.23%	0.06%
Major Head Injury	65	0.16%	107	0.24%	0.08%
Metastatic Cancer and Acute Leukemia	65	0.16%	101	0.23%	0.07%
Senility, Nonpsychotic Organic Brain Syndromes/Conditions	64	0.16%	74	0.17%	0.01%
Dementia	62	0.15%	75	0.17%	0.02%
Other and Unspecified Heart Disease	61	0.15%	76	0.17%	0.02%
Pleural Effusion/Pneumothorax	59	0.14%	71	0.16%	0.02%
Precerebral Arterial Occlusion and Transient Cerebral Ischemia	56	0.14%	60	0.13%	0.00%
Hemiplegia/Hemiparesis	55	0.13%	48	0.11%	-0.03%
Quadriplegia, Other Extensive Paralysis	54	0.13%	63	0.14%	0.01%
Intestinal Obstruction/Perforation	52	0.13%	91	0.20%	0.08%
Hip Fracture/Dislocation	52	0.13%	53	0.12%	-0.01%
Unstable Angina and Other Acute Ischemic Heart Disease	51	0.12%	63	0.14%	0.02%

Table 3.6 (continued)

	Number of Patients with HCC in Predictor Period				
	Calibration		Cross-Validation		
Hierarchical Condition Category	**N**	**Percent**	**N**	**Percent**	**Change**
Septicemia/Shock	50	0.12%	59	0.13%	0.01%
End-Stage Liver Disease	49	0.12%	59	0.13%	0.01%
Lung, Upper Digestive Tract, and Other Severe Cancers	46	0.11%	57	0.13%	0.02%
Vascular Disease with Complications	46	0.11%	61	0.14%	0.02%
Inflammatory Bowel Disease	46	0.11%	41	0.09%	-0.02%
Acute Myocardial Infarction	40	0.10%	40	0.09%	-0.01%
Vertebral Fractures	37	0.09%	43	0.10%	0.01%
Acute Liver Failure/Disease	37	0.09%	65	0.15%	0.06%
Rehabilitation	36	0.09%	174	0.39%	0.30%
Internal Injuries	35	0.08%	42	0.09%	0.01%
Aspiration and Specified Bacterial Pneumonias	35	0.08%	43	0.10%	0.01%
Significant Ear, Nose, and Throat Disorders	35	0.08%	48	0.11%	0.02%
Ectopic Pregnancy	35	0.08%	35	0.08%	-0.01%
Chemotherapy	34	0.08%	121	0.27%	0.19%
Proliferative Diabetic Retinopathy and Vitreous Hemorrhage	34	0.08%	24	0.05%	-0.03%
Major Congenital Cardiac/Circulatory Defect	31	0.08%	31	0.07%	-0.01%
Amputation Status, Lower Limb/Amputation Complications	31	0.08%	25	0.06%	-0.02%
Female Infertility	30	0.07%	35	0.08%	0.01%
Cerebrovascular Disease Late Effects, Unspecified	28	0.07%	43	0.10%	0.03%
Cerebral Hemorrhage	26	0.06%	32	0.07%	0.01%
Multiple Sclerosis	26	0.06%	40	0.09%	0.03%
Protein-Calorie Malnutrition	25	0.06%	25	0.06%	0.00%
Coma, Brain Compression/Anoxic Damage	23	0.06%	23	0.05%	0.00%
Respirator Dependence/Tracheostomy Status	22	0.05%	19	0.04%	-0.01%
Pneumococcal Pneumonia, Empyema, Lung Abscess	21	0.05%	36	0.08%	0.03%
Opportunistic Infections	21	0.05%	29	0.06%	0.01%
Cystic Fibrosis	20	0.05%	24	0.05%	0.01%
Reactive and Unspecified Psychosis	19	0.05%	37	0.08%	0.04%
Heart Infection/Inflammation, Except Rheumatic	19	0.05%	37	0.08%	0.04%
Other Respiratory and Heart Neoplasms	19	0.05%	20	0.04%	0.00%
Hypertensive Heart Disease	19	0.05%	27	0.06%	0.01%
Retinal Disorders, Except Detachment and Vascular Retinopathies	18	0.04%	8	0.02%	-0.03%
Major Eye Infections/Inflammations	17	0.04%	29	0.06%	0.02%
Decubitus Ulcer of Skin	17	0.04%	31	0.07%	0.03%
Major Organ Transplant Status	17	0.04%	16	0.04%	-0.01%
Traumatic Amputation	16	0.04%	11	0.02%	-0.01%
Very Low Birthweight Neonates	14	0.03%	14	0.03%	0.00%
Speech, Language, Cognitive, Perceptual Deficits	14	0.03%	23	0.05%	0.02%
Retinal Detachment	13	0.03%	13	0.03%	0.00%
Dialysis Status	13	0.03%	20	0.04%	0.01%
Cerebral Atherosclerosis and Aneurysm	13	0.03%	14	0.03%	0.00%
Legally Blind	12	0.03%	13	0.03%	0.00%
Muscular Dystrophy	12	0.03%	10	0.02%	-0.01%
Kidney Transplant Status	10	0.02%	20	0.04%	0.02%
Tuberculosis	9	0.02%	15	0.03%	0.01%
Parkinson's and Huntington's Diseases	9	0.02%	18	0.04%	0.02%
Profound Mental Retardation/Developmental Disability	9	0.02%	11	0.02%	0.00%
Other Organ Transplant/Replacement	9	0.02%	8	0.02%	0.00%
Personality Disorders	8	0.02%	24	0.05%	0.03%
Hypertensive Heart and Renal Disease or Encephalopathy	8	0.02%	22	0.05%	0.03%
Extremely Low Birthweight Neonates	7	0.02%	19	0.04%	0.03%
Severe Head Injury	6	0.01%	5	0.01%	0.00%
Respiratory Arrest	6	0.01%	1	0.00%	-0.01%
Cerebrovascular Disease, Unspecified	5	0.01%	14	0.03%	0.02%
Other Third-Degree and Extensive Burns	3	0.01%	6	0.01%	0.01%
Radiation Therapy	2	0.00%	10	0.02%	0.02%
Extensive Third-Degree Burns	2	0.00%	0	0.00%	0.00%
Amputation Status, Upper Limb	2	0.00%	2	0.00%	0.00%
Moderate Mental Retardation/Developmental Disability	1	0.00%	1	0.00%	0.00%
Severe Mental Retardation/Developmental Disability	0	0.00%	1	0.00%	0.00%
End Stage Renal Disease	0	0.00%	0	0.00%	0.00%

The most common CCs during the Predictor Period for both the Calibration Subgroup and the Cross-Validation Subgroup are for general screening (e.g., *Screening/Observation/Special Exams*) or for non-specific injuries or symptoms. An example of the latter would be a general diagnosis of dizziness or headache. Table 3.6 also shows the differences in the percentages of assigned CCs between the Calibration Subgroup and Cross-Validation Subgroup patients.

The largest differences between the two patient subgroups can be seen with the CCs *Hypertension* and *Other Endocrine/Metabolic/Nutritional Disorders*, each of which increased almost 3% in terms of percentage of the subgroup population. A total of 8.53 percent of patients (n = 3,520) in the Calibration subgroup had a diagnosis of hypertension, compared to 11.46 percent (n = 5,129) of the Cross-Validation Subgroup. Put another way, the Cross-Validation Subgroup contained approximately 69 percent more hypertensive patients than the Calibration Subgroup.

A total of 6.53 percent of patients (n = 2,694) in the Calibration subgroup had a diagnosis of non-diabetes endocrine, metabolic, or nutrition disorders, compared to 9.54 percent (n = 4,267) of the Cross-Validation Subgroup. Put another way, the Cross-Validation Subgroup contained approximately 63 percent more patients with non-diabetes endocrine, metabolic, or nutrition disorders than the Calibration Subgroup.

The relative risk (RR) score calculated by the DxCG model indicates the expected costliness of the patient; the higher the RR score, the more costly the patient is expected to be. The Predictor Period RR score was used as one of the independent variables for the

predictive models M1, M4, and M5. Table 3.7 shows how the Predictor Period RR score differs by gender for Calibration Subgroup patients.

Table 3.7 Descriptive Statistics of Relative Risk Scores for Calibration Subgroup Patients in the Predictor Period (2003) Controlling for Gender

	N	Mean	SD	Min	25%	Median	75%	Max
Female	29,319	0.298	0.293	0.031	0.216	0.218	0.394	10.020
Male	11,941	0.370	0.470	0.052	0.052	0.231	0.534	9.228
All	41,260	0.318	0.355	0.031	0.087	0.218	0.394	10.020
Rescaled*	41,260	1.000	1.115	0.097	0.273	0.685	1.237	31.468

*Relative risk scores were rescaled to have a mean of 1.00 by dividing each score with the mean score (0.318)

Males had a higher mean RR score than females (0.370 vs. 0.298) and also had more variability in their scores (SD = 0.470 vs. 0.293). Both males and females had a maximum score close to 10 (9.228 vs. 10.020). When the RR scores were rescaled to have a mean of 1.00 (by dividing each score by the mean score), the maximum score of 10.020 became 31.468, meaning the highest risk patient would be expected to account for over 31 times the mean cost of the patients in this group. The disparity between the mean and median RR scores (1.00 vs. 0.685, rescaled) illustrates that RR scores are skewed.

Table 3.8 shows the same descriptive breakdown of relative risk scores for the Cross-Validation Subgroup patients in their Predictor Period (2004). This value was used as one of the independent variables for prediction with models M1, M4, and M5.

Table 3.8 Descriptive Statistics of Relative Risk Scores for Cross-Validation Subgroup Patients in the Predictor Period (2004) Controlling for Gender

	N	Mean	SD	Min	25%	Median	75%	Max
Female	31,614	0.383	0.425	0.030	0.220	0.220	0.510	12.440
Male	13,124	0.509	0.636	0.050	0.060	0.390	0.700	12.850
All	44,738	0.420	0.500	0.030	0.220	0.220	0.550	12.850
Rescaled*	44,738	1.000	1.191	0.071	0.524	0.524	1.310	30.607

*Relative risk scores were rescaled to have a mean of 1.00 by dividing each score with the mean score (0.420)

The Cross-Validation Subgroup showed similar trends with relative risk scores and gender. Males had a higher mean relative risk score than females (0.509 vs. 0.383) and also had more variability in their scores (SD = 0.636 vs. 0.425). Both males and females had a maximum score close to 13 (12.850 vs. 12.440). When the relative risk scores were rescaled to have a mean of 1.00, the maximum score of 12.850 became 30.607, meaning the highest risk patient would be expected to account for almost 31 times the mean cost of the patients in this group.

Inpatient Hospital Utilization for Calibration Subgroup Patients

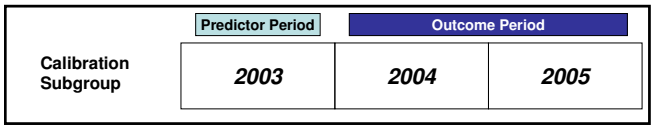


Table 3.9 shows the frequency distribution of 2003, 2004, and 2005 inpatient admissions for the Calibration Subgroup patients. For each Calibration Subgroup patient, the number of Predictor Period (2003) inpatient admissions was used as an independent variable for calibration of all models. The number of inpatient admissions for the Outcome Period (2004 and 2005) are shown both separately (i.e., the number of patients with *x* inpatient admissions in each year) and summed (i.e., the number of patients with *x*

inpatient admissions in the two-year period). The sum of 2004 and 2005 inpatient admissions was used as the dependent variable for calibration of all models.

Table 3.9 Number of Calibration Subgroup Patients with Number of Inpatient Admissions in the Predictor Period (2003) and the Outcome Period (2004-2005)

	Number of Patients by Year							
	Predictor Period		Outcome Period (Individual Years)*				Outcome Period (Combined)**	
Inpatient Admissions	2003		2004		2005			
0	36,687	88.92%	37,903	91.86%	38,740	93.89%	36,109	87.52%
1	3,838	9.30%	2,742	6.65%	2,025	4.91%	3,837	9.30%
2	512	1.24%	388	0.94%	310	0.75%	769	1.86%
3	131	0.32%	110	0.27%	87	0.21%	247	0.60%
4	49	0.12%	46	0.11%	41	0.10%	105	0.25%
5	20	0.05%	27	0.07%	21	0.05%	67	0.16%
6	6	0.01%	15	0.04%	11	0.03%	43	0.10%
7	5	0.01%	8	0.02%	1	0.00%	19	0.05%
8	4	0.01%	4	0.01%	7	0.02%	18	0.04%
9	2	0.00%	4	0.01%	5	0.01%	11	0.03%
10	0	0.00%	5	0.01%	3	0.01%	5	0.01%
11	1	0.00%	1	0.00%	2	0.00%	5	0.01%
12	1	0.00%	0	0.00%	2	0.00%	5	0.01%
13	0	0.00%	2	0.00%	0	0.00%	4	0.01%
14	1	0.00%	0	0.00%	1	0.00%	4	0.01%
15	0	0.00%	1	0.00%	0	0.00%	4	0.01%
16 - 49	2	0.00%	4	0.01%	4	0.01%	5	0.01%
>50	1	0.00%	0	0.00%	0	0.00%	3	0.01%
Total Patients	41,260	100.00%	41,260	100.00%	41,260	100%	41,260	100.00%

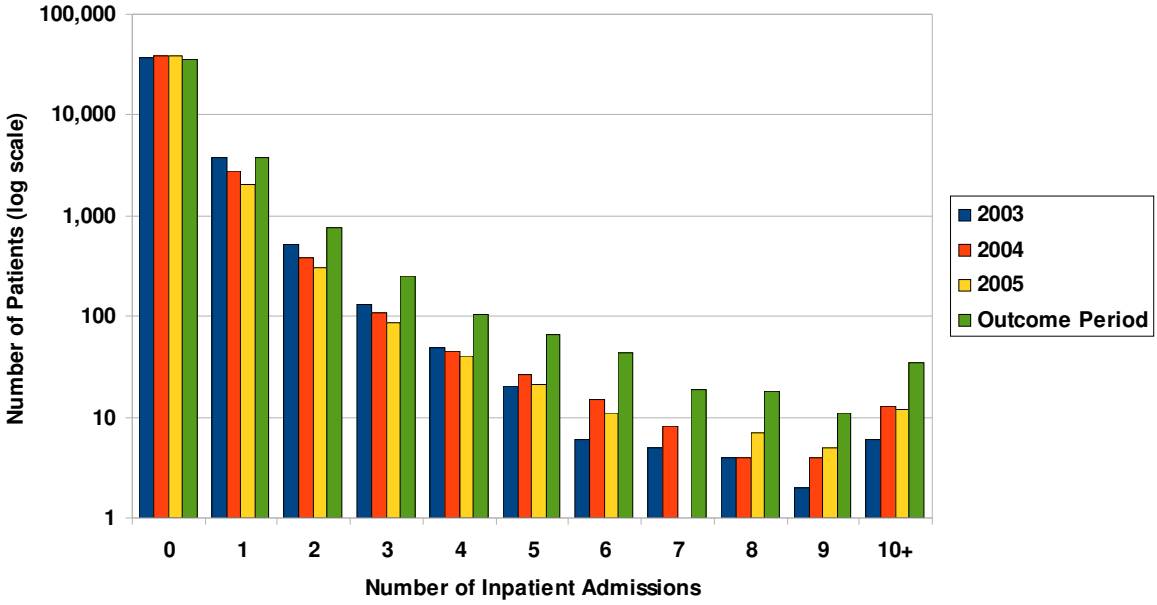
*These columns display the number of patients with y admissions in the specified year only.

**This column displays the number of patients with y admissions during the two year period of 2004 through 2005. Note that this is not the sum of the previous two columns.

The bulk of the patients had no impatient admissions with the frequency tapering quickly as inpatient admissions increases. This trend is illustrated in Figure 3.4. The number of patients with one or more inpatient admissions is highest in 2003 and lowest in 2005. This is likely an artifact of the inclusion criteria and not a trend of decreasing utilization. A patient with high utilization in 2003 is likely to have encounters in 2004 and 2005, making that patient eligible for subgroup inclusion. However, high utilization in 2005, for example, does not imply the same likeliness of encounters in the two

preceding years. This is discussed in greater detail in the Study Limitations section of Chapter.

Figure 3.4 Number of Calibration Subgroup Patients with Number of Inpatient Admissions in the Predictor Period (2003) and the Outcome Period (2004-2005)



The distribution of inpatient admissions is highly skewed for Calibration Subgroup patients for years 2003, 2004, and 2005. The distribution of admissions during the two-year Outcome Period follows a similar pattern although has a slightly higher frequency at each value. Table 3.10 shows the mean number of inpatient admissions per year controlling for gender.

Table 3.10 Descriptive Statistics of Calibration Subgroup Inpatient Admissions during the Predictor Period (2003) and the Outcome Period (2004 - 2005) by Gender

		Number of Inpatient Admissions							
		Predictor Period		Outcome Period (Individual Years)				Outcome Period (Combined)	
	N	2003		2004		2005			
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Male	11,941	0.17	0.67	0.12	0.72	0.09	0.61	0.21	1.20
Female	29,319	0.13	0.58	0.11	0.57	0.08	0.50	0.19	0.95
Total	41,260	0.14	0.61	0.11	0.62	0.09	0.54	0.20	1.03

The mean number of inpatient admissions for Calibration Subgroup patients dropped from 0.14 (SD = 0.61) in 2003 to 0.11 (SD = 0.62) in 2004 and 0.09 (SD = 0.54) in 2004. The mean combined 2004 and 2005 number of inpatient admissions for Calibration Subgroup patients was 0.20 (SD = 1.03).

The mean number of inpatient admissions for Calibration Subgroup patients in 2003 was 0.14 (SD = 0.61), with males having a slightly higher mean than females (0.17, SD 0.67 vs. 0.13, SD 0.58).

Table 3.11 shows the frequency of the number of Outcome Period (combined 2004 and 2005) inpatient admissions (the dependent variable in all models) for Calibration Subgroup patients by age category. Although the largest age category is Under 10 with 10,456 patients, the 20-29 age category has the most admissions (2,008).

Table 3.11 Frequency of the Number of Outcome Period (2004 – 2005) Inpatient Admissions for Calibration Subgroup Patients Controlling for Age Category

Inpatient Admissions	Age Category							
	Under 10	10-19	20-29	30-39	40-49	50-59	60-64	Total
0	9,872	4,306	8,827	5,246	4,003	2,822	1,033	36,109
1	454	518	1,288	610	500	339	128	3,837
2	77	119	183	119	120	112	39	769
3	19	29	32	43	58	54	12	247
4	8	5	19	22	19	22	10	105
5	12	7	12	5	16	8	7	67
6	4	3	6	7	12	7	4	43
7	0	1	0	5	6	7	0	19
8	5	2	2	2	2	4	1	18
9	1	1	1	3	2	3	0	11
10	1	1	1	1	1	0	0	5
11	1	1	1	0	1	1	0	5
12	1	0	1	1	1	0	1	5
13	0	1	1	0	1	1	0	4
14	0	1	0	3	0	0	0	4
15	1	1	1	0	1	0	0	4
20	0	0	0	1	0	0	0	1
>20	0	2	0	2	3	0	0	7
Total Patients	10,456	4,998	10,375	6,070	4,746	3,380	1,235	41,260
Total Admissions	878	1,070	2,008	1,402	1,466	1,027	361	8,212

Inpatient Hospital Utilization for Cross-Validation Subgroup Patients

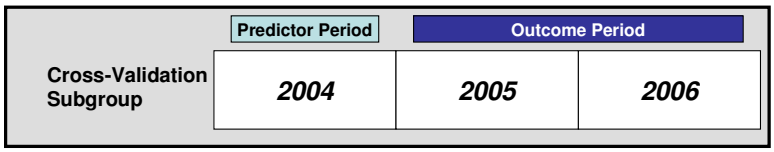


Table 3.12 shows the frequency distribution of 2004, 2005 and 2006 inpatient admissions for the Cross-Validation Subgroup patients. For each patient, the number of 2004 inpatient admissions was used as an independent variable in all models. The number of inpatient admissions for the Outcome Period (2005 through 2006) are shown both separately (i.e., the number of patients with *x* inpatient admissions in each year) and summed (i.e., the number of patients with *x* inpatient admissions in the two year period). The sum of 2005 and 2006 inpatient admissions was used as the dependent variable to evaluate the performance of the models.

The bulk of the patients had no impatient admissions with the frequency tapering quickly as inpatient admissions increases. This trend is illustrated in Figure 3.5. The number of patients with one or more inpatient admissions is highest in 2004. As with the Calibration Subgroup, this is likely an artifact of the inclusion criteria and not a trend of decreasing utilization. A patient with high utilization in 2004 is likely to have encounters in 2005 and 2006, making that patient eligible for subgroup inclusion. However, high utilization in 2006, for example, does not imply the same likeliness of encounters in the two preceding years. This is discussed in greater detail in the Study Limitations section of Chapter.

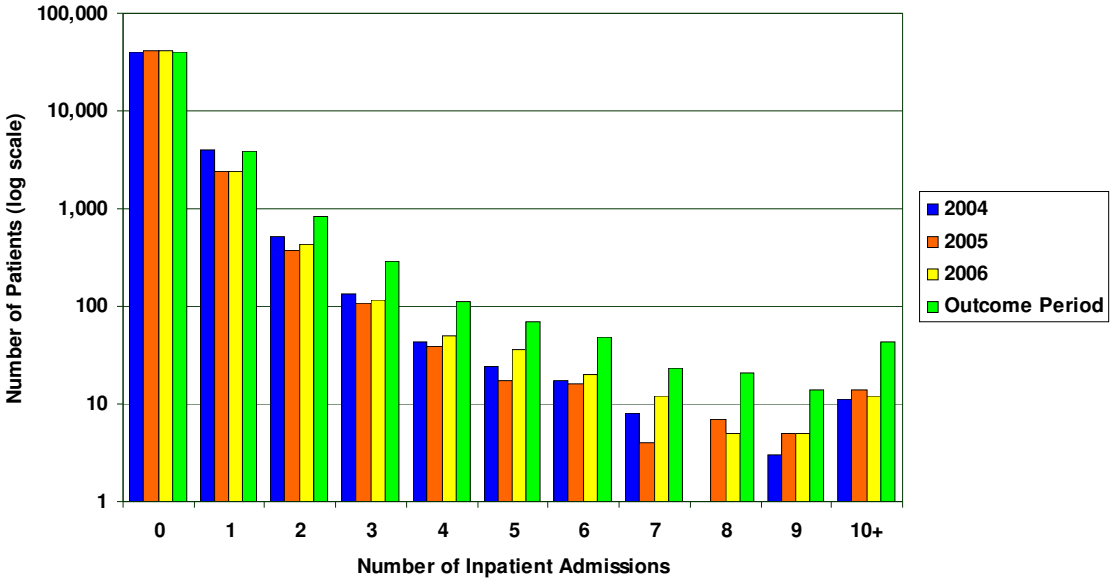
Table 3.12 Number of Cross-Validation Subgroup Patients with Number of Inpatient Admissions in the Predictor Period (2004) and the Outcome Period (2005 through 2006)

	Number of Patients by Year							
	Predictor Period		Outcome Period (Individual Years)*				Outcome Period (Combined)**	
Inpatient Admissions	2004		2005		2006			
0	40,023	89.46%	41,761	93.35%	41,678	93.16%	39,478	88.24%
1	3,958	8.85%	2,391	5.34%	2,366	5.29%	3,809	8.51%
2	518	1.16%	378	0.84%	437	0.98%	831	1.86%
3	132	0.30%	107	0.24%	117	0.26%	291	0.65%
4	43	0.10%	38	0.08%	50	0.11%	110	0.25%
5	24	0.05%	17	0.04%	36	0.08%	70	0.16%
6	17	0.04%	16	0.04%	20	0.04%	48	0.11%
7	8	0.02%	4	0.01%	12	0.03%	23	0.05%
8	1	0.00%	7	0.02%	5	0.01%	21	0.05%
9	3	0.01%	5	0.01%	5	0.01%	14	0.03%
10	3	0.01%	2	0.00%	2	0.00%	9	0.02%
11	2	0.00%	3	0.01%	1	0.00%	10	0.02%
12	0	0.00%	3	0.01%	0	0.00%	3	0.01%
13	2	0.00%	1	0.00%	2	0.00%	2	0.00%
14	0	0.00%	0	0.00%	1	0.00%	4	0.01%
15	0	0.00%	0	0.00%	0	0.00%	1	0.00%
16 - 49	4	0.01%	5	0.01%	4	0.01%	11	0.02%
>50	0	0.00%	0	0.00%	2	0.00%	3	0.01%
Total	44,738	100.00%	44,738	100.00%	44,738	100.00%	44,738	100.00%

*These columns display the number of patients with y admissions in the specified year only.

**This column displays the number of patients with y admissions during the two year period of 2005 through 2006. Note that this is not the sum of the previous two columns.

Figure 3.5: Number of Cross-Validation Subgroup Patients with Number of Inpatient Admissions in the Predictor Period (2004) and the Outcome Period (2005 through 2006)



The distribution of inpatient admissions is highly skewed for Cross-Validation Subgroup patients for years 2004, 2005, and 2006. The distribution of admissions during the two-year Outcome Period follows a similar pattern although has a slightly higher frequency at each value. Table 3.13 shows the mean number of inpatient admissions per year controlling for gender.

Table 3.13 Descriptive Statistics of Cross-Validation Subgroup Inpatient Admissions during the Predictor Period (2004) and the Outcome Period (2005 – 2006) by Gender

		Number of Inpatient Admissions							
		Predictor Period		Outcome Period (Individual Years)				Outcome Period (Combined)	
	N	2004		2005		2006			
		Mean	SD	Mean	SD	Mean	SD	Mean	SD
Male	13,124	0.17	0.74	0.10	0.69	0.12	0.86	0.22	1.43
Female	31,614	0.12	0.56	0.09	0.51	0.09	0.51	0.19	0.89
Total	44,738	0.14	0.62	0.09	0.57	0.10	0.64	0.20	1.08

The mean number of inpatient admissions for Cross-Validation Subgroup patients in 2004 was 0.14 (SD = 0.62), with males having a slightly higher mean than females (0.17, SD = 0.74 vs. 0.12, SD = 0.56). The mean number of inpatient admissions for Cross-Validation Subgroup patients dropped to 0.09 (SD = 0.57) in 2005 and 0.10 (SD = 0.64) in 2006. The mean combined 2005 and 2006 number of inpatient admissions for Cross-Validation Subgroup patients was 0.20 (SD = 1.08). The values in this table are extremely similar to those in Table 3.9 for the Calibration Subgroup patients.

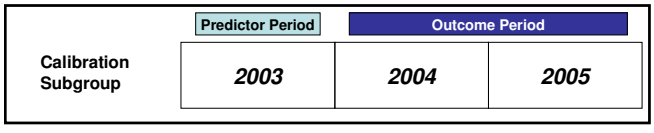
Table 3.14 shows the frequency of the number of Outcome Period (combined 2005 and 2006) inpatient admissions for Cross-Validation Subgroup patients by age category.

Table 3.14 Frequency of the Number of Outcome Period (2005 – 2006) Inpatient Admissions for Cross-Validation Subgroup Patients Controlling for Age Category

Inpatient Admissions	Age Category							
	Under 10	10-19	20-29	30-39	40-49	50-59	60-64	Total
0	10,064	4,769	9,831	5,981	4,497	3,249	1,087	39,478
1	431	556	1,191	602	539	375	115	3,809
2	65	97	195	152	148	129	45	831
3	31	21	54	48	60	57	20	291
4	10	9	18	13	24	28	8	110
5	3	6	9	15	18	13	6	70
6	3	5	3	5	18	10	4	48
7	4	1	2	8	4	2	2	23
8	2	4	4	4	2	5	0	21
9	1	2	1	2	4	4	0	14
10	1	1	0	1	1	2	3	9
11	1	0	3	2	1	3	0	10
12	2	0	0	0	0	0	1	3
13	0	1	0	0	1	0	0	2
14	1	0	1	0	1	1	0	4
15	0	0	0	1	0	0	0	1
16	1	0	0	0	1	0	0	2
17	0	0	2	0	0	0	0	2
18	0	0	0	0	0	0	0	0
19	0	1	0	1	0	0	0	2
20	0	0	0	0	0	0	0	0
>20	0	1	1	3	3	0	0	8
Total Patients	10,620	5,474	11,315	6,838	5,322	3,878	1,291	44,738
Total Admissions	855	1,040	2,100	1,516	1,624	1,198	407	8,314

For the Cross-Validation Subgroup of patients, the 20-29 age category contains the largest number of patients (11,315) and contributes the largest number of inpatient admissions (2,100).

Objective 1: Model Construction



For model construction, the 2003, 2004, and 2005 data for the Calibration Subgroup patients were loaded into R version 2.5.1, a multi-platform, open-source

statistics program.[180] Both the linear regression and generalized linear model functions of R were accessed via the R Commander graphical user interface (R add-on package "Rcmdr").

Model M1 is an ordinary least squares linear model that used Predictor Period inpatient admissions and Predictor Period relative risk (RR) score to predict Outcome Period inpatient admissions.

Model M2 is an ordinary least squares model that used Predictor Period inpatient admissions, Predictor Period age, and sex to predict Outcome Period inpatient admissions.

Model M3 is an ordinary least squares model that used only Predictor Period inpatient admissions to predict Outcome Period inpatient admissions.

Model M4 is a generalized linear model that used Predictor Period inpatient admissions, Predictor Period RR score, and a Poisson link to predict Outcome Period inpatient admissions.

Model M5 is a two-step model using Predictor Period inpatient admissions, Predictor Period RR score and to predict Outcome Period inpatient admissions. Model M5 uses the equations created by two generalized linear models: a logistic regression model that predicts the likelihood of any hospitalization, and a Poisson regression that predicts the number of hospitalizations for those with any hospitalizations. The results of both equations are multiplied to determine the predicted Outcome Period inpatient admissions.

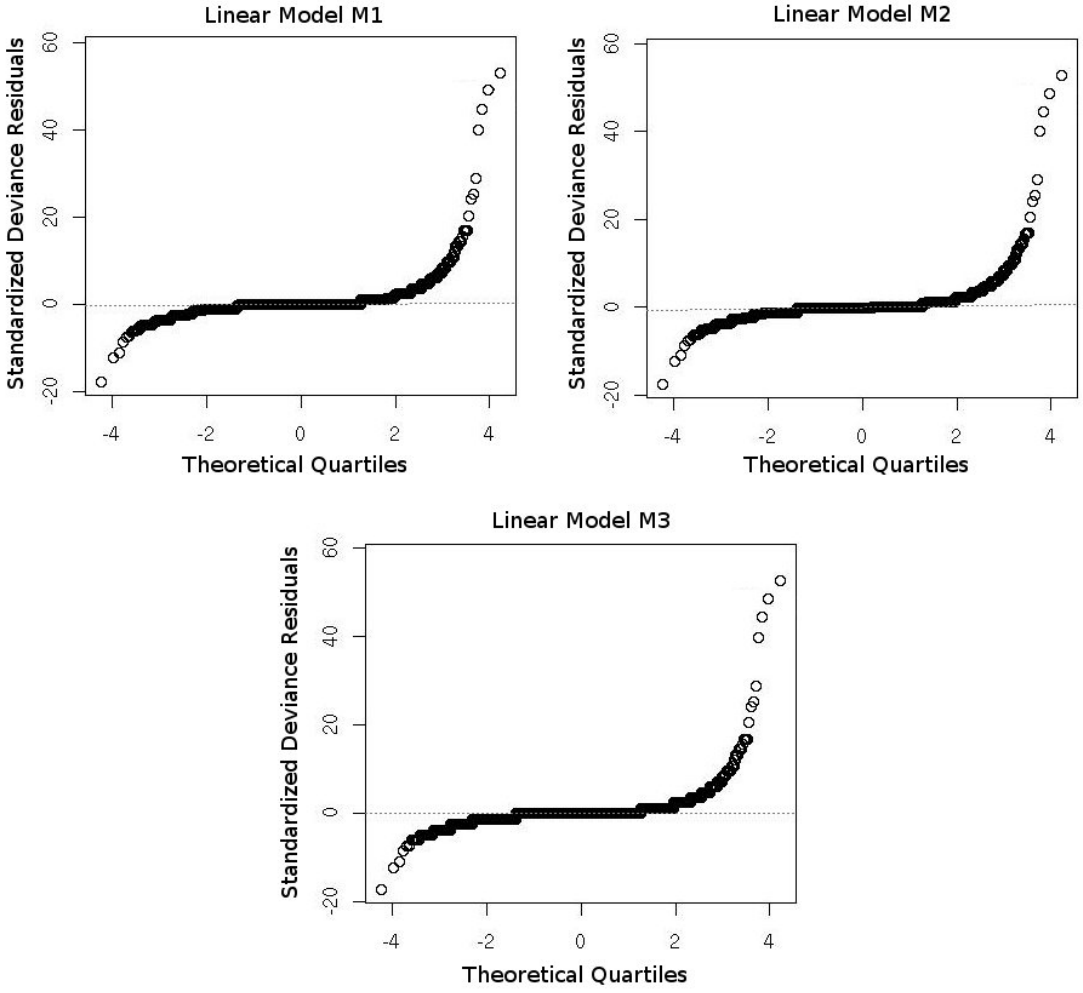
[180] R Development Core Team. R: A language and environment for statistical computing. Available at: http://www.R-project.org.

Models M1, M2, and M3 all use ordinary least squares (OLS) linear regression. Four key assumptions govern the use of linear regression: linearity, independence, constant variance, and normality. The assumption of linearity specifies that a linear relationship should exist between the predictor variables and the outcome variable. This assumption can be tested by evaluating scatterplots of the deviance residuals associated with the model. Figure 3.6 displays the deviance residuals associated with models M1, M2, and M3. In each case, the residuals fail to cluster evenly around the line at $y = 0$, forming instead a distinct S-curve pattern. Based on this, the assumption of linearity is not met.

The assumption of independence stipulates that the error terms associated with the regression variables are independent of each other. In this analysis, lack of independence of outcome variables is likely not a major concern, since the variables of age, sex, and number of inpatient admissions are not difficult to measure. Furthermore, since this is not a time-series analysis, autocorrelation effects are not a concern.

The assumption of constant variance (homoscedasticity) requires that the variance of errors remain constant across the levels of the predictor variables. Generally the assumption of homoscedasticity can be tested by visually inspecting the scatterplots of the deviance residuals. When the assumption is met, the residuals should reside randomly around $y = 0$ for all values of the outcome variable. In Figure 3.6, the lack of linearity makes the assumption of homoscedasticity hard to address.

Figure 3.6 Q-Q Plot of the Standardized Deviance Residuals for the Linear Models M1, M2, and M3



Note: All three models use Predictor Period inpatient admissions as an independent variable. Model M1 additionally uses Predictor Period relative risk score. M2 additionally uses age and sex. M3 uses no other independent variables.

Finally, the assumption of normality implies that variables in OLS linear regression should be normally distributed. The Kolmogorov-Smirnov (K-S) test was used to test the hypothesis that the distribution of the outcome variable was normal in nature.

The resulting K-S D statistic of 0.466 ($p < 0.05$) indicates that the distribution of the outcome variable differs significantly from the normal distribution.

In short, the data do not meet the assumptions for the use of OLS linear regression. As discussed in the Methods chapter of this manuscript, the expected failure of the data in this regard (based on theory) is the rationale for using the non-normal generalized linear models of M4 and M5. (Please refer to the Methods chapter for detailed theoretical support for the use of the techniques associated with M4 and M5). However, models M1, M2, and M3 will still be used for two main reasons. First, although theory and tests of the regression assumptions suggest that non-normal models are preferred for utilization data, the practical difference in results between the methods is unknown. Second, the use of OLS linear regression is quite common for predicting utilization despite the failure to meet the assumptions.

Table 3.15 shows how each model equation was constructed.

Table 3.15 Model Parameter Coefficients, Coefficient Descriptive Statistics, and Equations

Model	Parameter	Parameter Coefficient Estimate	Std. Error	*t*/*z**	*p*	R-squared	Equation
M1	Intercept	0.01	0.005	1.789	0.07	0.354	$\hat{y}$ = 0.01 + 0.99*Admissions + 0.15*RR
	2003 Inpatient Admissions	0.99	0.007	147.022	<0.01		
	2003 Relative Risk Score	0.15	0.012	13.157	<0.01		
M2	Intercept	-0.04	0.016	-2.752	0.01	0.358	$\hat{y}$ = -0.04 + 0*Age + 1.01*Admissions - 0.01*Sex
	Age	0.00	0.000	20.192	<0.01		
	2003 Inpatient Admissions	1.01	0.007	150.529	<0.01		
	Sex	-0.01	0.009	-1.071	0.28		
M3	Intercept	0.06	0.004	13.53	<0.01	0.352	$\hat{y}$ = 0.06 + 1.00*Admissions
	2003 Inpatient Admissions	1.00	0.007	149.56	<0.01		
M4	Intercept	-1.83	0.013	-145.84	<0.01	N/A**	$\hat{y}$ = exp^(-1.83 + 0.12*Admissions + 0.42RR)
	2003 Inpatient Admissions	0.12	0.001	86.76	<0.01		
	2003 Relative Risk Score	0.42	0.012	35.37	<0.01		
M5 pt. 1	Intercept	-2.31	0.021	-109.99	<0.01	N/A**	$\hat{y} = \frac{\exp^{(-2.31 + 0.81*\text{Admissions} + 0.58\text{RR})}}{1 + \exp^{(-2.31 + 0.81*\text{Admissions} + 0.58\text{RR})}}$
	2003 Inpatient Admissions	0.81	0.027	29.36	<0.01		
	2003 Relative Risk Score	0.58	0.037	15.63	<0.01		
M5 pt. 2	Intercept	0.34	0.013	25.816	<0.01	N/A**	$\hat{y}$ = exp^(0.34 + 0.08*Admissions + 0.12RR)
	2003 Inpatient Admissions	0.08	0.001	56.476	<0.01		
	2003 Relative Risk Score	0.12	0.015	7.945	<0.01		

Note: M1, M2, and M3 are linear models. M4 is a Poisson regression generalized linear model. M5 is a two-step generalized linear model (logistic and Poisson, respectively).

*For linear models M1, M2, M3 the *t statistic* is displayed. For generalized linear models M4 and M5, the *z* statistic is displayed.

**R-squard does not exist for generalized linear models

Judging by the parameter coefficients and the range of the predictor variables, models M1, M2, and M3 relied almost exclusively on the number of Predictor Period inpatient admissions to estimate the number of Outcome Period inpatient admissions. In model M1, the relatively low coefficient for relative risk score, 0.15, limits the extent to which this variable can impact predicted Outcome Period inpatient admissions. For example, the maximum Predictor Period (2003) relative risk score of 10.02 would only add approximately 1.5 to the predicted number of Outcome Period inpatient admissions. On the other hand, the maximum number of Predictor Period inpatient admissions, 86, would add approximately 86 to the predicted number of Outcome Period inpatient admissions.

Models M4 and M5 both relied more heavily on relative risk score than did models M1, M2, and M3. In model M4, relative risk score had a coefficient of 0.42 compared to 0.12 for number of current year inpatient admissions. Model M5 used the number of current year inpatient admissions slightly more than relative risk score to determine the likelihood of next year hospitalization (step 1). However model M5 used relative risk score slightly more than the number of current year inpatient admissions to determine the magnitude of next year hospitalizations (step 2).

Table 3.16 Quartiles of Deviance Residuals by Model Equation

	Min	25%	Median	75%	Max
M1	-14.57	-0.09	-0.04	-0.02	40.94
M2	-14.28	-0.13	-0.05	0.04	40.61
M3	-14.12	-0.06	-0.06	-0.06	40.74
M4	-8.24	-0.62	-0.59	-0.57	16.68
M5 pt. 1	-4.78	-0.50	-0.46	-0.44	2.18
M5 pt. 2	-4.39	-0.44	-0.40	-0.09	11.96

Table 3.16 shows the quartiles of deviance residuals for the equations of each model. The median deviance residuals for the equations of models M1, M2, and M3 were lower than those for models M4 and M5, although the range of deviance residuals was greater for models M1, M2, and M3.

Testing of Objective 1 Hypotheses

With the predictive models created and relevant model statistics calculated, the hypotheses for Objective 1 can be evaluated.

H1: Model fit will vary between the five models.

The fit of the models can be compared by assessing the deviance residuals associated with each model's predicted scores. Table 3.16 shows the median deviance residuals for each model equation. The linear models (M1, M2, and M3) have similar median deviance residuals: -0.04, -0.05, and -0.06, respectively. The generalized linear models have much higher median deviance residuals: -0.59 for M4 and -0.46 and -0.40 for the two equations of M5. R-squared is a common measure of the goodness of fit for linear models. R-squared values for the three linear models were very similar: 0.354, 0.358, and 0.352 for M1, M2, and M3, respectively. Unfortunately, no true analog of R-squared exists for generalized linear models, such as the logistic and Poisson regression models used in M4 and M5.

Perhaps the most straightforward method of determining whether model fit differs between the models is by comparing plots of the standardized deviance residuals. The Q-Q plots in Figure 3.6, used previously to test the OLS linear regression assumptions, show the distributions of standardized deviance residuals by theoretical quartile for the three linear models, M1, M2, and M3.

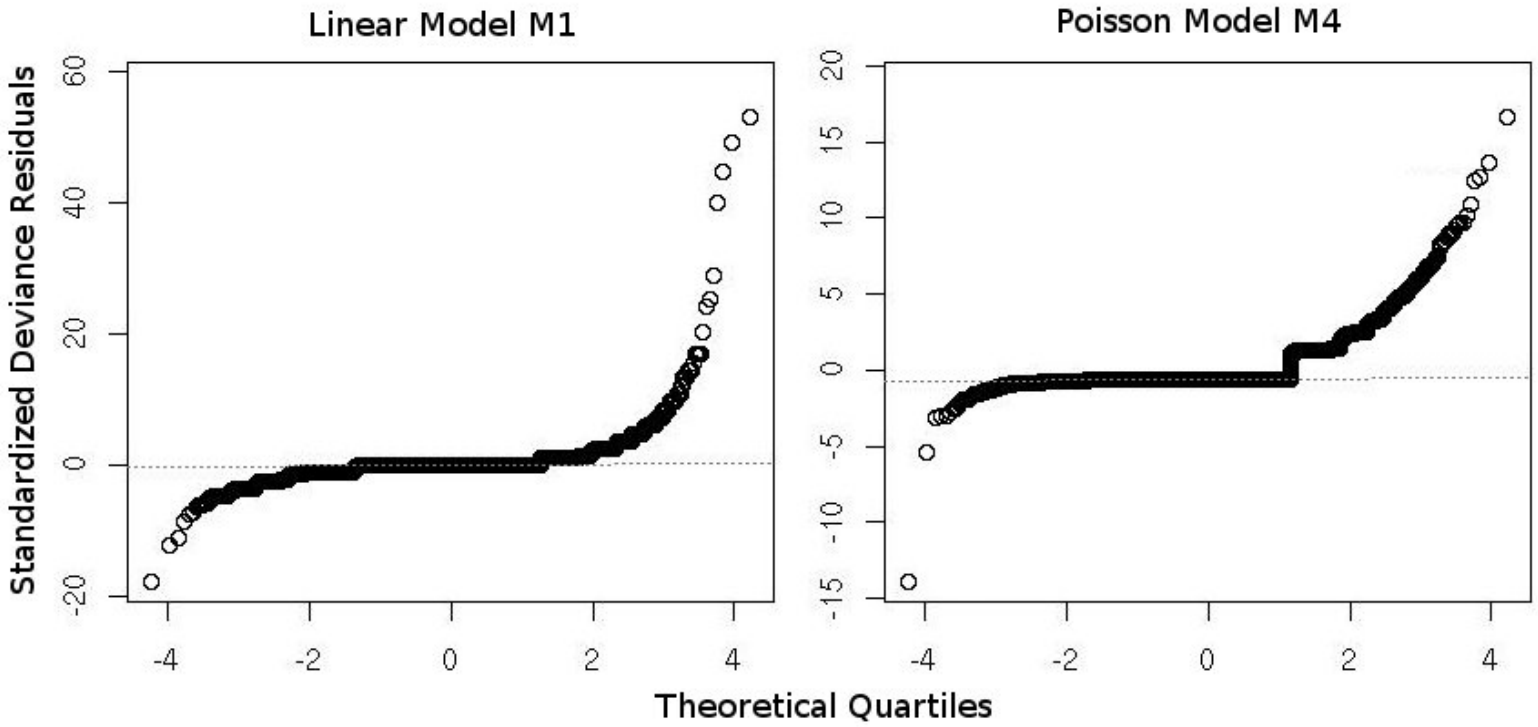
The Q-Q plots for models M1, M2, and M3 are virtually identical. This in conjunction with the fact that the median deviance residuals are also extremely similar for these three models leads one to conclude that very little difference in model fit exists between the three models. Therefore, hypothesis H1 is only tenable if the generalized linear models show sufficient difference in terms of model fit with their linear model counterparts.

Models M1 and M4 both use the same variables; however, M1 was created via linear regression while M4 was created via Poisson regression. The two models have the largest difference between their respective median deviance residuals, but the magnitude of difference is difficult to determine without a visual representation. Figure 3.7 shows the standardized deviance residuals by theoretical quartiles for models M1 and M4.

The Q-Q plots for models M1 and M4 show how differently the two models fit the underlying data. While M1 has a lower median deviance residual, its range is much greater, extending from -14.57 to 40.94, compared to the narrower range of M4: -8.24 to 16.68. Model M1's deviance residuals appear more evenly distributed between under- and over-prediction than that of M4, which is more heavily associated with over-prediction.

The difference in fit between M1 and M4 is enough to support hypothesis H1 that differences in model fit exist between the models, even if that difference is only apparent in comparisons of linear and generalized linear models. Based on this difference, hypothesis H1 is supported.

Figure 3.7 Q-Q Plot of the Standardized Deviance Residuals for Models M1 and M4



Note: Models M1 and M4 both use the number of Predictor Period (2003) inpatient admissions and relative risk score as independent variables.

H2: Models M1, M4, and M5 will have lower deviance residuals than models M2 and M3.

Models M4 and M5 have higher median deviance residuals than models M2 and M3, as seen in Table 3.16. Benefits of using the DCG model are not apparent when looking solely at model fit. If such benefits exist, they must be found in the cross-validation analyses of Objective 2. Hypothesis H2 is rejected.

H3: Model M1 will have a higher R^2 than the square of the Pearson product moment correlation between the 2003 DCG RR scores and the number of inpatient admissions in 2004 and 2005.

The R^2 value for model M1 is 0.354, while the square of the Pearson product moment correlation between the DCG RR scores and number of Outcome Period inpatient admitted days is 0.020. This disparity is consistent with the difference in model M1 coefficients for these two independent variables. Clearly, adding prior utilization data increases the explained variance of the dependent variable. Hypothesis H3 is supported.

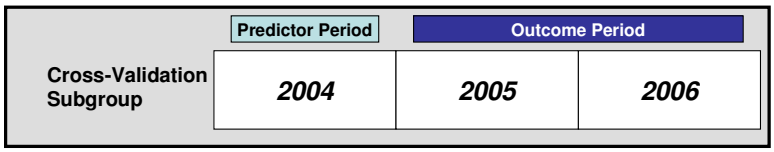
H4: Model M4 will have lower deviance residuals than model M1.

Despite the theoretical rationale for using a non-normally distributed model for the count data of the dependent variable, the median deviance residual for the linear model is much lower: -0.04 for model M1 vs. -0.59 for model M4. Hypothesis H4 is rejected.

H5: Model M5 will have the lowest deviance residuals of all models.

Despite the theoretical rationale for using a two-step model, model M5 had a much higher median deviance residual than all three of the linear models: -0.40 for model M5 vs. -0.04, -0.05, and -0.06 for models M1, M2, and M3, respectively. Based on the deviance residual scores, hypothesis H5 is rejected.

Objective 2: Model Performance



After the model equations were created using the Calibration Subgroup patients' data, Predictor Period (2004) data for Cross-Validation Subgroup patients were inserted into the equations to produce predicted inpatient admissions for the Outcome Period (2005 through 2006). By comparing this predicted number of admissions with the actual number of admissions in this patient subset, model accuracy and performance can be assessed.

Although Objectives 2 and 3 in the Methods section of this document call for the high-use subset to be defined as the patients with the top 2% of actual inpatient use, a cut-off at precisely 2% is not possible, since large blocks of patients have the same number of Outcome Period inpatient admissions. The 1,451 patients with two or more inpatient admissions during the Outcome Period represent 3.24 percent of the Cross-Validation Subgroup. The 620 patients with three or more inpatient admissions in the Outcome Period represent approximately 1.5% of the Cross-Validation Subgroup. The latter subset will be considered the actual high-use patient subset for analysis purposes. Similarly, for each model, the 620 patients with the highest predicted utilization were considered its identified high-risk patient subset. Model M3 assigned large numbers of patients the same predicted utilization scores; because of this lack of differentiation, a subset of the highest-risk 620 patients was impossible to identify. Therefore, comparisons of high-risk subsets involving model M3 were impossible.

Table 3.17 compares various attributes of the 620 patients rated highest risk by each of the models.

Table 3.17 Summary of Performance of Predictive Models in Cross-Validation Subgroup of Patients

Patient Subset Descriptive Measure	Actual High Use Subset*	Predictive Models				
		M1**	M2**	M3***	M4	M5
Number of actual high use patients contained in predicted subset˚	620	166	165	NA	110	153
Positive predictive value (PPV)˚˚	NA	26.8%	26.6%	NA	17.7%	24.7%
Mean absolute prediction error (MAPE)˚˚˚	NA	0.29	0.30	0.29	0.33	0.23
Number of actual high use patients contained in predicted subset who had no Predictor Period˚˚˚˚ inpatient admissions	304	0	0	0	12	6
Total Outcome Period˚˚˚˚ inpatient admissions of predicted subset	3,269	1,597	1,624	NA	1,171	1,513
Mean number of Outcome Period inpatient admissions per correctedly identified patient	5.27	7.78	7.99	NA	8.38	7.98
Standard deviation of Outcome Period inpatient admissions per correctly identified patient	6.85	12.23	12.40	NA	14.60	12.70
Mean Predictor Period rescaled relative risk score per patient contained in predicted subset	2.58	3.56	3.19	NA	7.43	6.16
Mean 2005 rescaled relative risk score per patient contained in predicted subset	3.20	2.64	2.50	NA	4.93	4.20
Mean 2006 rescaled relative risk score per patient contained in predicted subset	3.68	2.75	2.61	NA	5.05	4.34
Mean number of assigned Predictor Period Hierarchical Condition Categories (HCCs) per patient contained in predicted subset	8.11	12.15	11.99	NA	10.59	11.87
Mean number of assigned Outcome Period Hierarchical Condition Categories (HCCs) per patient contained in predicted subset	19.96	13.84	13.79	NA	12.94	13.99
Number of patients contained in predicted subset with a pregnancy-related ICD9 code in the Predictor Period	42	61	108	NA	9	19
Number of patients contained in predicted subset with a pregnancy-related ICD9 code in the Outcome Period	91	46	74	NA	12	18

*The actual high use patient subset includes the 620 patients who had 3 or more inpatient admissions in the Outcome Period.

**With both M1 and M2, two patients tied for the 620th highest predicted utilization score. Therefore the subsets for both M1 and M2 have 621 patients instead of 620.

***Model M3 exhibited insufficient discriminatory capability to identify a predicted subset. I.e., very large groups of patients received the same predicted utilization.

˚The predicted subset for each model are the 620 patients with the highest number of predicted Outcome Period inpatient admissions based on Predictor Period data.

˚˚PPV is calculated by dividing the number of correctly predicted high use patients with the total number of high use patients.

˚˚˚˚The Predictor Period for the Cross-Validation Subgroup is calendar year 2004. The Outcome Period for the Cross-Validation Subgroup is 2005 through 2006.

˚˚˚MAPE is average of the absolute values of all predicted number of Outcome Period inpatient days subtracted from the actual number of Outcome Period inpatient days.

Models M1 correctly identified the largest number of actual high use patients, 166, equating to a positive predictive value (PPV) of 26.8% (166/620). Model M2 correctly identified 165 actual high use patients, equating to a PPV of 26.6%. Model M5 correctly identified 153 actual high use patients, equating to a PPV of 24.7%. Model M4 correctly identified the lowest number of actual high use patients, 110, equating to a PPV of 17.7%.

Mean absolute prediction error (MAPE) is another summary measure of the accuracy of predictive models. MAPE is the average of the absolute differences between predicted admissions and actual admissions. Model M5 had the lowest MAPE, 0.23. Models M1 and M3 had a MAPE of 0.29, compared with 0.30 for M2 and 0.33 for M4.

Of the 620 actual high use patients, 304 had no 2004 inpatient admissions. Models M1 and M2, which relied almost exclusively on Predictor Period (2004) inpatient admissions to predict Outcome Period inpatient admissions, identified none of these patients. Model M4, which had a higher relative weighting of relative risk score than the other models, identified 12 of these patients. Model M5 identified 6 of these patients.

To simplify comparisons involving relative risk score, the Predictor Period relative risk score was rescaled to have a mean of one. The rescaled descriptive statistics are shown in Table 3.7. Not surprisingly the high-risk patients identified by models that focused more heavily on the number of Predictor Period inpatient admissions (M1 and M2) had lower mean Predictor Period rescaled relative risk scores than those identified by the models that more heavily weighted relative risk score (M4 and M5). The mean Predictor Period rescaled relative risk score for high-risk patients identified by model M4

was 7.43, compared to 6.16 for M5, 3.56 for M1, 3.19 for M2, and 2.58 for the actual high-use patients.

A total of 42 actual high-use patients had a pregnancy-related ICD-9 diagnosis in the Predictor Period. Inpatient admissions associated with pregnancy in the Predictor Period would increase a patient's predicted Outcome Period utilization for models that heavily weight Predictor Period inpatient utilization. Consistent with this, the high-risk patients identified by model M1 included 61 patients with a pregnancy-related ICD-9 diagnosis in the Predictor Period, compared to 108 for M2, 9 for M4, and 19 for M5. A total of 91 actual high-use patients had at least one pregnancy related ICD-9 diagnosis in the Outcome Period, compared to 61 for M1, 74 for M2, 14 for M4, and 20 for M5.

The overall accuracy of each model can be depicted graphically by plotting how many actual high-use patients are contained in *x*% of patients with highest predicted utilization. The resulting curves for each model can be compared with each other, with a hypothetical perfect model, and with null model that assigns risk randomly. This type of chart, called a cumulative accuracy profile (CAP), is commonly used in the credit industry where models are designed to predict the risk of foreclosure.[181] Figure 3.8 shows the CAPs of models M1, M2, M4, and M5. Each curve is summarized into an accuracy ratio, which is the area under the curve and above the null model divided by that of the perfect model. The closer an accuracy ratio is to 1.0 the more accurate the model. All models had relatively close accuracy ratios, the highest of which, 0.82, belonged to M1, while the lowest, 0.77, belonged to M2.

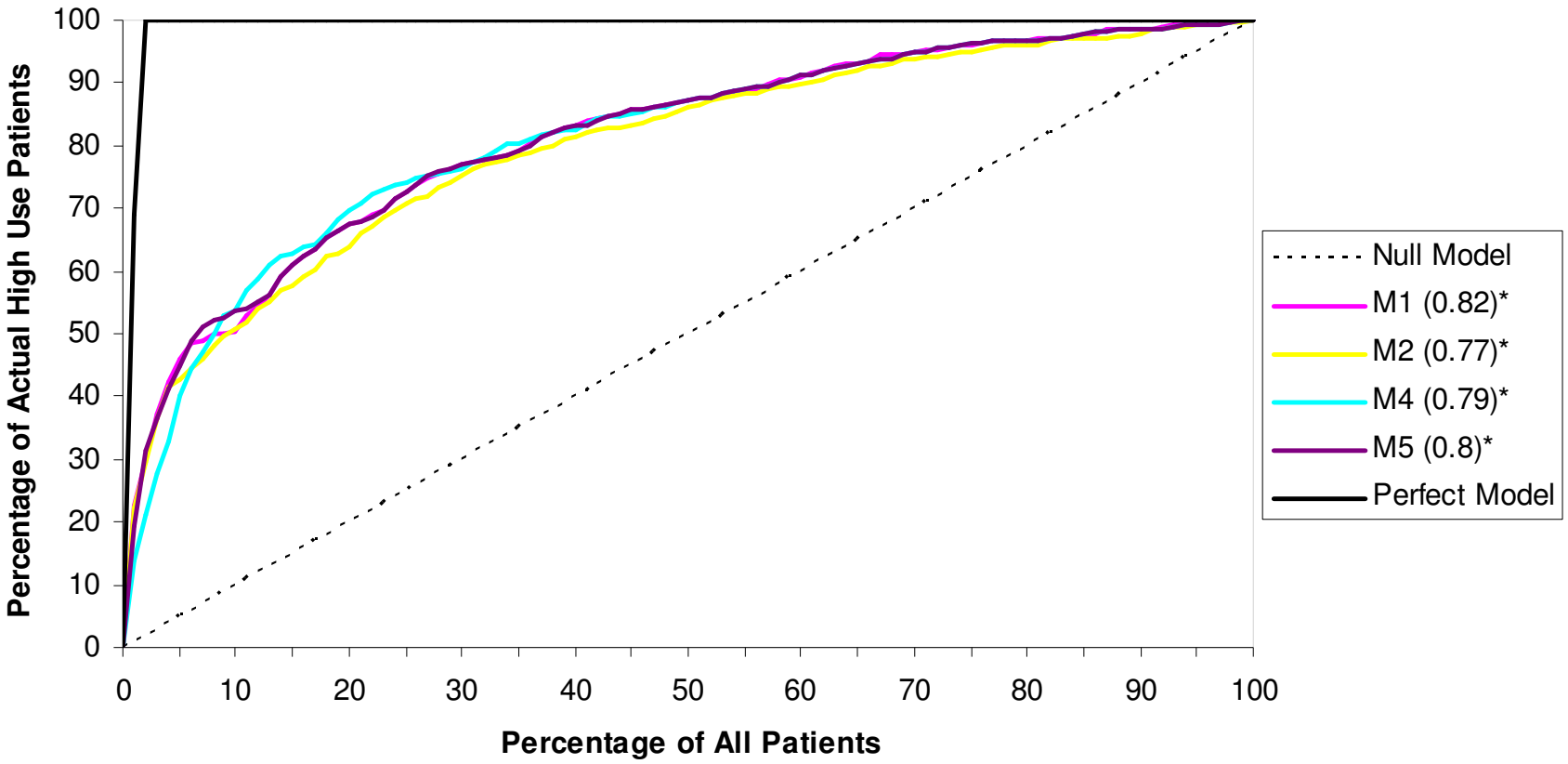
[181] Oderda, Dacorogma and Jung. Credit risk models - do they deliver their promises? A quantitative assessment. *Economic Notes.* 2003;32(2):177-195.

The CAP of model M4 lags behind those of the other models through approximately 50 percent of the actual high use subset, at which point it switches to lead the other models through approximately 80 percent of the actual high use subset. Beyond 80 percent, all model CAPs converge and follow the same curve. The CAP of model M2 lags those of all other models beyond approximately 45 percent of the actual high use subset.

Although the CAPs of all models appear relatively similar, some important information can be gleaned from their differences. Since model M2 relies almost exclusively on Predictor Period inpatient admissions and model M4 relies heavily on Predictor Period relative risk score, the point at which their respective CAPs cross likely indicates the point at which prior inpatient admissions ceases to be the most important predictor of future admissions.

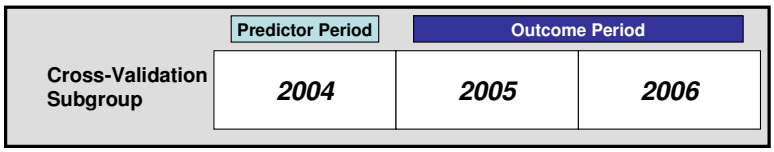
Despite the differences in median deviance residuals for the linear and generalized linear models, the CAP curves are highly similar.

Figure 3.8 Cumulative Accuracy Profiles of Models on Cross-Validation Subgroup Patients



This chart illustrates, for each model, the percentage of actual high use patients y(x) contained in the subset of x% of patients with highest predicted use.
The Null Model represents a model that assigns risk randomly.
The Perfect Model represents a model that perfectly predicts whether each patient will be categorized as high use or not.
*The numbers in parentheses are Accuracy Ratios, which are the ratio of each model's area under the curve but above the Null Model to the Perfect Model's area under the curve but above the Null Model. The closer an Accuracy Ratio is to one the better.

Comparison of Models and Utilization Method



The Utilization Method (UM) is the process of identifying potential high-use patients through a query of patients' prior utilization data. The UM identifies patients as high risk when, in the Predictor Period, they have an inpatient or emergency department (ED) diagnosis of any of the following: diabetes, hypertension, congestive heart failure, asthma, obesity, and tobacco abuse. Furthermore, the patient must have at least one Predictor Period inpatient admission or 12 or more ED encounters.

The UM identified 344 high-risk Objective 2 patients. Unlike models M1 through M5, the UM does not assign a risk score to all patients, meaning creating a subset of the highest risk 620 (or any other number than 344) patients based solely on the UM is not possible. Therefore for comparison purposes, the highest risk 344 patients were selected using each model except M3, which lacked the granularity to create this cut-off point. Table 3.18 shows how each model's high-risk subgroup of patients compared with the subset identified by the UM.

Table 3.18 Summary Comparison of Performance of Models and Utilization Method Among Cross-Validation Subgroup Patients

Patient Subset Descriptive Measure	Utilization Method*	Predictive Models				
		M1	M2**	M3***	M4	M5
Number of actual high use patients contained in predicted subset	104	108	111	NA	69	106
Positive predictive value (PPV)˚	30.23%	31.40%	32.27%	NA	20.06%	30.81%
Number of actual high use patients contained in predicted subset who had no Predictor Period˚˚ inpatient admissions	6	0	0	NA	8	1
Total Outcome Period˚˚˚ inpatient admissions of predicted subset	913	1,194	1,201	NA	788	1,117
Mean number of Outcome Period inpatient admissions per correctly identified patient	8.78	9.35	9.24	NA	9.77	9.04
Standard deviation of Outcome Period inpatient admissions per correctly identified patient	15.06	14.86	14.68	NA	18.08	14.84
Mean Predictor Period rescaled relative risk score per patient contained in predicted subset	3.43	4.59	3.39	NA	8.98	7.26
Mean 2005 rescaled relative risk score per patient contained in predicted subset	2.78	3.35	2.76	NA	6.03	4.93
Mean 2006 rescaled relative risk score per patient contained in predicted subset	2.85	3.42	2.82	NA	6.20	4.88
Mean number of assigned Predictor Period Hierarchical Condition Categories (HCCs) per patient contained in predicted subset	13.74	13.34	13.34	NA	10.68	13.15
Mean number of assigned Outcome Period Hierarchical Condition Categories (HCCs) per patient contained in predicted subset	16.58	15.37	15.15	NA	12.73	15.37

*The Utilization Method identified 344 patients. To compare models M1 - M5 with the Utilization Method, the 344 patients with the highest predicted utilization were selected for each model.

**With M2, three patients tied for the 344st highest predicted utilization score. Therefore the subset for M2 has 346 patients instead of 344.

***Model M3 exhibited insufficient discriminatory capabilitity to identify a predicted subset. I.e., very large groups of patients received the same predicted utilization.

˚PPV in this table is the percentage of the subset of 344 predicted high use patients who were among the 644 actual patients with 3 or more inpatient admission in the Outcome Period.

˚˚The Predictor Period for the Cross-Validation Subgroup is calendar year 2004.

˚˚˚The Outcome Period for the Cross-Validation Subgroup is 2005 through 2006.

As shown in Table 3.18, the high-risk patient subgroups identified by the UM, M1, M2, and M5 all correctly identified a similar number of actual high-use patients, ranging from 104 with the UM to 111 with model M2. Model M4 identified the lowest number of actual high use patients, 69. However, M4 correctly identified the highest number of patients who had no Predictor Period (2004) inpatient admissions (8). The UM identified six such patients and M5 identified one. M4's subset of high-risk patients also had the highest mean number of Outcome Period (2005 through 2006) inpatient

admissions at 9.77 (SD = 18.08), while the UM had the lowest at 8.78 (SD = 15.06). M4's subset of high-risk patients also had the highest mean rescaled relative risk score at 8.98, with M5 next highest at 7.26, and the UM last at 3.43. The subsets identified by the UM and by M1, M2, and M5 all had an average of approximately 13 HCCs assigned to each patient. M4's subset had an average of just under 11 Predictor Period HCCs assigned per patient.

Another method to compare the performance of the models and the UM is to calculate mean and median percentile rankings of their high-risk patient subsets. To do this, all patients were ranked in order of actual Outcome Period inpatient utilization, with ties resolved by the mean of 2005 and 2006 DCG relative risk scores. For example, if two patients each had seven Outcome Period inpatient admissions, the one with the higher mean 2005 and 2006 relative risk score would be ranked ahead of the other patient. Then, each patient is given a percentile score, which is the percentage of patients in the Cross-Validation subset who have a lower ranking. The highest-use patient ranked in this fashion would receive a percentile score of near 100 percent, while the lowest ranked patient would receive a percentile score of near zero percent. The mean and median percentile scores of the high-risk subsets identified by each model and the UM were then calculated. Table 3.19 displays this information for subsets with the full high-risk patient subsets of 620 patients and for the smaller, 344 patient subset.

Although model M4 had the lowest PPV of all models for the 620 high-risk patient subsets, it had the highest mean percentile score (88.3%) and lowest percentile score standard deviation (14.6%). Overall, models M4 and M5 identified the patients

with the highest mean and median percentile scores. Among the top identified 344 patients, the UM had nearly as high mean and median percentile scores as M5 but with slightly less variability.

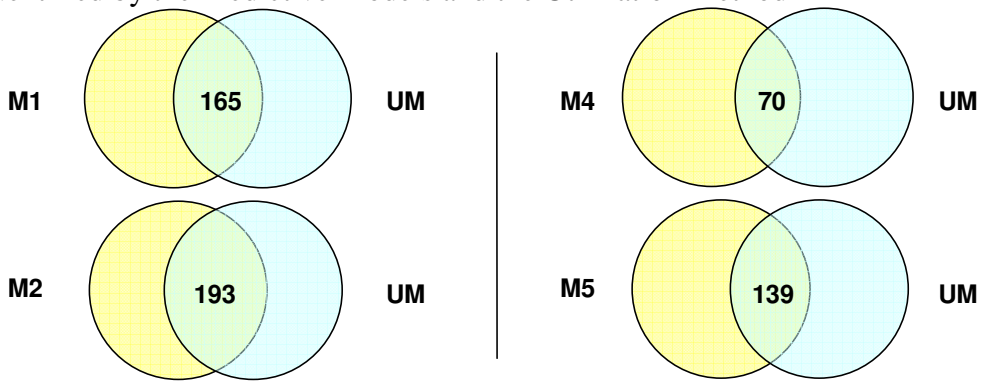
Table 3.19 Mean and Median Percentile Ranking of Patients Identified as High-Risk by the Models and Utilization Method

	M1	M2	M4	M5	UM
Top 620 Patients					
Mean Percentile	84.1%	83.6%	**88.3%**	87.3%	NA
Standard Deviation	23.2%	22.9%	**14.6%**	19.1%	NA
Median Percentile	95.1%	94.7%	88.2%	**95.5%**	NA
Top 344 Patients					
Mean Percentile	87.5%	88.0%	88.3%	**90.3%**	89.8%
Standard Deviation	20.9%	19.7%	15.1%	15.4%	**13.6%**
Median Percentile	**96.4%**	96.3%	88.2%	**96.4%**	96.2%

Note: Percentile was determined by ranking patients in descending order by Outcome Period (2005 through 2006) inpatient utilization followed by the mean of their individual year 2005 and 2006 relative risk scores. Best scores are bolded.

Figure 3.9 uses Venn diagrams to show the number of shared patients identified by the UM and each model with the exception of M3. This comparison uses the 344-patient subsets of high-risk patients identified by each model and the UM. The smaller patient subsets (as opposed to the 620-patient subsets) are used to determine whether the highest-risk patients as identified by the UM are the same as the highest-risk patients as identified by the predictive models. M4 showed the lowest amount of agreement with the other models and the UM, sharing 70 patients (20.1%) with the UM. The UM shared a maximum of 165 patients (44.5%) with M1.

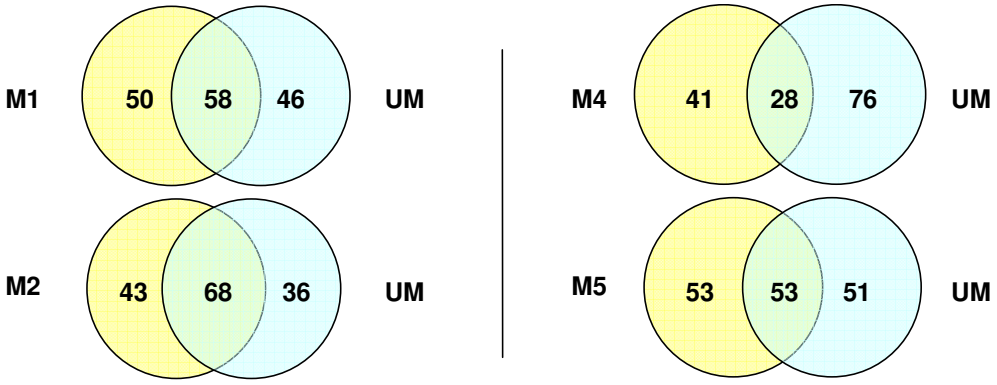
Figure 3.9 Venn Diagrams Depicting the Number of Shared Patients in the High-Risk Subsets Identified by the Predictive Models and the Utilization Method



The Venn diagrams show the number of patients in the high-risk patient subsets of both the Utilization Method and the predictive model. For example, 165 patients were among both sets of 344 patients identified as high-risk by the Utilization Method and model M1.

Figure 3.10 is similar to Figure 3.9 but shows only correctly identified actual high-use patients (i.e., those patients who were among the 344 high-risk patients of a particular model or the UM who were among the actual high-use patients). Model M2 shared the largest number of correctly identified high-use patients (68) with the UM. In other words, of the 111 actual high-use patients correctly identified by M2, 68 were also among the 104 actual high-use patients correctly identified by the UM. Of the 104 actual high-use patients correctly identified by the UM, only 28 were among the 69 actual high-use patients correctly identified by M4.

Figure 3.10 Venn Diagrams Depicting the Number of Shared Patients in the High-Risk Subsets Identified by the Predictive Models and the Utilization Method who were Actual High-Use Patients



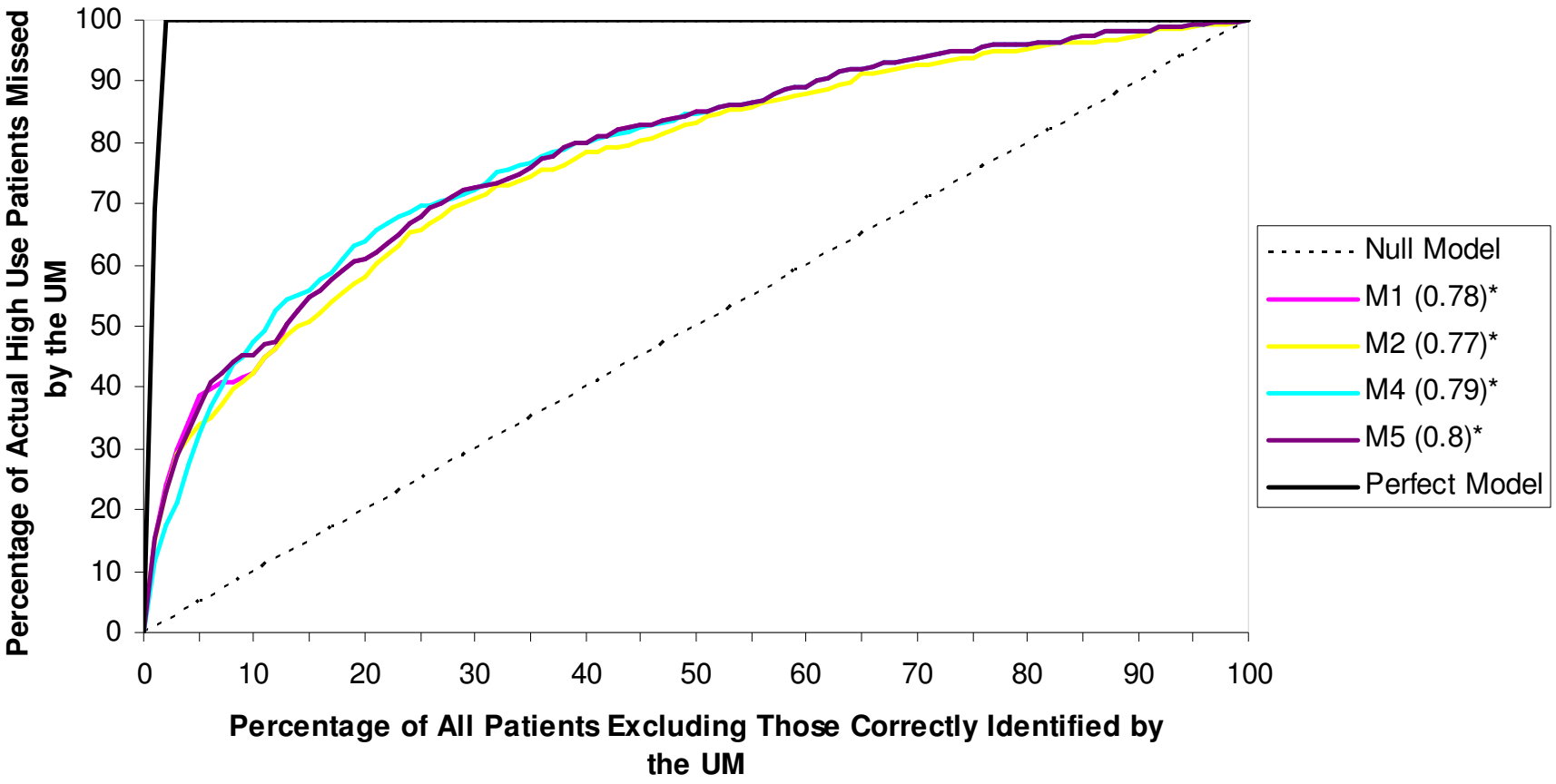
The Venn diagrams show the number of patients in the high-risk patient subsets of both the Utilization Method and the predictive model who were actual high use patients. For example, model M1 correctly identified 108 actual high use patients. Of these, 58 were among the 104 actual high use patients correctly identified by the Utilization Method.

Since the predictive models give risk scores to all patients while the UM identifies only a limited subset, it is helpful to look at the ability of the models to identify actual high-use patients who do not meet the criteria of the UM. Figure 3.11 shows the cumulative accuracy profiles of models M1, M2, M4, and M5 for actual high-use Cross-Validation Subgroup patients who were not identified by the UM.

Models M1, M2, M4, and M5 all had Accuracy Ratios of near 0.80 for this subset of patients. Each model's top 10% highest-risk patients included 40% or more of actual high-use Cross-Validation Subgroup patients missed by the UM. The cumulative accuracy profile curve of model M2 begins to lag the models that use RR scores (M1, M4, and M5) after identifying approximately 30% of the high-use patients. Models M1, M4, and M5 identified 31 actual high-use patients that were missed by both the Utilization Method and model M2 (which does not use RR scores). Specifically, model

M4 identifies 29 such patients, with models M5 and M1 identifying 20 and 4, respectively. These specific patients, representing 5.0 percent of the total actual high-use patient subset, are the contribution of the RR score as a predictor variable.

Figure 3.11 Cumulative Accuracy Profiles of Models on Cross-Validation Subgroup Patients Not Identified by the Utilization Method



This chart illustrates, for each model, the percentage of actual high use patients y(x) contained in the subset of x% of patients with highest predicted use.
This chart focuses on patients who were not identified by the Utilization Method.
The Null Model represents a model that assigns risk randomly.
The Perfect Model represents a model that perfectly predicts whether each patient will be categorized as high use or not.
*The numbers in parentheses are Accuracy Ratios, which are the ratio of each model's area under the curve but above the Null Model to the Perfect Model's area under the curve but above the Null Model. The closer an Accuracy Ratio is to one the better.

Testing of Objective 2 Hypotheses

H6: The Utilization Method and Models M1-M5 will identify different patients as highest-risk.

Figure 3.9 shows that the highest number of shared identified patients between the UM and one of the predictive models was 165 (44.5%), the number shared with model M1. The UM shared only approximately 20 percent of identified patients with model M4, the lowest of its pairings with the predictive models. Although some overlap exists, the majority of the highest-risk patients identified by the UM and the predictive models are different. Based on this observed difference, hypothesis H6 supported.

H7: The number of 2005 and 2006 inpatient admissions of the j patients identified by Models M1, M4, and M5 will be higher than the number of 2005 and 2006 inpatient admissions of the j patients identified by models M2 and M3.

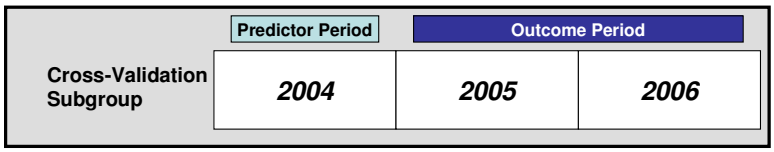
Although model M4 had the highest mean number of Outcome Period inpatient admissions for correctly identified actual high-use patients (8.38, as seen in Table 3.17), the total number of Outcome Period inpatient admissions is highest for the high-risk patients identified by model M2 (1,624). When considering the entire subset designated as high-risk (as is called for by hypothesis H7), model M2 has the highest mean number of Outcome Period inpatient admissions (1,624/621 = 2.62). Therefore, hypothesis H7 is rejected.

H8: *The data-driven models (M1 – M5) will identify case management candidates with a higher mean number of comorbid chronic conditions than those candidates identified by the Utilization Method.*

As shown in Table 3.18, the mean number of Outcome Period conditions per patient contained in the high-risk identified subset was higher for the UM than any of the predictive models (16.58 compared to 15.37 for M1, 15.15 for M2, 12.73 for M4, and 15.37 for M5). However, hypothesis H8 focuses on chronic conditions rather than all conditions in general. The specific chronic conditions that were counted were the six specific preventable chronic conditions identified by the World Health Organization: heart disease, stroke, cancer, asthma, chronic obstructive pulmonary disease, and diabetes.[182] The 344 patients identified by the UM had an average of 1.73 of the above chronic conditions. The 620 patients identified by models M1, M2, M4, and M5 had an average of 1.24, 1.23, 0.99, and 1.15 of the above chronic conditions, respectively. If the top 344 patients identified by each model are considered instead, this number increases to 1.28 for M1, 1.48 for M2, 0.88 for M4, and 1.22 for M5. The patients identified by the predictive models have, on average, fewer assigned HCCs in total and fewer assigned HCCs of the above chronic conditions, which suggests that hypothesis H8 be rejected.

[182] World Health Organization. *Preventing Chronic Diseases: A Vital Investment* 2005.

Objective 3: The Actual High-Use Patient Subgroup



This section provides a description of the 620 Cross-Validation Subgroup patients with 3 or more inpatient admissions during the Outcome Period (2005 through 2006)—the patients that made up the actual high-use patient subgroup.

Table 3.20 shows the proportion of high-use patients with the top 40 most commonly assigned Predictor Period (2004) HCCs, broken down by whether or not the patient was correctly identified by any model or method. Almost all HCCs were more common among the 210 patients who were correctly identified by one or more of the models and the UM. Only the HCCs *Other Female Genital Disorders* and *Uncompleted Pregnancy with No or Minor Complications* were more common among the 410 patients not identified by any model or method, with differences in relative proportions of approximately 2% and 4%, respectively.

Table 3.20 Top 40 Most Commonly Assigned Predictor Period (2004) Hierarchical Condition Codes Among Actual High Use Cross-Validation Subgroup Patients, Categorized by Not Identified vs. Identified Subgroups

	All (n=620)		Not Identified (n=410)		Identified (n=210)		
Hierarchical Condition Category	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Diff***
Major Symptoms, Abnormalities	267	43%	144	35%	123	59%	23%
Minor Symptoms, Signs, Findings	234	38%	128	31%	106	50%	19%
Hypertension	190	31%	108	26%	82	39%	13%
Screening/Observation/Special Exams	189	30%	114	28%	75	36%	8%
Diabetes with No or Unspecified Complications	180	29%	117	29%	63	30%	1%
Other Musculoskeletal and Connective Tissue Disorders	177	29%	96	23%	81	39%	15%
Other Gastrointestinal Disorders	176	28%	86	21%	90	43%	22%
Other Injuries	135	22%	77	19%	58	28%	9%
Disorders of Fluid/Electrolyte/Acid-Base Balance	129	21%	44	11%	85	40%	30%
Drug/Alcohol Abuse, Without Dependence	123	20%	64	16%	59	28%	12%
Other Endocrine/Metabolic/Nutritional Disorders	111	18%	60	15%	51	24%	10%
Other Ear, Nose, Throat, and Mouth Disorders	106	17%	48	12%	58	28%	16%
Post-Surgical States/Aftercare/Elective	100	16%	43	10%	57	27%	17%
History of Disease	93	15%	33	8%	60	29%	21%
Iron Deficiency and Other/Unspecified Anemias and Blood Disease	88	14%	31	8%	57	27%	20%
Cellulitis, Local Skin Infection	84	14%	40	10%	44	21%	11%
Other Infectious Diseases	82	13%	35	9%	47	22%	14%
Congestive Heart Failure	76	12%	34	8%	42	20%	12%
Other Lung Disorders	75	12%	37	9%	38	18%	9%
Urinary Tract Infection	70	11%	28	7%	42	20%	13%
Other Female Genital Disorders	63	10%	45	11%	18	9%	-2%
Depression	60	10%	27	7%	33	16%	9%
Seizure Disorders and Convulsions	59	10%	14	3%	45	21%	18%
Type I Diabetes Mellitus	55	9%	26	6%	29	14%	7%
Chronic Obstructive Pulmonary Disease	54	9%	30	7%	24	11%	4%
Asthma	54	9%	23	6%	31	15%	9%
Other Dermatological Disorders	54	9%	27	7%	27	13%	6%
Uncompleted Pregnancy With No or Minor Complications	50	8%	39	10%	11	5%	-4%
Other Psychiatric Disorders	48	8%	22	5%	26	12%	7%
Other Urinary Tract Disorders	45	7%	19	5%	26	12%	8%
Coagulation Defects and Other Specified Hematological Disorders	42	7%	17	4%	25	12%	8%
Chronic Hepatitis	40	6%	21	5%	19	9%	4%
Other Hepatitis and Liver Disease	40	6%	18	4%	22	10%	6%
Peptic Ulcer, Hemorrhage, Other Specified Gastrointestinal Disorders	40	6%	11	3%	29	14%	11%
Viral and Unspecified Pneumonia, Pleurisy	40	6%	7	2%	33	16%	14%
Coronary Atherosclerosis/Other Chronic Ischemic Heart Disease	37	6%	17	4%	20	10%	5%
Drug/Alcohol Dependence	36	6%	11	3%	25	12%	9%
Poisonings and Allegic Reactions	36	6%	14	3%	22	10%	7%
Diabetes with Acute Complications	35	6%	13	3%	22	10%	7%
Severe Hematological Disorders	35	6%	7	2%	28	13%	12%

Note: A patient was considered "Identified" if one of the predicitve models or Utilization Method contained the patient in their high-risk patient subset.

*Diff represents the increased proportion of patients with this condition in the Identified subgroup as opposed to the Not Identified subgroup. A negative number indicates the condition was more common in the Not Identified subgroup than in the Identified subgroup.

The most commonly assigned Predictor Period HCCs were the broad categorizations of *Major Symptoms, Abnormalities* and *Minor Symptoms, Signs, Findings*, which were assigned to 43% and 38% of all actual high-use patients, respectively. Both were substantially more commonly assigned to patients who were correctly identified by one or more of the models or the UM. *Hypertension* was the most common disease-specific HCC assigned to actual high-use patients, at 31% for all patients, 26% for patients who were not correctly identified, and 39% for patients

correctly identified by one or more models or the UM. Table 3.21 shows the same information but for the top 60 HCCs assigned in 2005 and 2006. This table displays more HCCs since more HCCs were assigned over the two-year period of the Outcome Period than for the one-year period of the Predictor Period.

As seen with Predictor Period HCCs, the most commonly assigned HCCs in the Outcome Period were assigned to a higher proportion of patients who had been correctly identified by the models or the UM than for patients who had been missed. Notable exceptions included pregnancy-related HCCs, which were substantially more common among patients who were not identified as compared to patients who were correctly identified. The largest difference between the Not Identified and Identified subsets of patients was with the HCC *Uncompleted Pregnancy with No or Minor Complications* at 14%. Both *Completed Pregnancy Without Complications (Normal Delivery) and Uncompleted Pregnancy with Complications* were 12% more common among patients in the Not Identified subgroup than in the Identified subgroup.

Also as seen with Predictor Period HCCs, the most commonly assigned HCCs were broad categorizations. A total of 77% of all actual high-use patients were *assigned Major Symptoms, Abnormalities* as opposed to 43% in the Predictor Period. This difference in part reflects the two-year time horizon in Table 3.21. Again, *Hypertension* was the most assigned HCC for a specific condition, at 52%.

However, unlike the Predictor Period HCCs, the difference of the percentage of HCCs assigned between Not Identified and Identified patients was smaller with Outcome Period HCCs. The top 60 most commonly assigned HCCs for the Outcome Period were

only 4% more frequently assigned to correctly high use identified patients than high-use patients who were not identified by any predictive model or the UM. For Predictor Period HCCs, this difference was 9%.

Table 3.21 Top 60 Most Commonly Assigned Outcome Period (2005 through 2006) HCCs Among Actual High Use Cross-Validation Patients, Broken Down by Not Identified vs. Identified Subgroups

	All (n=620)		Not Identified (n=410)		Identified (n=210)		
Hierarchical Condition Category	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Diff***
Major Symptoms, Abnormalities	476	77%	305	74%	171	81%	7%
Minor Symptoms, Signs, Findings	451	73%	289	70%	162	77%	7%
Other Gastrointestinal Disorders	407	66%	252	61%	155	74%	12%
Screening/Observation/Special Exams	404	65%	262	64%	142	68%	4%
Disorders of Fluid/Electrolyte/Acid-Base Balance	379	61%	234	57%	145	69%	12%
Post-Surgical States/Aftercare/Elective	337	54%	202	49%	135	64%	15%
Other Musculoskeletal and Connective Tissue Disorders	329	53%	205	50%	124	59%	9%
Hypertension	322	52%	210	51%	112	53%	2%
History of Disease	322	52%	197	48%	125	60%	11%
Iron Deficiency and Other/Unspecified Anemias and Blood Disease	316	51%	203	50%	113	54%	4%
Other Injuries	287	46%	180	44%	107	51%	7%
Other Endocrine/Metabolic/Nutritional Disorders	286	46%	190	46%	96	46%	-1%
Drug/Alcohol Abuse, Without Dependence	273	44%	179	44%	94	45%	1%
Other Infectious Diseases	269	43%	169	41%	100	48%	6%
Diabetes with No or Unspecified Complications	256	41%	175	43%	81	39%	-4%
Other Ear, Nose, Throat, and Mouth Disorders	226	36%	131	32%	95	45%	13%
Other Lung Disorders	193	31%	103	25%	90	43%	18%
Cellulitis, Local Skin Infection	192	31%	125	30%	67	32%	1%
Urinary Tract Infection	190	31%	123	30%	67	32%	2%
Depression	189	30%	122	30%	67	32%	2%
Other Urinary Tract Disorders	157	25%	100	24%	57	27%	3%
Renal Failure	154	25%	102	25%	52	25%	0%
Viral and Unspecified Pneumonia, Pleurisy	152	25%	87	21%	65	31%	10%
Congestive Heart Failure	146	24%	93	23%	53	25%	3%
Coagulation Defects and Other Specified Hematological Disorders	134	22%	85	21%	49	23%	3%
Poisonings and Allegic Reactions	133	21%	76	19%	57	27%	9%
Other Psychiatric Disorders	124	20%	74	18%	50	24%	6%
Other Hepatitis and Liver Disease	120	19%	80	20%	40	19%	0%
Peptic Ulcer, Hemorrhage, Other Specified Gastrointestinal Disorde	120	19%	73	18%	47	22%	5%
Other Dermatological Disorders	120	19%	73	18%	47	22%	5%
Cardio-Respiratory Failure and Shock	117	19%	73	18%	44	21%	3%
Major Depressive, Bipolar, and Paranoid Disorders	117	19%	66	16%	51	24%	8%
Asthma	114	18%	67	16%	47	22%	6%
Chronic Obstructive Pulmonary Disease	107	17%	69	17%	38	18%	1%
Drug/Alcohol Dependence	107	17%	61	15%	46	22%	7%
Other Female Genital Disorders	106	17%	77	19%	29	14%	-5%
Other Complications of Medical Care	104	17%	67	16%	37	18%	1%
Seizure Disorders and Convulsions	104	17%	42	10%	62	30%	19%
Coronary Atherosclerosis/Other Chronic Ischemic Heart Disease	96	15%	68	17%	28	13%	-3%
Major Complications of Medical Care and Trauma	96	15%	53	13%	43	20%	8%
Type I Diabetes Mellitus	90	15%	57	14%	33	16%	2%
Chronic Hepatitis	86	14%	55	13%	31	15%	1%
Severe Hematological Disorders	83	13%	37	9%	46	22%	13%
Uncompleted Pregnancy With No or Minor Complications	82	13%	74	18%	8	4%	-14%
Other Heart Rhythm and Conduction Disorders	82	13%	52	13%	30	14%	2%
Diabetes with Acute Complications	81	13%	55	13%	26	12%	-1%
Other Circulatory Disease	81	13%	44	11%	37	18%	7%
Completed Pregnancy Without Complications (Normal Delivery)	78	13%	68	17%	10	5%	-12%
Mononeuropathy, Other Neurological Conditions/Injuries	77	12%	48	12%	29	14%	2%
Uncompleted Pregnancy With Complications	74	12%	66	16%	8	4%	-12%
Polyneuropathy	67	11%	42	10%	25	12%	2%
Other Significant Endocrine and Metabolic Disorders	67	11%	38	9%	29	14%	5%
Cirrhosis of Liver	65	10%	38	9%	27	13%	4%
Pancreatic Disease	65	10%	37	9%	28	13%	4%
Diabetes with Neurologic or Peripheral Circulatory Manifestation	64	10%	42	10%	22	10%	0%
Other Eye Disorders	63	10%	41	10%	22	10%	0%
Vascular Disease	62	10%	38	9%	24	11%	2%
Completed Pregnancy With Complications	59	10%	53	13%	6	3%	-10%
Osteoporosis and Other Bone/Cartilage Disorders	59	10%	35	9%	24	11%	3%
Diabetes with Renal Manifestation	58	9%	42	10%	16	8%	-3%

Note: A patient was considered "Identified" if one of the predicitve models or Utilization Method contained the patient in their high-risk patient subset.

*Diff represents the increased proportion of patients with this condition in the Identified subgroup as opposed to the Not Identified subgroup. A negative number indicates the condition was more common in the Not Identified subgroup than in the Identified subgroup.

Table 3.22 shows the proportion of high-use patients with the most commonly assigned Predictor Period Aggregate Condition Codes (ACCs), categorized by whether or not the patient was correctly identified by any model or method. The ACC level is the most summarized level of the DCG condition categorization methodology. The 183 HCCs are summarized into 30 ACCs. The most commonly assigned ACCs were the two general categorizations of *Symptoms, Signs and Ill-Defined Conditions* and *Screening/History*, at 59% and 48%, respectively. The ACC *Diabetes* was 10% more common among Not Identified patients than for Identified patients, at 20% and 10% respectively.

Table 3.22 Assignment of Predictor Period (2004) ACCs Among Actual High Use Cross-Validation Subgroup Patients, Broken Down by Not Identified vs. Identified Subgroups

	All (n=620)		Not Identified (n=410)		Identified (n=210)		
Aggregate Condition Category	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Diff***
Symptoms, Signs and Ill-Defined Conditions	364	59%	204	33%	160	26%	-7%
Screening/History	295	48%	158	25%	137	22%	-3%
Heart	234	38%	136	22%	98	16%	-6%
Nutritional and Metabolic	212	34%	98	16%	114	18%	3%
Musculoskeletal and Connective Tissue	209	34%	111	18%	98	16%	-2%
Gastrointestinal	202	33%	96	15%	106	17%	2%
Diabetes	187	30%	123	20%	64	10%	-10%
Injury, Poisoning, Complications	181	29%	94	15%	87	14%	-1%
Lung	175	28%	81	13%	94	15%	2%
Hematological	139	22%	50	8%	89	14%	6%
Urinary System	138	22%	67	11%	71	11%	1%
Substance Abuse	135	22%	68	11%	67	11%	0%
Skin and Subcutaneous	134	22%	65	10%	69	11%	1%
Infectious and Parasitic	111	18%	44	7%	67	11%	4%
Ears, Nose, and Throat	109	18%	50	8%	59	10%	1%
Neurological	108	17%	34	5%	74	12%	6%
Mental	106	17%	48	8%	58	9%	2%
Genital System	86	14%	62	10%	24	4%	-6%
Liver	81	13%	47	8%	34	5%	-2%
Pregnancy-Related	60	10%	46	7%	14	2%	-5%
Vascular	53	9%	25	4%	28	5%	0%
Malignant Neoplasm	45	7%	19	3%	26	4%	1%
Eyes	36	6%	22	4%	14	2%	-1%
Benign/In Situ/Uncertain Neoplasm	34	5%	22	4%	12	2%	-2%
Cardio-Respiratory Arrest	31	5%	5	1%	26	4%	3%
Transplants, Openings, Other V-Codes	28	5%	8	1%	20	3%	2%
Cerebro-Vascular	21	3%	3	0%	18	3%	2%
Developmental Disability	18	3%	5	1%	13	2%	1%
Cognitive Disorders	16	3%	3	0%	13	2%	2%
Neonates	14	2%	13	2%	1	0%	-2%

Note: A patient was considered "Identified" if one of the predicitve models or Utilization Method contained the patient in their high-risk patient subset.

*Diff represents the increased proportion of patients with this condition in the Identified subgroup as opposed to the Not Identified subgroup. A negative number indicates the condition was more common in the Not Identified subgroup than in the Identified subgroup.

The mean difference in proportion of assignment of all Predictor Period ACCs between Not Identified and Identified patients is approximately 0%. Table 3.23 shows this same information but for ACCs assigned in the Outcome Period.

Table 3.23 Assignment of Outcome Period (2005 through 2006) ACCs Among Actual High-Use Cross-Validation Subgroup Patients, Categorized by Not Identified vs. Identified Subgroups

	All (n=620)		Not Identified (n=410)		Identified (n=210)		
Aggregate Condition Category	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Diff***
Screening/History	565	91%	368	90%	197	94%	4%
Symptoms, Signs and Ill-Defined Conditions	558	90%	368	90%	190	90%	1%
Nutritional and Metabolic	481	78%	311	76%	170	81%	5%
Gastrointestinal	443	71%	278	68%	165	79%	11%
Injury, Poisoning, Complications	386	62%	245	60%	141	67%	7%
Hematological	384	62%	239	58%	145	69%	11%
Heart	376	61%	242	59%	134	64%	5%
Musculoskeletal and Connective Tissue	365	59%	224	55%	141	67%	13%
Lung	357	58%	218	53%	139	66%	13%
Urinary System	341	55%	224	55%	117	56%	1%
Infectious and Parasitic	299	48%	188	46%	111	53%	7%
Substance Abuse	281	45%	183	45%	98	47%	2%
Mental	275	44%	169	41%	106	50%	9%
Skin and Subcutaneous	267	43%	171	42%	96	46%	4%
Diabetes	259	42%	178	43%	81	39%	-5%
Ears, Nose, and Throat	233	38%	135	33%	98	47%	14%
Neurological	212	34%	115	28%	97	46%	18%
Liver	184	30%	123	30%	61	29%	-1%
Genital System	150	24%	111	27%	39	19%	-9%
Vascular	138	22%	81	20%	57	27%	7%
Cardio-Respiratory Arrest	120	19%	75	18%	45	21%	3%
Eyes	97	16%	63	15%	34	16%	1%
Benign/In Situ/Uncertain Neoplasm	94	15%	68	17%	26	12%	-4%
Pregnancy-Related	93	15%	81	20%	12	6%	-14%
Malignant Neoplasm	87	14%	53	13%	34	16%	3%
Transplants, Openings, Other V-Codes	86	14%	47	11%	39	19%	7%
Cerebro-Vascular	65	10%	36	9%	29	14%	5%
Cognitive Disorders	39	6%	20	5%	19	9%	4%
Developmental Disability	37	6%	17	4%	20	10%	5%
Neonates	7	1%	6	1%	1	0%	-1%

Note: A patient was considered "Identified" if one of the predicitve models or Utilization Method contained the patient in their high-risk patient subset.

*Diff represents the increased proportion of patients with this condition in the Identified subgroup as opposed to the Not Identified subgroup. A negative number indicates the condition was more common in the Not Identified subgroup than in the Identified subgroup.

In contrast to the Predictor Period ACCs, the Outcome Period ACCs were 4% more commonly assigned to Identified patients than to Not Identified patients. Only 5 of the 30 ACCs were more commonly assigned to Not Identified patients, with the largest discrepancy belonging to *Pregnancy-Related* at 14%.

Table 3.24 shows a summary of the Predictor Period and Outcome Period categorization of selected conditions.

Table 3.24 Summary of Predictor Period (2004) and Outcome Period (2005 through 2006) Selected Condition Assignment Among Actual High-Use Cross-Validation Subgroup Patients, Categorized by Not Identified vs. Identified Subgroups

	2004							2005+2006						
	All (n=620)		Not Identified (n=410)		Identified (n=210)			All (n=620)		Not Identified (n=410)		Identified (n=210)		
Condition	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Diff****	**Count**	**%**	**Count**	**%**	**Count**	**%**	**Diff****
Hypertension	190	31%	108	26%	82	39%	13%	322	52%	210	51%	112	53%	2%
Diabetes*	187	30%	123	20%	64	10%	-10%	259	42%	178	43%	81	39%	-5%
Substance Abuse*	135	22%	68	11%	67	11%	0%	281	45%	183	45%	98	47%	2%
Congestive Heart Failure	76	12%	34	8%	42	20%	12%	146	24%	93	23%	53	25%	3%
Depression	60	10%	27	7%	33	16%	9%	189	30%	122	30%	67	32%	2%
Pregnancy-Related*	60	10%	46	11%	14	7%	-5%	93	15%	81	20%	12	6%	-14%
Asthma	54	9%	23	6%	31	15%	9%	114	18%	67	16%	47	22%	6%
COPD	54	9%	30	7%	24	11%	4%	107	17%	69	17%	38	18%	1%

*These conditions were identified at the ACC level, while all other conditions were identified at the HCC level.

**Diff represents the increased proportion of patients with this condition in the Identified subgroup as opposed to the Not Identified subgroup. A negative number indicates the condition was more common in the Not Identified subgroup than in the Identified subgroup.

All conditions were more commonly assigned in the two years of the Outcome Period (2005 through 2006) than in the one year of the Predictor Period (2004). In both the Predictor Period and in the Outcome Period, *Hypertension* is the most commonly assigned condition for actual high-use patients. *Diabetes* is second in the Predictor Period but overtaken by *Substance Abuse* in the Outcome Period, with 45% of all actual high-use patients categorized with Substance Abuse at least once during that two-year period.

These conditions are further examined in Table 3.25, which displays model and Utilization Method (UM) performance by condition. Mean average prediction error (MAPE) was calculated for each model except M3, since M3's predicted utilization was not granular enough to differentiate a subset of 620 highest-risk patients. MAPE also could not be calculated for the UM, since it does not generate predicted utilization. Table 3.25 also displays, by condition and model, the number of individual high-use patients correctly identified, the number of false-positive identifications (i.e., patients predicted as high-use who were not in the high use subset), and the number of actual high-use patients who were not identified.

During the Outcome Period, 322 actual high-use patients had a diagnosis of hypertension, compared to 190 for the Predictor Period. A total of 173 patients had a diagnosis of hypertension in the Predictor Period and the Outcome Period. Model M5 had the lowest MAPE for hypertension (0.48) while M2 and M4 tied for the highest (0.55). M2 had the highest percentage correct (28%), while M4 had the lowest (15%). The UM had the lowest false-positive percentage (66%). M2 had the lowest percentage of missed high-use patients (72%).

Table 3.25 Model and Utilization Method Performance for Specified Conditions Among Cross-Validation Subgroup Patients

		Predictive Models								Utilization Method	
		M1		M2		M4		M5			
Hypertension											
n = 322 (2004: n = 190; Both: n = 173)	MAPE	0.50		0.55		0.55		0.48		NA	
	Identified	89	28%	91	28%	49	15%	78	24%	77	24%
	False-Positive	182	67%	184	67%	170	78%	163	68%	151	66%
	Not Identified	233	72%	231	72%	273	85%	244	76%	245	76%
Substance Abuse											
n = 281 (2004: n = 135; Both: n = 121)	MAPE	0.70		0.71		0.76		0.72		NA	
	Identified	83	30%	84	30%	52	19%	77	27%	55	20%
	False-Positive	114	58%	116	58%	118	69%	109	59%	92	63%
	Not Identified	198	70%	197	70%	229	81%	204	73%	226	80%
Diabetes											
n = 259 (2004: n = 187; Both: n = 179)	MAPE	0.56		0.60		0.60		0.55		NA	
	Identified	68	26%	69	27%	37	14%	57	22%	60	23%
	False-Positive	116	63%	117	63%	99	73%	109	66%	96	62%
	Not Identified	191	74%	190	73%	222	86%	202	78%	199	77%
Depression											
n = 189 (2004: n = 60; Both: n = 43)	MAPE	0.57		0.60		0.64		0.57		NA	
	Identified	49	26%	51	27%	35	19%	52	28%	37	20%
	False-Positive	69	58%	71	58%	78	69%	71	58%	60	62%
	Not Identified	140	74%	138	73%	154	81%	137	72%	152	80%
Congestive Heart Failure											
n = 146 (2004: n = 76; Both: n = 68)	MAPE	1.56		1.57		1.63		1.89		NA	
	Identified	45	31%	46	32%	29	20%	41	28%	42	29%
	False-Positive	56	55%	55	54%	34	54%	42	51%	37	47%
	Not Identified	101	69%	100	68%	117	80%	105	72%	104	71%
Asthma											
n = 114 (2004: n = 54; Both: n = 45)	MAPE	0.42		0.43		0.46		0.37		NA	
	Identified	32	28%	31	27%	29	25%	36	32%	31	27%
	False-Positive	69	68%	61	66%	65	69%	69	66%	50	62%
	Not Identified	82	72%	83	73%	85	75%	78	68%	83	73%
Chronic Obstructive Pulmonary Disease											
n = 107 (2004: n = 54; Both: n = 44)	MAPE	0.94		0.97		1.02		0.98		NA	
	Identified	31	29%	31	29%	18	17%	28	26%	25	23%
	False-Positive	38	55%	39	56%	35	66%	40	59%	33	50%
	Not Identified	76	71%	76	71%	89	83%	79	74%	82	77%
Pregnancy-Related											
n = 93 (2004: n = 60; Both: n = 45)	MAPE	0.46		0.46		0.49		0.44		NA	
	Identified	8	9%	9	10%	3	3%	3	3%	3	3%
	False-Positive	41	84%	71	89%	9	75%	16	84%	7	70%
	Not Identified	85	91%	84	90%	90	97%	90	97%	90	97%

Note: MAPE = mean average prediction error; Identified = the number of actual high use patients with this condition correctly identified in the model's high-risk subset; False-positive = the number of the model's high-risk subset with this condition who were not in the actual high use patient subset; Not Identified = the number of actual high use patients with this condition who were not identified in the model's high-risk subset.'

In the Outcome Period, 281 actual high-use patients had a diagnosis of substance abuse, compared to 135 for the Predictor Period. A total of 121 patients had a diagnosis of substance abuse in the Predictor Period and the Outcome Period. Model M1 had the

lowest MAPE for substance abuse (0.70) while M4 had the highest (0.76). M2 had the highest percentage Identified (30%), while M4 had the lowest (19%). M1 had the lowest false-positive percentage (58%). M1 had lowest percentage of Not Identified high-use patients (70%).

In the Outcome Period, 259 actual high-use patients had a diagnosis of diabetes, compared to 187 for the Predictor Period. A total of 179 patients had a diagnosis of diabetes in the Predictor Period and the Outcome Period. Model M5 had the lowest MAPE for diabetes (0.55) while M4 and M2 tied for the highest (0.60). M2 had the highest percentage Identified (27%), while M4 had the lowest (14%). The UM had the lowest false-positive percentage (62%). M2 had lowest percentage of Not Identified high-use patients (73%).

In the Outcome Period, 189 actual high-use patients had a diagnosis of depression, compared to 60 for the Predictor Period. A total of 43 patients had a diagnosis of depression in the Predictor Period and the Outcome Period. Models M1 and M5 had the lowest MAPE for depression (0.57) while M4 had the highest (0.64). M5 had the highest percentage Identified (28%), while M4 had the lowest (19%). M1 had the lowest false-positive percentage (58%). M5 had lowest percentage of Not Identified high-use patients (72%).

In the Outcome Period, 146 actual high-use patients had a diagnosis of congestive heart failure, compared to 76 for the Predictor Period. A total of 68 patients had a diagnosis of congestive heart failure in the Predictor Period and the Outcome Period. Model M1 had the lowest MAPE for congestive heart failure (1.56) while M5 had the

highest (1.89). M2 had the highest percentage Identified (32%), while M4 had the lowest (20%). M1 and M4 had the lowest false-positive percentage (58%). M2 had lowest percentage of Not Identified high-use patients (68%).

In the Outcome Period, 114 actual high-use patients had a diagnosis of asthma, compared to 54 for the Predictor Period. A total of 45 patients had a diagnosis of asthma in the Predictor Period and the Outcome Period. Model M5 had the lowest MAPE for asthma (0.37) while M4 had the highest (0.46). M5 had the highest percentage Identified (32%), while M4 had the lowest (25%). The UM had the lowest false-positive percentage (62%). M5 had lowest percentage of Not Identified high-use patients (68%).

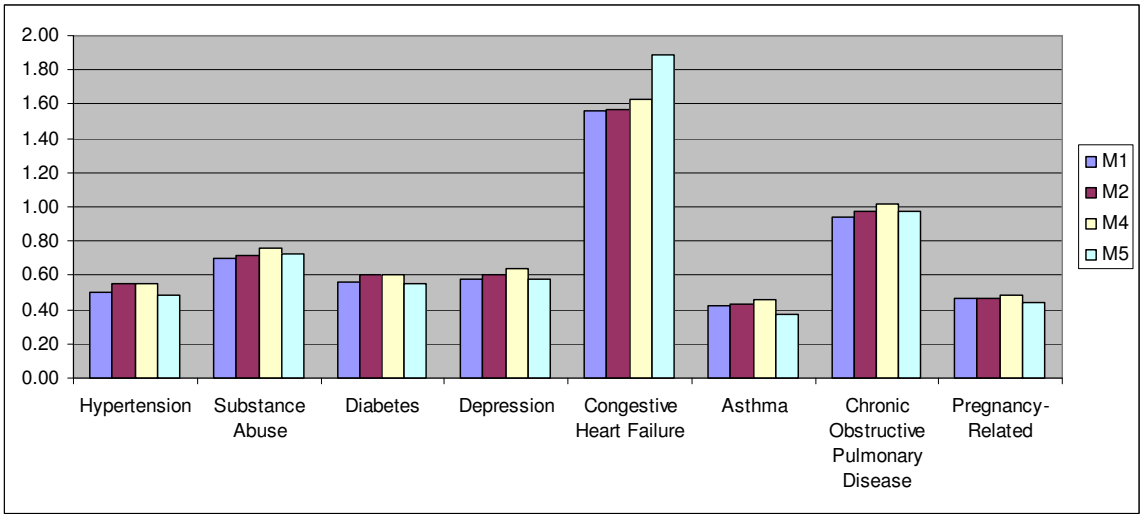
In the Outcome Period, 107 actual high-use patients had a diagnosis of chronic obstructive pulmonary disease (COPD), compared to 54 for the Predictor Period. A total of 44 patients had a diagnosis of COPD in the Predictor Period and the Outcome Period. Model M1 had the lowest MAPE for COPD (0.94) while M4 had the highest (1.02). M1 and M2 had the highest percentage Identified (29%), while M4 had the lowest (17%). The UM had the lowest false-positive percentage (50%). M1 and M2 had lowest percentage of Not Identified high-use patients (71%).

In the Outcome Period, 93 actual high-use patients had pregnancy-related diagnoses, compared to 60 for the Predictor Period. A total of 45 patients had pregnancy-related diagnoses in the Predictor Period and the Outcome Period. Model M5 had the lowest MAPE for asthma (0.44) while M4 had the highest (0.49). M2 had the highest percentage Identified (10%), while M4, M5, and the UM had the lowest (3%). The UM

had the lowest false-positive percentage (70%). M2 had lowest percentage of Not Identified high-use patients (90%).

Figure 3.12 shows the MAPE for each model, by condition.

Figure 3.12 Mean Average Prediction Error by Model and By Selected Condition Among Cross-Validation Subgroup Patients



The MAPE was highest for all models for congestive heart failure and lowest for asthma.

Figure 3.13 shows the percentage of patients that each model and the UM identified correctly, by condition. All models and the UM correctly identified at least 25% of high-use asthma patients. Only M1 and M2 correctly identified more than 5% of high-use patients with pregnancy-related diagnoses. M4 correctly identified the lowest percent of patients for each condition.

Figure 3.13 Percentage of Patients Identified Correctly by Each Model and the Utilization Method for Selected Conditions Among Cross-Validation Subgroup Patients

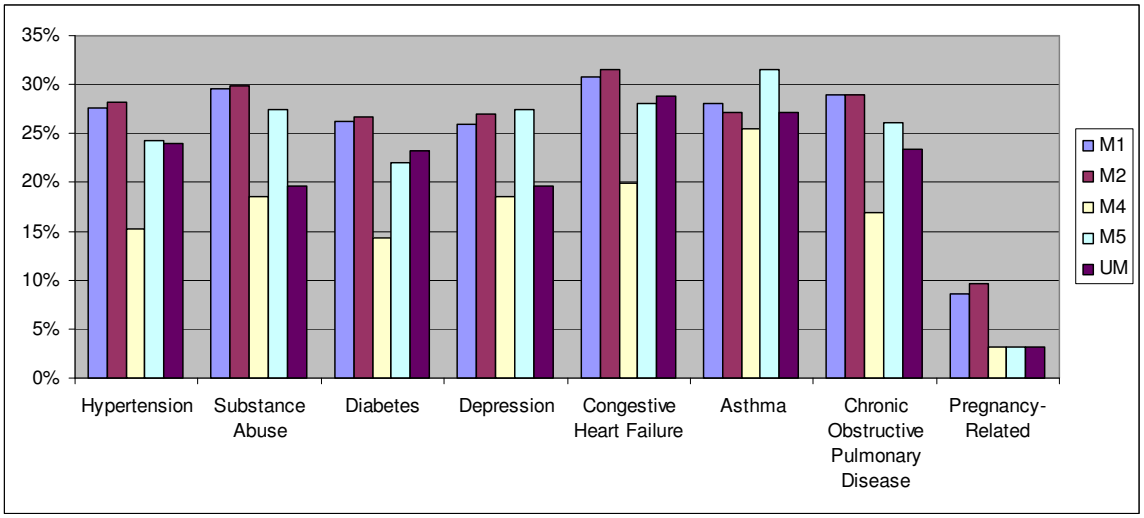
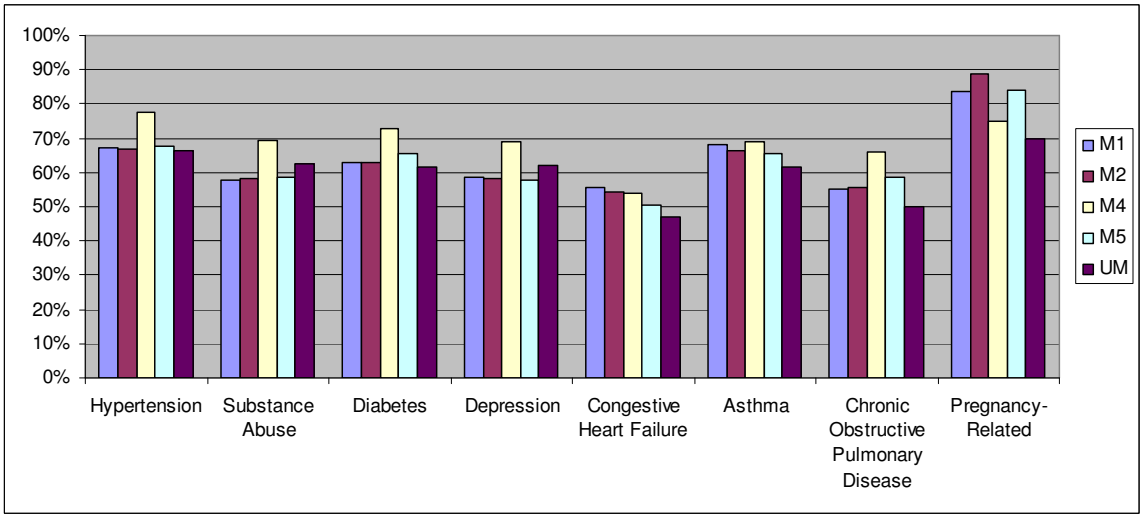


Figure 3.14 shows the percentage of false-positive identifications made by each model and the UM, by condition.

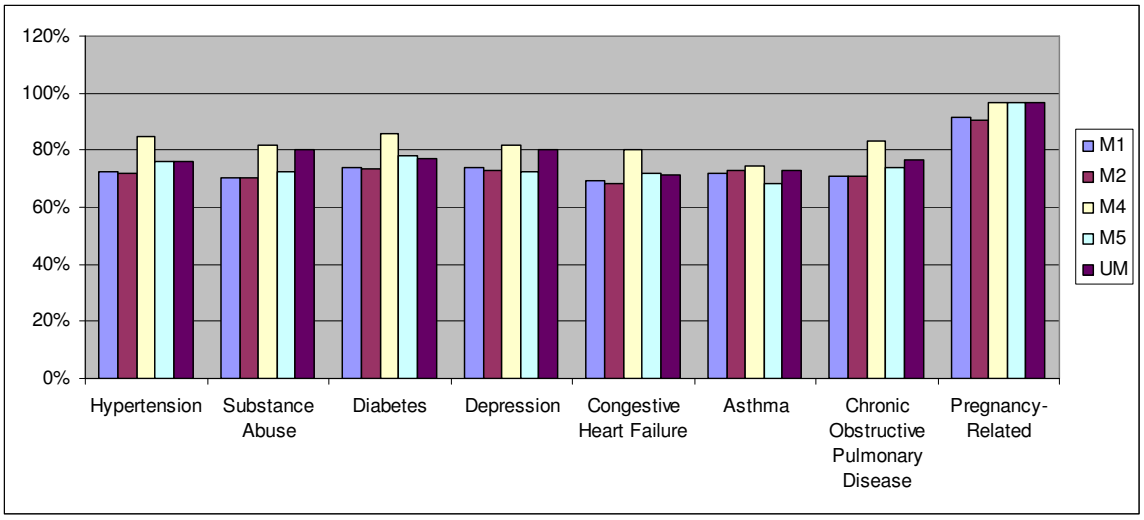
Figure 3.14 Percentage of Patients Incorrectly Identified as High-Use by Each Model and the Utilization Method for Selected Conditions Among Cross-Validation Subgroup Patients



The UM tended to have the lowest false-positive percentage, although it identified a small number of patients. Congestive heart failure was the condition associated with the fewest number of false positives by all models and the UM.

Figure 3.15 shows the percentage of actual high-use patients missed by each model and the UM, by condition.

Figure 3.15 Percentage of Actual High-Use patients Not Identified by Each Model and the Utilization Method for Selected Conditions Among Cross-Validation Subgroup Patients



M4 consistently missed the highest percentage of actual high-use patients. All models and the UM missed 90% or more of actual high-use patients with pregnancy-related diagnoses.

Table 3.26 shows descriptive differences between the subgroup of patients who were not identified vs. identified.

Table 3.26 Descriptive Differences between Cross-Validation Subgroup Patients who were Not Identified by Any Method vs. Patients who were Identified by Any Method

	Not Identified	Identified
N	410	210
Mean Predictor Period rescaled relative risk score	1.6	4.5
Mean number of Predictor Period HCCs	5.84	12.42
Mean number of Outcome Period HCCs	19.04	21.76
Mean Predictor Period inpatient admissions	0.298	3.57
Mean age	36.74	35.66
Percent female	58.3	53.3

The Predictor Period for the Cross-Validation Subgroup is calendar year 2004.
The Outcome Period for the Cross-Validation Subgrou is 2005 through 2006.

On average, actual high-use patients who were not identified by any of the models or the UM tended to have a lower Predictor Period rescaled relative risk score (1.6 vs 4.5), fewer diagnosed comorbidies (5.84 vs. 12.42 assigned HCCs), and fewer Predictor Period inpatient admissions (0.298 vs. 3.57). Both mean age and percent female were similar between groups.

Table 3.27 looks in detail at individual actual high-use patients who were not identified by any model or the UM. The table shows outcome data, baseline data, and rankings by each model for patients with seven or more Outcome Period inpatient admissions. The actual high-use patients not identified by all models and the UM tended to have zero or one Predictor Period inpatient admissions, a lower than average (in terms of actual high-use patients) Predictor Period relative risk score, and fewer than average assigned Predictor Period HCCs.

The Not Identified actual high use patient with the greatest number of Outcome Period inpatient admissions was a 28 year-old female, with no 2004 inpatient admissions, 2 assigned Predictor Period HCCs, and a rescaled relative risk score of 0.521. This

patient was ranked in the bottom half of all patients by models M1, M2, and M5, with M4 ranking her above the 52nd percentile.

Many of the actual high-use patients who were not in the highest-risk subset of any model or identified by the UM nevertheless were still ranked highly by the models. Most patients with one Predictor Period inpatient admission were ranked above the 90th percentile by all models. When those patients additionally had a 2004 rescaled relative risk score above 2.00, they were ranked in the 95th percentile by all models. For example, a 54 year-old female had 10 Outcome Period inpatient admissions, 1 Predictor Period inpatient admission, and a Predictor Period rescaled relative risk score of 4.142. This patient received a ranking above the 97th percentile by all models, with a maximum of above the 98th percentile by M5.

Table 3.27 Details of High-Use Cross-Validation Subgroup Patients with Seven or More Outcome Period (2005 through 2006) Inpatient Admissions Not Identified by Any Models or the Utilization Method

Outcome Period				Baseline Data 2004					Model Rank* Predicted Score Percentile				
Rank	Combined Inpatient Admits	2005 Relative Risk	2006 Relative Risk	Relative Risk	Inpatient Admits	Assigned HCCs	Age	Sex	Mean	M1	M2	M4	M5
1	17	5.340	22.175	0.521	0	2	28	2	< 50	< 50	< 50	52.4	< 50
2	16	3.712	7.400	2.171	0	6	49	2	84.9	84.9	85.2	77.6	91.6
3	14	4.498	3.501	1.006	0	4	42	2	65.9	65.9	63.8	70.6	65.2
4	14	0.658	2.905	0.688	0	2	9	1	< 50	< 50	50.6	< 50	50.6
5	13	1.155	4.842	1.376	0	8	43	1	74.5	74.5	71.7	78.8	76.0
6	12	5.879	3.266	2.620	0	5	0	2	67.4	67.4	86.9	< 50	94.3
7	12	2.482	2.038	1.639	0	1	60	1	81.3	81.3	76.9	89.2	82.2
8	12	0.673	1.158	0.325	0	2	3	2	< 50	< 50	< 50	< 50	< 50
9	11	3.703	3.727	2.465	1	4	54	1	**97.0**	**97.0**	**97.4**	**98.1**	**95.7**
10	11	2.352	4.351	0.941	0	2	39	2	62.9	62.9	60.9	67.4	62.4
11	11	1.422	1.344	0.941	0	1	38	2	62.2	62.2	60.4	66.4	61.8
12	10	5.014	4.760	4.142	1	6	54	2	**97.9**	**97.9**	**97.9**	**97.5**	**98.2**
13	10	2.235	2.696	0.939	1	8	30	1	91.6	91.6	94.9	**96.2**	81.8
14	10	1.603	3.132	1.677	0	3	48	1	80.7	80.7	78.2	82.9	83.5
15	10	1.984	2.638	1.345	1	5	13	1	93.4	93.4	**95.8**	93.5	89.6
16	10	2.651	1.836	3.215	1	12	61	2	**97.5**	**97.5**	**97.7**	**98.0**	**96.7**
17	9	1.641	11.125	2.847	1	20	55	1	**97.3**	**97.3**	**97.6**	**98.1**	**96.2**
18	9	5.301	5.086	4.166	1	22	50	1	**98.0**	**98.0**	**97.9**	**97.8**	**98.3**
19	9	2.089	7.998	1.209	1	9	43	1	93.8	93.8	**95.5**	**97.2**	88.2
20	9	2.016	6.413	1.469	0	8	41	1	77.1	77.1	75.6	77.0	80.1
21	9	2.746	4.665	2.221	1	12	56	1	**96.8**	**96.8**	**97.2**	**98.2**	**95.1**
22	9	3.758	2.619	0.521	0	3	33	2	50.4	50.4	< 50	60.3	< 50
23	8	4.753	9.044	1.677	0	3	48	1	80.6	80.6	78.1	82.8	83.4
24	8	4.520	8.973	3.473	0	9	41	1	87.6	87.6	87.8	77.1	**96.2**
25	8	5.518	6.193	4.464	1	9	54	2	**98.1**	**98.1**	**98.0**	**97.5**	**98.5**
26	8	3.509	7.446	2.616	0	5	51	2	86.9	86.9	86.9	79.4	94.2
27	8	2.979	2.994	1.299	0	2	33	1	69.8	69.8	68.6	69.1	72.8
28	8	2.315	2.849	2.725	1	10	59	1	**97.3**	**97.3**	**97.6**	**98.3**	**96.1**
29	8	2.781	0.522	0.141	0	1	8	1	< 50	< 50	< 50	< 50	< 50
30	8	1.247	1.808	0.521	0	2	30	2	< 50	< 50	< 50	55.7	< 50
31	8	1.598	1.378	0.516	1	5	23	2	86.5	86.5	92.9	94.1	68.0
32	8	1.032	0.697	0.872	0	3	10	1	< 50	< 50	52.6	< 50	53.9
33	8	0.427	1.074	0.208	0	1	13	2	< 50	< 50	< 50	< 50	< 50
34	7	8.038	13.425	1.620	0	5	62	2	80.7	80.7	76.7	87.4	81.8
35	7	8.011	5.770	0.521	0	3	30	2	< 50	< 50	< 50	55.5	< 50
36	7	1.938	9.540	2.405	1	10	38	2	**96.5**	**96.5**	**97.3**	**96.3**	**95.6**
37	7	3.438	4.245	2.451	0	11	45	2	84.8	84.8	86.1	73.8	93.2
38	7	3.925	3.181	1.765	1	4	0	1	94.5	94.5	**96.6**	92.1	93.2
39	7	2.778	3.641	2.255	0	11	54	1	87.6	87.6	85.6	87.1	92.2
40	7	2.797	2.362	1.209	1	2	62	2	94.0	94.0	**95.5**	**98.1**	88.2
41	7	2.491	2.547	1.209	0	7	36	1	68.1	68.1	65.4	72.1	69.4
42	7	2.091	1.760	0.602	1	1	15	2	87.8	87.8	94.2	92.5	72.4
43	7	1.032	2.614	1.304	0	1	35	2	68.8	68.8	69.2	63.3	73.4
44	7	1.603	2.051	1.677	0	5	48	1	80.8	80.8	78.3	83.0	83.6
45	7	1.756	1.674	1.307	1	13	21	2	93.3	93.3	**95.7**	93.7	89.0
46	7	1.103	1.687	1.087	0	4	27	1	64.1	64.1	64.1	62.4	65.8

Note: All relative risk scores in this scale have been rescaled to have a mean of 1.00. Shaded cells indicate percentile above 95%.
*All model ranks of 50 or less are reported as < 50, since models may not meaningfully differentiate low-ranked patients.

Objective 3 Hypothesis

H9: The subset of patients with the top 2 percent actual 2005 and 2006 inpatient utilization will contain a subgroup of patients with 1 or fewer 2004 inpatient encounters and at least 1 chronic condition HCC in 2004.

In evaluating this hypothesis, the chronic conditions that are considered include: diabetes, hypertension, asthma, chronic obstructive pulmonary disease, substance abuse, congestive heart failure, and depression. A total of 157 out of the top 620 high use patient subset had no Predictor Period inpatient admissions but had one or more Predictor Period assigned chronic condition HCCs.

Table 3.28 Actual High-Use Cross-Validation Subgroup Patients with No Predictor Period Inpatient Admissions and One or More Predictor Period Diagnoses of Selected Chronic Conditions

Total N	157
Number of Patients Diagnosed with Chronic Diseases*	
Diabetes	93
Hypertension	79
Substance Abuse	31
Chronic Obstructive Pulmonary Disease	19
Asthma	18
Non-major Depression	17
Congestive Heart Failure	14
Number of Patients with Number of Listed Condtions	
One	81
Two	49
Three	18
Four	7
Five	2
Number of Patients Included in the High-Risk Subset of Model or Method	
M1	0
M2	0
M4	7
M5	4
Utilization Method	6

*Patients may be counted next to each condition with which they are diagnosed.

Table 3.28 shows a breakdown of the 157 actual high-use Cross-Validation Subgroup Patients with no Predictor Period inpatient admissions and at least one Predictor Period diagnosis of the above listed chronic conditions. Diabetic and hypertensive patients were the most common among this subset of patients, with counts of 93 and 79, respectively. Over half of the patients (81) had diagnoses for only one of the listed conditions, compared to 49 diagnosed with two, 18 diagnosed with three, 7 diagnosed with four, and 2 diagnosed with five. No individual patients had diagnoses for six or more of the listed conditions. The Utilization Method (UM) used the criterion of 12 or more emergency department visits to identify 6 of these patients. While model M1 used the Predictor Period relative risk (RR) score as an independent variable, its coefficient weighting was not high enough to identify high-use patients without Predictor Period inpatient admissions. Model M2 did not use RR score as an independent variable and identified no high-use patients without Predictor Period inpatient admissions. Models M4 and M5, each of which weighted Predictor Period RR score heavier than M1, identified seven and two patients, respectively, who had no Predictor Period inpatient admissions and one or more chronic conditions listed above. Although hypothesis H9 was supported, the bulk of these patients remained unidentifiable with the predictive models and the UM.

Chapter 4: Discussion

This chapter includes a summary and interpretation of the results, an example of how the predictive models could be used to focus on patients with a specific disease state, a sensitivity analysis, a discussion of limitations, and a conclusion.

Overview

This study sought to test various methods of identifying future high-cost patients among the Austin, Texas, medically indigent population. Specifically, theory-based statistical models were compared with a straightforward, commonsense approach of selecting patients based on simple prior-year utilization criteria. The analysis herein provides both a means to distinguish the strengths and weakness between possible methods of high-risk patient identification and also a descriptive framework with which to assess to what extent high-use patients are actually identifiable *a priori*.

All data were submitted by the safety net providers to the Indigent Care Collaboration (ICC), who aggregates, standardizes, and stores all area safety net data in their patient database, the MPI/CDR. The ICC provided the researchers all patient data for the study period of January 1, 2003, to December 31, 2006.

Summary of Study Patients

During the study period of January 1, 2003, through December 31, 2006, a total of 494,551 patients had 3,556,034 recorded encounters at ICC partner provider locations.

The study used two subsets of these patients: a Calibration Subgroup used to create the predictive models and a Cross-Validation Subgroup used to evaluate the performance of the predictive models. To be in the Calibration Subgroup, a patient must have had an encounter of type "Inpatient", "Emergency Room", "Outpatient", "Clinic Visit", or "Office Visit" in each of years 2003, 2004, and 2005. To be in the Cross-Validation Subgroup, a patient must have had a valid encounter type in each of years 2004, 2005, and 2006. Furthermore, all patients age 65 or older were excluded from the two analysis subgroups.

After exclusions, the Calibration Subgroup consisted of 41,260 patients. After exclusions, the Cross-Validation Subgroup consisted of 44,738 patients. A total of 29,663 patients met the inclusion criteria for, and were included in, both subgroups.

The mean age for Calibration Subgroup patients was 25.07 (SD = 17.60), compared to 25.68 (SD = 17.49) for the Cross-Validation Subgroup ($t = -5.14$, $p < 0.01$). In both the Calibration Subgroup and the Cross-Validation Subgroup, approximately 35 percent of patients were 18 years of age or younger. In 2004, the entire ICC population consisted of approximately 36 percent of patients age 18 or younger. At the national level, the Kaiser Family Foundation (KFF) estimated that, in 2004, 20 percent of the uninsured were 18 or younger.[183]

The percentage of patients who were female was similar between Calibration Subgroup patients and Cross-Validation Subgroup patients (52.2 vs. 51.9, chi-squared =

[183] Kaiser Commission on Medicaid and the Uninsured. 50 State Comparisons: Introduction to Health Coverage & Uninsured. Available at: http://www.statehealthfacts.kff.org/cgi-bin/healthfacts.cgi?action=compare&welcome=1&category=Health+Coverage+%26+Uninsured.

0.50, p = 0.48). In 2004, approximately 61 percent of the entire ICC population were female patients. The KFF estimated the 2004 national gender distribution of the uninsured at 46% female and 54% male.[184] Although the age and gender distributions between the national uninsured population, as described by the KFF, and the study patients in the two subgroups varied considerably, the age and sex distribution of the study subgroup patients more closely reflects that of the entire ICC population. Furthermore, the difference in age and gender distributions between the ICC population and the national population that served as the basis for the KFF estimates is largely explained by the presence of Medicaid and State Children's Health Insurance Program patients in the ICC population.

[184] Ibid.

Results of Tests of Hypotheses

Table 4.1 shows each hypothesis that was tested. Objectives 1 and 2 and their respective hypotheses are discussed in the section Summary of Model Performance. Objective 3 and its one hypothesis is discussed in the section Summary of the High-Use Patient Subgroup.

Table 4.1 Results of the Tests of Hypotheses

Hypothesis		Supported?
Objective 1		
H1	Model fit will vary between the five models.	Yes
H2	Models M1, M4, and M5 will have lower deviance residuals than models M2 and M3.	No
H3	Model M1 will have a higher R2 than the square of the Pearson product moment correlation between the 2003 DCG RR scores and the number of inpatient admissions in 2004 and 2005.	Yes
H4	Model M4 will have lower deviance residuals than model M1.	No
H5	Model M5 will have the lowest deviance residuals of all models.	No
Objective 2		
H6	The Utilization Method and Models M1-M5 will identify different patients as highest-risk.	Yes
H7	The mean number of 2005 and 2006 inpatient admissions of the j patients identified by Models M1, M4, and M5 will be higher than the 2005 and 2006 inpatient admissions of the j patients identified by models M2 and M3.	No
H8	The data-driven models (M1 - M5) will identify case management candidates with a higher mean number of comorbid chronic conditions than those candidates identified by the Utilization Method.	No
Objective 3		
H9	The subset of patients with the top 2 percent actual 2005 and 2006 inpatient utilization will contain a subgroup of patients with 1 or fewer 2004 inpatient encounters and at least 1 chronic condition HCC in 2004.	Yes

Summary of Model Performance

Objective 1 was to create predictive models using patient data from year one (the Predictor Period) to predict the number of inpatient admissions in years two and three (the Outcome Period). Models M1, M2, and M3 were ordinary least squares (OLS) linear models, while models M4 and M5 were generalized linear models. Since the outcome variable (the number of Outcome Period inpatient admissions) is not normally distributed, a basic assumption of linear regression is not met. This was, in part, the basis for using the generalized linear model with a Poisson link for both M4 and one of the equations of M5. Despite this violation, OLS linear models M1, M2, and M3 were included in the analysis because of the common use of linear models for this type of analysis and for comparison purposes. Each model was calibrated on the Calibration Subgroup, using 2003 data to predict inpatient admissions in 2004 and 2005.

All five of the Objective 1 hypotheses dealt with how well the models fit the Calibration Subgroup data (i.e., the data used to create them). Once the models were created, the deviance residuals for each model could be calculated. Deviance residuals are the difference between the observed values and the values predicted by the model. Generally, low deviance residuals indicate high goodness of fit of a model. The linear models (M1, M2, and M3) had considerably lower median deviance residuals (-0.04, -0.05, and -0.06, respectively) compared with the generalized linear models (-0.59 for M4 and -0.46 and -0.40, respectively, for the two equations used in M5). This difference supports hypothesis H1 (that model fit will differ for the various models), but does not support hypotheses H2, H4, and H5 (that theory-based models using the DCG relative

risk score will have lower deviance residuals), since the highest residuals were seen with the theory-based models M4 and M5, both of which used the DCG relative risk score.

Although deviance residuals give some indication of a model's goodness-of-fit, the use of deviance residuals to compare linear and non-linear models may be problematic, since deviance residuals for Gaussian and Poisson models, for example, are not calculated in the same way (see Table 2.6). From a practical standpoint, a model with low deviance residuals may not be as useful as a model with higher deviance residuals, depending on the use of the models. For example, in the current project, a hypothetical linear model that predicted near-zero inpatient admissions for all patients would have relatively low mean deviance residuals, since the vast majority of patients have no inpatient admissions. However, this hypothetical model would not be useful, despite its low deviance residuals, since it would not discriminate between low-risk and high-risk patients. Similarly, a hypothetical model with high deviance residuals could be very useful if the highest-risk patients were, largely, actual high-use patients, even if the actual predictions themselves were inaccurate. Put another way, for identifying high-use patients *a priori*, it is the accuracy of the ranking of patient risk that is most telling about a model's utility, rather than simply the accuracy of its predictions. This reasoning is why the largest focus of this study is on the performance of the models in the Cross-Validation Subgroup.

Hypothesis H3 tested whether prior inpatient admissions and the DCG relative risk score predicted Outcome Period inpatient admissions better than the DCG relative risk score alone. Hypothesis H3 was supported by the substantially higher R^2 of model

M1 when compared to the square of the Pearson product moment correlation of the DCG relative risk score and the Outcome Period inpatient admissions.

Objective 2 sought to cross-validate the models created in Objective 1. The models created with the Calibration Subgroup population were tested on the Cross-Validation Subgroup population. The main measurement of model performance in the Cross-Validation Subgroup of patients was positive predictive value (PPV), which is the percentage of high-risk patients who are in the actual high-use subset. The size of the high-risk subsets (620 patients) was chosen because 620 patients had three or more inpatient admissions during the Outcome Period of 2005 through 2006. To calculate the PPV of each predictive model, five patient subsets (one for each model) of the highest-risk 620 patients were identified based on predicted inpatient admissions. The predictions of model M3 lacked the level of detail needed to create a cutoff point at exactly 620 patients, and therefore no PPV was calculated for model M3. Model M1 had the highest PPV (26.8%), followed by M2 (26.6%), M5 (24.7%), and M4 (17.7%).

Hypothesis H7 tested whether the models using the DCG relative risk score (M1, M4, and M5) identified patients with a greater number of Outcome Period inpatient admissions than those identified by models M2 and M3. Since model M2 identified patients with the greatest number of Outcome Period inpatient admissions, hypothesis H7 was rejected.

Hypotheses H6 and H8 concerned comparing the predictive models with the Utilization Method (UM). The Utilization Method is a non-empirically-based method of identifying high-risk patients by looking for high prior year inpatient utilization and

diagnoses of chronic diseases. The performance of the predictive models was compared with that of the UM in terms of PPV. However, the UM only identified 344 high-risk patients. If the PPV for the UM were calculated in the same manner as previously described for the predictive models (i.e., using a denominator of 620), the UM's correct identification of 104 actual high-use patients would equate to a PPV of 17.1%, lower than that of any predictive model. However, this figure is not a true PPV, since it no longer shows the proportion of high-risk patients who are actual high-use patients. Therefore, a second PPV was calculated for each predictive model, using the highest-risk 344 patients, and for the UM, using all 344 identified patients. The actual high use patient group remained the same. At the 344 patient level, model M2 had the highest PPV (32.27%), followed by M1 (31.40%), M5 (30.81%), the UM (30.23%) and M4 (20.06%).

Calculating PPV at both the 620 and 344 patient levels provides two insights into the relative performance of the predictive models and the UM. First, since the UM does not rank all patients, it is capable of only identifying a set number of high-risk patients, irrespective of the size of the desired candidate pool. The PPV calculation at the 620 level, where the UM comes in last, reflects this shortcoming. The PPV calculation at the 344 patient level shows that the UM's accuracy is comparable to that of the predictive models when the number of high-risk patients identified by the predictive models is limited to the number found by the UM. This second conclusion is supported by the mean and median percentile rankings displayed in Table 3.19. At the 344 patient level, the patients identified by the UM have higher mean Outcome Period inpatient admissions

than 89.8 percent of all Cross-Validation Subset patients. The highest mean patient percentile ranking by one of the predictive models is 90.3 percent for M5.

Hypothesis H6 suggested that the UM and the predictive models would identify different patients as highest-risk. This was supported by the comparisons illustrated in Figures 3.9 and 3.10. Hypothesis H8 supposed that the patients identified by the predictive models would have a greater number of comorbid chronic conditions than those patients identified by the UM. The data did not support this conjecture.

One practical way to evaluate the utility of the predictive models is to examine how they perform on patients who were not identified by the UM. Figure 3.11 shows the cumulative accuracy profiles of the predictive models on the actual high-use Cross-Validation Subgroup patients who were not identified by the UM.

Although each model performed similarly in total, with all accuracy ratios near 0.80, two main conclusions can be reached from examining the cumulative accuracy profiles. First, models that focus heavily on prior inpatient admissions (M1, M2, and M5) are most accurate initially, with approximately a third of actual high-use patients scored in the top three percent of risk scores for all patients. However, beyond this point, almost complete reliance on prior inpatient admissions (as seen with M2) becomes problematic, since the remaining population of patients have one or no prior inpatient admissions, and distinctions between patients without prior inpatient admissions is therefore more or less random. The benefit of using the relative risk score as a predictor variable is apparent at this point. Models M1, M4, and M5 identified 31, or approximately 5 percent of, actual high-use patients that were missed by both the UM and M2. Model M4, which relies

most heavily on relative risk score, identified 29 such patients, with models M5 and M1 identifying 20 and 4, respectively.

Overall, the predictive utility of the DCG relative risk score appears disappointing. The Pearson product moment correlation between the 2003 relative risk score and the number of inpatient admissions during the 2004 through 2005 period was 0.020. This poor correlation may underestimate the value of the relative risk score, since the vast majority of patients had no inpatient admissions, and all patients had non-zero relative risk scores. Thus, a weak correlation is virtually assured, even if the relative risk score provides valuable insight for future high-use patients. Furthermore, this project's evaluation method largely considers only binary outcomes: a patient was considered high use or was not considered high use. A patient with a high DCG relative risk score has, by definition, a higher number of comorbid conditions associated with high cost. However, the two-year Outcome Period may not be long enough to see this increased level of risk translated into increased actual use. This is discussed in greater detail in the Limitations section of this chapter.

Finally, even with its low parameter coefficient estimate in model M1, the relative risk score does provide a unique function. Using the UM or predictive models M2 or M3, which all rely primarily on prior inpatient admissions, there is no clear way to rank the risk of patients with equal prior use. For example, if 20 diabetic patients each had one Predictor Period inpatient admission, they would each receive approximately the same predicted use from models M2 and M3, respectively, and they would each be identified

via the UM, but in no particular order. Predictive models M1, M4 and M5 would all rank these patients in order of their relative risk scores.

Summary of the High-Use Patient Subgroup

In the analysis of Objective 3, the high-use patient subset was defined as the Cross-Validation Subgroup patients with three or more Outcome Period (2005 through 2006) inpatient admissions. A total of 620 patients fit this description, 351 of which were female. These 620 patients had a mean age of 36.37 (SD = 17.3) and a mean Predictor Period (2004) rescaled relative risk score of 2.58 (SD = 2.93), which means these patients would be expected to cost 2.58 times as much as the average patient. A total of 157 actual high-use patients had no Predictor Period inpatient admissions and one or more Predictor Period diagnoses of chronic conditions. This fact supported hypothesis H9, which supposed the existence of such patients.

A total of 210 of these patients were among the highest-risk patients based on the predictions of one or more of the predictive models or by selection by the Utilization Method (UM). The other 410 patients were not among the highest-risk patients of any of the predictive models or the UM. Following the nomenclature used in the Results chapter, these patients are referred to as the Identified and Not Identified subgroups, respectively.

The predictive models largely focus on Predictor Period inpatient admissions and Predictor Period relative risk, both of which varied greatly between the Identified and Not Identified subgroups. Identified patients had a higher mean Predictor Period rescaled

relative risk score (4.5 vs. 1.6) and a higher mean Predictor Period number of inpatient admissions (3.57 vs. 0.298). This difference makes sense: any models that rely on prior inpatient admissions and high relative risk scores will not identify actual high-use patients who have neither.

Chronic conditions that were common among the 620 high-use Cross-Validation Subgroup patients included hypertension (n = 322), substance abuse (n = 281), diabetes (n = 259), depression (n = 189), congestive heart failure (n = 146), asthma (n = 114), and chronic obstructive pulmonary disease (n = 107). In all conditions except hypertension, the majority of high-use patients had a diagnosis for the condition in either but not both the Predictor Period and the Outcome Period. For example, of the 146 actual high-use patients with diagnosed congestive heart failure (CHF), only 76 had a diagnosis for CHF in 2004. Some of the CHF diagnoses in 2005 and 2006 for these patients may represent the first clinical diagnosis of the disease. However, for some patients, a past diagnosis (e.g., 2003) for a condition may have been made and a Predictor Period diagnosis for that condition may be absent. In fact, 46 of the 620 high-use Cross-Validation Subgroup patients had a diagnosis of CHF in 2003 but not in 2004. This may help explain why only a third of actual high-use CHF patients were among the highest-risk subsets of the models and UM.

A total of 93 high-use patients had a pregnancy-related diagnosis between 2004 and 2006, only three of which were identified by M4, M5, and the UM, while M1 identified eight and M2 identified nine. Pregnancy poses a unique challenge to identifying high-use patients *a priori*. Normal pregnancy-related hospitalizations are not

a preventable outcome in the context of care management. For this reason, the UM removes from consideration all patients who have a pregnancy-related diagnosis during the Predictor Period. Removing all such patients from consideration may miss legitimate candidates for individualized case management intervention. However, allowing patients with Predictor Period pregnancy-related diagnoses can cause problems for risk models that focus on prior utilization and do not focus on overall patient health status, since they will overstate the risk of pregnant patients in general.

Example of Model Use for Targeting Hypertension Patients

The analysis in the Results chapter of this manuscript focused on the performance of the predictive models and the Utilization Method (UM) on ICC patients in general, scoring each patient (in the cases of the predictive models). While a broad focus may be appropriate for identifying the highest-risk patients and creating individualized intervention strategies (i.e., a case management approach), another approach would be to focus only on patients with a specific diagnosis (i.e., a disease management approach).

This section describes how the models and the UM would be utilized in the disease management approach, focusing specifically on patients with a diagnosis of hypertension in 2004. To more closely simulate the real-world usage of the models and UM, care management patient candidates were selected as if the current date were January 1, 2005. The 2004 data were used for input into the predictive models and for use with the UM process, but inclusion criteria regarding 2005 and 2006 encounters were not imposed, since such information would be unknown.

A total of 175,677 patients had a diagnostic encounter in 2004. Of these, 10,447 had a diagnosis of hypertension. These patients were the total pool from which candidates for care management efforts would be selected. Models M1, M2, M3, M4, and M5 each were loaded with 2004 data to predict the number of inpatient admissions for the following two years (January 1, 2005, through December 31, 2006). Similarly, 2004 data was used with the UM to identify high-risk patients, although the UM produces only binary results: a patient is either high-risk or not.

The top 500 patients were selected based on predicted inpatient admissions using each model. Due to insufficient discriminatory capabilities, a 500-patient subset could not be identified using model M3. Using models M1 and M5, a high-risk subset of 500 patients could be identified. Using models M2 and M4, the identified subsets contained 503 and 502 patients, respectively, due to patients receiving the same scores near the cut-off point. The UM identified 405 patients.

Table 4.2 displays the descriptive statistics and the distribution of 2005-2006 inpatient admissions among the patients identified by models M1, M2, M4, and M5 and the UM.

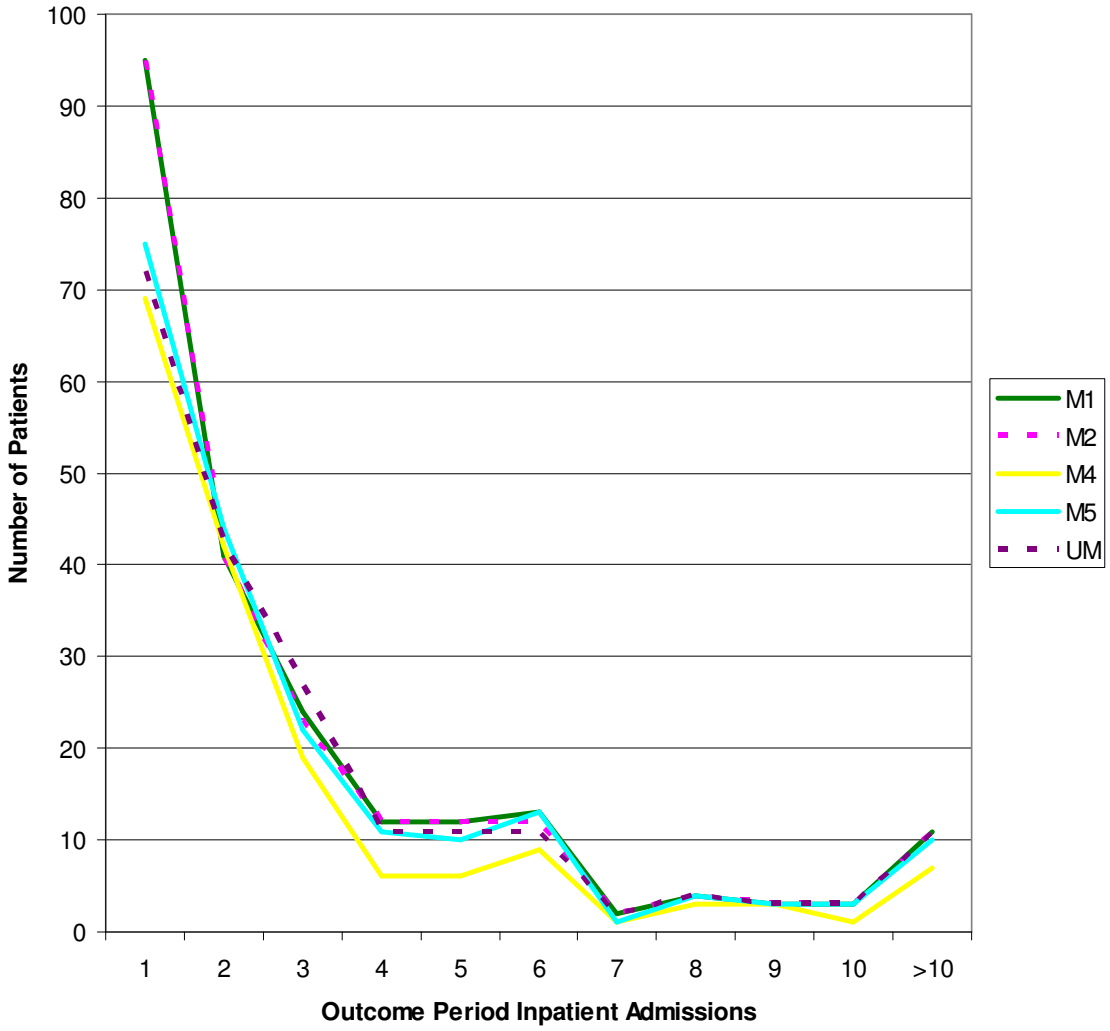
Table 4.2 Distribution of 2005-2006 Inpatient Admissions Among Hypertension Patients Designated as High-Risk by the Predictive Models M1, M2, M4 and M5 and the Utilization Method

	Number of Patients in High-Risk Subsets with Count of Inpatient Admissions				
	Predictive Models				Utilization
Number of Inpatient Admissions	M1	M2	M4	M5	Method
0	280	285	336	304	208
1	95	95	69	75	72
2	41	41	42	44	42
3	24	23	19	22	27
4	12	12	6	11	11
5	12	12	6	10	11
6	13	12	9	13	11
7	2	2	1	1	2
8	4	4	3	4	4
9	3	3	3	3	3
10	3	3	1	3	3
>10	11	11	7	10	11
Total*	500	503	502	500	405
Inpatient Admissions Descriptive Statistics					
Total	939	930	740	887	906
Mean	1.88	1.85	1.47	1.77	2.24
Standard Deviation	7.43	7.41	7.33	7.42	8.21

*The outputs of models M2 and M4 lacked the granuality to create a cut-off point at exactly 500 patients.

Over half of the identified patients, irrespective of model or method, had no inpatient admissions during the 2005 through 2006 time period. The patients identified by model M1 had the highest number of inpatient admissions, 939, followed by M2 at 930, the UM at 906, M5 at 887, and M4 at 740. Patients identified by the UM had the highest mean number of inpatient admissions, 2.24. The distribution of the number of inpatient admissions is shown graphically in Figure 4.1.

Figure 4.1 Distribution of 2005-2006 Inpatient Admissions Among Hypertension Patients Designated as High-Risk by the Predictive Models M1, M2, M4 and M5 and the Utilization Method



Note: This chart shows the number of Oucome Period inpatient admissions for the 500 patients designated high-risk by each of the predictive models M1, M2, M4, and M5 and the Utilization Method (UM). Each high-risk subset included patients with no Outcome Period inpatient admissions (M1: 280; M2: 285; M4: 336; M5: 304; UM: 208). The high-risk subset identified by the UM included only 405 patients.

As seen in Figure 4.1, model M4 tended to identify the fewest patients at each number of inpatient admissions. Models M1 and M2 identified the largest number of

patients with a single inpatient admission, although beyond one admission, the performances of models M1, M2, and M5 and the UM were virtually the same.

Although the UM identified only 405 patients, the additional patients identified by the predictive models resulted in little identifiable benefit. In fact, models M4 and M5 identified fewer actual high-use hypertension patients than the UM. Although models M1 and M2 identified a larger number of patients with one admission between 2005 and 2006, these patients could have been identified by the UM if the criteria were more flexible (i.e., including more diagnosed conditions and/or relaxing the pregnancy exclusion criteria).

However, the UM performed relatively well in part because the number of patients it selects (405) was similar to the number of high-risk patients chosen using each predictive model (500, an arbitrary choice). Changes in either number could impact the UM's performance relative to that of the predictive models. For example, if 2003 were the Predictor Period (instead of 2004), the UM would have only selected 171 hypertension patients, 234 fewer than in 2004. Using the UM on 2005 data selects 240, while only 198 patients are selected using 2006 data.

Sensitivity Analysis

By limiting the two study groups (the Calibration Subgroup and the Cross-Validation Subgroup) to patients with encounters in three contiguous years, a large percentage of ICC population patients were excluded from the analyses. During the study period of January 1, 2003, through December 31, 2006, nearly 500,000 unique patient

identification numbers were associated with encounters at ICC partner provider locations. However, both study groups contained less than 10 percent of this number.

To test whether the exclusions skewed the results, some of the Objective 2 analyses can be recalculated using relaxed inclusion criteria. In 2004, 175,677 patients had one or more encounters of type “Inpatient”, “Emergency Room”, “Outpatient”, “Clinic Visit”, or “Office Visit”. In Objective 2, these patients were required to have one or more diagnostic encounters in years 2005 and 2006 as well. This constraint eliminated over 75 percent of the starting group of 175,677 patients. For this sensitivity analysis, that constraint was removed, leaving all 175,677 patients in the analysis group (referred to henceforth as the Sensitivity Analysis Subgroup). The year 2004 remains the Predictor Period, and the combined two-year period of 2005 and 2006 remains the Outcome Period.

The Sensitivity Analysis Subgroup included 718 patients with three or more inpatient admissions (i.e., actual high-use patients) during the Outcome Period. Although over four times the number of patients were included in the Sensitivity Analysis Subgroup than in the Cross-Validation Subgroup, only 98 more patients were categorized as high use (718 vs. 620).

To calculate the positive predictive value (PPV) of the predictive models and Utilization Method (UM) in the Sensitivity Analysis Subgroup, the highest risk 718 patients were selected with each method. The risk scores produced by model M3 lacked the level of detail needed to create a subset of 718 patients (i.e., too many patients received the same risk scores). The subset for model M2 included 720 patients due to

some patients receiving the same score near the cut-off point. The UM identified only 644 patients.

Table 4.3 shows the summary performance of the predictive models and the UM for the Sensitivity Analysis Subgroup. Included in Table 4.3 (in parentheses) are the equivalent summary performance measures from the Cross-Validation Subgroup, as detailed in Objective 2 of the Results chapter of this manuscript. Using the larger subset of patients (the Sensitivity Analysis Subgroup) caused the number of correctly identified actual high-use patients to increase for the UM. Since the UM is simply a filter for patients who meet its requirements, increasing the size of the patient pool cannot result in fewer actual high-use patients correctly identified. However, with the predictive models, increasing the size of the patient pool increases the number of false-positives (i.e., the number of patients identified as high-risk who were not actually high-use).

Table 4.3 Summary of Performance of Predictive Models and Utilization Method in Sensitivity Analysis Subgroup of Patients

Patient Subset Descriptive Measure	UM		M1		M2		M3*	M4		M5	
Number of actual high use patients contained in predicted subset	122	(104)	131	(166)	127	(165)	NA	81	(110)	120	(153)
Positive predictive value (PPV)**	16.99%	(16.77%)	18.25%	(26.8%)	17.69%	(26.6%)	NA	11.28%	(17.7%)	16.71%	(24.7%)
Number of actual high use patients contained in predicted subset who had no Predictor Period inpatient admissions	6	(6)	0	(0)	0	(0)	NA	9	(12)	4	(6)
Total Outcome Period inpatient admissions of predicted subset	1,225	(913)	1,358	(1,597)	1,333	(1,624)	NA	888	(1,171)	1,207	(1,513)
Mean number of Outcome Period inpatient admissions per correctly identified patient	8.14	(8.78)	8.44	(7.78)	8.66	(7.99)	NA	9.06	(8.38)	8.40	(7.98)
Mean 2005 rescaled relative risk score per patient contained in predicted subset	3.23	(2.78)	4.03	(2.64)	3.16	(2.50)	NA	7.03	(4.93)	5.94	(4.20)
Mean 2006 rescaled relative risk score per patient contained in predicted subset	3.44	(2.85)	4.26	(2.75)	3.24	(2.61)	NA	7.41	(5.05)	6.26	(4.34)

Figures in parentheses represent the same measures in the Cross-Validation Subgroup of patients.

UM = Utilization Method

Models M1, M4, and M5 identified 718 high-risk patients. Model M2 identified 718 high-risk patients. The UM identified 644 high-risk patients.

The Predictor Period is calendar year 2004.

The Outcome Period is 2005 through 2006.

*Model M3 exhibited insufficient discriminatory capabilitity to identify a predicted subset. I.e., very large groups of patients received the same predicted utilization.

**PPV in this table is the percentage of the predicted high use patients who were among the 718 actual patients with 3 or more inpatient admission in the Outcome Period.

While models M1, M2, and M5 had nearly 10 percent higher PPV than the UM in the Cross-Validation Subgroup, this difference dropped substantially in the Sensitivity Analysis Subgroup. In the Sensitivity Analysis Subgroup, model M1 had the highest PPV at 18.25%, followed by M2 at 17.69%, the UM at 16.99%, M5 at 16.71%, and M4 at 11.28%. However, in the Sensitivity Analysis Subgroup, the actual high-use subset of 718 patients represents only approximately 0.4 percent of all patients, compared to approximately 1.5 percent in the Cross-Validation Subgroup.

Relaxing the definition of "high-use" to two or more Outcome Period inpatient admissions (as opposed to three) provides a subset of 1,779 actual high-use patients, representing slightly over one percent of the Sensitivity Analysis Subgroup, which is

closer to the size of the high-use subset used in analyzing the Cross-Validation Subgroup (approximately 1.5 percent). By selecting the highest-risk 1,779 patients, as calculated by each predictive model, PPV can be recalculated in a manner that more closely resembles how the models would actually be used. Table 4.4 shows the summary performance of the predictive models and the UM for the Sensitivity Analysis Subgroup when the high-use subset has been redefined to include patients with two or more inpatient admissions during the Outcome Period.

Table 4.4 Summary of Performance of Predictive Models and Utilization Method in Sensitivity Analysis Subgroup of Patients using Expanded High-Use Subset

Patient Subset Descriptive Measure	UM	M1	M2	M3*	M4	M5
Number of actual high use patients contained in predicted subset	184	300	298	NA	244	327
Positive predictive value (PPV)**	10.34%	16.86%	16.75%	NA	13.72%	18.38%
Number of actual high use patients contained in predicted subset who had no Predictor Period inpatient admissions	13	0	0	NA	32	13
Total Outcome Period inpatient admissions of predicted subset	1,225	1,912	1,909	NA	1,502	1,912
Mean number of Outcome Period inpatient admissions per correctly identified patient	6.07	5.53	5.54	NA	5.27	5.14

UM = Utilization Method

Model M1 identified 1,764 patients. Model M2 identified 1,797 patients. Model M4 identified 1,779 patients. Model M5 identified 1,773 patients. The UM identified 644 high-risk patients.

The Predictor Period is calendar year 2004.

The Outcome Period is 2005 through 2006.

*Model M3 exhibited insufficient discriminatory capabilitity to identify a predicted subset. I.e., very large groups of patients received the same predicted utilization.

**PPV in this table is the percentage of the predicted high use patients who were among the 1,779 actual patients with 2 or more inpatient admission in the Outcome Period.

As shown in Table 4.4, when PPV is calculated using the larger high-use subset of patients, advantages of the predictive models become apparent. The UM can only select more high-risk patients if the criteria are relaxed. This would require adding more eligible diagnoses, reducing the number of required emergency department visits, or eliminating

the inpatient admission requirement, all of which may cause unexpected and undesirable results. In contrast, since the predictive models score all patients, the high-risk subset can be whatever size is needed.

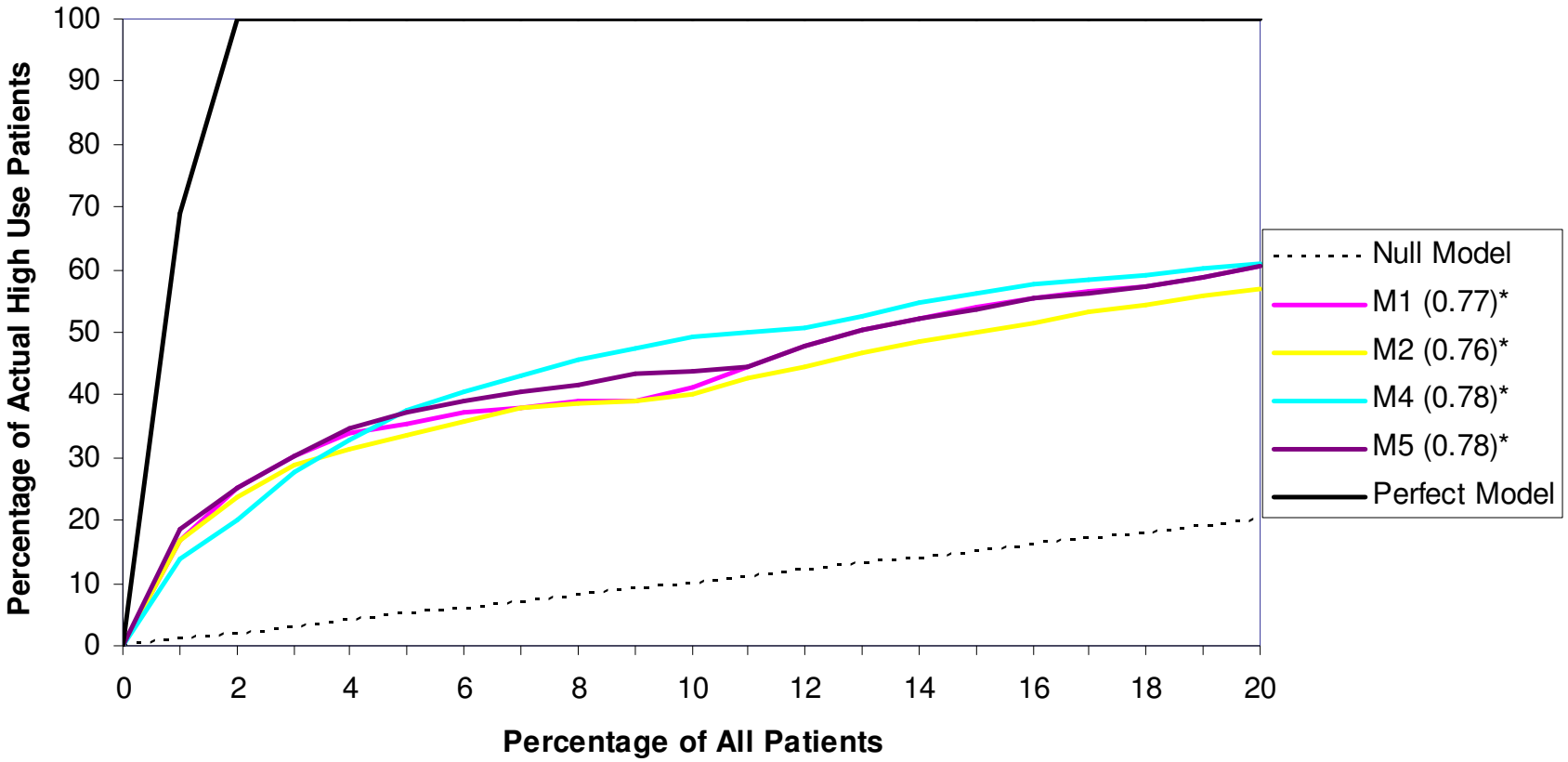
Interestingly, when the high-use criterion was relaxed, the PPV for the two models that relied most heavily on prior inpatient admissions (M1 and M2) both decreased: from 18.25 to 16.86 percent for M1, and from 17.69 to 16.75 percent for M2. The models that more heavily utilized relative risk score both showed improved PPV when the high-use criterion was relaxed. Model M4 increased from 11.28 to 13.72 percent, while model M5 increased from 16.71 to 18.38 percent, the highest of all methods. Model M5 correctly identified almost 10 percent more high-use patients than the second-best performing model, M1 (327 vs. 300, respectively).

The relative changes in PPV as the criteria were relaxed provide insight into how an effective predictive model or method would designate high-risk patients. The most obvious high-risk patients are those with a history of high use. As seen in the analysis of the Cross-Validation Subgroup and again in the Sensitivity Analysis Subgroup with the more strict high-use criteria (three or more Outcome Period inpatient admissions), the UM seems to identify these patients as well as the predictive models. However, once the supply of obvious high-risk patients is exhausted, a model or method must base risk scores on other predictors of use. Model M4's over-reliance on relative risk score causes it to miss the more obvious high-risk patients—those with a history of high use. Model M5 appears to provide a balance of the two approaches. This is most apparent in Figure 4.2, which shows the cumulative accuracy profile (CAP) for the predictive models, using

the most relaxed exclusion criteria (i.e., the Sensitivity Analysis Subgroup, with the high-use patient subset including those with two or more Outcome Period inpatient admissions) and removing all actual high-use patients that were correctly identified by the UM. Essentially, Figure 4.2 is the same as Figure 3.9 in the Results chapter but with the relaxed exclusion criteria. Furthermore, Figure 4.2 only shows the top 20 percent of patients with highest predicted use, whereas previous CAP charts have shown all 100 percent. This smaller frame of focus is used here for two main reasons. First, beyond 20 percent very little difference in model performance exists. Second, when looking at the entire patient population, screening candidates beyond the top 5 to 10 percent is probably impractical due to the number of patients that entails.

The CAPs show model M5 correctly identifying the most actual high-use patients within the top five percent of highest-risk patients. Model M4, relying primarily on the DCG relative risk score, performs the worst through the top five percent but then performs the strongest from five percent through approximately twenty percent. Model M2, relying primarily on prior inpatient admissions performs the worst of all models at all percentages. Model M1 underperforms model M5 through 10 percent, at which point the CAPs converge.

Figure 4.2 Partial Cumulative Accuracy Profile of Models on Sensitivity Analysis Subgroup Patients Not Identified by the Utilization Method with Actual High Use Subset Including All Patients with Two or More Outcome Period Inpatient Admissions



This chart illustrates, for each model, the percentage of actual high use patients y(x) contained in the subset of x% of patients with highest predicted use.
The Null Model represents a model that assigns risk randomly.
The Perfect Model represents a model that perfectly predicts whether each patient will be categorized as high use or not.
*The numbers in parentheses are Accuracy Ratios, which are the ratio of each model's area under the curve but above the Null Model to the Perfect Model's area under the curve but above the Null Model. The closer an Accuracy Ratio is to one the better.

Study Limitations

This study had the following limitations:

1. The number of inpatient admissions was used instead of the number of inpatient days;
2. The two-year Outcome Period may be insufficient to see the outcomes of the progression of uncontrolled chronic diseases;
3. The single-year Predictor Period may be missing diagnoses of high-risk conditions that have been previously diagnosed;
4. Model evaluation techniques lack measures of statistical significance;
5. Certain provider data are omitted from the ICC database;
6. Coding errors and inaccuracy may exist in the data; and
7. The exclusion criteria may have impacted the results.

The use of the number of inpatient admissions instead of inpatient admitted days is a serious limitation. Length of stay may provide more information about the health status of a patient than the number of admissions. While insufficient length of stay data were available to use this variable throughout the study, 2005 and 2006 length of stay data were nearly complete, allowing for the calculation of the Pearson product moment correlation (*r*) between the 2004 relative risk (RR) score and the number of inpatient days between 2005 and 2006. This *r* value (0.14) is nearly identical to the *r* value calculated as part of the testing of hypothesis H3, which determined that the Pearson product moment correlation between the 2003 RR score and the number of inpatient admissions was also 0.14. At the aggregate level, it appears that the measure of health status available in this analysis (the DCG RR score) is not a better predictor of inpatient days than inpatient admissions. However, performance at the aggregate level is not as important as performance with the actual high-use subset. A model could have higher discriminatory capabilities without having higher overall accuracy. This can only be empirically tested as more length of stay data becomes available.

The two year Outcome Period may be insufficient to see the outcomes of the progression of uncontrolled chronic diseases. A patient without Outcome Period inpatient admissions may still have experienced progression in his or her chronic disease(s) that, without intervention, may lead to future hospitalizations. However, without data on intermediate outcomes (e.g., indicators of glycemic control, physiology, and lifestyle for diabetic patients), such progression cannot be evaluated within the current analysis framework. As the ICC continues to collect data, similar analysis projects could be conducted with outcome periods longer than two years.

The reliance by the models and the UM on the single-year Predictor Period may not maximize the potential historical utility of the ICC database. A total of 46 patients had a diagnosis of congestive heart failure (CHF) in 2003, no CHF diagnosis in 2004, and three or more inpatient admissions between 2005 and 2006. Using 2004 data alone, the calculation of the DCG relative risk score would not consider CHF, a condition associated with high risk.

The statistical significance of differences between summary measures of high-risk patients identified by the various methods is unknown. The lack of exclusivity of patient selection by models (i.e., patients could be included in the high-risk subset by multiple models) precludes the use of analysis of variance, since the independence requirement is not met. The use of repeated measures analysis of variance also does not fit the data, since the high-risk patient subset changes from model to model, while the various patient measures (e.g., number of comorbid conditions) stay the same (i.e., these measures are neither random nor fixed effects). Similarly, the summary model performance measures

(e.g., positive predictive value and accuracy ratio) have no corresponding tests of significance.

As a matter of policy, some ICC partner members have withheld certain types of patients and/or encounters from the data files submitted to the ICC. In these cases, patient history or entire patients may be missing from the ICC database and subsequently from this analysis. HIV diagnoses were omitted from St. David's hospitals data and from the Austin Travis County HIV specialty clinic. The People's Community Clinics omitted all encounters with minors.

Coding errors and data inaccuracy may be a limitation of this analysis. Since most of the encounter data on non-Medicaid and State Children's Health Insurance Program (SCHIP) patients is never used for billing purposes, data entry staff and supervising personnel may be less strict on data accuracy.

As briefly discussed in the Results chapter, the exclusion criteria likely resulted in subgroups where patients had more inpatient admissions during the first year (i.e., the Predictor Period) than in the two subsequent years (i.e., the Outcome Period). A patient with high utilization in the Predictor Period is likely to have encounters in the two subsequent years, making that patient eligible for subgroup inclusion. However, high utilization in the last year of the Outcome Period, for example, does not imply the same likeliness of encounters in the two preceding years. This artifact of the analysis methods may cause some misperception regarding trends in inpatient admissions during the study period. However, its impact on the results of this analysis is likely minor, since the results

of the preceding sensitivity analysis (where no utilization exclusion criteria were imposed) were consistent with the results found in the Results chapter of this manuscript.

Chapter 5: Conclusion

Conclusion

Organizations caring for the medically indigent population must find ways to control costs and improve outcomes through proactive, preventative medicine. Disease and case management programs are well suited to these goals; however, key to the success of any interventionary care management program is the identification of high-risk patients who would benefit from increased focus and care.

Published studies evaluating methods of predicting non-disease-specific hospitalization for medically indigent patients are rare. Few organizations like the ICC exist to gather and maintain the data necessary to perform this analysis. Even with pioneering organizations like the ICC, the amount of historically available data may be limited.

Among this historically underserved population, some patients clearly stand out as "common sense" targets for care management intervention. These patients typically have a high number of prior inpatient admissions and known diagnoses of chronic diseases.

The current research compared the performance of theory-based, data-driven models with the "common sense" approach of screening for history of high utilization and chronic disease in identifying future high-cost patients in the medically indigent population of Central Texas. On an individual patient basis, those identified by the "common sense" approach had a similar number of admissions as those identified by the

statistical models. However, the model approach offered two distinct benefits: the ability to identify any number of candidates (instead of only those who met the "common sense" screening criteria), and the ability to rank all patients based on expected utilization.

When care management intervention efforts are small, the "common sense" approach may be adequate. In this situation, the theoretical benefits of the model-based approach may not translate into any practical benefits. However, when the desired candidate pool is larger than what the "common sense" approach can provide and the manpower to contact and evaluate candidates is limited, the risk-ranked candidate pool supplied by the model-based approach may be the right choice.

Implications for Future Research

Based on the analysis herein, the two primary weaknesses of the Utilization Method when compared to the predictive models are:

1. The Utilization Method identifies only a set number of patients and cannot identify more without weakening the query criteria, while the predictive models score all patients; and
2. The Utilization Method provides no explicit ordering of patients, while the risk scores provided by the predictive models allows for ordering of all patients.

To compensate for these weaknesses, the Utilization Method (UM) could be converted into a predictive model, with a predictor variable replacing each query requirement. The diagnostic requirements would become binary variables (e.g., diagnosis of diabetes represented by a 1 for yes or a 0 for no), and the number of inpatient and emergency department encounters both as continuous variables.

Prior to performing the analysis in the current project, it was thought that diagnostic variables such as those described above would provide less predictive utility than the DCG relative risk score, which, in theory, incorporates this information and more. The greatest potential issue with a binary variable approach is that it may produce unexpected and undesirable results when model calibration creates negative coefficients for diagnosis-based predictor variables—in effect, selecting against the very patients upon which the UM was designed to focus. However, this approach may overcome one major limitation of the predictive models and the UM seen in the current analysis. Rather than using a single year as the Predictor Period, “history of disease” binary variables could be used for the conditions of interest, allowing for the use of past diagnoses.

In an evaluation of prediction models used to identify future high-cost cases, Cohen et al. found that adding longitudinal data beyond one year added little discrimination capability.[185] The authors suggest that a single year of prior information is sufficient. Unfortunately, this generalization, made from analysis on detailed Medical Expenditure Panel Survey data, may not be extendable to analysis utilizing data sources such as that maintained by the ICC. The membership of the medically indigent population is in a constant state of flux, resulting in holes in the data. Patients may travel from neighboring communities for hospital use but receive primary care from clinics and outpatient centers that are not ICC data-sharing members. Some uninsured patients, especially when immigration status is of concern, may pay cash for services and/or use another individual's Social Security Number. For these reasons, the generalizations made by Cohen et al. should not be applied to this population.

As more length-of-stay data become available in the ICC database, this project could be repeated with length-of-stay as the dependent variable instead of number of inpatient admissions and with history of disease binary variables.

[185] Cohen, Ezzati-Rice and Yu. The utility of extended longitudinal profiles in predicting future health care expenditures. *Med Care.* May 2006;44(5 Suppl):I45-53.

Works Cited

Lawsuits imply EMTALA requires EDs to admit all uninsured patients. *ED Manag.* Aug 2004;16(8):85-87.

Aaron HJ. Should public policy seek to control the growth of health care spending? *Health Aff (Millwood).* Jan-Jun 2003;Suppl Web Exclusives:W3-28-36.

Adams RJ, Smith BJ, Ruffin RE. Factors associated with hospital admissions and repeat emergency department visits for adults with asthma. *Thorax.* July 1, 2000 2000;55(7):566-573.

Anderson JW, Konz EC. Obesity and Disease Management: Effects of Weight Loss on Comorbid Conditions. *Obesity Res.* November 1, 2001 2001;9(90004):326S-334.

Arrow KJ. Uncertainty and the Welfare Economics of Medical Care. *The American Economic Review.* 1963;53(5):941-973.

Arrow KJ, Debreu G. Existence of an Equilibrium for a Competitive Economy. *Econometrica.* 1954/07 1954;22(3):265-290.

Ash A, Ellis RP, Yu W, et al. *Risk Adjustment for the Non-Elderly.* Boston University. Boston, MA: Final report submitted to the Health Care Financing Organization under cooperative agreement No. 18-C-90462/1-02; Jun 1998.

Ash A, Porell F, Gruenberg L, Sawitz E, Beiser A. Adjusting Medicare capitation payments using prior hospitalization data. *Health Care Financ Rev.* Summer 1989;10(4):17-29.

Ash AS, Zhao Y, Ellis RP, Schlein Kramer M. Finding future high-cost cases: comparing prior cost versus diagnosis-based methods. *Health Serv Res.* Dec 2001;36(6 Pt 2):194-206.

Ayanian JZ, Weissman JS, Schneider EC, Ginsburg JA, Zaslavsky AM. Unmet health needs of uninsured adults in the United States. *Jama.* Oct 25 2000;284(16):2061-2069.

Beddhu S, Bruns FJ, Saul M, Seddon P, Zeidel ML. A simple comorbidity scale predicts clinical outcomes and costs in dialysis patients. *Am J Med.* Jun 1 2000;108(8):609-613.

Berk ML, Monheit AC. The concentration of health care expenditures, revisited. *Health Aff (Millwood).* Mar-Apr 2001;20(2):9-18.

Bernstein RH. New arrows in the quiver for targeting care management: high-risk versus high-opportunity case identification. *J Ambul Care Manage.* Jan-Mar 2007;30(1):39-51.

Blough DK, Madden CW, Hornbrook MC. Modeling risk using generalized linear models. *Journal of Health Economics.* 1999/4 1999;18(2):153-171.

Blustein J. Medicare coverage, supplemental insurance, and the use of mammography by older women. *N Engl J Med.* Apr 27 1995;332(17):1138-1143.

Bodenheimer T. High and rising health care costs. Part 2: technologic innovation. *Ann Intern Med.* Jun 7 2005;142(11):932-937.

Borger C, Smith S, Truffer C, et al. Health Spending Projections Through 2015: Changes On The Horizon. *Health Aff.* March 1, 2006 2006;25(2):w61-73.

Bradbury RC, Golec JH, Steen PM. Comparing uninsured and privately insured hospital patients: admission severity, health outcomes and resource use. *Health Serv Manage Res.* Aug 2001;14(3):203-210.

Bruns FJ, Seddon P, Saul M, Zeidel ML. The cost of caring for end-stage kidney disease patients: an analysis based on hospital financial transaction records. *J Am Soc Nephrol.* May 1998;9(5):884-890.

Buchmueller TC, Feldstein PJ. Consumers' sensitivity to health plan premiums: evidence from a natural experiment in California. *Health Aff (Millwood).* Spring 1996;15(1):143-151.

Cameron E, Peloso GF. Risk management and the precautionary principle: a fuzzy logic model. *Risk Anal.* Aug 2005;25(4):901-911.

Chen A, Brown R, Archibald N, Aliotta S, Fox PD. Best practices in coordinated care. Available at: http://www.mathematica-mpr.com/PDFs/bestpractices.pdf. Accessed 3/16/2007.

Chernew M, Cutler DM, Keenan PS. Increasing health insurance costs and the decline in insurance coverage. *Health Serv Res.* Aug 2005;40(4):1021-1039.

Cohen J. Multiple regression as a general data-analytic system. *Psychological Bulletin.* 1968;70:426-443.

Cohen SB, Ezzati-Rice T, Yu W. The utility of extended longitudinal profiles in predicting future health care expenditures. *Med Care.* May 2006;44(5 Suppl):I45-53.

Congressional Budget Office. An analysis of the literature on disease management programs. Available at: http://www.cbo.gov/ftpdocs/59xx/doc5909/10-13-DiseaseMngmnt.pdf. Accessed 2/9/2007.

Crystal S, Sambamoorthi U, Walkup J, Akincigil A. Diagnosis and treatment of depression in the elderly Medicare population: predictors, disparities, and trends. *Journal of the American Geriatrics Society.* Dec. 2003;51(12):1718.

Davis K, Doty M, Ho A. *How High is Too High? Implications of High-Deductible Health Plans*: The Commonwealth Fund; April, 2005.

Davison AC, Gigli A. Deviance residuals and normal scores plots. *Biometrika.* June 1, 1989 1989;76(2):211-221.

DeLong ER, Peterson ED, DeLong DM, Muhlbaier LH, Hackett S, Mark DB. Comparing risk-adjustment methods for provider profiling. *Statistics in Medicine.* 1997;16(23):2645-2664.

Derlet RW, Kinser D, Ray L, Hamilton B, McKenzie J. Prospective identification and triage of nonemergency patients out of an emergency department: a 5-year study. *Ann Emerg Med.* Feb 1995;25(2):215-223.

Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *Journal of Clinical Epidemiology.* 1992;45(6):613-619.

Diehr P, Yanez D, Ash A, Hornbrook M, Lin DY. Methods for analyzing health care utilization and costs. *Annu Rev Public Health.* 1999;20:125-144.

Dowd BE. The Logic of Moral Hazard: A Game Theoretic Illustration. *The Journal of Risk and Insurance.* 1982;49(3):443-447.

DxCG Inc. *DxCG Risk Smart Stand Alone Version 2.2 Help*; 2007.

DxCG Inc. *DxCG RiskSmart Models and Methodologies Guide*; Nov 2002.

Eggert GM, Friedman B. The need for special interventions for multiple hospital admission patients. *Health Care Financ Rev.* Dec 1988;Spec No:57-67.

Ellis RP, Ash A. Refinements to the Diagnostic Cost Group (DCG) model. *Inquiry.* Winter 1995;32(4):418-429.

Ellis RP, Pope GC, Iezzoni L, et al. Diagnosis-based risk adjustment for Medicare capitation payments. *Health Care Financ Rev.* Spring 1996;17(3):101-128.

Engel CC, von Korff M, Katon WJ. Back pain in primary care: predictors of high health-care costs. *Pain.* May-Jun 1996;65(2-3):197-204.

Enthoven AC. Employment-based health insurance is failing: now what? *Health Aff (Millwood).* Jan-Jun 2003;Suppl Web Exclusives:W3-237-249.

Fan VS, Curtis JR, Tu S-P, McDonell MB, Fihn SD. Using Quality of Life to Predict Hospitalization and Mortality in Patients With Obstructive Lung Diseases. *Chest.* August 1, 2002 2002;122(2):429-436.

Felt-Lisk S, Mays GP. Back to the drawing board: new directions in health plans' care management strategies. *Health Aff (Millwood).* Sep-Oct 2002;21(5):210-217.

Fields WW, Asplin BR, Larkin GL, et al. The Emergency Medical Treatment and Labor Act as a Federal Health Care Safety Net Program. *Acad Emerg Med.* November 1, 2001 2001;8(11):1064-1069.

Fihn SD, Callahan CM, Martin DC, McDonell MB, Henikoff JG, White RH. The Risk for and Severity of Bleeding Complications in Elderly Patients Treated with Warfarin. *Ann Intern Med.* June 1, 1996 1996;124(11):970-979.

Frasure-Smith N, Lesperance F, Talajic M. Depression and 18-month prognosis after myocardial infarction. *Circulation.* Feb 15 1995;91(4):999-1005.

Gibson TB, Ozminkowski RJ, Goetzel RZ. The effects of prescription drug cost sharing: a review of the evidence. *Am J Manag Care.* Nov 2005;11(11):730-740.

Gilmer TP, Dolder CR, Lacro JP, et al. Adherence to Treatment With Antipsychotic Medication and Health Care Costs Among Medicaid Beneficiaries With Schizophrenia. *Am J Psychiatry.* April 1, 2004 2004;161(4):692-699.

Goetzel RZ, Ozminkowski RJ, Villagra VG, Duffy J. Return on Investment in Disease Management: A Review. *Health Care Financing Review*

Health Care Financing Review J1 - Health Care Financing Review. 2005///Summer 2005;26(4):1-19.

Guirguis LM, Taylor EW. The complications of cholelithiasis caused by state authorization delays. *Surg Endosc.* Sep 1995;9(9):974-976.

Hadley J. Sicker and poorer--the consequences of being uninsured: a review of the research on the relationship between health insurance, medical care use, health, work, and income. *Med Care Res Rev.* Jun 2003;60(2 Suppl):3S-75S; discussion 76S-112S.

Hadley J, Holahan J. The Cost of Care for the Uninsured: What Do We Spend, Who Pays, and What Would Full Coverage Add to Medical Spending? *Kaiser Commission on Medicaid and the Uninsured*. Available at: http://www.kff.org/uninsured/7084.cfm.

Harris JM, Jr. Disease management: new wine in new bottles? *Ann Intern Med.* May 1 1996;124(9):838-842.

Harris LE, Luft FC, Rudy DW, Kesterson JG, Tierney WM. Effects of multidisciplinary case management in patients with chronic renal insufficiency. *The American Journal of Medicine.* 1998/12 1998;105(6):464-471.

Haskell WL, Berra K, Arias E, et al. Multifactor cardiovascular disease risk reduction in medically underserved, high-risk patients. *Am J Cardiol.* Dec 1 2006;98(11):1472-1479.

Herring B. The effect of the availability of charity care to the uninsured on the demand for private health insurance. *Journal of Health Economics.* 2005/3 2005;24(2):225-252.

Hoaglin D. A Poissonness plot. *The American Statistician.* 1980;34(3):146-149.

Holloway F, Oliver N, Collins E, Carson J. Case management: a critical review of the outcome literature. *European Psychiatry.* 1995;10:113-128.

Holton GA. Defining Risk. *Financial Analysts Journal.* 2004;60(6):19-25.

Hsu J, Price M, Huang J, et al. Unintended Consequences of Caps on Medicare Drug Benefits. *N Engl J Med.* June 1, 2006 2006;354(22):2349-2359.

Iezzoni L. *Risk Adjustment for Measuring Health Care Outcomes*. 3rd ed; 2003.

Institute of Medicine Committe on Utilization Management by Third Parties. *Controlling Costs and Changing Patient Care? The Role of Utilization Management*. Washington, D.C.: National Academy Press; 1989.

Kaiser Commission on Medicaid and the Uninsured. Medicaid Outpatient Prescription Drug Benefits: Findings from a National Survey, 2003. Available at:

http://www.kff.org/medicaid/upload/Medicaid-Outpatient-Prescription-Drug-Benefits-Findings-from-a-National-Survey-2003.pdf.

Kaiser Commission on Medicaid and the Uninsured. 50 State Comparisons: Introduction to Health Coverage & Uninsured. Available at: http://www.statehealthfacts.kff.org/cgi-bin/healthfacts.cgi?action=compare&welcome=1&category=Health+Coverage+%26+Uninsured.

Keeler EB. Effects of cost sharing on use of medical services and health. *Journal of Medical Practice Management.* 1994;8:317-321.

Keenan SC, Sobehart JR. Performance measures for credit risk models, Research report# 1-10-10-99. *Risk Management Services, Moody's Investors Service*. 1999.

Knabel T, Louwers J. Intervenability: another measure of health risk: by coupling predictive modeling with evidence-based medicine, health plans can identify patients who will benefit the most from care management intervention. *Health Management Technology.* 2004;25(7):36(33).

Kongstvedt PR. *Managed Care: What It Is and How It Works*. 2nd Edition ed. Gaithersburg, Maryland: Aspen Publishers; 2002.

Koopmanschap MA, van Ineveld BM. Towards a new approach for estimating indirect costs of disease. *Soc Sci Med.* May 1992;34(9):1005-1010.

Ku L, Coughlin TA. Sliding-scale premium health insurance programs: four states' experiences. *Inquiry.* Winter 1999;36(4):471-480.

Kushel MB, Colfax G, Ragland K, Heineman A, Palacio H, Bangsberg DR. Case management is associated with improved antiretroviral adherence and CD4+ cell counts in homeless and marginally housed individuals with HIV infection. *Clinical Infectious Diseases: An Official Publication Of The Infectious Diseases Society Of America.* 2006/07/15/ 2006;43(2):234-242.

Latour CHM, Bosmans JE, van Tulder MW, et al. Cost-effectiveness of a nurse-led case management intervention in general medical outpatients compared with usual care: An economic evaluation alongside a randomized controlled trial. *Journal of Psychosomatic Research.* 2007/3 2007;62(3):363-370.

Lee KH, Davenport L. Can case management interventions reduce the number of emergency department visits by frequent users? *Health Care Manag (Frederick).* Apr-Jun 2006;25(2):155-159.

Leslie RC. *Cost Analysis of the Uninsured, Indigent Population of Central Texas*, Master's Thesis. College of Pharmacy, The University of Texas at Austin; 2006.

Liebowitz S. Why health care costs too much. *Cato Institute*. Available at: http://www.cato.org/pubs/pas/pa211.html. Accessed 1/21/2007.

Linden AL. What will it take for disease management to demonstrate a return on investment? New perspectives on an old theme. Am J Manag Care. 2006;12(4):217-22.

Lu K-Y, Lin P-L, Tzeng L-C, Huang K-Y, Chang L-C. Effectiveness of case management for community elderly with hypertension, diabetes mellitus, and hypercholesterolemia in Taiwan: a record review. *International Journal Of Nursing Studies.* 2006/11/27/ 2006;43(8):1001-1010.

Lurk JT, Dejong DJ, Woods TM, Knell ME, Carroll CA. Effects of changes in patient cost sharing and drug sample policies on prescription drug costs and utilization in a safety-net-provider setting. *American Journal of Health-System Pharmacy.* 2004/02/01/ 2004;61(3):267-272.

McCulloch P, Nelder JA. *Generalized Linear Models*. 2nd ed. Cambridge: Chapman & Hall; 1991.

McNutt L-A, Wu C, Xue X, Hafner JP. Estimating the Relative Risk in Cohort Studies and Clinical Trials of Common Outcomes. *Am. J. Epidemiol.* May 15, 2003 2003;157(10):940-943.

Medco Health Solutions. *Drug Trend Report 2005.* Franklin Lakes, N.J: Medco; 2005.

Meenan RT, O'Keeffe-Rosetti C, Hornbrook MC, et al. The sensitivity and specificity of forecasting high-cost users of medical care. *Med Care.* Aug 1999;37(8):815-823.

Mueser KT, Bond GR, Drake RE, Resnick SG. Models of community care for severe mental illness: a review of research on case management. *Schizophr Bull.* 1998;24(1):37-74.

Mullins CD, DeVries AR, Hsu VD, Meng F, Palumbo FB. Variability And Growth In Spending For Outpatient Specialty Pharmaceuticals. *Health Aff.* July 1, 2005 2005;24(4):1117-1127.

Murdoch SH, Armstrong D. County Estimates of the Uninsured for Texas, 2005. *Texas State Data Center*. Available at: http://txsdc.utsa.edu/download/pdf/presentations/2006_12_01_County_Estimates_of_the_Uninsured_for_Texas_Tables.pdf.

National Chronic Care Consortium. Case Management: Methods and Issues. Available at: http://www.nccconline.org/pdf/CaseManagement.pdf. Accessed 3/16/2007.

Nelder JA, Wedderburn RWM. Generalized linear models. *J. R. Statist. Soc.* 1972;A135:370-384.

Nelson AA, Jr., Reeder CE, Dickson WM. The effect of a Medicaid drug copayment program on the utilization and cost of prescription services. *Med Care.* Aug 1984;22(8):724-736.

Norris SL, Nichols PJ, Caspersen CJ, et al. The effectiveness of disease and case management for people with diabetes: A systematic review. *American Journal of Preventive Medicine.* 2002/5 2002;22(4, Supplement 1):15-38.

Oderda G, Dacorogma M, Jung T. Credit risk models - do they deliver their promises? A quantitative assessment. *Economic Notes.* 2003;32(2):177-195.

Okin RL, Boccellari A, Azocar F, et al. The effects of clinical case management on hospital service use among ED frequent users. *The American Journal of Emergency Medicine.* 2000/9 2000;18(5):603-608.

Paterson DL, Swindells S, Mohr J, et al. Adherence to Protease Inhibitor Therapy and Outcomes in Patients with HIV Infection. *Ann Intern Med.* July 4, 2000 2000;133(1):21-30.

Pauly MV. The Economics of Moral Hazard: Comment. *American Economic Review.* 1968;58(3):531-537.

Phillips GA, Brophy DS, Weiland TJ, Chenhall AJ, Dent AW. The effect of multidisciplinary case management on selected outcomes for frequent attenders at an emergency department. *The Medical Journal Of Australia.* 2006/06/19/ 2006;184(12):602-606.

Pope GC, Ellis RP, Ash AS, et al. Principal inpatient diagnostic cost group model for Medicare risk adjustment. *Health Care Financ Rev.* Spring 2000;21(3):93-118.

Pradhan AD, Manson JE, Rossouw JE, et al. Inflammatory Biomarkers, Hormone Replacement Therapy, and Incident Coronary Heart Disease: Prospective Analysis From the Women's Health Initiative Observational Study. *JAMA.* August 28, 2002 2002;288(8):980-987.

R Development Core Team. R: A language and environment for statistical computing. Available at: http://www.R-project.org.

Radcliff TA, Cote MJ, Duncan RP. The identification of high-cost patients. *Hosp Top.* Summer 2005;83(3):17-24.

Rapp CA. The active ingredients of effective case management: a research synthesis. *Community Ment Health J.* Aug 1998;34(4):363-380.

Reinhardt UE. Does the aging of the population really drive the demand for health care? *Health Aff.* Nov-Dec 2003;22(6):27-39.

Simon GE, Manning WG, Katzelnick DJ, Pearson SD, Henk HJ, Helstad CP. Cost-effectiveness of systematic depression treatment for high utilizers of general medical care. *Arch Gen Psychiatry.* February 1, 2001 2001;58(2):181-187.

Smalley WE, Griffin MR, Fought RL, Sullivan L, Ray WA. Effect of a prior-authorization requirement on the use of nonsteroidal antiinflammatory drugs by Medicaid patients. *N Engl J Med.* Jun 15 1995;332(24):1612-1617.

Solomon P. The efficacy of case management services for severely mentally disabled clients. *Community Ment Health J.* Jun 1992;28(3):163-180.

Soumerai S, Avorn J, Ross-Degnan D, Gortmaker S. Payment restrictions for prescription drugs under Medicaid. Effects on therapy, cost, and equity. *N Engl J Med.* August 27, 1987 1987;317(9):550-556.

Soumerai S, Ross-Degnan D, Avorn J, McLaughlin T, Choodnovskiy I. Effects of Medicaid drug-payment limits on admission to hospitals and nursing homes. *N Engl J Med.* October 10, 1991 1991;325(15):1072-1077.

Spillane L, Lumb E, Cobaugh D, Wilcox S, Clark J, Schneider S. Frequent users of the emergency department: can we intervene? *Acad Emerg Med.* June 1, 1997 1997;4(6):574-580.

Stewart S, Marley JE, Horowitz JD. Effects of a multidisciplinary, home-based intervention on planned readmissions and survival among patients with chronic congestive heart failure: a randomised controlled study. *The Lancet.* 1999/9/25 1999;354(9184):1077-1083.

Stuart B, Simoni-Wastila L, Chauncey D. Assessing The Impact Of Coverage Gaps In The Medicare Part D Drug Benefit. *Health Aff.* April 19, 2005 2005:167.

Stuart B, Zacker C. Who bears the burden of Medicaid drug copayment policies? *Health Aff.* March 1, 1999 1999;18(2):201-212.

Thorpe KE. The rise in health care spending and what to do about it. *Health Aff.* Nov-Dec 2005;24(6):1436-1445.

U.S. Department of Health and Human Services CfMaMS. National Health Expenditures Aggregate, Per Capita Amounts, Percent Distribution, and Average Annual Percent Growth, by Source of Funds: Selected Calendar Years 1960-2005. Available at: http://www.cms.hhs.gov/NationalHealthExpendData/downloads/tables.pdf. Accessed 1/22/2007.

United States General Accounting Office. *Medicaid and SCHIP: States' Premium and Cost Sharing Requirements for Beneficiaries* March 2004.

Villagra V. Strategies to control costs and quality: a focus on outcomes research for disease management. *Med Care.* Apr 2004;42(4 Suppl):III24-30.

Watson Wyatt. *Managing Health Care Costs in a New Era: Tenth Annual National Business Group on Health/Watson Wyatt Survey Report 2005.* Washington: Watson Wyatt; 2005.

Weiner JP, Starfield BH, Lieberman RN. Johns Hopkins Ambulatory Care Groups (ACGs). A case-mix system for UR, QA and capitation adjustment. *HMO Pract.* Mar 1992;6(1):13-19.

Weiner S. "I can't afford that!". Dilemmas in the care of the uninsured and underinsured. *Journal of General Internal Medicine.* 2001;16(6):412-418.

Weiss KB, Gergen PJ, Hodgson TA. An economic evaluation of asthma in the United States. *N Engl J Med.* Mar 26 1992;326(13):862-866.

Weiss KB, Sullivan SD. The economic costs of asthma: a review and conceptual model. *Pharmacoeconomics.* Jul 1993;4(1):14-30.

World Health Organization. *Preventing Chronic Diseases: A Vital Investment* 2005.

Zhao Y, Ash AS, Ellis RP, et al. Predicting pharmacy costs and other medical costs using diagnoses and drug claims. *Med Care.* Jan 2005;43(1):34-43.

Zhao Y, Ash AS, Ellis RP, Slaughter JP. Disease burden profiles: an emerging tool for managing managed care. *Health Care Manag Sci.* Aug 2002;5(3):211-219.

Zook CJ, Moore FD. High-cost users of medical care. *N Engl J Med.* May 1 1980;302(18):996-1002.

Zou G. A modified poisson regression approach to prospective studies with binary data. *Am J Epidemiol.* Apr 1 2004;159(7):702-706.

Vita

Ryan Christopher Leslie is the son of Steven and Denese Leslie. In August 1998, he graduated with a Bachelor of Business Administration degree from the McCombs School of Business at the University of Texas at Austin. After graduation, he spent almost five years as a management consultant with Deloitte & Touche, an international accounting and consulting firm. In 2003 he began graduate school in pharmacoeconomics and outcomes research at the College of Pharmacy at the University of Texas at Austin, getting his M.S. in 2006. During graduate school he has provided consulting services to a variety of health care organizations. He is married and has a strikingly handsome son.

Permanent Address: 6709 Beauford Drive

Austin, Texas 78750

This dissertation was typed by the author.