UNIVERSITY OF CALIFORNIA

Los Angeles

Cost, Quality, and Access: A Foundation for Managerial and Policy Analysis in the Health Care System

A dissertation submitted in partial satisfaction of the requirements for the degree Doctor of Philosophy in Management

by

Rahul Asthana

1997

UMI Number: 9804601

The dissertation of Rahul Asthana is approved.

Stuart O. Schweitzer

Marvin B. Lieberman

Donald G. Morrison

Reza H. Ahmadi, Committee Chair

University of California, Los Angeles

1997

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

Acknowledgments

I am very deeply indebted to a number of people for the assistance and encouragement they have given me during the years I did this research and wrote this dissertation. I would like to use this opportunity to thank all of them.

I want to express my deepest appreciation to advisor, Dr. Reza H. Ahmadi, whose teachings and guidance will remain with me for many years. More, though, I would like to thank Professor Ahmadi for his kindness and support, the level of which very often surprised me, during the four years I was his student.

I am extremely grateful to Professors Stuart Schweitzer, Don Morrison, and Marvin Lieberman for both agreeing to serve on my committee, and for their very helpful suggestions and remarks on my research.

I am grateful to the UCLA Graduate School, the Center for Technology Management at the Anderson School, the John M. Olin Foundation, the Doctoral Program Office, and Professors Steve Lippman, Bill Yost and Reza Ahmadi for their fellowship and scholarship support.

The help of my fellow doctoral students has gone a long way in making my stay at UCLA productive. In particular, I am deeply indebted to Clare Chow who served as my comrade - in - arms during the first two years of the Ph.D. program. I am very grateful to Hosun Rhim for many valuable comments on my research, and to Akihiro Inoue for providing me with help in statistical analysis and simulation modeling. I am also very grateful to Illyung Kim for many helpful and constructive comments in the early stages of my research.

The friendship of my colleagues at the Anderson School, especially that of Akihiro Inoue, Hosun Rhim, Clare Chow, David Choi, Lieven Deemester, Rachel

Yang, Hongbo Wang, Juin Chong, Jovan Grahovac, Shigeru Asaba, Jina Kang, Tisha Greenwood, and Bashyam T.C.A. helped make the years I spent in the Ph.D. program very enjoyable.

My friends outside of UCLA deserve special mention for their help and encouragement over the years. I am extremely grateful to Linda Yang for her unfailing support and friendship. I shall be eternally indebted to her. I am extremely grateful, also, to Celia Chang for her considerable help in writing the necessary programming code for this dissertation. I would like to thank Nobi Mori for doing much to make my life outside of UCLA pleasant and enjoyable. I would also like to thank Raghu Rangarajan, without whom this dissertation would have been completed years ago.

Finally, but most importantly, I am deeply grateful to my family for their constant support and encouragement. I thank my brother, Praveen, for providing me with much of the strength necessary to complete a Ph.D. I thank my sister, Seema, and my sister-in-law, Mona, for their encouragement. I would like to thank my parents for their unselfish love, infinite support and faith. Theirs is a debt I can never repay. It is to my family that I dedicate this dissertation.

VITA

January 4, 1967	Born, Debre Berhan, Ethiopia
1984 - 1988	B.S.E.E., Electrical Engineering Princeton University Princeton, New Jersey
1988 - 1989	Research Scientist Almaden Research Center IBM Corporation San Jose, California
1989	U.S. Patent Award: Detecting Polarization State of an Optical Wavefront IBM Corporation San Jose, California
1989 - 1991	M.S.E.E., Electrical Engineering University of Southern California Los Angeles, California
1991 - 1992	Candidate, Doctor of Philosophy in Engineering University of Southern California Los Angeles, California
1992 - 1996	Ph.D. student The Anderson School at UCLA University of California at Los Angeles Los Angeles, California
1994	Named Doctoral Fellow of the Olin Center for Policy The Anderson School at UCLA University of California at Los Angeles Los Angeles, California

ABSTRACT OF THE DISSERTATION

Cost, Quality, and Access: A Foundation
for Managerial and Policy Analysis
in the Health Care System

by

Rahul Asthana
Doctor of Philosophy in Management
University of California, Los Angeles, 1997
Professor Reza H. Ahmadi, Chair

The American health care system is beset by significant problems. Chief among these are the continuing rapid growth in costs, and a denial of access to a large portion of the population. Amidst the problems, however, remains the fact that the U.S. system offers a very high quality of care to those with access. To address the problems in the health care system, yet preserve quality standards, a considerable number of reform proposals have been advanced. These proposals have, however, largely failed to pass legislative muster and the problem of providing universal access without decreasing quality or increasing cost remains unsolved.

Exclusive to the legislative attempts at reform, the health care system in the United States is undergoing a paradigm shift from fee-for-service to managed care. The defining characteristic of managed care is the capitated budget; consequently, the provision of health care will increasingly be subject to budget constraints. It is against the landscape of the problems in the health care system, and the paradigm shift in health care delivery that we present the two essays that constitute the research in this dissertation.

In Essay 1, we demonstrate a methodology that, for a *single* disease and within modeling assumptions, shows that it is possible to achieve universal access and *increase* service quality without increasing cost. The central tenet of the methodology is to relate quantitatively the three measures of cost, quality, and access. The methodology is developed as follows: We define two measures of quality. First, 'national health quality,' or NHQ, that defines the quality of health of the population and, second, service quality, or SQ, at medical treatment facilities. We use two concurrent queueing systems to model the relationships between each measure of quality with cost and access.

The relationships among cost, quality, and access serve as the foundation for single disease managerial and policy analysis. In particular, we demonstrate that there are several policies that achieve, at constant cost, universal access and increase in quality. The relationships developed in this paper allow for the selection of the supremum policy.

In Essay 2 we consider the formulation of managerial policy, for a *single* chronic disease, under constrained budgets. The policy we formulate has three objectives. The first objective is the maximization of a societal level health index called National Health Quality (NHQ). The second objective is the maximization of service quality (SQ) at the health care facilities. The third is to maximize access without cost to either SQ or NHQ.

Our concern is to allocate available capitated resources such that the constituent measures of the policy are maximized. The method is to formulate a one disease product form multi class queueing network model of the health care system. With the measures NHQ and SQ forming the objective functions, we develop optimization models to allocate resources throughout the health care system, and over the age groups that suffer the chronic disease. The results of this paper demonstrate the necessary resource allocations, explore the dynamics of capitated health care systems as a function of access, and provide managerial implications. Finally, the paper develops trade off curves for the elements that constitute the policy. This serves as a tool for decision makers.

Chapter 1

Introduction

1.1. Reform of a Health Care System in Crisis

There is a sense of urgency in the need to reform the US health care system. This urgency is driven by three factors. The first is the magnitude of health care expenditures, estimated at over $ 1 trillion in 1995 (Figure 1), and its increasing dominance (presently over 14% of GDP is dedicated to medical care) of the US economy (Levit *et al*, 1996; Monaco *et al*, 1995). The second is the more than decade long and continuing rapid rise in health care costs at an average annual inflationary growth of 9.8% (Levit *et al*, 1994) in an economy that has averaged an annual GDP economic growth of 6.3%, an average annual inflation rate (measured by the Consumer Price Index) of 3.6% (Figure 2), and an average annual population growth of only 1.04 % in the same time period (Statistical Abstract of the United States, 1995). The third factor is the large number of Americans without health care insurance or financial protection from catastrophic illness: It is estimated that presently over 17.4% of the non elderly population, or 38.5 million people (Vincenzino, 1994; Davis *et al*, 1995; Rowland *et al*., 1994) lack health coverage. These numbers are likely to increase if the cost of care continues to spiral at an economically unsustainable rate.

The twin concerns of rising health care costs and diminished access brought forward numerous legislative proposals aimed at reform (White House Domestic Policy Council, 1993; Mitchell, 1994; Moynihan, 1994; Chafee, 1993). These proposals had, as their primary goals, the provision of universal access to health care, the maintenance of costs, and the increase of health care quality (Heclo, 1995).

It remains a fact of history that legislative attempts at health care reform that began with the inauguration of the Clinton Administration in 1992 failed. The reasons

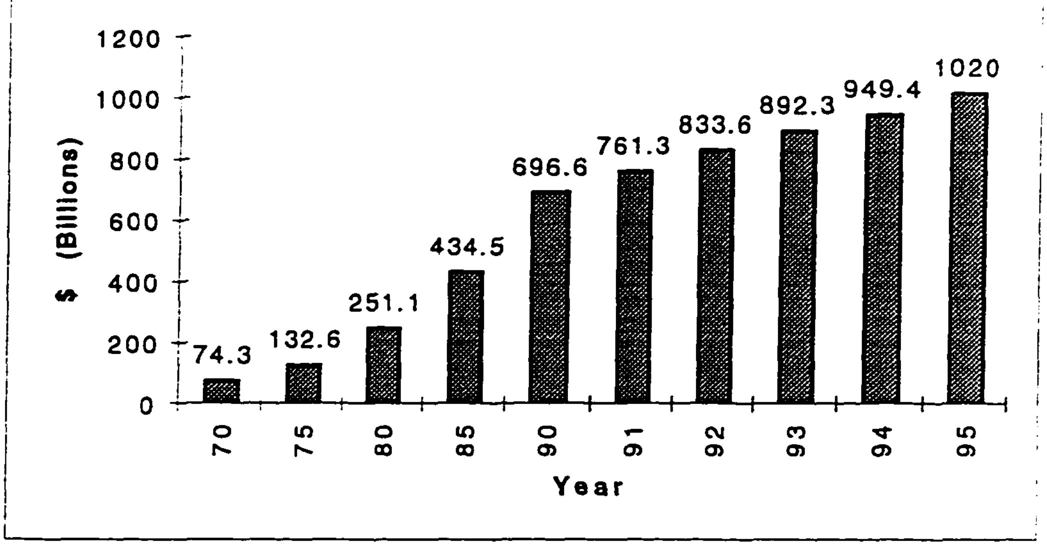

Figure 1.1. National Health Expenditures 1970 - 1995. (Source: Levit *et al.* 1996)

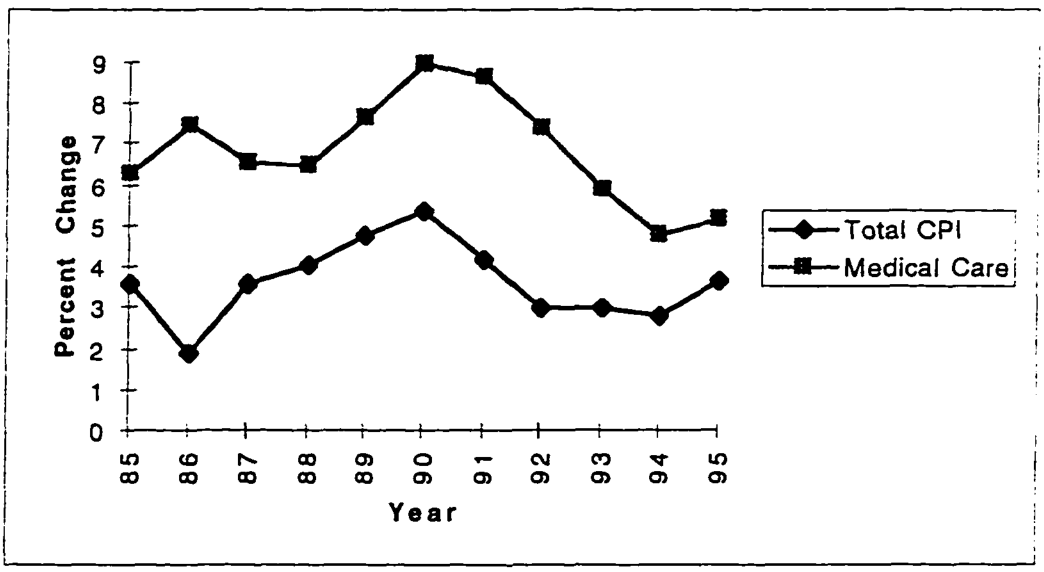

Figure 1.2. Annual Average Percent Change From Previous Year in Medical Care and Consumer Price Index, 1985 - 1995. (Source: Vincenzino, 1996)

for the failure have been noted (Skocpol, 1995) to include, among others, a distrust of governmental involvement in health care, and of consumer concerns about their quality of health care under a reformed system. Stated in different terms, the proposed reform plans failed to pass legislative muster because the plans failed to convince a majority of Americans that their health care, and their cost for health care, would be better under the new plans (Blendon R. J. *et al*, 1995). This is a startling result given that the intent of the legislative proposals was to control costs and improve the health care arrangements for a majority of Americans.

The demise of legislative efforts to reform health care did not, however, mean the end of evolution of the health care system. Indeed, without and independent of legislative efforts, a quiet revolution has been unfolding in the delivery of health services in the United States, causing the demise of traditional medical provision and leading to the rise of managed care.

1.2. Managed Care

Traditionally, health care in the United States has been delivered under the fee-for-service paradigm. The fee-for-service system is characterized by three distinct components: the insurers, the providers, and the consumers (the patients). The nature of the interactions between the three components has meant the cost unconscious delivery of health care (Enthoven, 1985), and was one of the most important components in the rapid rise in health care costs that began in the early 80's and

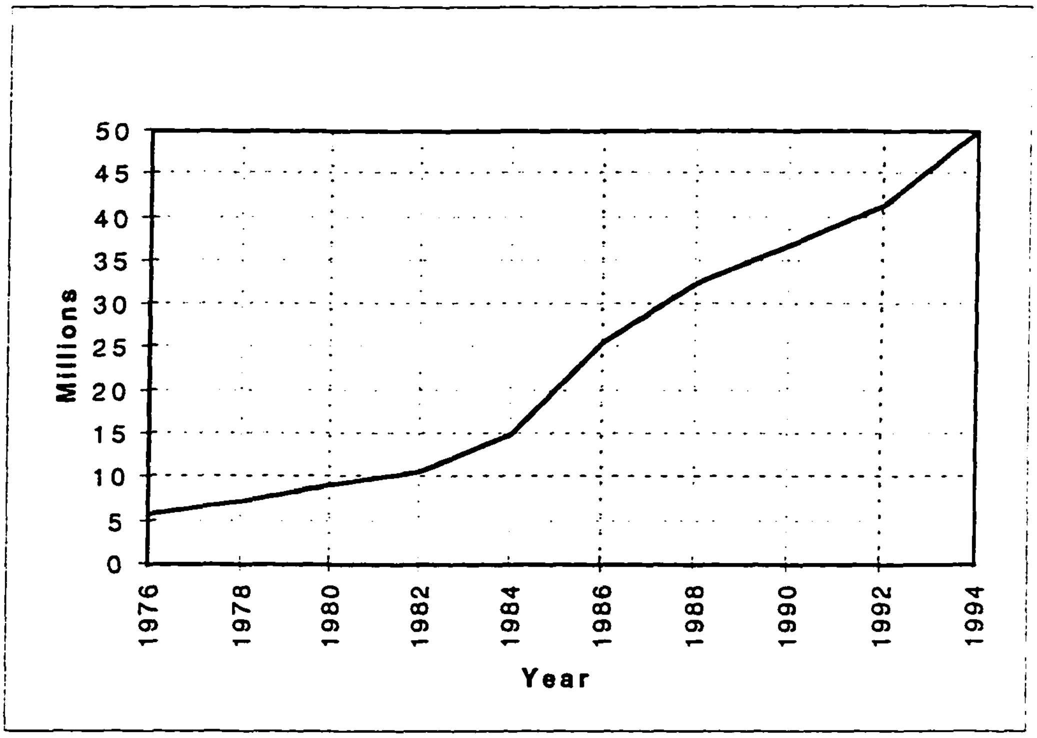

Figure 1.3. Health Maintenance Organization (HMO) Enrollment, 1976 to 1994

continued for more than a decade. The market place response (where the market is typified by corporations and other purchasers of medical insurance) to the rapid rise in costs has been to move away from the traditional fee-for-service health care delivery systems, and towards cheaper managed care programs. Consequently, managed care organizations have seen their share of the insurance market rise from 4.7% in 1980 to 20% in 1994 (enrolling almost 50 million people) with growth continuing at a significant rate (Vincenzino, 1995).

Managed care organizations (hereby referred to as MCO's) are characterized by operation under a capitated budget; as MCO's rise in dominance the United States health care system will move increasingly to a system of health care delivery operating under budget constraints. Thereby, it is important to ask how policy formulation should be cast within the new paradigm of health care delivery to achieve societal goals. In addition to budget constraints, public policies formulated under managed care plans must also consider the effect of chosen polices on consumer service quality. For example, part of the resistance to adopting a Canadian style single payer health care system in the United States was due to the perceived negative effect on service quality (Danzon, 1992). In addition, the need to consider service quality follows from the fact that managed care plans have come under considerable criticism for perceived (or real) poor quality of service (Coughlin, 1993; Hillman, 1987). Subsequently, any proposed policy must account for its effect upon service quality: For any pair of policies, the policy that delivers the higher level of service quality is more likely to be accepted.

The rising influence, and possible future dominance, of managed care also presents new opportunities for the application of decision analytic techniques to the internal management of health maintenance organizations. In particular, in this dissertation, we are interested in examining the issue of physician compensation.

1.3. Research Agenda and Dissertation Overview

The research presented in this dissertation may be considered as being constituted of two different but related streams:

I. The Quantitative Relationships Among Cost, Quality, and Access

The premise that forms the foundation for the research in this dissertation is the observation that the desirable objective of any policy reform in the health care system is to provide universal access to high quality medical care, and achieve these goals at low monetary cost. Further, the effectiveness of any proposed health care reform plan can be gauged by its effect on these three dimensions.

In this dissertation, the first part of the research agenda concerns itself with developing the quantitative relationships among cost, quality, and access in the health care system. The motivation for developing these relationships follows from the observation that the effect of a proposed health care policy which aims to affect one or more of the dimensions cost, quality, or access will be difficult to gauge unless a clear quantitative understanding exists of the relationships among the three dimensions. Consequently, the quantitative relationships among cost, quality, and access (which will henceforth be called CQA) form the basis for policy formulation and management of the health care system.

Our purpose, in the first part of the research agenda, is to take the first steps towards obtaining this quantitative relationship and thereby provide a unifying foundation for the strategic analysis of the health care system. For example, if a policy is advocated to increase access by a certain percentage, then a quantitative understanding of CQA will indicate commensurate changes in cost and quality (positive or negative) in the system.

Different policies and managerial paradigms will be explored as part of a strategic analysis, and it will be demonstrated the method by which CQA can serve as an aid to decision makers in the adoption of policies. In particular, we will explore the hypothesis that the goals of cost control, increased quality, and universal access are simultaneously realizable. We demonstrate, within the assumptions of our model, that the hypothesis is realizable by specific resource allocation policies and for diseases of specific characteristics. We will quantify these resource allocations and search among them for the optimal allocation scheme. The resource allocation policies to demonstrate the hypothesis are developed both in steady state and via a simulation model. The purpose of the simulation model is to study the cash flows and pay-back periods when an investment in screening is undertaken.

In this dissertation, we will develop the relationship CQA for a *single* chronic disease, using breast cancer as a case study. While it is true that a complete analysis of the health care system must incorporate the study of multiple diseases and their possible relationships, this is a very analytically difficult and complex task. We believe that the study of a single disease is a necessary first step in the development of CQA that incorporates diseases in their totality. The single disease study will serve as a "demonstration of concept" and provide the framework for further analysis with multiple diseases.

The development of the relationships CQA will be as follows: In Chapter 2, we define the quantities cost, quality, and access, and highlight their importance and relationship to the overall health care system. In Chapter 3, we provide an overview of breast cancer, the disease we have chosen to illustrate our research, and its relationship to the American health care system. Chapter 3 also provides the necessary grounding for the health care system representation we present in Chapter 4. The representation details our abstraction of the health care system into a form suitable for modeling. Chapter 5 presents the quantification of the relationships among cost, quality, and access. In Chapter 6 we apply the derived relationships CQA to explore the hypothesis that cost control, increased quality, and universal access are simultaneously realizable for diseases of specific characteristics. Further, we use CQA to illustrate the effects of different possible health care policies and managerial paradigms as part of a strategic analysis, thereby demonstrating the efficacy of the relationships CQA as an aid to decision makers in the adoption of policies. In Chapter 6 we also run the relationships CQA using parameters for cervical cancer.

II. Resource Allocation in Managed Care Systems

In the second part of the research agenda, we consider, *for a single disease*, the fundamental question in the management of capitated budgets systems: If resources are fixed, how should they be optimally allocated across the health care system to create the most desired outcomes? Desired outcomes, in this research, include maximizing the health of the population, and maximizing the service quality available at the health care facilities. In addition, we are interested in understanding the dynamics of a capitated health care system as a function of access level. We ask the following questions: What are the population health outcomes as a function of increased access? Will increased access mean that more people may be treated, but the intensity of treatment for all the population will decline such that no net gain in the health of the population is obtained? What are the service quality implications of increased access? Will increased access necessitate a decrease in service quality? Will increased access lead to a shift in resources within the health care system? A resource shift may have implications for the distribution of the nature of specialists and equipment. For example, in the case of breast cancer, a greater emphasis on preventative services may require a greater number of radiologists but fewer oncologists, and greater expenditure on magnetic resonance imaging machines and less on chemotherapy, surgery, and radiation equipment.

In exploring the dynamics of a capitated health care system as a function of access, we are primarily concerned with the problem of the medically uninsured. The uninsured presently number approximately 38.5 million, or 17.4% of the non-elderly population (Rowland *et al*, 1994). Health insurance is obtained primarily as an employment based benefit, but is often not available to employees of benefit poor employers such as service industries and small businesses (Levit *et al*, 1992). With an increasing proportion of the United States economy constituted by the service sector, the numbers of medically uninsured are likely to increase. The problem of the uninsured remains intractable -- there does not presently exist a macro-economic or political solution to the problem of the uninsured. We ask, however, if the problem cannot be solved within the health care industry. In particular, we consider whether fixed resources can be allocated within the health care system such that access is increased, without sacrificing service quality.

In summary, in the second part of our research agenda, our purpose is to develop managerial policies that will allow the achievement of the following four goals: increasing access to the health care system, containing costs, increasing the quality of health of the population, and increasing health service quality. The primary importance of these four goals to a health care system have been established in the first part of our research agenda. However, we make the additional statement that the first three goals may be considered as societal level goals, and the last as important to determine the acceptability of the proposed managerial policies.

Our method, in the second part of the dissertation, is to develop, for an age group, separate models to optimally allocate resources to maximize the population's quality of health, and to maximize service quality at the facilities. Further, we develop models to optimally allocate resources across age groups. We develop the optimization models with respect to the access level of the population, enabling us to explore the dynamics of the health care system with respect to increased access. If both societal quality goals and service quality goals are in synch with respect to access, our analysis ends -- both goals will be simultaneously achievable. If, on the other hand, national health and service quality diverge with access, it is necessary to develop to trade off curves between the two types of quality. In sum, understanding the system dynamics as a function of access will allow us to make policy recommendations for the management of capitated budget health care systems.

We will develop the managed care resource allocation models as follows: In Chapter 7, based on the health care system representation of Chapter 4, we present models for the optimal distribution of resources to achieve the maximization of societal health. In Chapter 8, our purpose is to develop models for the maximization of age group specific service quality, and we present optimization models with this aim. Chapter 9 explores the age specific system dynamics as a function of access, and the resulting managerial implications. Further, in Chapter 9 we complete the single disease optimal allocation of resources across age groups.

1.4. Literature Review

The research proposed in this study carries with it several interwoven themes. some of which have been addressed in the literature and are reviewed here. We present the literature review by examining in order the contributions to: i. the theory of mass screening, and ii. the analysis and effectiveness of health care systems

I. Mass Screening

Mass screening and preventative services have received considerable attention from researchers in the decision analytic, mathematics and public health communities. Zelen and Feinleib (1969) presented one of the earlier papers on disease screening. Their model presents a method for obtaining the lead time (lead time is defined as the difference in time between the detection of a disease by screening measures and the detection of diseases without screening) from a screening procedure by noting the incidence rate, the prevalence rate, and the average duration of the illness.

The design of a screening procedure to obtain desired lead times was presented in papers by Kirch and Klein (1974) and Prorok (1976(a) and (b)), among others. Kirch and Klein (1974) present one of the most often referenced works in mass screening. Their paper is concerned primarily with developing an optimal schedule for screening examinations of age dependent diseases. The formulation is developed by minimizing the detection delay for a disease subject to a constraint on the number of screening exams. Prorok (1976 (a) and 1976 (b)) developed a theoretical framework for the determination of lead time, and the proportion detected. The design of a screening procedure requires, however, knowledge of the sojourn time of the disease in its undetected state, and knowledge of the probability that a screening test will detect a disease in its preclinical state.

The economic aspects of mass screening were considered by Schweitzer (1974) and Eddy (1983). Schweitzer (1974) presented a framework for the cost effective evaluation of mass screening tests. The economic costs of the tests are based, aside from the cost of the tests themselves, on disease incidence, probabilities of test error, and the treatment for found cases. Benefits are obtained from economic value of additional life years for those cured of the disease. Eddy (1983) developed a model for the timing of repeated medical tests taking into account the explicit costs of the medical examinations, and economic benefits of a program of mass screening. The model yields a frequency for monitoring and the probability of detecting a disease as a function of the disease incidence rate.

Models for the development of breast cancer have been developed by Schwartz (1978) and Eddy (1985). The paper by Schwartz (1978) develops a model that is based on parameter estimation based on published data. In an authoritative work, Eddy (1985) developed a comprehensive theory for the screening of non-contagious diseases (e.g. cancer) that explicitly considers the etiology of disease development in the human body. This work interweaves the etiology of the disease with prognosis for the disease and the reliability of tests used to detect the disease and derives expressions for the true positive rate, and true negative rate for different screening strategies. A paper by Voelker and Pierskalla (1976) examines test selection for a mass screening program, an important issue when there are several available tests that differ in their efficacy and cost. Butler, Furnival, and Hart (1993) presented a paper that uses the theory of multiproduct cost functions to estimate the treatment cost function for a disease that advances through progressively worsening stages. This is a useful paper in that it provides parameter values for use in this dissertation.

The primary point of separation between the work in this dissertation and previous work on screening is that we consider the implications of screening in a system wide context. Thereby, decisions that are made with respect to the allocation of resources to screening are made to strive for a global optimum. If decisions on screening are made independent of the effect of screening on the rest of medical care, the optimal decisions that are reached are, by necessity, local optimums. It is not necessarily true that the global and local optima will be the same. Additionally, the difference between the literature and the models presented in this dissertation is the following: the literature is primarily concerned with the operations of a screening program (e.g. the frequency of screening tests) while this dissertation has, among its goals, the optimal allocation of resources to a screening or prevention program.

II. Analysis and Resource Allocation in Health Care Systems

Smallwood, Sondik, and Offensend (1970) presented a paper that largely qualitatively describes the need to consider health care at the system level when formulating decisions. The idea of a system wide view is similar to the work that has been undertaken in this dissertation. The difference is that the work in this dissertation has attempted to bring a level of quantification to the problem of making system wide decisions in medical care. Packer and Shellard (1970) provide a method of evaluating the effectiveness of a health care system. Their method is based on evaluating an individual's state of health over the part of his life or her during which the health care system is in effect. While their method captures health care quality for an individual, it does not speak to the costs, or to the level of access accorded to the society in which the health care system exists. In this dissertation, it is contented that it is important to view, in addition to the quality of the health care system, the overall costs of and the level of access to the system.

There have been a significant number of works presented in addressing the problem of resource allocation and planning in the health sector. Gillis (1992) and Eyles and Birch (1993) examines frameworks for health care resource allocation. The aims of the work by Gillis and by Eyles and Birch and this work are similar, with the difference being that the work by Gillis and by Eyles and Birch are qualitative. Jonsson (1994) proposes the use of clinical trials to study the safety and efficacy of therapeutic interventions. While Jonsson examines only economic evaluations of three therapeutic areas in cardiology, we believe that an extension to other diseases will provide useful information on parameters for use in the models proposed in this dissertation. Robinson (1993) reviews the use of cost utility analysis in the allocation of resources. Cost utility analysis compares the outcomes of different procedures using a measure called QALY. QALY is a measure of quality adjusted life years. In this dissertation, we do not adopt the use of QALY for the primary reason that the measure does not reflect the economic return to the health care providers of any policy measures they adopt. Our purpose in this dissertation is to facilitate the adoption of policies that will prove economically viable for health care providers.

A paper by Murray, Kreuser, and Whang (1994) presents an encompassing study of the cost effectiveness of various investments in the allocation of health care resources. The objectives of their work and the research presented in this dissertation is the same. However, the primary focus of the work by Murray et al is to examine the allocation of resources to improving of physical or human infrastructure in the health care system. The outcome of their research is to provide a method for the direct comparison of an investment in the delivery system or an investment in the purchase of more specific interventions in the health care system. In this dissertation, our focus is limited to examining a disease management approach to the allocation of resources. The possible relationship between the work of Murray et al and this work is that our research follows in the hierarchy of decisions in the allocation of health care resources. That is, once the infrastructure of medical has been constructed, what should then be the allocation of health care resources? In this dissertation we propose a disease management approach to the allocation of resources after the infrastructure is constructed.

1.5. Summary of Key Contributions

The main contributions of this dissertation are:

- The development of an intuitive methodology that takes the first steps in quantitatively relating the three important and defining measures in a health care system: Cost, quality, and access.

- The development of the relationships among the dimensions cost, quality, and access (CQA) allows for the measure, within modeling assumptions, of the change in one or more of the dimensions if the remaining dimension, or dimensions, are altered. The relationships CQA will form a quantitative foundation for the analysis and/or comparison of national policies in health care delivery.

- We demonstrate the efficacy of CQA as a policy tool by showing, within the limitations of the model, that it is possible by a multiple of resource allocation policies in a policy set to achieve universal access without sacrificing quality, or increasing cost. Further, we provide a method to distinguish the supremum policy from within the policy set.

- We consider one of the fundamental questions in the management of capitated budget systems -- that of resource allocation. We develop a three tiered optimization system for resource allocation which answers the following questions:

- What is the optimal division of resources as a function of age? The importance of this problem follows from the age specific characteristics of many chronic diseases.
- What is the optimal division of resources among the different stages of a disease that is characterized by progressively more severe stages of the disease?
- What is the optimal division of resources between the treatment of a disease and prevention and mass screening programs?
- In a mass screening program, what is the optimal division of resources between those allocated to the detection of persons who may be ill with a disease, and resources devoted to the early treatment of persons detected ill?

Chapter 2

A Primer on Cost, Quality, and Access in the Health Care System

2.1. Introduction

In this chapter, we present a short exposition on cost, quality, and access. Our purpose is three fold: The first is to illustrate the prime importance of these three quantities in any strategic analysis of the health care system. The second purpose is to highlight the interdependencies among cost, quality, and access. The third purpose is to describe the *status quo* in the present health care system with respect to these quantities.

2.2. Cost

Costs, as defined by resources expended to treat illnesses, are compensated for by a number of means in the American medical system. The breakdown for 1993 is the following: 34% via private insurance, 18% paid out of pocket, 43% paid by government programs (Medicare, Medicaid, etc.) and 5% other (Levit *et al.*, 1994). Out of pocket expenses are generally used to meet the costs of over-the-counter medications, insurance deductibles, and co-payments. The primacy of insurance payments (private and government) as a means of covering health care costs reflects the often considerable resources required to treat ailments. For example, coronary artery bypass graft operations cost, in 1992, up to $61,990 at some hospitals in the U. S. (Statistical Bulletin, 1994).

For most Americans, insurance is obtained as a employment based benefit. However this is a benefit that has come increasingly under siege as health care costs continue to rise. For some corporations, providing health insurance as a benefit is a large part of their operating budgets, and has significant implications for their global competitiveness. For example, the price of every General Motors car sold in 1992 carried with it $1,100 in cost of medical benefits for GM workers. In comparison, the price of Toyota cars included only about $550 in the cost of medical benefits in the same time period (The White House Domestic Policy Council, 1993). Insurance premiums continue to grow for companies in the U. S., with the growth reaching 20 and 30 percent in some recent years (Levit and Cowan, 1990). In response to rising health related costs, corporations have reduced their benefit packages, or done away with the provision of health insurance altogether. This is especially true in service based industries (Levit *et al.*, 1992). The implications of rising health insurance premiums and the reduction in employer based health coverage benefits are very significant for the issue of access.

2.2. Access

Access to health care is defined as possessing the ready means to pay for health care without significant financial hardship. Because of the importance of insurance as a method of covering moderate to large medical expenses, access is almost universally synonymous with having insurance. However, insurance coverage may be limited by exclusionary clauses for pre-existing conditions. This defines partial access. There are approximately 78 million Americans without full access to the medical system: *38 million* without access, and *40 million* with only partial access (Starr, 1994). This large number of medically uninsured is considered a serious social problem.

The uninsured face significant financial barriers to obtaining care. Consequently, they are more than twice as likely as the insured to be without care (Congressional Research Service, 1988; also see Hafner-Eaton, 1993). It is likely that the uninsured seek care only when compelled by the severity of their condition. Treatment outcomes, thus, are poorer than they would have been if medical attention had been sought earlier. The uninsured do have some options open to them with regard to obtaining medical attention. Many uninsured can use state and county funded public hospitals and clinics. In addition, private hospitals are required by law not to turn away patients with immediate life threatening ailments (for example, massive cardiac arrests), insured or not, from their emergency rooms and trauma centers. However, private hospitals may transfer emergency patients to other facilities (e.g. county or local hospitals) after the patients have stabilized. Private hospitals are under no legal obligations to treat non-emergency patients who lack the means to pay for their care.

Social problems associated with access are likely to worsen due to a twin set of factors. The first is that since the early 1980's the numbers of uninsured has been increasing. The other factor is that rising health care costs have made it more difficult to lend charitable support to the uninsured, be it through public programs or uncompensated care provided by private hospitals.

2.3. Quality

In this dissertation quality is defined at two levels. The first is to define the quality of a health care system by the aggregate measures of health of the population. This measure of quality will be called "national health quality," and the relationships among cost, access and national health quality will be denoted by CQA(NHQ). The second definition of quality, which will be called "service" quality and denoted by CQA(SQ), stems from both the level of care provided by medical personnel to patients, and from patient perceptions of their quality of care. Thus, components of service quality include the state of the technology used in curative procedures, the intensity of treatment, and the amount of time spent with a doctor on an office visit. Anecdotal evidence suggests that service quality is directly correlated to resource expenditure: More expensive fee-for-service physicians perform more services and spend more time developing a rapport with their patients. HMO's which have capitated budgets are often faulted by enrollees for very quick doctor visits, and for often long waiting times for appointments (see "Cutting Medical Costs - or Corners?"*Los Angeles Times*, May 5, 1995; and Coughlin, 1993).

It is perceived changes in service quality that plays the greatest part in consumer acceptance of new policies or reform measures. For example, it has been suggested (see Woolhandler and Himmelstein, 1991) that the US can mitigate its health care problems without sacrificing quality by emulating the Canadian style single payer health care system since they spend less, but enjoy higher aggregate measures of health: In 1991 per-capita health care expenditures in the United States and Canada were $2,868 and $1,915 respectively. In contrast, the life expectancy of the nations was 72 and 73.8 for men, and 78.8 and 80.4 for women respectively. Further, the infant mortality rates (per 1000 live births) were 9.1 and 6.8, respectively (Schieber *et. al.*, 1993). In the Canadian system, provincial governments act as the sole payers in their respective systems. This provides for considerable administrative savings, and gives the provincial governments strong negotiating, and hence cost containment, powers when setting prices with private physicians. However, evidence of considerable waiting times for elective (i.e. for non life threatening ailments such as hip replacement) surgeries has been documented (Danzon 1992) and has driven consumer resistance to the adoption of the Canadian style system, especially since these waiting times are not present in the US system. (It has been suggested (Reagan, 1992) that, in essence, the U. S. rations care by denying many of its citizens access to health care, while other industrialized nations, such as Canada, ration care by decreasing the quality of care, or by increasing the waiting times for procedures.)

In summary, the challenge facing decision makers who attempt to provide answers to the US health care conundrum is to provide solutions which preserve that which is best about the *status quo* (the cutting edge technology, high level of care, and short waiting times, to list a few), while containing costs and attempting to achieve universal access. The methodology we present in this dissertation, by quantifying the relationships among cost, quality, and access, will provide decision makers with a tool by which to compare policy choices, consequently arriving at a policy that best meets societal goals.

Chapter 3

A Case Study: Breast Cancer

3.1. Introduction

In our research, breast cancer serves as the illustrative disease of choice. Breast cancer is characterized by a relatively long asymptomatic phase where early detection of the disease is possible and lends itself to superior, both in terms of cost and outcomes, treatments over measures taken later in the disease's natural history. It thus serves as a natural candidate for the models and managerial paradigms presented in this dissertation. Breast cancer is also an important disease. The disease strikes one of nine women in their lifetime; it is the second leading cause, after lung cancer, of cancer related mortality in women: In 1993, breast cancer occurred in 182,000 women and caused approximately 46,000 deaths (Sondik 1994), many of which could have been prevented with the early detection of the disease. Further, breast cancer is representative of a class of diseases that are similarly characterized as having clinical courses that are progressive and chronic. Examples include other cancers, heart disease and diabetes -- diseases that are some of the leading causes of mortality. Thus, managerial insight gained from the study of breast cancer will extend to other important diseases, and provide considerable additional benefit.

We introduce our case study of breast cancer by first presenting the natural history, and epidemiology of breast cancer. Following the description of the disease's pathology, we describe treatment methods, including the role of screening, with respective costs and outcomes. Finally, we describe the interactions breast cancer patients have with the health care system. This treatment will lay the foundation for understanding the health care system representation we present in Chapter 4.

3.2. Natural History and Epidemiology of Breast Cancer

The details on the natural history and epidemiology of breast cancer can be obtained elsewhere (see Miller *et al.*, 1994; and Henney and DeVita, 1987). A brief overview, however, is the following: Breast cancer is a disease that primarily afflicts women. There is a strong age dependency in the incidence rates of breast cancer. Breast cancer is almost non existent in females under the age of 20. After age 20 there is a rising incidence rate that peaks at age 57 and then falls. Breast cancer begins, in most cases, in the mammary ducts. The disease develops at rates that differ among women but that follows a similar path of development. From the mammary ducts the cancer may spread to the surrounding tissue. At this stage of development, the cancer is said to be *localized*. From the localized stage, the cancer cells can spread to the regional lymph nodes, and the breast cancer is said to have *regional* spread. In the most serious cases of breast cancer, the cancer mestatasizes and forms distant colonies in other organs such as the lungs and the liver, and in the bones. At this stage the disease is said to be *distant*.

Figure 3.1 (Adapted from Love 1991, and Morrison 1992) schematizes the natural history of breast cancer. The biological onset of the disease is at a time denoted by point O. The disease becomes detectable at point D_M. However, the cancerous lesion is small enough that the disease can only technologically be detected by mammography. D_P is representative of a point where the disease advancement is such that detection is additionally possible by physical exam from an experienced practitioner. D_S represents disease advancement to the point where self examination will make the cancer evident.

Up to the point where the disease symptoms and signs are not evident to the medically untrained patient, the disease is referred to as preclinical. The point of evidence, without screening measures, of the cancer is known as the clinical surfacing of the disease. This point is represented by C. Clinical surfacing of the disease may occur at either the localized, regional, or distant phase of the cancer, depending upon

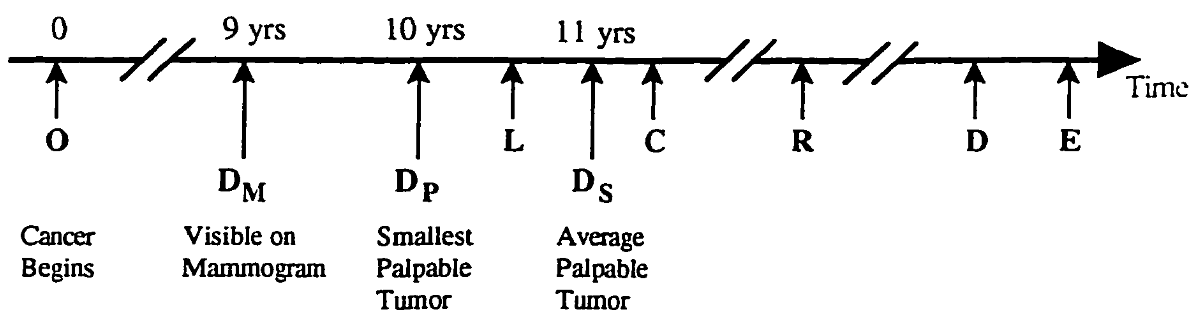

Figure 3.1: Natural History of Breast Cancer

the sensitivity of the patient to symptoms and on the rapidity of disease development. The localized, regional, and distant phases of the disease are denoted respectively by L, R, and D. Finally, death due to breast cancer is denoted by E, on the time line.

3.3. Treatment Methods, Costs and Outcomes

Treatment methods and options are a function of the stage of the breast cancer when presented for treatment. A complete discussion can be obtained from (Henderson *et al.*, 1989). Briefly, early detection of breast cancer allows for the use of the least invasive methods including lumpectomy (the removal of the cancerous lump and immediately surrounding tissue) followed by radiation therapy. Simple mastectomy (removal of the breast and immediately adjacent lymph nodes), radical mastectomy (removal of the breast with the dissection carried up the armpit to remove the axiallary lymph nodes and the pectoralis muscle underlying the breast) and modified radical mastectomy (radical mastectomy with preservation of the pectoralis muscle) are surgical methods used for more advanced localized and regional cancer. The surgical procedures are supplemented with radiation therapy, chemotherapy (anti-cancer drugs) and endocrine therapy. Distant cancer is incurable, more often than not. However, surgical and systemic treatments are applied as treatment measures to prolong survival or palliate symptoms.

Standard measures of treatment success are defined as being disease free for multiple periods of five years after treatment. "Cure" is defined as returning to the same life expectancy as would have been present without the advent of the disease.

3.4. Screening for Cancer

The benefits of treating breast cancer as early as possible, and especially in its pre-clinical phase, have been confirmed by numerous studies (see Section 1 in Miller *et al.*, 1991). However, the detection of pre-clinical cancer carries with it a cost. Further, the sensitivity (true positive detection rate) and specificity (true negative detection rate) are cost proportional: Mammography, which is the most sensitive of all test methods is the most expensive. Professional physical exams are considerably less expensive, self exams even less so. However, the sensitivity and specificity of these tests, on average, have correspondingly lower efficacies of detection. Multiple, or combinations of, tests can be performed, at higher cost, but with possibly greater accuracy of state detection than any one individual test. However, the effect will saturate with oft repeated or combination tests providing no gain in detection probability (Eddy, 1983).

3.5. Cancer Patient Interactions with the Health Care System

It is well documented that insurance status is a strong determinant in the level of interaction of breast cancer patients with the health care system (see Ayanian *et al.*, 1993). Insured patients have greater suggestions from their primary care physicians to undertake screening measures. If asymptomatic breast cancer is detected, the disease is immediately treated. For the uninsured, often little encouragement is offered to undergo screening procedures, especially in light of their inaccessibility (Zapka *et al.*, 1989, Fox *et al.*, 1991). Further, the uninsured are likely to forgo treatment for early stage cancer especially since mild symptomatic conditions are often ignored until the condition advances to the point where symptoms produce sufficient discomfort or a life threatening situation arises. Consequently, uninsured patients have been observed, statistically, to obtain treatment later in the development of breast cancer than insured patients, and more likely present after the disease has clinically surfaced or when the disease is in its most advanced stage (Ayanian *et al.*, 1993; Leitch and Garvey 1994). Uninsured patients with breast cancer, therefore, are more likely to require care in intensive care units at high cost (see, for example, Schapira *et al.*, 1993, for the ICU costs of treating cancer).

The case study of breast cancer has served to illustrate the age, severity of disease, and insurance status dependencies of breast cancer patient interactions with the medical system. In the next chapter, we present a representation of the health care system which, along with the characteristics of breast cancer patients and their interactions with the medical system, will form the basis for the models we present in this dissertation.

Chapter 4

Health Care System Representation

4.1. Introduction

This chapter presents a health care system representation that captures the movements of breast cancer patients within the health care system, in accord with the severity of their disease, their insurance status, and the quality of treatment available at the medical centers. The representation captures the effect of different policies by noting their effects on the transitions of patients in the health care system, and noting the resulting desirability of the patient movements. The representation consists of a network of interconnected components: The components represent medical facilities, and the interconnections serve to represent the flow of individuals to, from, and between, the components.

The schematic representation is provided in Figure 4.1. The components of the representation are obtained by considering the provision of care to breast cancer patients and identifying only the relevant components of the medical system charged with treating breast cancer. These components include screening services, treatment services for preclinical local stage, clinically surfaced local stage, regional stage, and distant stage breast cancer. While it is true that these services may be provided at the same medical facility (for example, at a hospital), the distinction is made as if screening services, and the treatment of the different stages of cancer are provided at different facilities. The validity of the approach follows from noting that the prime determinant in the course of action a physician or surgeon embarks upon in treating the disease is the stage at which the cancer is presented for medical attention. Consequently, the treatment of the different breast cancer stages can be considered separable.

A further distinction is made among the components according to their accessibility, by insurance status, to breast cancer patients. Thus, a distinction is made between private sector health care facilities and government facilities. The different components are defined as:

1. **RMS:** This node represents the aggregation of all the private sector medical facilities such as ambulatory centers, physicians offices, clinics, and hospitals that are accessible to the medically insured. In the representation, this node serves to treat clinically surfaced local stage, and regional stage breast cancer.

2. **GMS:** This node is the public sector counterpart to the node RMS. It includes all municipal, county, state and federal medical facilities that are available to uninsured and indigent populations.

3. **EMS:** This node captures the aggregate provision of medical service for distant stage breast cancer in the health care system. It is the most expensive form of treatment available because of the intensity of care required to treat the very advanced stage of the disease. No distinction is made between private and public sector emergency facilities because treatment is accessible to all, regardless of insurance status.

4. **EDC:** This node is a private sector facility and reflects the provision of screening for assymptomatic breast cancer. The screening is accomplished either by clinical breast exams, or by more specific diagnostic tests such as mammography.

5. **ETS:** This node serves as a private sector facility dedicated to the early treatment of preclinical breast cancer uncovered at node EDC.

6. **T**: This node is an abstraction in the model. It serves as the gateway into the medical system and further has the role of guiding people to different parts of the system dependent upon their needs. An assessment is made of the health and insurance status of individuals presented to this node. Those who are evidently ill and who can afford, or are insured for, medical services are directed to seek treatment at regular health centers. Those who are ill, but without the means to access regular medical care will either remain outside of the medical system, or will seek services at government facilities and emergency centers. The emergency centers also draw insured persons from node T who require emergency care. There will also be persons who are completely well, and the asymptomatically ill. At this node these individuals will either remain outside the medical system, or will be directed to obtain screening services. Finally, those individuals who have passed on will make the transition from node T to the exit node, Z.

7. **POPULATION**: This node represents the population of individuals who have either remained outside the health care system with respect to breast cancer or, if they did enter the health care system, the path was through node EDC with a subsequent return to node POPULATION after a negative screening result.

8. **C**: In the representation, this node collects previously ill individuals who have transited through the treatment nodes ETS, RMS, GMS, or EMS and have been deemed cured, or disease free.

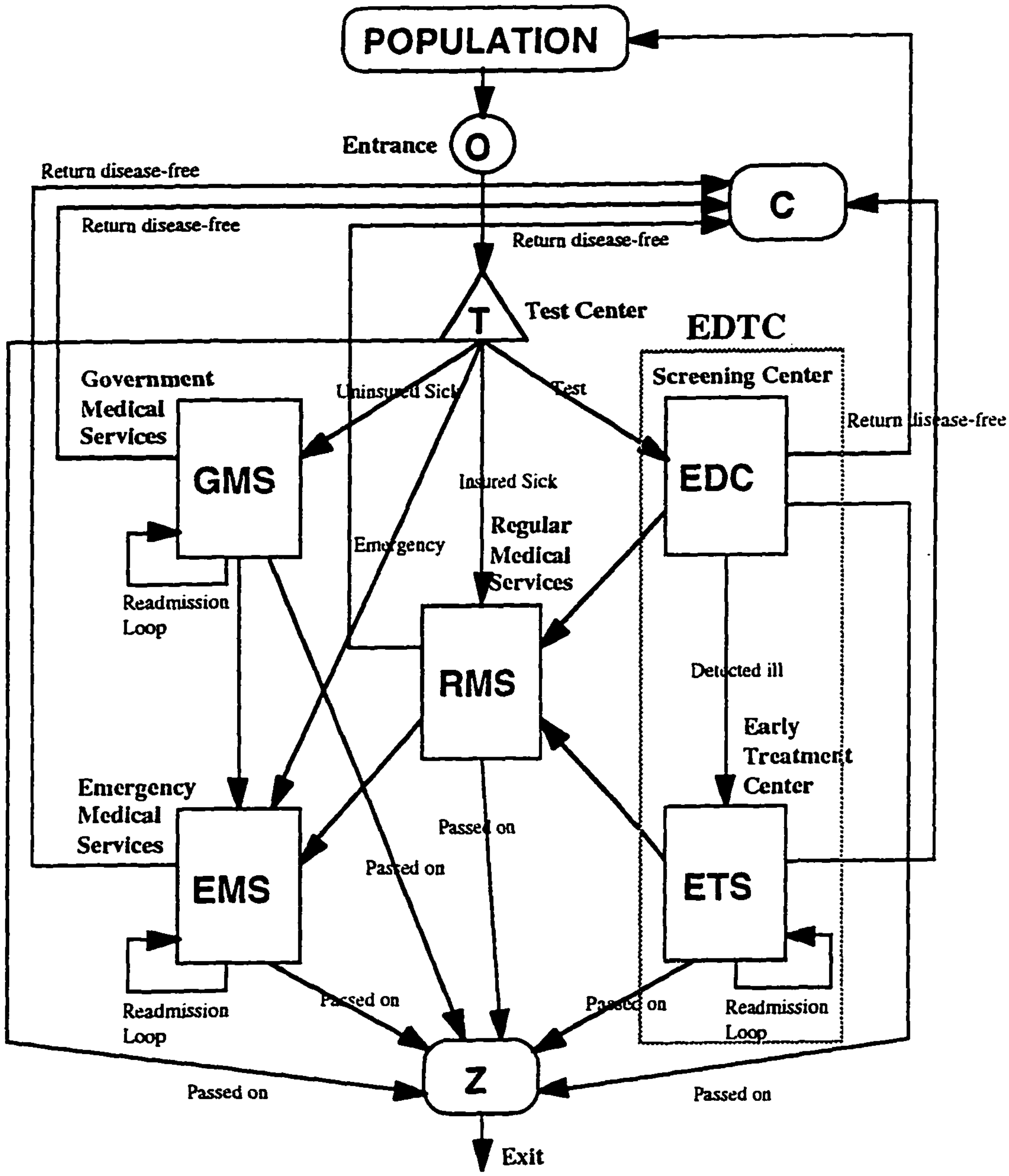

Figure 4.1. A Network Representation (simplified) of the Health Care System

9. **O**: This is the entrance into the system. Entrance may be limited to persons with specific characteristics --such as age -- that are determined by the purpose and scope of the study.

10. **Z**: This is the exit node of the system and indicates those who have passed on.

The interconnections between the nodes in Figure 4.1 represent possible flow paths of individuals within the health care system according to their health status, and access. Flows T→RMS and T→GMS represent medically insured and uninsured patients with clinically surfaced local and regional stage cancer respectively. The flow T→EMS represents the traffic of distant stage individuals, regardless of ability to pay, who seek medical care. The flow T→EDC includes both the medically insured healthy and asymptomatic populations that seek screening services.

The flow of individuals from node EDC are a function of the individual's characteristics (asymptomatic or not, age), the breast cancer stage in the patient, and the sensitivity and specificity of screening tests used at the node. Asymptomatic persons correctly diagnosed with breast cancer will transit either to node ETS, RMS, or EMS depending upon the stage of the detected cancer. Disease free individuals (healthy) will return from node EDC to the population. The transition probabilities out of node EDC will have a strong age dependency that follows the age specific prevalence of breast cancer. For example, the prevalence of breast cancer is lower in women aged thirty than in women aged 55. Screening a given number of thirty year old women will yield fewer women afflicted with breast cancer than would screening the same number of women aged fifty five. Therefore, a larger fraction of 30 year old than 55 year old women can be expected to return to the population from node EDC; similarly a smaller fraction of 30 year old women will transit to treatment facilities.

Transitions of individuals out of the treatment centers ETS, RMS, GMS, and EMS will be a function of the quality of treatment received at these centers. Generally, poorer quality treatment will lead to higher breast cancer recurrence among patients or to a larger number patients dying from the disease, than would higher quality treatment. The quality of treatment provided at the treatment centers will be reflected, then, in the relative rates of flows on the readmission loops, in the transitions out of the treatment centers to the node Z, and in the returns to the node C. Transitions between treatment nodes represent the flow of individuals who were initially diagnosed with a certain stage of breast cancer, treated, but then the cancer recurred at a different stage. For example, the transition RMS→EMS represents the flow of individuals treated at node RMS for local or regional stage breast cancer, but the cancer recurs and presents itself with distant mestatatis. The transition ETS→RMS is similarly defined.

Chapter 5

The Mathematical Relationships CQA

5.1. Introduction

In this chapter, we utilize the natural history of breast cancer discussed in Chapter 3, and the health care representation of Chapter 4 to derive the mathematical relationships among cost, quality, and access. Our approach is to utilize two concurrent multi-class Markovian queueing systems to model, respectively, service quality at health care facilities (SQ), and national health quality (NHQ). Both queueing systems are based upon a common transition matrix, and the purpose of this chapter is to derive this transition matrix.

This chapter will be developed as follows: We introduce first the notation that will define quantities of interest in this and in subsequent chapters. Secondly, we list and justify the assumptions that underlie the models presented henceforth. Thirdly, we derive the transition matrix that is the foundation for the relationships in CQA(NHQ) and CQA(SQ).

5.2. Notation and Modeling Assumptions

I. Product Form Network

We model the health care system as a product form network of Markovian queues in equilibrium, based upon the representation shown in Figure 4.1. The choice of a Markovian system allows for analytical tractability. The assumptions for this modeling choice are the following:

i) A patient's entrance or desire to seek care at a treatment center is not a function of the number of people who may already be seeking cure at the treatment center. This assumption gives that the system is without balking.

ii) There is effectively an infinite capacity buffer at the treatment centers where persons who cannot get immediate treatment can wait outside the system.

iii) Patients do not balk when they are in line for medical service, especially if the condition is chronic and life threatening.

iv) The routing within the jump chain is Markovian, with the present health state of a person encapsulating the patients entire history to that point in time.

v) Arrivals to a particular node can be modeled according to a Poisson distribution. Arrivals to a particular node are independent of all other arrivals within the system. This assumption is justified by noting that for non-contagious diseases such as cancer, whether a person falls ill with the disease is almost entireiy independent of whether anyone else has cancer. The only causal relationship between two persons having cancer may be the hereditary relationship. However, for modeling purposes this effect is assumed away.

vi) Service requirements for patients with cancer are independent of the service requirements for any other patients with cancer.

II. Health Classes

We classify the population of interest according to health status, and assign a respective class to each sub-population. Thus, we establish a class for the healthy sub-population, *h*, and for the following disease states: preclinical local stage, *pcl*, clinically surfaced local stage, *csl*, regional stage, *r*, and distant stage, *d*. These are defined as the disease classes. We establish a class for the previously ill who are successfully treated and cured, *c*. Finally, we establish a class for those who have passed on, *po*..

Further, we will define:

DC: Set of disease classes, i.e., (pcl, csl, r, d) $\in$ **DC**.

HC: Set of all health classes, i.e., (h, pcl, csl, r, d, po) $\in$ **HC**.

III. Age Groups

The population will be further categorized according to age as persons, in the aggregate, exhibit age specific statistical characteristics with regard to the acquisition of disease. For example, it is more likely that a 50 year old woman than a 20 year old woman will be stricken with breast cancer. Further, persons may be grouped into age groups if persons within that age group exhibit identical statistical characteristics.

For modeling purposes, we define:

I: Set of age groups, with $i \in \mathbf{I} = \{1, \ldots, I_{max}\}$

where I_{max} is the number of age groups in the time window of study.

IV. Arrival Rates

In the model we assume that the arrival rates of the different categorized sub-populations are exponentially distributed with the following rates:

λ_i^v: Arrival rate of class v ($v \in$ **HC**), within the i th age group.

γ_i^c: Arrival rate of successfully treated individuals from classes pcl, csl, r, or d.

It will be convenient to define the vector of arrival rates as:

$$\mathbf{A}_i = (\lambda_i^h, \lambda_i^{pcl}, \lambda_i^{csl}, \lambda_i^r, \lambda_i^d, \lambda_i^{po}, \gamma_i^c) \qquad \forall\, i \in \mathbf{I}. \qquad (5.1.1)$$

V. Access, and Compliance

Let:

α: Proportion of population with access to health care, $0 \leq \alpha \leq 1$.

ε: Proportion of entrants with access to, and seeking preventative services; $0 \leq \varepsilon \leq 1$.

κ^{csl}: Proportion of uninsured with clinically surfaced breast cancer who seek medical care at public medical centers; $0 \leq \kappa^{csl} \leq 1$.

κ^r: Proportion of uninsured with regional stage breast cancer who seek medical care at public medical centers; $0 \leq \kappa^r \leq 1$.

VI. Mutations

Mutations occur between classes. These mutations follow natural processes and, to reflect reality, are uni-directional. In particular the mutation probability from class *po* (passed on) to any other class is zero. Similarly, we assign a zero probability of a natural mutation from a sick state to a healthy state. (Natural remission of cancer is possible; however, this effect is small and assumed away.) Any transitions from a worse health state to a better state can be accomplished only by way of medical intervention. We define:

$\delta_i^{a,b}$: Denotes the single period mutation probability from period i to period i+1 between health classes a and b; a,b = (h, pcl, csl, r, d, po).

It is assumed that the mutation rates between disease classes capture the probability of an individual at a pre-specified stage of one level mutating to a pre-specified level of another level. For example, in Figure 3.1, the quantity $\delta_i^{pcl,r}$ captures the mutation rate from point L of breast cancer in its latent stage to point R of breast cancer in its manifest regional phase.

VII. Resource Expenditures and Cure Probabilities

We assume that the probability of cure for an individual who enters one of the treatment nodes ETS, RMS, GMS, and EMS is a function of the per-patient resource expense at that node. We define the per-patient expenditures according to:

R_i^{edc}: Per-patient resource expenditure at the early detection center EDC in the i th age group.

R_i^x: Class x ($x \in$ DC) per-patient resource expenditure in the i th age group. Associated with the per-patient resource expenditures at the detection and treatment centers, we have the following detection sensitivity and probability of cure:

$\Omega(\bullet)$: Detection sensitivity at node EDS, as a function of resources allocated per patient.

$\theta(\bullet)$: Probability of cure for pre-clinical breast cancer patients as a function of per-patient allocation of resources.

$\Xi(\bullet)$: Class 1 detection and removal efficacy as a function of resources allocated to super node EDTS.

$\psi(\bullet)$: Probability of cure for clinically surfaced local stage breast cancer patients as a function of per-patient allocation of resources.

$\Phi(\bullet)$: Probability of cure for regional stage breast cancer patients as a function of per-patient allocation of resources.

$\xi(\bullet)$: Probability of cure for distant stage breast cancer patients as a function of per-patient allocation of resources.

A patient not cured at a treatment center, for example at the nodes RMS or GMS with probability $[1-\Psi(R_i^{csl})]$ for clinical local stage cancer, will be subject to re-entrance into the same node for treatment, or will mutate to a more severe disease state. The mutation to a more severe disease state may include the patient passing on. Once a patient is "cured" at a treatment center, they will no longer be candidates for the disease and will exit the queueing network at node C, as defined in Figure 4.1.

5.3 The Transition Matrix Foundation of CQA

Derivation: The transition matrix that forms the basis of the relationships in CQA is derived as follows: Consider a population of size n at a beginning time of interest. The initial population is introduced into the network of Figure 4.1. At the test node T, and based on statistical evidence, the patients are classified according to their health status, i.e., according to the presence or absence of cancer and, if present, the severity of the cancer. The appropriate incidence rates through node T are noted, and the following state vector for the arrival rates at the beginning of the initial period is obtained:

$$\mathbf{A}_0 = \left(\lambda_0^{\mathrm{h}}, \lambda_0^{pcl}, \lambda_0^{\mathrm{csl}}, \lambda_0^{\mathrm{r}}, \lambda_0^{\mathrm{d}}, \lambda_0^{\mathrm{po}}, \gamma_0^{\mathrm{c}} = 0\right) \tag{5.3.1}$$

where the index 0 denotes the initial period of interest and the arrival rate γ_0^c is assigned the value 0 (as no treatment measures have been undertaken by the beginning of the first period of study). The evolution of the initial vector is observed through the I_{max} periods of the study according to transitions between treatment nodes in Figure 4.1 and the transition scheme between health classes depicted in Figure 5.1. Healthy individuals may mutate out of the class *h* and into classes *pcl, csl, r, d,* and *po* with age specific probabilities given by $\delta_i^{h,x}$ (x= *pcl, csl, r, d, po*). The mutation probabilities from classes pcl and csl reflect the probability that an individual with untreated local stage cancer will develop cancer in its advanced localized or regional phase. Similarly, the mutation probabilities of individuals out of class r or class d reflect the probability that an untreated or, if treatment measures were attempted, an uncured individual will enter a more advanced disease state. There are no mutations out of class po as this represents an absorbing state.

Individuals can be removed from the four disease classes (*pcl, csl, r,* and *d*) by medical treatment measures. Additionally, screening measures can be applied that may detect asymptomatic latently ill persons in class *pcl*, and persons subsequently treated.

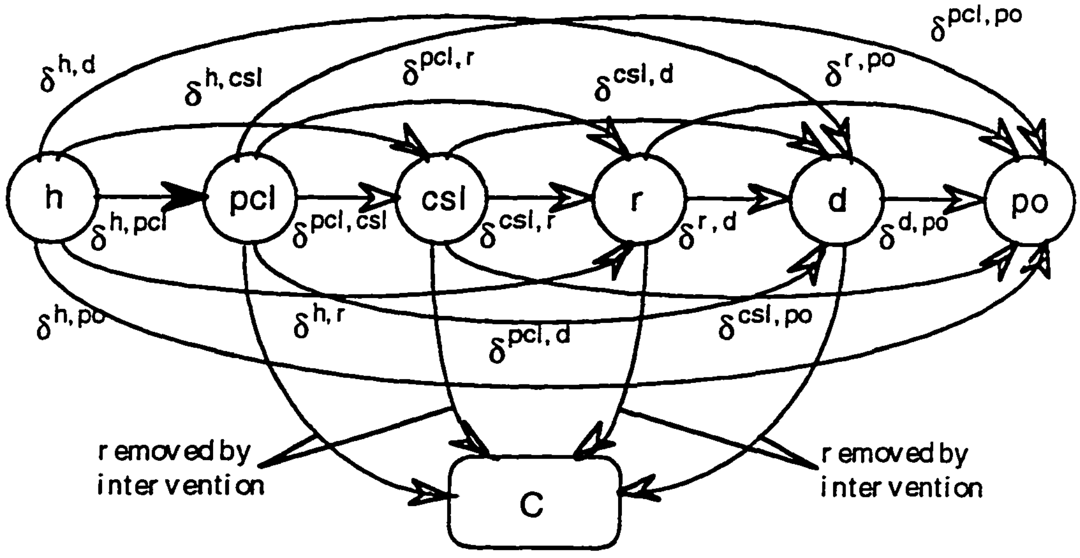

Figure 5.1: The Transition Scheme Between Health Classes

Individuals removed, by medical intervention, from the disease classes are not subject to mutations to other classes. The flow of these individuals is captured by transitions to the node "C."

Thus, the evolution in time of a population is subject to both mutation probabilities, and to the intensity of intervention measures. The one period evolution in time from period i-1 to period i, is summarized by the following equations:

The one-period evolution of healthy, or class *h*, individuals is given by:

$$\lambda_i^h = (1-\delta_{i-1}^h)\lambda_{i-1}^h, \qquad \forall i \in \{1, ..., I_{max}-1\} \qquad (5.3.2)$$

where $\delta_{i-1}^h = \delta_{i-1}^{h,pcl} + \delta_{i-1}^{h,csl} + \delta_{i-1}^{h,r} + \delta_{i-1}^{h,d} + \delta_{i-1}^{h,p}$.

Equation (5.3.2) is obtained by noting that the size of a fixed initial healthy population is non increasing and, further, if individuals mutate out of the health class this serves only to decrease the arrival rate of class *h* individuals in the next period.

The one-period evolution of pre-clinical local stage, or class *pcl*, individuals is given by:

$$\lambda_i^{pcl} = \delta_{i-1}^{h,pcl}\lambda_{i-1}^{h} + (1-\delta_{i-1}^{pcl})\left(1-\alpha\varepsilon\Xi(R_{i-1}^{edts})\right)\lambda_{i-1}^{pcl}, \quad \forall i \in \{1, \ldots, I_{max}-1\} \qquad (5.3.3)$$

where $\delta_{i-1}^{pcl} = \delta_{i-1}^{pcl,csl} + \delta_{i-1}^{pcl,r} + \delta_{i-1}^{pcl,d} + \delta_{i-1}^{pcl,p}$.

Equation (5.3.3) is derived as follows: From Figure 4.1 we note that individuals in class pcl can be removed if they transit first to the node EDC for screening tests and, if detected with cancer, onto node ETS for treatment. The successful removal of class *pcl* individuals arriving at node EDTS, the combination of nodes EDC and ETS, is given by the function $\Xi(R_{i-1}^{edts})$. Further, the proportion of class *pcl* individuals who transit from node T to the nodes EDC and ETS is $\alpha\varepsilon$. Individuals not removed from the class *pcl* will either remain asymptomatic or mutate to more advanced disease states, as shown in Figure 4.1. In Equation (5.3.2), the first term captures the numbers of asymptomatic ill in period i-1 who remain in class *pcl* in period i. The second term reflects the increase in the number of latently ill due to arrivals from period i-1 class h individuals who mutated into class *pcl* by the beginning of period i.

The one-period evolution of clinically surfaced local stage, or class *csl*, individuals is given by:

$$\lambda_i^{csl} = \delta_{i-1}^{h,csl}\lambda_{i-1}^{h} + \delta_{i-1}^{pcl,csl}\left(1-\alpha\varepsilon\Xi(R_{i-1}^{edts})\right)\lambda_{i-1}^{pcl} + (1-\delta_{i-1}^{csl})\left(1-[\alpha+(1-\alpha)\kappa]\Psi(R_{i-1}^{csl})\right)\lambda_{i-1}^{csl},$$

$$\forall i \in \{1,\ldots,I_{max}-1\} \qquad (5.3.4)$$

where $\delta_{i-1}^{csl} = \delta_{i-1}^{csl,r} + \delta_{i-1}^{csl,d} + \delta_{i-1}^{csl,p}$.

Equation (5.3.4) is obtained by considering the following flows in Figure 4.1. Of all individuals in class *csl* at time i, persons with access to the system transit to the node RMS at rate $\alpha\lambda_{i-1}^{csl}$. Individuals without access seek treatment at county facilities, node GMS, at rate $(1-\alpha)\kappa\lambda_{i-1}^{csl}$. However, some uninsured class *csl* individuals will remain outside the health care system; this fraction is $(1-\alpha)(1-\kappa)$. The probability of an individual's cure at the treatment centers RMS, and GMS is given by the functional relationship $\Psi(R_i^{csl})$, and is dependent upon the per-patient expense. Thus, the first term in Equation (5.3.3) collects i-1 th period class *csl* individuals who remain in class *csl* because of unsuccessful or no treatment and who do not mutate to another health state by period i. The second term reflects the increase in the number of cancer patients due to arrivals from period i-1 class *h* and *pcl* individuals who mutate into class *csl* by the beginning of period i.

The one-period evolution of regional stage, or class *r*, individuals is given by:

$$\lambda_i^r = \delta_{i-1}^{h,r}\lambda_{i-1}^h + \delta_{i-1}^{pcl,r}\left(1-\alpha\varepsilon\Xi(R_{i-1}^{edis})\right)\lambda_{i-1}^{pcl} + \delta_{i-1}^{csl,r}\left(1-[\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_{i-1}^{csl})\right)\lambda_{i-1}^{csl}$$
$$+(1-\delta_{i-1}^r)\left(1-\left[\alpha+(1-\alpha)\kappa^r\right]\Phi(R_{i-1}^r)\right)\lambda_{i-1}^r, \quad \forall i \in \{1,\ldots,I_{max}-1\} \qquad (5.3.5)$$

where $\delta_{i-1}^r = \delta_{i-1}^{r,d} + \delta_{i-1}^{r,p}$.

The derivation of Equation (5.3.5) is similar to that of Equation (5.3.4) with two exceptions. The first is that the probability of cure at the nodes RMS and GMS for class *r* individuals is given by $\Phi\left(R_{i-1}^r\right)$; and the second is that in addition to mutations from classes *h* and *pcl*, the equation collects mutations from untreated or uncured i-1 th period class *csl* individuals.

The one-period evolution of distant stage, or class d, individuals is given by:

$$\lambda_i^d = \delta_{i-1}^{h,d}\lambda_{i-1}^h + \delta_{i-1}^{pcl,d}\left(1-\alpha\varepsilon\Xi(R_{i-1}^{edts})\right)\lambda_{i-1}^{pcl} + \delta_{i-1}^{csl,d}\left(1-[\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_{i-1}^{csl})\right)\lambda_{i-1}^{csl}$$
$$+\delta_{i-1}^{r,d}\left(1-\left[\alpha+(1-\alpha)\kappa^{r}\right]\Phi(R_{i-1}^{r})\right)\lambda_{i-1}^{r} + (1-\delta_{i-1}^{d})\left(1-\xi(R_{i-1}^{d})\right)\lambda_{i-1}^{d},$$
$$\forall i \in \{1,\ldots,I_{max}-1\} \qquad (5.3.6)$$

Equation (5.3.6) follows by noting that emergency care is available to all patients who require care, by law, regardless of ability to pay. Therefore, all breast cancer patients with distant mestatatis are allowed to traffic from node T to node EMS in Figure 3.1. The treatment success at this node is given by $\xi\left(R_{i-1}^{d}\right)$. Thus, the first term in Equation (5.3.6) collects the i-1 th period distant stage individuals who failed treatment but whom remain in class d by period i. The further terms in the equation collect mutations into class d from i-1th period h, pcl, csl, and r classes.

The one period evolution of individuals who have passed on, or class po, individuals is given by:

$$\lambda_i^p = \delta_{i-1}^{h,p}\lambda_{i-1}^h + \delta_{i-1}^{pcl,p}\left(1-\alpha\varepsilon\Xi(R_{i-1}^{edts})\right)\lambda_{i-1}^{pcl} + \delta_{i-1}^{csl,p}\left(1-[\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_{i-1}^{csl})\right)\lambda_{i-1}^{csl}$$
$$+\delta_{i-1}^{r,p}\left(1-[\alpha+(1-\alpha)\kappa]\Phi(R_{i-1}^{r})\right)\lambda_{i-1}^{r} + \delta_{i-1}^{d,p}\left(1-\xi(R_{i-1}^{d})\right)\lambda_{i-1}^{d},$$
$$\forall i \in \{1,\ldots,I_{max}-1\} \qquad (5.3.7)$$

Equation (5.3.7) sums the traffic from the treatment nodes ETS, RMS, GMS, EMS, and T to node Z in Figure 4.1. This sum includes that proportion of individuals who entered period i-1 as class h but have an aggressive enough cancer that they mutate into class po by time i. In addition, the equation includes the proportion of undetected or uncured i-1 th period class pcl, csl, r, and d individuals with cancers such that they mutate into class p by the beginning of the i th age group.

The collection of individuals into a disease free or cured state, c, is given by:

$$\lambda_i^c = \alpha\varepsilon\Xi(R_{i-1}^{edts})\lambda_{i-1}^{pcl} + [\alpha + (1-\alpha)\kappa^{csl}]\Psi(R_{i-1}^{csl})\lambda_{i-1}^{csl}$$

$$+ [\alpha + (1-\alpha)\kappa^{r}]\Phi(R_{i-1}^{r})\lambda_{i-1}^{r} + \xi(R_{i-1}^{d})\lambda_{i-1}^{d} \qquad \forall i \in \{1, \ldots, I_{max} - 1\} \qquad (5.3.8)$$

Equation (5.3.8) collects the proportion of traffic through the nodes RMS, GMS, EMS, and the combined node EDTS, that are successfully treated for breast cancer, i.e., those removed to node C.

Expressing the relationships (5.3.2) - (5.3.8) in matrix form as

$$\mathbf{A}_i = \mathbf{M}_{i-1}\mathbf{A}_{i-1} \qquad \forall i \in I \qquad (5.3.9)$$

where,

$$\mathbf{M}_{i-1} = \begin{bmatrix} a_{1,1;i-1} & 0 & 0 & 0 & 0 & 0 & 0 \\ a_{2,1;i-1} & a_{2,2;i-1} & 0 & 0 & 0 & 0 & 0 \\ a_{3,1;i-1} & a_{3,2;i-1} & a_{3,3;i-1} & 0 & 0 & 0 & 0 \\ a_{4,1;i-1} & a_{4,2;i-1} & a_{4,3;i-1} & a_{4,4;i-1} & 0 & 0 & 0 \\ a_{5,1;i-1} & a_{5,2;i-1} & a_{5,3;i-1} & a_{5,4;i-1} & a_{5,5;i-1} & 0 & 0 \\ a_{6,1;i-1} & a_{6,2;i-1} & a_{61,3;i-1} & a_{6,4;i-1} & a_{6,5;i-1} & 0 & 0 \\ 0 & a_{7,2;i-1} & a_{7,3;i-1} & a_{7,4;i-1} & a_{71,5;i-1} & 0 & 0 \end{bmatrix}$$

and the i-1 period matrix elements are given by:

$a_{1,1;i-1} = (1-\delta_{i-1}^{h})$; $a_{2,1;i-1} = \delta_{i-1}^{h,pcl}$; $a_{3,1;i-1} = \delta_{i-1}^{h,csl}$; $a_{4,1;i-1} = \delta_{i-1}^{h,r}$; $a_{5,1;i-1} = \delta_{i-1}^{h,d}$; $a_{6,1;i-1} = \delta_{i-1}^{h,p}$;

$a_{2,2;i-1} = (1-\delta_{i-1}^{pcl})\left(1-\alpha\varepsilon\Xi(R_i^{edts})\right)$; $a_{3,2;i-1} = \delta_{i-1}^{pcl,csl}\left(1-\alpha\varepsilon\Xi(R_i^{edts})\right)$; $a_{4,2;i-1} = \delta_{i-1}^{pcl,r}\left(1-\alpha\varepsilon\Xi(R_i^{edts})\right)$;
$a_{5,2;i-1} = \delta_{i-1}^{pcl,d}\left(1-\alpha\varepsilon\Xi(R_i^{edts})\right)$; $a_{6,2;i-1} = \delta_{i-1}^{pcl,p}\left(1-\alpha\varepsilon\Xi(R_i^{edts})\right)$; $a_{7,2;i-1} = \alpha\varepsilon\Xi(R_i^{edts})$

$a_{3,3;i-1} = (1-\delta_{i-1}^{csl})\left(1-[\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_i^{csl})\right)$; $a_{4,3;i-1} = \delta_{i-1}^{csl,r}\left(1-[\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_i^{csl})\right)$;
$a_{5,3;i-1} = \delta_{i-1}^{csl,d}\left(1-[\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_i^{csl})\right)$; $a_{6,3;i-1} = \delta_{i-1}^{csl,p}\left(1-[\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_i^{csl})\right)$;
$a_{7,3;i-1} = [\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_i^{csl})$;

$$a_{4,4;i-1} = (1-\delta^{r}_{i-1})\left(1-[\alpha+(1-\alpha)\kappa^{r}]\Phi(R^{r}_{i})\right);\ a_{5,4;i-1} = \delta^{r,d}_{i-1}\left(1-[\alpha+(1-\alpha)\kappa^{r}]\Phi(R^{r}_{i})\right);$$

$$a_{6,4;i-1} = \delta^{r;p}_{i-1}\left(1-[\alpha+(1-\alpha)\kappa^{r}]\Phi(R^{r}_{i})\right);\ a_{7,4;i-1} = [\alpha+(1-\alpha)\kappa^{csl}]\Phi(R^{r}_{i});$$

$$a_{5,5;i-1} = (1-\delta^{d}_{i-1})\left(1-\xi(R^{d}_{i})\right);\ a_{6,5;i-1} = \delta^{d;p}_{i-1}\left(1-\xi(R^{d}_{i})\right);\ a_{7,5;i-1} = \xi(R^{d}_{i}).$$

Thereby, we obtain the transition matrix foundation of CQA.

In summary, this section has presented the time evolution of different health classes both in accordance with natural processes (the mean mutation probabilities), and due to artificial processes. These artificial processes include the access level into the health care system, the level of compliance with preventative measures, and the per-patient expenditures at the treatment centers. Consequently, policies which affect these artificial processes will alter the evolution of the health classes. However, before the exploration of policies can be conducted, it is necessary to define the method by which the transition rates of Equation (5.3.2) - (5.3.8) can be used to obtain CQA. This is the subject of the next sub-section.

5.4. Cost, Access, and Quality

Access in the system is captured by the variable $\alpha \in (0, 1)$. Cost in the health care system for the treatment of breast cancer in a given age group is captured by:

$$Cost_i = \sum_{\substack{x \in \mathbf{DC} \\ y \in \mathbf{TS}}} \lambda^{x}_{i} \bullet R^{y}_{i}, \tag{5.4.1}$$

where **TS** defines the set of medical centers (i.e. **TS** = {EDC, ETS, RMS, GMS, EMS}). Equation (5.4.1) captures the summation of total traffic through each of the medical centers multiplied by the per-patient expenditures at the centers. The total resources in the health care system follows as:

$$TC = \sum_{i \in I} Cost_i, \qquad (5.4.2)$$

The relationships of cost and access with both measures of quality are derived as:

A) National Health Quality

National health quality is defined as the weighted sum of the mean number of individuals within different classes. Thus, the level of quality in an age group i, NHQ_i, is defined as:

$$NHQ_i = M_1 N_i^h + M_1 N_i^c + M_2 N_i^{pcl} + M_3 N_i^{csl} + M_4 N_i^r + M_5 N_i^d + M_6 N_i^p, \quad \forall i \in I \qquad (5.4.3)$$

where the N_i^x's ($x=h, pcl, csl, r, d, p$) denote the mean number of class x, in age group i, in the population, and M_1, ..., M_6 are the weights of the different classes. The quantities N_i^x's ($x=h, pcl, csl, r, d, po$), are obtained by modelling individuals as existing in an M/G/∞ queue with class specific service times given according to the following scheme:

I) Service time in class *h*: Life expectancy of i th age group individuals.

II) Service time in class *pcl*: Average time from smallest detectable cancer tumor to development as clinically surfaced local stage cancer.

III) Service time in class *csl*: Average time an untreated clinically surfaced localized cancer remains in the local stage before progressing to the regional stage.

IV) Service time in class *r*: Average development time an untreated cancer remains in the regional phase before progressing to distant mestatasis.

V) Service time in class *d*: Average life expectancy of an individual with untreated distant mestatasis.

VI) Service time in class *po*: Infinite.

VII) Service time of individuals returned to node C: The same as (I) above.

Let the service rate in class x and time i be generally distributed with mean denoted by $E[S_i^x]$. Define:

L = Steady state length of queue

n = queue length

p_n = steady state probability of n in the system at an arbitrary point in time.

The steady state queue length, L, is given by:

$$L = \sum_{n=0}^{\infty} np_n \qquad (5.4.4)$$

For the M/G/∞, the steady state probability of n in the system is given by:

$$p_n = \frac{e^{-\rho}\rho^n}{n!} \quad (5.4.5)$$

where

$$\rho = \lambda E[S] \quad (5.4.6)$$

and

λ = Arrival rate

S = Service time

From Equations (5.4.4) and (5.4.5) we obtain:

$$L = \sum_{n=0}^{\infty} n \frac{e^{-\rho}\rho^n}{n!} = 0 \bullet \frac{e^{-\rho} \bullet 1}{1} + e^{-\rho} \sum_{n=1}^{\infty} n \frac{\rho^n}{n!} \quad (5.4.7) - (5.4.8)$$

and

$$L = \rho e^{-\rho} \sum_{n=1}^{\infty} \frac{\rho^{n-1}}{(n-1)!}. \quad (5.4.9)$$

Note

$$\sum_{n=1}^{\infty} \frac{\rho^{n-1}}{(n-1)!} = e^{\rho}. \quad (5.4.10)$$

Therefore,

$$L = \rho e^{-\rho} e^{\rho} = \rho. \quad (5.4.11) - (5.4.12)$$

Finally,

$$N_i^x = \lambda_i^x E[S_i^x] \quad (5.4.13)$$

(where we make the notation change L = N).

We define the vector $\mathbf{N}_i = \{N_i^h, N_i^{pcl}, N_i^{csl}, N_i^r, N_i^d, N_i^p, N_i^c\}$ and the diagonal matrix with diagonal elements listed as:

$$\mathbf{m}_i = \{E[S_i^h], E[S_i^{pcl}], E[S_i^{csl}], E[S_i^r], E[S_i^d], E[S_i^{po}], E[S_i^c]\} \;\; \forall i \in \mathbf{I} \quad (5.4.14)$$

and note the following vector relationship $\mathbf{N_i} = \mathbf{m_i} \, \mathbf{H_i}$.

The choice of weights $M_1, \ldots, M_6$, will reflect the relative desirability to society of the different health categories. The weights will follow the health class preference order established as

$$h > pcl > csl > r > d > po \qquad (5.4.15)$$

Thus, the societal goal to maximize the number of healthy people in its populace will require that the weight M_1 be significantly larger than the other weights. Similarly, societal values will seek to minimize the magnitude of the weight M_6. In the models we present, the values given to the weights $M_1, \ldots, M_6$, are chosen arbitrarily; however, we maintain the following relationship:

$$M_1 >> M_2 >> M_3 >> M_3 >> M_4 >> M_5 >> M_6, \qquad (5.4.16)$$

in accordance with the inequalities (5.4.15). We define the vector of the weights $M_1, \ldots, M_6$ as:

$$W = \{M_1, M_2, M_3, M_4, M_5, M_6\} \qquad (5.4.17)$$

The total national health quality will be defined as:

$$\text{Total NHQ} = \sum_{i \in I} NHQ_i(\alpha, TC). \qquad (5.4.18)$$

To calculate Total NHQ as a function of access level, α, and total cost, TC, it is necessary to establish specified initial operating conditions, including an initial health status vector, H_0, and specified system wide cure probabilities represented by the vector

$$\mathbf{Q}_i^o = \left\{ \Omega(R_i^{o\,ed}), \Theta(R_i^{o\,pcl}), \Psi(R_i^{o\,csl}), \Phi(R_i^{o\,r}), \xi(R_i^{o\,d}) \right\} \quad \forall\, i \in I \qquad (5.4.19)$$

due to specific per patient resource allocations among the nodes. The calculation of NHQ_i is obtained according to the following scheme: From Eq. (5.3.9), it is noted that the health status vector $\mathbf{H}_1$ is obtained from $\mathbf{H}_0$ by $\mathbf{H}_1=\mathbf{M}_0\mathbf{H}_0$. It follows that $\mathbf{H}_2=\mathbf{M}_1\mathbf{M}_0\mathbf{H}_0$, with the consequent relationship $\mathbf{H}_n=\mathbf{M}_{n-1}\mathbf{M}_{n-1}...\mathbf{M}_1\mathbf{M}_0\mathbf{H}_0$. Associated with the health status vectors are vectors of the expected numbers of individuals in each class in each age group, as given in Eq. (5.4.4). Thereby, we obtain an alternate expression for total national health quality as:

$$Total\ NHQ = \mathbf{W}\sum_{i\in I}\prod_{n=1}^{i}\mathbf{m}_i\mathbf{M}_{n-1}\mathbf{H}_0 \qquad (5.3.9')$$

where **W** represents the weight vector of Eq. (5.4.8), and **m** represents the service time matrix of Eq. (5.4.5). Specific policy choices are imbued into the $\mathbf{M}_{n-1}$ matrices, consequently giving a return on Total NHQ as a function of policy choice.

B) Service Quality.

We argue here that per-patient resource expenditure can serve as a surrogate for service quality received at treatment centers: The availability of resources defines the choice set of possible diagnostic and treatment procedures. Further, the resources available to a medical institution serve as a determinant in the physician-patient ratio at the facility. This ratio, in turn, guides the waiting times for procedures, and determines both the length of time a physician has available to spend with a patient, and the level of detailed attention a physician can pay a patient (strong determinants in patient perceptions of quality).

Per-patient resources, in the model we have presented, are obtained as inverses of the functions presented in Modelling Assumption VII in Section 5.2. That is, $R_i^x = \vartheta^{-1}(R_i^x)$, where x = {h, pcl, csl, r, d, po}, and $\vartheta = (\Omega,\Theta,\Psi,\Phi,\xi)$. It will be convenient, often times, to operate in the space of the functions ϑ's, instead of the space of the per-patient resource amounts, R_i^x. This is accomplished if we assume, further on Modelling Assumption VII, that the functions $\vartheta(\bullet)$ measure the quality of service received for the detection or treatment of a specific stage of breast cancer, according to a specific per-patient resource expenditure. The measure of quality is such that if two per individual per-patient resource expenditure instances R_i^x and $R_i'^x$ yield the function values $\vartheta(R_i^x)$ and $\vartheta(R_i'^x)$ respectively with the property

$$\vartheta(R_i^x) > \vartheta(R_i'^x) \tag{5.4.20}$$

then the instance R_i^x yields higher quality of service than the instance $R_i'^x$.

Following the inequality (5.4.11), we model the $\Omega(R_i^{edc})$, $\Phi(R_i^d)$, $\theta\ (R_i^{pcl})$, $\Psi(R_i^{csl})$, and $\xi(R_i^d)$, as strictly concave non-decreasing functions that are bounded between 0 and 1. And, henceforth, the functions $\vartheta(\bullet)$, $\vartheta = (\Omega, \Theta, \Psi, \Phi, \xi)$, will be referred to as 'resource-quality' functions.

In summary, this chapter has provided the dynamic relationships between breast cancer health classes, and also defined the measures of service and national health quality. The relationships are particularly significant in capturing the effects of access levels, and per-patient expenditures on the arrival rates of the different health classes. Both access levels and per-patient expenditures are are affected by policy choice, the subject of which is addressed in the next chapter.

Chapter 6

Policy Experiments

6.1. Introduction

In this chapter our primary purpose is to demonstrate the utilization of the framework formed by the relationships that constitute CQA, and the quantification of national health quality, as a policy formulation aid for decision makers. The demonstration has two parts: First, we test the hypothesis that cost control, increased access and quality are simultaneously realizable for diseases of specific characteristics. Second, we introduce four managerial policy experiments with the intent of using CQA to measure the effect of a given policy on both national health quality and service quality. The experiments we present in this chapter are not all inclusive; however, our intent is to demonstrate the quantitative comparison of different policies in a policy choice set.

This chapter will be developed as follows: We first present a set of initial conditions that are common to the experiments. The four experiments are then explained, and experimental results presented. Finally, we discuss the implications of each experiment. The policy experiments are based upon a common *constructed* data set. The data set and its construction is presented in the Appendix.

6.2. Initial Conditions

We assume a *status quo* operation of the health care system defined by an initial access level at α^o, and a given level of per-patient expenditure at each medical facility for the treatment of each of the diseases classes, i.e. we define $R^{x,y}$ where $x \in$ **DC**, $y \in$ **TS**. Note that the initial per-patient expenditure does not distinguish an age-group dependence. The given per-patient expenditures serve to determine the values of the resource-quality functions, and thereby to determine the transition matrices $\mathbf{M}_i$, for $\forall i$. With a given arrival rate vector, $\mathbf{H}_0$, the matrices $\mathbf{M}_i$ determine the initial arrival rate vectors $\mathbf{H}_1, \ldots, \mathbf{H}_{Imax}$ by Eq. (5.3.9). Finally, the *status quo* budgets B_i^y are established by:

$$B^y = R^{x,y} \sum_{i=1}^{I} \lambda_i^x \qquad \forall i \in I \qquad (6.2.1)$$

6.3. Experiments

Experiment 1: The Cost, Quality, Access Hypothesis

The premise underlying this experimental test of the hypothesis, $H^{[CQA]}$, that increased access can be obtained without increasing cost or sacrificing quality is the following: If we have a cost expenditure rate, $C(\alpha^*)$ for a status quo access level, α^*, and further a cost expenditure rate, $C(\alpha')$, for an access level, α', such that $\alpha' \neq \alpha^*$, then a constant cost with change in access simply requires:

$$C(\alpha') - C(\alpha^*) = 0. \qquad (6.3.1)$$

The purpose of this experiment is to search for conditions such that Eq. (6.3.1) holds when $\alpha' > \alpha^*$.

The first step in the experiment is to establish the cost expenditure rate for the treatment of breast cancer across all medical centers. This is given by:

$$\sum_{i=0}^{i=I} C_i \tag{6.3.2}$$

where

$$C_i = \alpha\Big(\varepsilon\big(\lambda_i^h + \lambda_i^{pcl}\big)R^{edc} + \varepsilon\Omega\big(R^{edc}\big)R^{ets}\lambda_i^{pcl} + \lambda_i^{csl}R^{csl}\big(1-\kappa^{csl}\big) + \lambda_i^r R^r\big(1-\kappa^r\big)\Big)$$
$$+\kappa^{csl}\lambda_i^{csl}R^{csl} + \kappa^r\lambda_i^r R^r + \lambda_i^d R^d \tag{6.3.3}$$

From Equations (6.3.2) and (6.3.3) define G= C(α') - C(α*). It will be noted that G is a function of access level, and of the per-patient expenditures, i.e. the vector **R** = (R^{ets}, R^{pcl}, R^{csl}, R^r, R^d). If the hypothesis $H^{[CQA]}$ is valid, then the equation G = 0 may be solved by allowing the elements of the vector **R** to vary. In our solution procedure, we keep four of the five elements of **R** fixed at the status quo level of Equation (A.2), and search the range of the free per-patient expenditure for the solution to Equation (6.3.1). This approach allows us to explore five resource allocation policies. (Note, we do not consider any "mixed" policies where more than one element in **R** is allowed to vary. In our solution procedure, we keep four of the five elements of **R** fixed at the status quo level of Equation (A.2), and search the range of the free per-patient expenditure for the solution to Equation (6.3.1). This approach allows us to explore five resource allocation policies. (Note, we do not consider any "mixed" policies where more than one element in **R** is allowed to vary.) Let each of the policies be indexed by P^k, k = (edc, pcl, csl, r, d) where the notation identifies the class k per-patient expenditure, R^k, as the free variable, with all other per-patient expenditures being fixed. Further, let $P^k \in \mathbf{P}$, the set of policies.

In the following propositions we define several properties of the function G that will aid in characterizing the solutions to Equation (6.3.1). In the propositions we assume the following first order stochastic dominance property of the resource quality functions: $\Theta(\bullet) \geq \Psi(\bullet) \geq \Phi(\bullet) \geq \xi(\bullet)$.

Proposition 6.1. The function G is convex in policy P^k.

Proof. Since P^k gives that only R^k in $\mathbf{R}$ is variable, the proof can be established by showing $\partial^2 G/\partial R^{k2} > 0$. We omit the details.

For the following propositions we define the sub-vector $\mathbf{R}_s = (R^{ets}, R^{pcl}, R^{csl})$. Further, we define a policy P_s^k as one in which a per - patient expenditure $R_s^k \in \mathbf{R}_s$ is allowed to vary while the per patient expenditures $R_s^l \in \mathbf{R}_s$, $l \neq k$, are fixed. Let $P_s^k \in \mathbf{P}_s$.

Corollary 6.2. For fixed access level and for policy P_s^k, the function G has a global minimum.

Proof. From the convexity of G established in Proposition 6.1, it follows that any minimum will be a global minimum. Thereby, the proof is established if

$$\frac{\partial G}{\partial R_s^k} = 0 \qquad (6.3.4)$$

has a solution. Equation (6.3.4) can be solved numerically for $\forall \alpha, \kappa^{cst}, \kappa^r \in (0,1)$ and for the resource quality functions given by Equation (A.8). The details are omitted.

Proposition 6.3. Let R_s^{k*} = argmin G. Then, for $0 < R_s^k < R_s^{k*}$, G is a monotonically decreasing function in R_s^k. For $R_s^k > R_s^{k*}$, G is a monotonically increasing function of R_s^k.

Proof. We show $\partial G/\partial R_s^k < 0$ for $0 \le R_s^k \le R_s^{k*}$, and $\partial G/\partial R_s^k > 0$ for $R_s^k \ge R_s^{k*}$. Details are omitted.

Corollary 6.4. If argmin G = Rsk* and min G ≤ 0, then there exist two solutions to the constant cost condition.

Proof. It suffices to demonstrate the following properties: For $0 \le R_s^k \le R_s^{k*}$, there exists a value of R_s^k such that $G \ge 0$. Further, for $R_s^k \ge R_s^{k*}$, there exists a value R_s^k such that $G \ge 0$. We omit the details.

Corollary 6.5. Argmin G is a decreasing function of access level.

Proof. The proof is established by showing $\partial(\text{argmin } G)/\partial\alpha < 0$. Details are omitted.

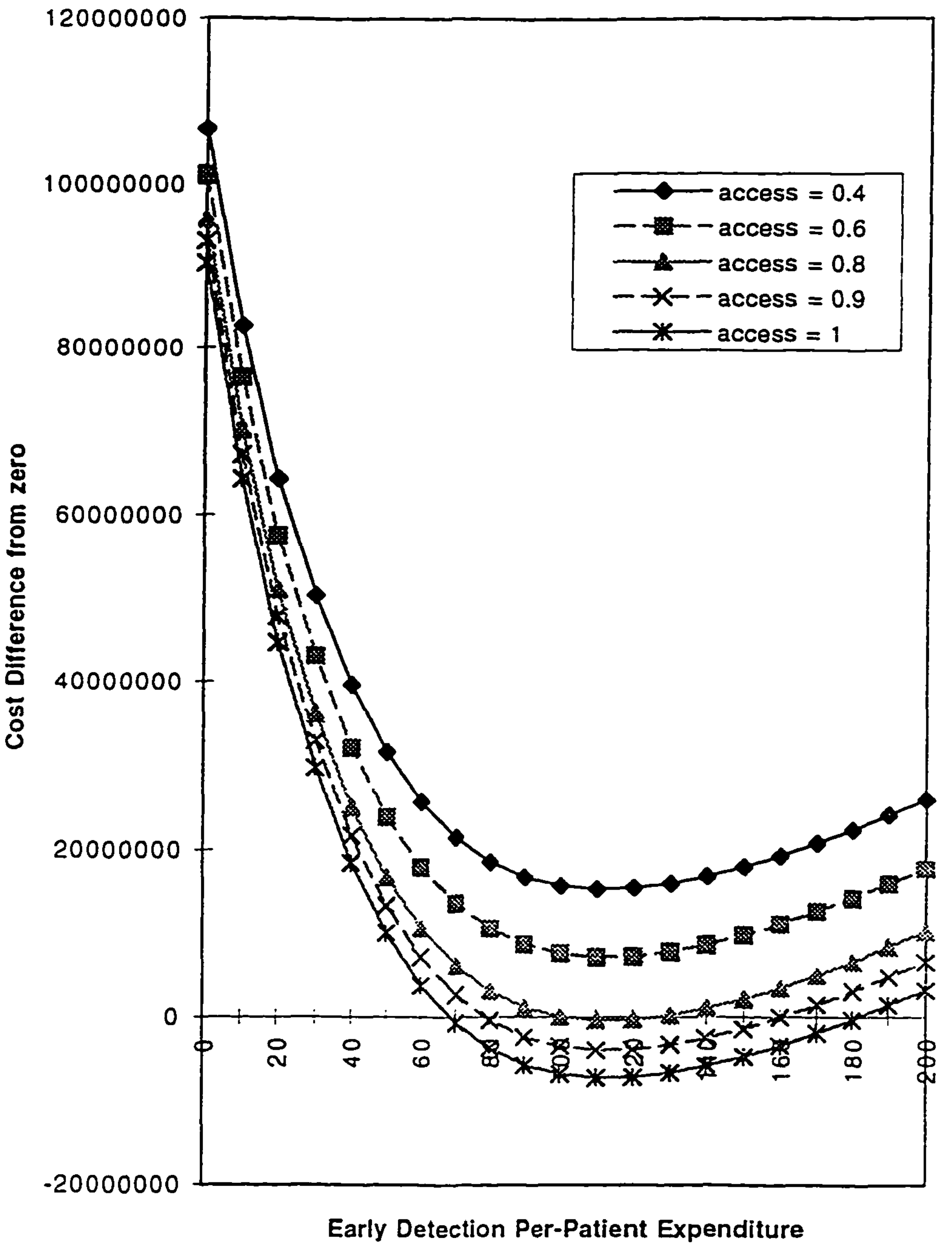

Figure 6.1. Cost (Access) - Cost (Access = 0.8) Difference from Zero for Early Detection Per-Patient Expenditure

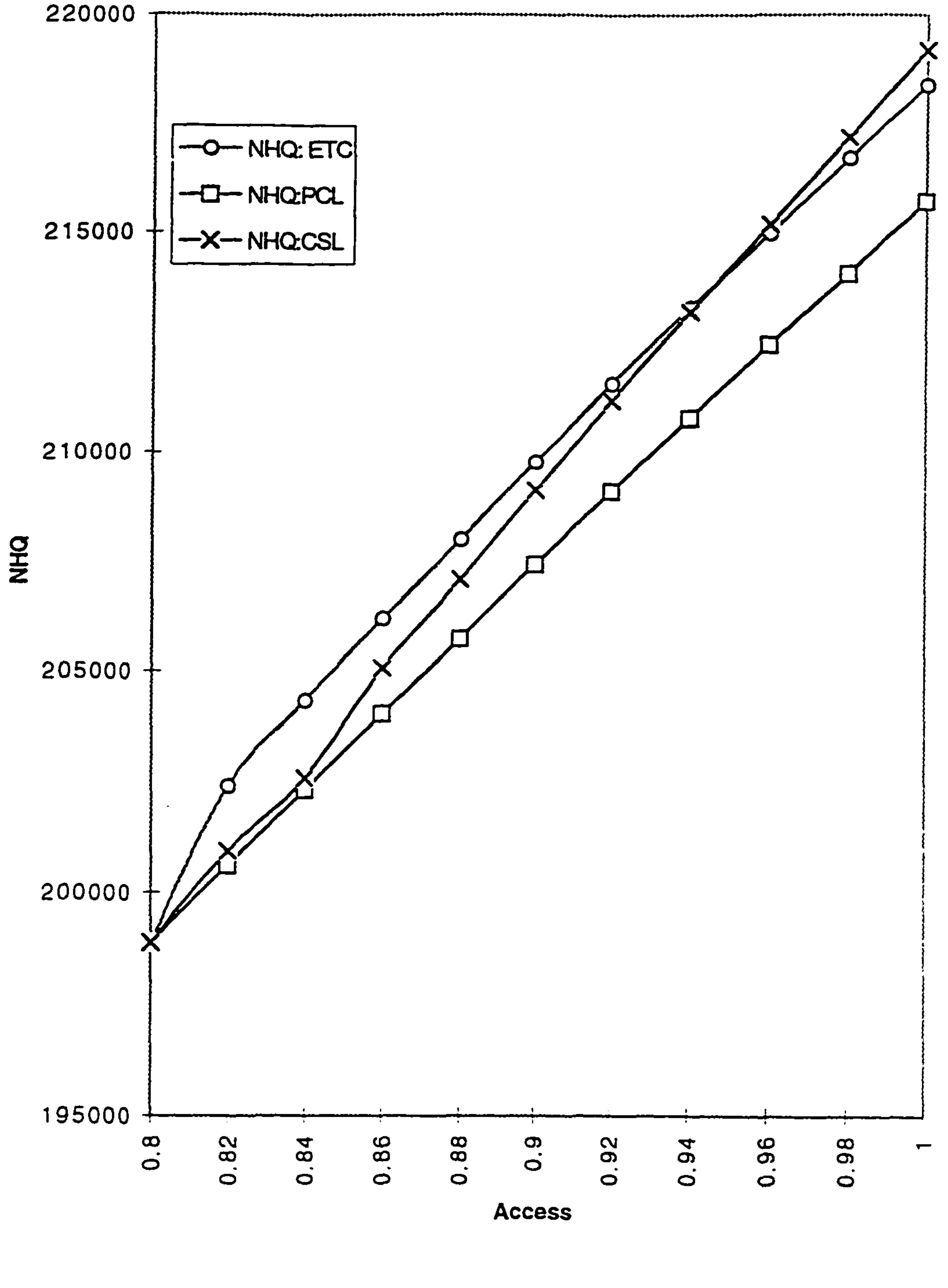

Figure 6.2. National Health Quality as a Function of Resource Allocation Policy

The consequence of propositions 6.1, 6.3, and 6.5 and corollaries 6.2, and 6.4 is the following. Suppose the health care system operates at a status quo access level α^0. For α^0, define the associated functional G^0 with per-patient expenditure $R^{k0} \in \mathbf{R}_s$. Let argmin G^0 equal zero. Proposition 6.5 gives that the minimum of a functional G for any access level $\alpha > \alpha^0$ is less than zero. Further, corollary 6.4 gives that there exist two solutions to Equation (6.3.3) for a functional G with $R_1^k < R^{k0}$, but $R_2^k > R_k^0$. We note, further from Proposition 6.5 that for $\alpha < \alpha^0$, there does not exist a solution to Equation (6.3.3). These results are displayed graphically in Figure 6.1.

The presence of two solutions to Equation 6.3.3 is explained as follows: There are two cost components to the solutions R_i^k, i=1, 2. The first component is the traffic into the health care system. The second is the per-patient expenditure. The first solution R_1^k results from a sufficient reduction in per-patient expenditure as to compensate for the rise in traffic that results from reducing the per-patient expenditure (as given by the transition equations (5.3.2) to (5.3.6)). Similarly, the second solution R_2^k results from an increase in per-patient expenditure that is compensated for by a reduction in traffic into the health care system. It is to be noted, by the definition of the resource-quality function given by Eq. (A.8) in the Appendix, that commensurate with the solution R_1^k is a reduction in service quality for detection or treatment of class k (k= ets, pcl, csl) breast cancer. Similarly, the solution R_2^k yields higher service quality than the status quo level. Given that our interest is the increase in quality, we adopt the solution R_2^k in the comparison of policies below. In determining the optimal resource allocation policy from within the policy choice set, $\mathbf{P}_s$, we compute the national health quality measure, NHQ, for each allocation policy. In Figure 6.2 we graphically display NHQ as a function of policy and access. Given that we are interested in higher measures of NHQ, it is seen that a policy of increased breast cancer detection resources is the optimal policy.

In summary, this experiment has demonstrated that it is possible, within modelling assumptions, to realize universal access and increased quality, while maintaining cost for diseases of specific characteristics.

Experiment 2: The Hypothesis $H^{[CQA]}$ for Cervical Cancer

This experiment mirrors Experiment 1 except the hypothesis $H^{[CQA]}$ is tested for cervical cancer data. The data is obtained as base clinical parameter estimates from (Fahs et al, 1992) and is presented below in Table 6.1.

Table 6.1. Cervical Cancer Model Assumptions

Variable	Clinical Parameter Estimate
Pap smear sensitivity, %	75
Pap smear specificity, %	95
Prevalence rate per 1000	
CIN	4.80
CIS	2.39
EICC	2.8
LICC	2.8
Incidence of progression, %	
Healthy to CIN	0.33
CIN to CIS	17.8

CIS to EICC	26.10
EICC to LICC	39.00

In addition, the five year survival rates are given as (Cook and Dresser, 1996): Stage CIS: 100%, stage CIN: 89%, stage EICC: 40% and stage LICC: 14%.

The results of this experiment are presented in Figures 6.3 - 6.5. The figures give the difference for different levels of access in cost from the status quo access level of 80%. The figures plot the difference in cost as a function of per - patient expenditures at the different stages of the disease. The results of the test of the hypothesis $H^{[CQA]}$ in Experiment 2 stand in stark contrast to the results of Experiment 1. In Experiment 2, the hypothesis could not be validated. The primary reason is the comparitively low incidence level of cervical cancer as compared to the incidence level of breast cancer. Consequently, in Figure 6.3, the only solution to Equation 6.3.3 as access levels are increased above 80% is obtained by reducing the per-patient expenditure. Contrary to Experiment 1, there is not sufficient of an impact made on future arrivals to compensate for the extra. Further, in Figure 6.3, we note that order of progression as a function of access level in the plotted curves is reversed from that in Figure 6.1.

Similar results hold for Figures 6.4, and 6.5 as for Figure 6.3. Of interest in Figure 6.5 is that the trade off between the change in consumption pattern and per - patient expenditure exhibits the characteristic U of Figure 6.1. However, a notable difference is that the intersection of the cost difference curve with the X axis for higher levels of access above the status quo level is below the status quo per patient expenditure.

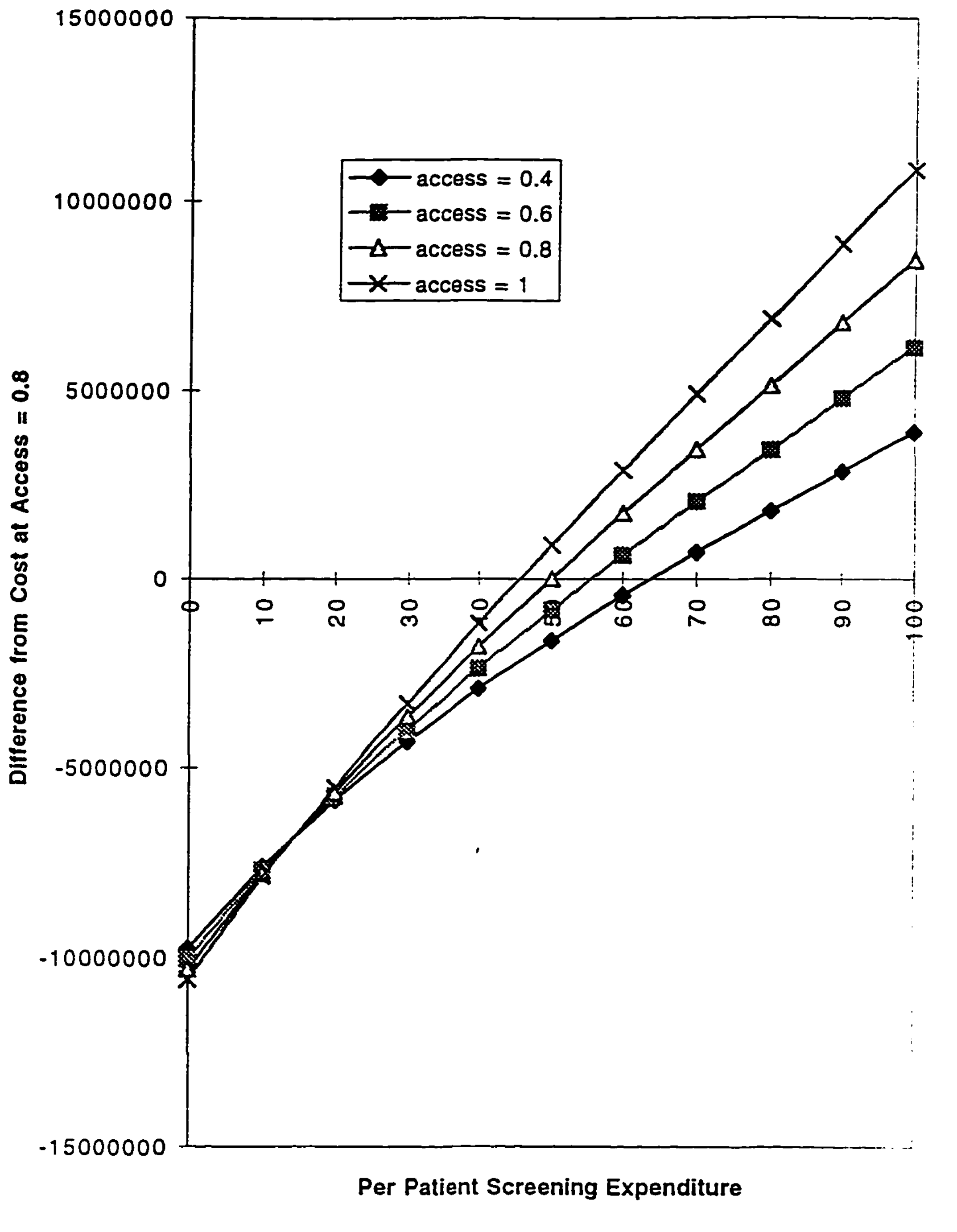

Figure 6.3. Cervical cancer cost differential from status quo as a function of access for per - patient screening expenditure

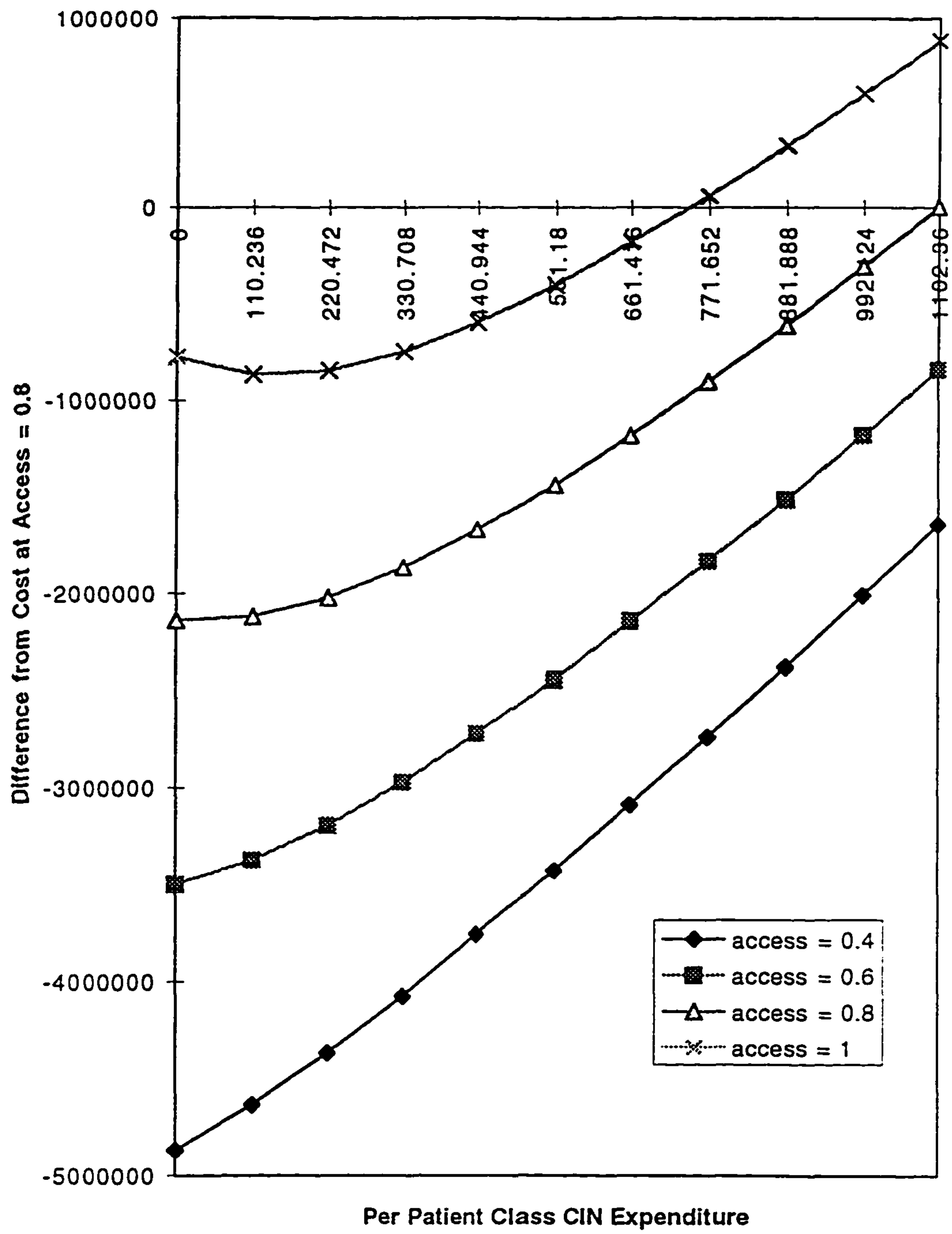

Figure 6.4. Cervical cancer cost differential from status quo as a function of access for class CIN per - patient expenditure.

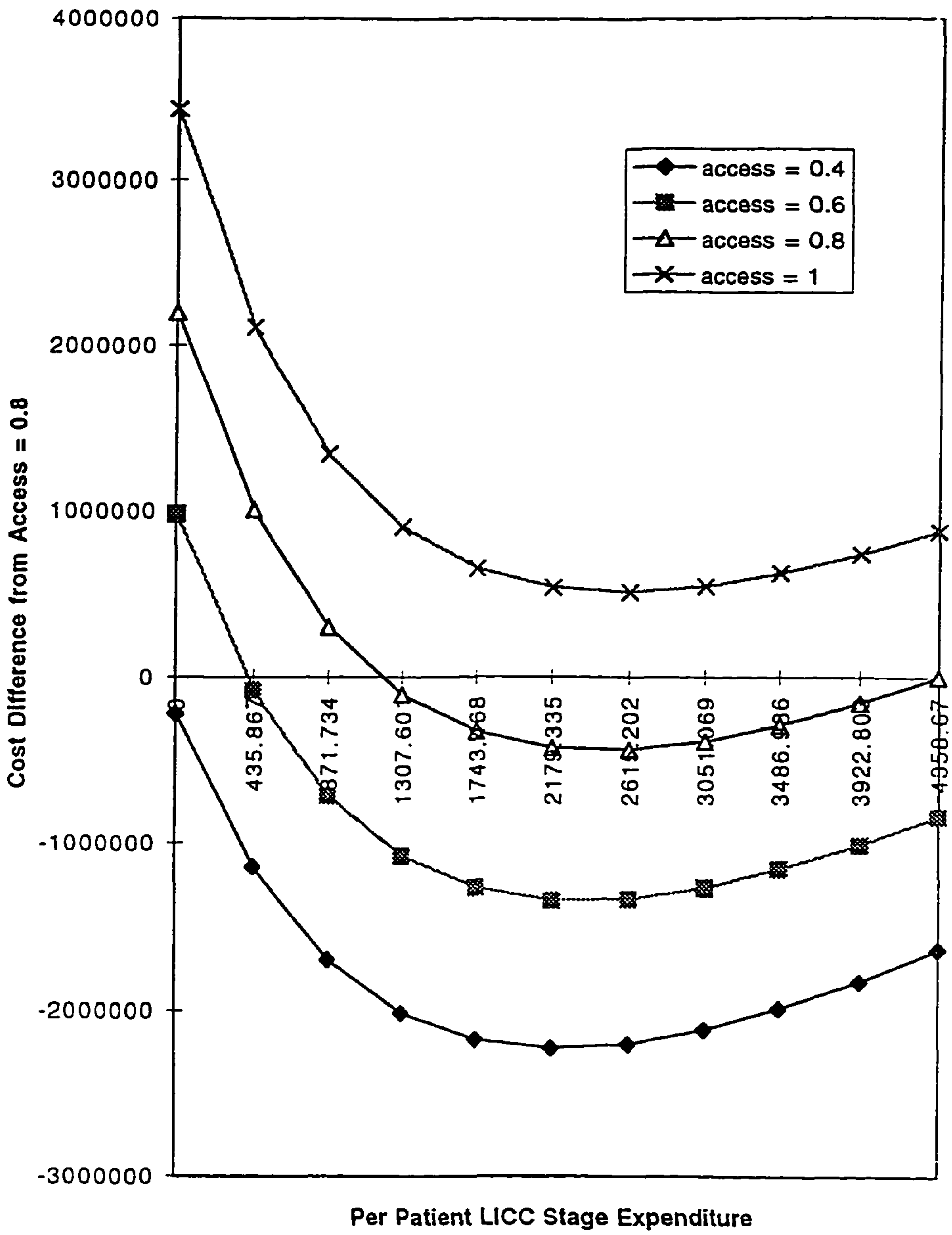

Figure 6.5. Cervical cancer cost differential from status quo as a function of access for stage LICC per - patient expenditure.

Experiment 3: Cash Flow Determination via Simulation

In Experiments 1 and 2 it was shown that the hypothesis $H^{[CQA]}$ was verified for diseases of specific characteristics (i.e. a sufficiently high incidence rate). The experimental results were derived, however, assuming steady state conditions. In this experiment we extend the method of Experiment 1 but our interest is to note the cash flows that result from a screening program in a health organization in transient state. The purpose of the experiment is to demonstrate to HMO's the economic viability of undertaking a screening program.

The experiment is purformed as follows: Using a built in Poisson random number generator, a stream of numbers was generated independently for each disease class. The number of healthy persons was presumed to be one hundred thousand minus the sum of persons in the disease classes. Once the Poisson distributed stream of arrivals in the different health classes in age period 0 had been generated, the arrivals were subject to the transition equations defined by Equations (5.3.2) - (5.3.8). Using the transition equations, the arrivals in periods 1 to 5 were obtained as a function of a resource allocation policy. The specific allocation policy we were interested in was to experiment with the period 0 per patient screening expenditure such that the resulting consumption pattern change possibly pays for the screening expenditure. The experiment was conducted in two parts. In part 1, we do not presume an inflation rate. In part 2 we presume an annual 4% inflation rate. The results are presented with a 90% confidence interval as follows:

Table 6.2. Arrival Rate in Period 0 of Different Health Classes

Healthy:	96698.53 ± 31.78
Pre-Clinical:	1522.97 ± 12.96
Clinically Surfaced:	801.6 ± 9.29
Regional:	838.3 ± 9.56
Distant:	108.6 ± 2.92

Table 6.3. Average Period 0 per patient screening expenditure

Part 1:	352.96 ± 1.87
Part 2:	140.99 ± 14.55

Table 6.4. Pay Back Period:

Part 1:	25 years
Part 2:	10 years

Table 6.5. Extra screening expenditure over status quo of access level

	Part 1:	Part 2:
Period 0:	$1.25 m ± 0.1	$5.9 ± 0.46
Period 1:	$(1.7) ± 0.05	$(0.581) ± 0.1
Period 2:	$(4.02) ± 0.041	$(5.1) ± 0.78
Period 3:	$(3.04) ± 0.01	-
Period 4:	$(1.02) ± 0.0008	-
Period 5:	$(2.14) ± 0.0005	-

From Table 6.3 and Table 6.4 we note that consideration of a 4 % inflation rate considerably reduces both the required screening per - patient expenditure, and the pay back period. Managerially, this strengthens the arguments for screening as a cost effective method of improving the health of the population. Table 6.5 gives the cash flows both with and without consideration of inflation. The important insight for health organizations that we obtain is that under inflation, the majority of the payback occurs later rather than sooner. This insight will allow health organizations to plan their financial investments accordingly.

Experiment 4: Reactive Policy, Free Resource Flow within Age Group

In this experiment we examine the following policy: We assign a priority list to the disease states of the following order: d > r > csl > pcl, i.e. assign greatest priority to treating the most severe stages of the disease, among all cases present in a specific age group. We assume a fixed budget, or availability of resources, per stage and do not allow for the flow of resources across age groups. In the priority scheme, a fixed 'token' amount is spent on treatment for a given class subject to available resources leading, in effect, to a to the free flow of resources within an age group. If the available resource amount is insufficient to meet the 'token' amount, the treatment expenditure rate is given by the divisor of the resource amount with the arrival rate of the disease class.

To run this experiment we establish a set of available age specific budgets B_1, B_2, ..., B_{Imax} according to the assumed initial conditions . With the given arrival rate vector, H_0, we observe the effect of the given policy on the arrival rates of the different disease classes as a function of access according to Equations (5.3.2) - (5.3.8).

Subsequently, we obtain the functional dependence of national health quality on access according to Equation (5.4.9). The service quality at the treatment centers as a function of access is obtained by observing the per-patient resource expenditures at the treatment centers, and the consequent effect on the respective resource-quality functions.

Experiment 5: Pro-active Policy, Free Resource Flow within Age Group.

The conditions for this experiment are very similar to Experiment 4 except that the treatment policy is reversed. Priority is now given to the detection and treatment of class pcl individuals, followed by treatment of class csl, r, and d respectively, i.e. the priority scheme is pcl > csl > r > d. As in Experiment 4 above, the detection and treatment amounts are fixed 'tokens', subject to available covering resources, or are given by the divisor of available resources with arrival rates of the disease class.

Experiment 6: Pro-active Policy, Increased Compliance.

This experiment is an extension of Experiment 5 with the only change being increased compliance with mass screening measures for those with access to the medical system. This experiment asks whether increased compliance with preventative measures will have a multiplicative effect in decreasing the arrivals of the more severe disease states. Further, we ask whether the decreased arrivals are sufficient to pay for increased access without sacrificing quality.

Experiment 7: Simultaneous Distribution of Resources

In this experiment, we observe the effect of the following policy: Let the resource amount available to treat a disease class be fixed and equal to a budget obtained from the assumed initial conditions. We specify that any change in per-patient resource expenditure that may ensue as a realization of a change in access (away from that of the initial conditions) is equally distributed among all age groups. Consequently, this experiment differs from Experiments (4) - (6) above in that there is no age discrimination in the receipt of available resources.

In order to obtain the equitable distribution of resources across age groups, we proceed as follows: Let the treatment centers operate under budgets of the specified initial conditions. As the access level is increased from the initial specified level, α^o, to a level α^* we may obtain class and age specific arrival rates, $\lambda *_i^x$, (x $\in$ **DC**) at the treatment centers different from that of the specified initial conditions. Under α^*, we obtain per-patient resource expenditures, $R^{*x,y}$ (y $\in$ **TS**) as:

$$R^{*x,y} = B^y \Big/ \sum_{i=1}^{I} \lambda *_i^x \qquad (6.2.6)$$

where B^y is obtained from Equation 6.2.1. We note that $R^{*x,y}$ may be different from the per-patient expenditures of the initial conditions. However, the change in available resources simultaneously alters the resource quality functions. Consequently, the class and age specific arrival rates at the treatment centers, and the per-patient resource expenditure as a function of access are obtained by the simultaneous solution of Equation (5.2.9) with the functional equations $\vartheta(R^{*x,y}) = f(R^{*x,y})$, where $f(\bullet)$ denotes the functional form of the resource-quality functions.

6.4. Results for Experiments Four to Seven

The experimental results are presented as a separate comparison of Total National Health Quality, and of service quality at the treatment centers, as a function of access. In all comparisons, we assume available resources in the health care system are fixed.

In Figure 6.6, we plot a comparison, as a function of access, of the national health qualities that result from adoption of the four policies above. We note one policy dominates for all values of access -- the pro-active, full compliance policy of Experiment 4. Experiments 2 and 3 yield similar values for total national health quality, but diverge in value after $\alpha = 0.8$. It will be noted that $\alpha = 0.8$ was defined in the

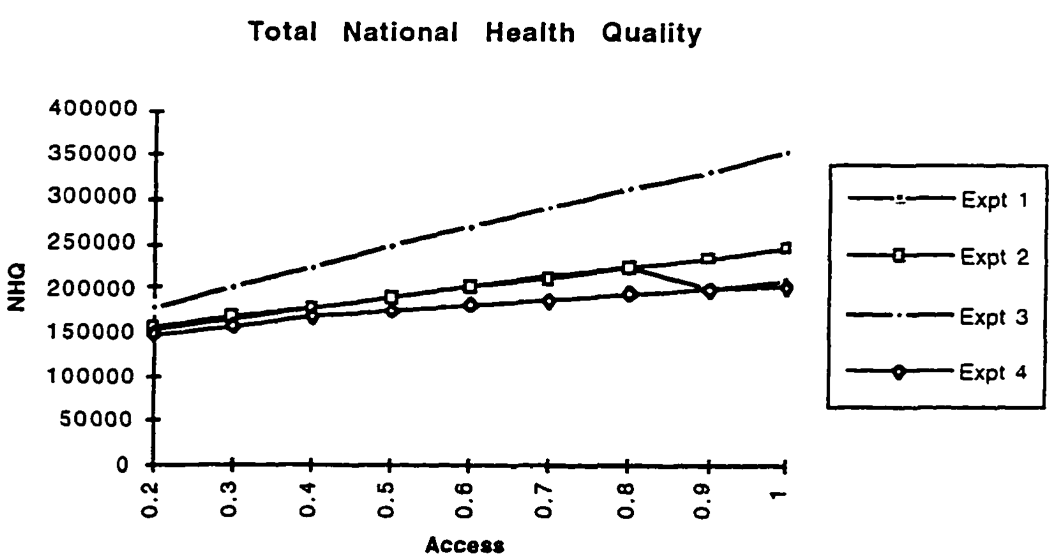

Figure 6.6: Policy Dependent Total National Health Quality as a Function of Access

Appendix as the status quo access level. Finally, it is noted that the policy of Experiment 5 performs the poorest with respect to national health quality.

The service quality effects of the four policies are indicated by presenting both the per-patient resource expenditures at the medical facilities, and the respective resource-quality functions. The purpose of both quantities, even though they are related, is that what appears to be a large discrepancy in per-patient expenditure may correspond to only a small or no change in service quality, as specified by the resource quality functions.

The service quality comparison is presented as follows: In Figures 6.7(a)to 6.7(e), we present the experiment specific levels of service quality available at the treatment centers, as a function of access.

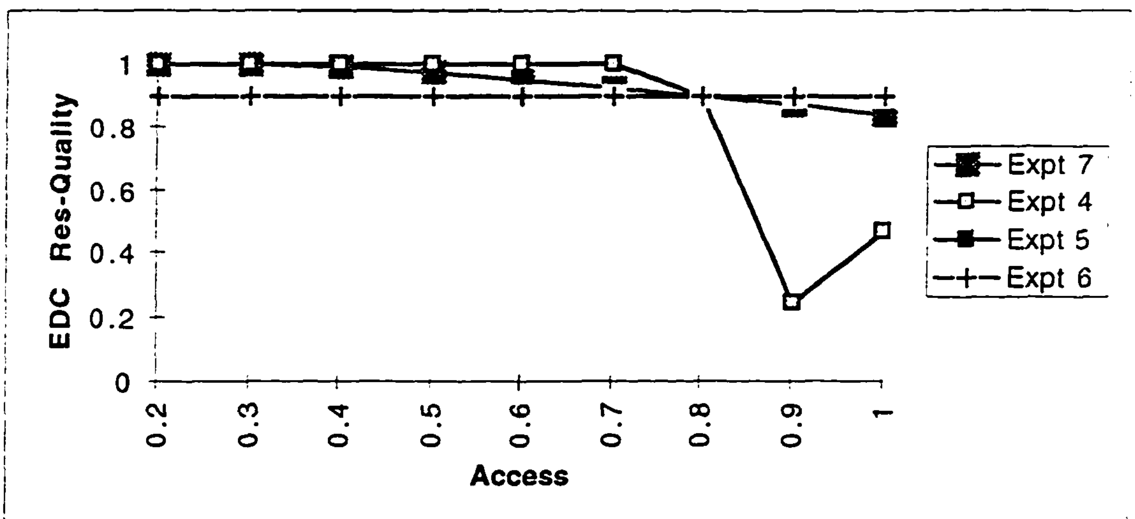

Figure 6.7 (a): Early Detection Resource-Quality as a Function of Access

Figure 6.7.(a) presents the sensitivity at node EDC as a function of policy experiment. It will be noted that while policy experiments 4 and 7 provide for

diminishing sensitivity with access. Consequently, adoption of these policies will lead to diminished service quality (from the status quo reference amount) at access levels

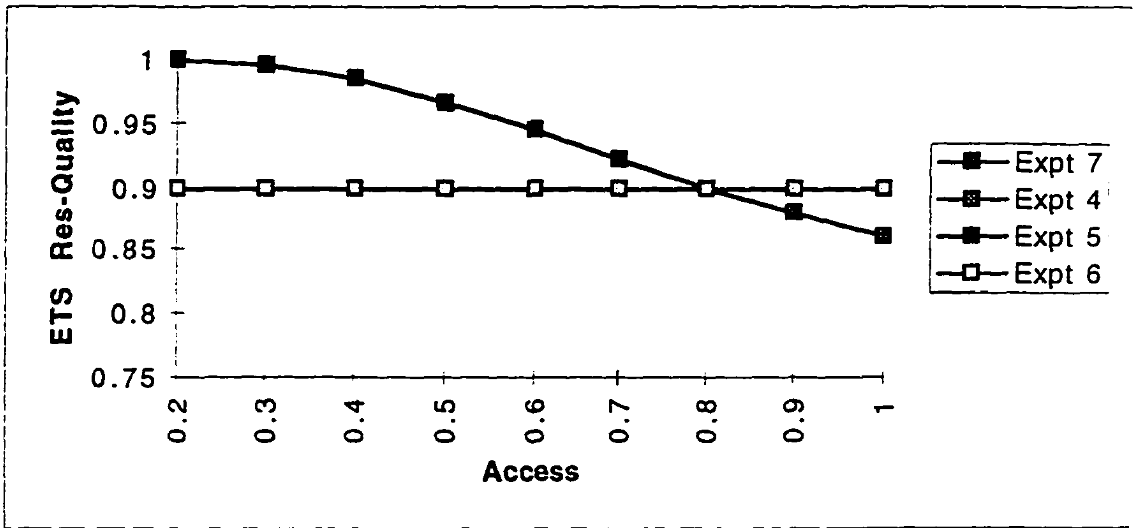

Figure 6.7 (b): Early Treatment Resource-Quality as a Function of Access

above $\alpha = 0.8$. Further, it is noted in Experiment 5 that the sensitivity is saturated at one for $0 \leq \alpha \leq 0.68$. Consequently, much of the resources available at low-levels of access are ineffectively utilized. As is to be expected, the pro-active policies reflect no change in per-patient resource amount.

Figure 6.7.(b) provides that there is no change in per-patient early treatment expenditures for all but the Experiment 5. The result is expected for the pro-active policies of Experiments 5 and where priority is given to the early treatment of the disease. Figure 6.7(b) indicates that the re-active policy of Experiment 4 retains

sufficient resources to provide the 'token' amount for early treatment, even at access levels above that of the initial level.

Figure 6.7 (c) illustrates that both the reactive and pro-active policies have little effect on either the amount spent or the service quality of treatment for clinical stage breast cancer.

In Figure 6.7(d), we note that both the re-active and pro-active policies of

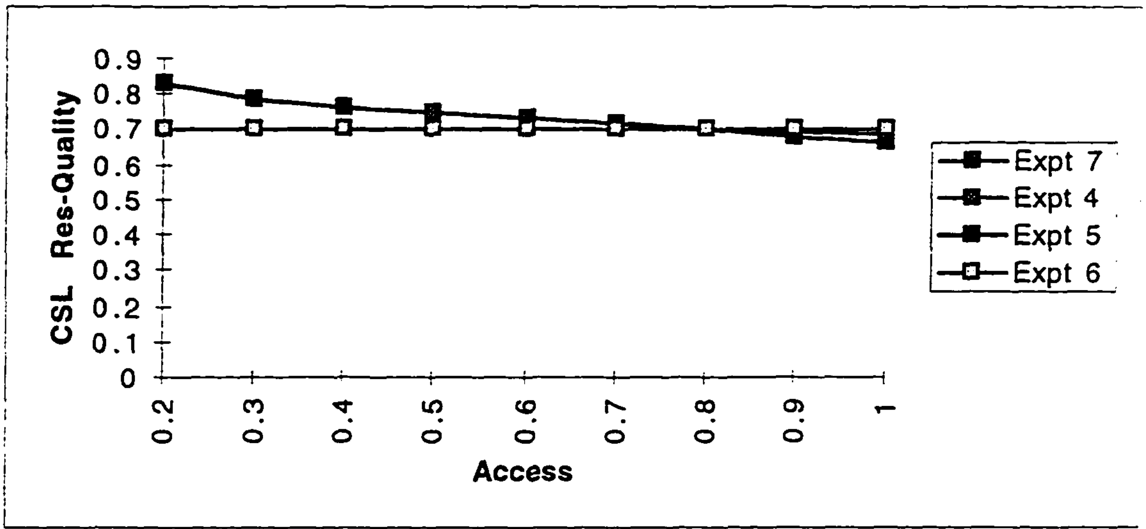

Figure 6.7 (c): CSL Stage Resource- Quality as a Function of Access

Experiments (4) - (6) provide for sufficient resources that the 'token' amount for regional stage treatment is met for all levels of access. In the policy of Experiment 7, we note the inverse relationship between service quality, and level of access.

From Figure 6.7(e), we note the decline in service quality for distant stage patients under the pro-active policies of Experiments (3) and (4) for levels of access above the initial level. For the full compliance policy (Expt. 4), no resources remain after $\alpha \geq 0.9$ to cover distant stage treatment expenses. Experiment 2 yields no access level dependence on resource-quality, as expected from the priority policy of this experiment. Experiment 5 provides for increasing resource-quality with access. This is

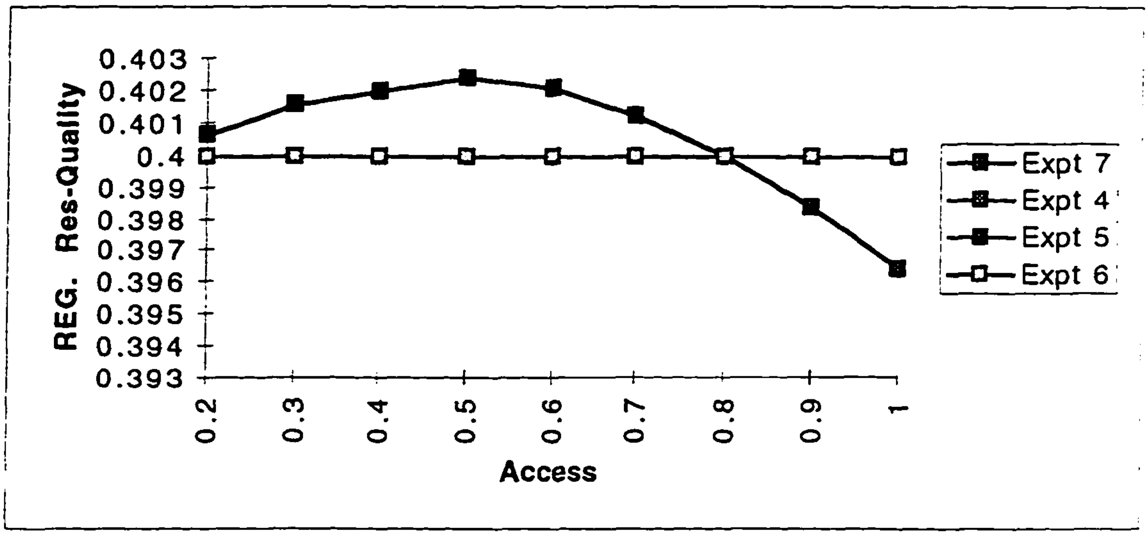

Figure 6.7 (d): Regional Stage Resource-Quality as a Function of Access

due to the decreasing distant stage arrivals that result with providing greater access to persons with less severe forms of breast cancer.

In sum, while a pro-active policy, especially with full compliance, yields superior values of national health quality, these same policies yield inferior distant stage service quality measures (see Figure 6.7(e)). It is to be emphasised that the purpose of the experiments above are to *quantify* the pros of a policy (for example the superior national health quality of the pro-active policy) versus the cons of the same policy (for

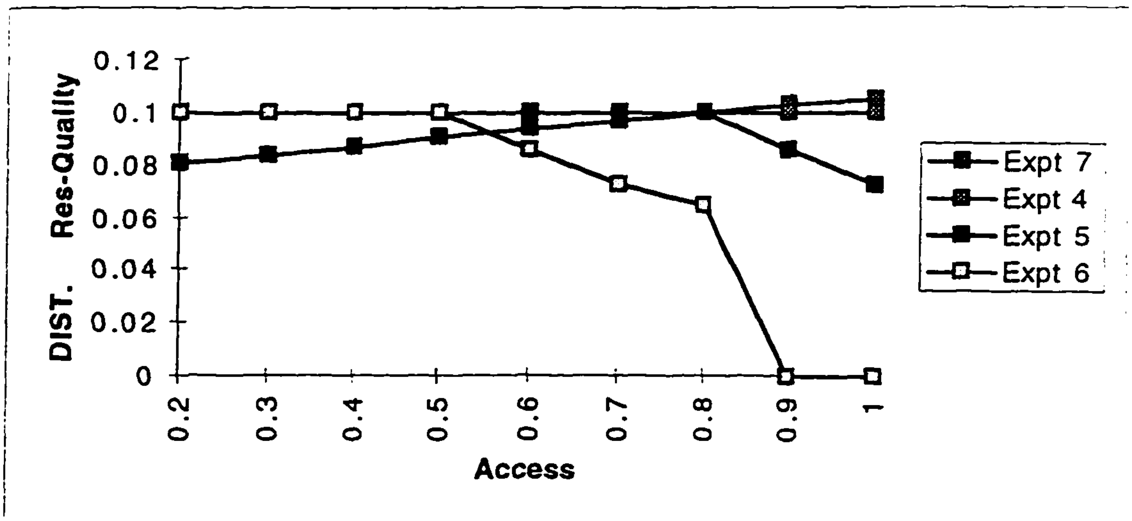

Figure 6.7 (e): Distant Stage Resource Quality as a Function of Access

example, the pro-active policy produces inferior service quality results). Having quantified the effects of policies on different dimensions, the results can be interpreted by policy makers to arrive at decisions according to priorities by which they are governed.

Chapter 7

The National Health Quality Resource Allocation Models

7.1. Introduction

This chapter develops optimization models that are based upon the dynamics of the health care system presented Chapters 4 and 5. The motivation for the optimization models is the following: In the new world of managed care, medicine is practised under captitated budgets. As managed care stakes an increasing proportion of the market for health care, there will be increasingly budget constraints placed upon the provision of medical care. Consequently, it is important to understand how best to allocate available resources to best provide medical care and meet specified goals.

In this dissertation, we consider the division of resources as falling along the following lines: First, constrained budget health care providers must make decisions on the age specific allocation of resources. The need for age specific resources follows from the age specificity of epidemiological data on the incidence rates of progressive chronic diseases. Of course, this decision does not preclude a policy on the allocation of resources independent of age. Such a decision is included in the possible set of allocations. More specifically, the meta-decision on allocation of resources across periods is motivated by equations (5.3.2) - (5.3.8) where it is noted that the i th period arrival rate vector, A_i, is dependent upon the arrival rate vector of the previous period, A_{i-1}. In particular, a change in the numbers of less severe disease classes affects the numbers downstream, in time, of the numbers in the more severe disease classes. The optimal allocation of resources across periods will globally maximize the health of all age groups in the time window of study.

Second, decisions must be made for each age group on the stage specific allocation of resources for progressive chronic diseases. The need for a stage specific allocation of resources can be seen by considering the following examples of possible policies: In Policy 1, we consider an allocation of resources that emphasizes preventitive care; In Policy 2 we have a "rescue medicine" allocation of resources. Assuming the same initial conditions on distributions of the different health classes in the population, it is conceivable to note that policies may produce differing outcomes in the health of the respective populations. Consequently, it is important to optimally allocate resources across stages to obtain desired goals. More specifically, we consider allocating a given budget, assigned to period i, across treatment centers such that the value of national health quality in that period, NHQ_i, is maximized.

Further along the decisions to allocate resources across stages is the decision on the relative split between resources allocated to treatment, and resources allocated to prevention. Within the realm of resources allocated to prevention, we need to consider the division of resources between the detection of persons who may be latently ill, and the subsequent treatment of persons who are detected ill. Conceivably, if a majority of resources are allocated to detection, a large number of persons will be detected. However, very few resources will remain to treat the large number of persons detected ill. It is important to allocate the resources so that the optimal outcome is obtained, as specified by desired goals. More specifically, a preclinical individual is considered removed by medical intervention from class *pcl* only if the pre-clinical breast cancer is detected *and* successfully treated. The decision arises, then, as to the optimal division of resources between detection and cure that will remove the largest number of class *pcl* individuals from that class. For example, is a policy to devote a large fraction of available resources to detection, and a smaller fraction of resources to cure better than a policy that reverses the resource allocation? Our concern is to obtain the optimal division of resources.

In this chapter, we develop optimization models that allocate resources: i. across age groups; ii. within an age group, across disease stages; and iii. within disease stages, the optimal division between treatment and prevention. These allocations will constitute a hierachy of decisions. The hierachy is illustrated in Figure 7.1.

This chapter is developed by presenting the hierachical models in reverse order. (The need to present the models in reverse hierachical order follows from the direction in which results from one model are required in another.) We assume an i th period resource allocation of R_i^{edts} to early detection and treatment and develop a model, called MSRA, to optimally allocate between the early detection and treatment components. The efficient allocation of resources between these components yields the maximum proportion of class *pcl* individuals detected for a specific resource allocation R_i^{edts}. By extension, the model yields the functional form of the proportion of class *pcl* individuals cured per allocated resource unit. This functional form is required in the next model, called HCSRA, we present to optimally divide resources among different health care centers in the system.

One of the results of model HCSRA is the derivation of the functional relationship between resources allocated to a period i and the return on national health quality, NHQ_i. This functional form is used in our final model on the allocation of resources across time periods, which will be considered in Chapter 9.

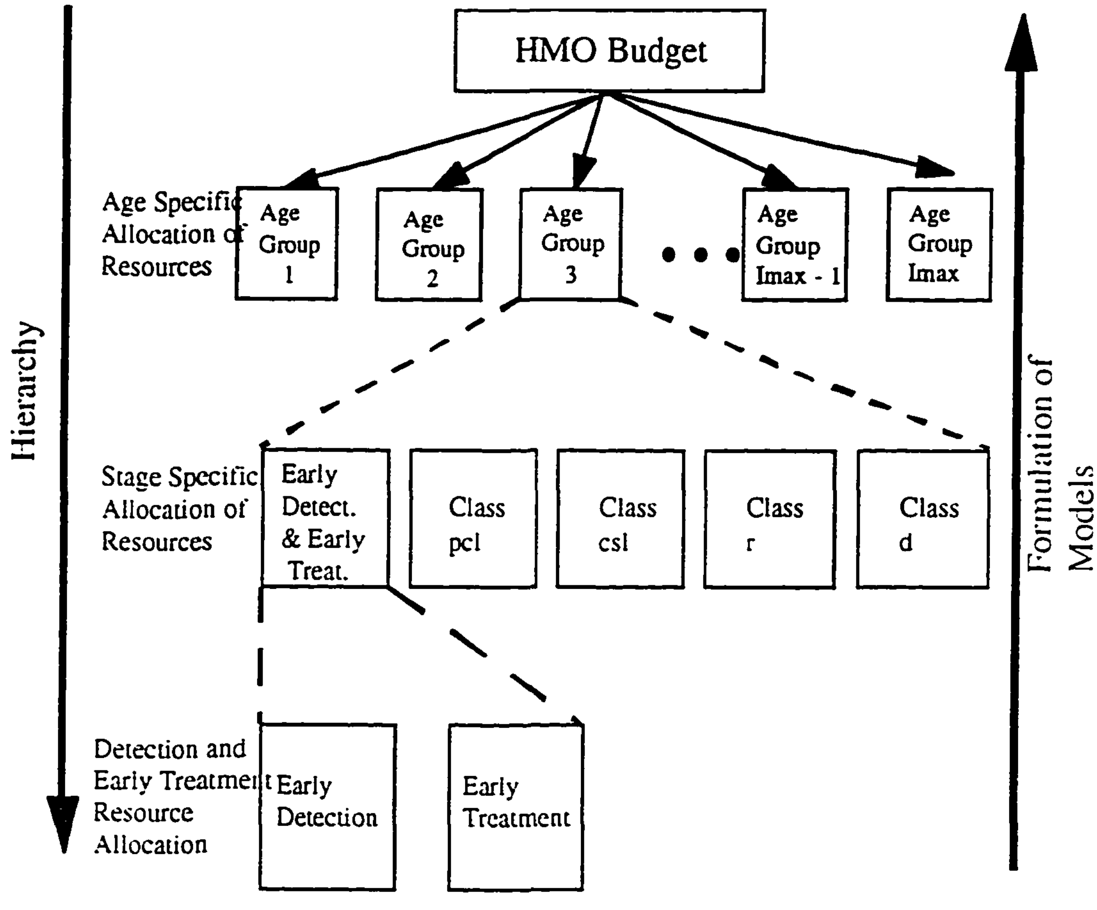

Figure 7.1. Resource allocation hierachy of decisions in managed care systems

7.2. The Mass Screening - Early Treatment Model

The detection of pre-clinical breast cancer requires testing both the healthy and asymptomatic latently ill populations. Without loss of generality, we presume a population with characteristics (e.g. age) such that the rates of healthy and latently ill are generated as λ_i^h and λ_i^{pcl}. Intervention measures are applied through the health care system, and specifically by the early detection and treatment centers (nodes EDC and ETS in Figure 1), upon these populations to detect and cure the latently ill. This process is illustrated in Figure 7.2.

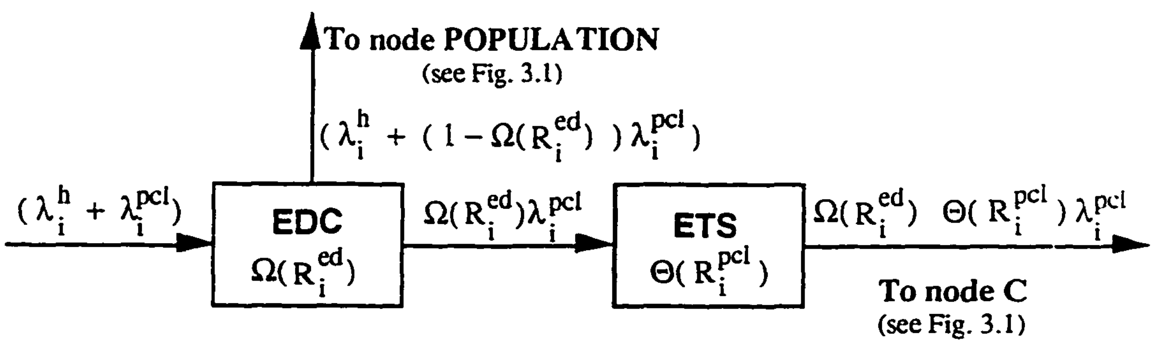

Figure 7.2: Traffic Rates Through Nodes EDC and ETS

Traffic into node EDC has rate $(\lambda_i^h+\lambda_i^{pcl})$ by the superposition principle of Poisson processes (Çinlar, 1975). Traffic from node EDC to node ETS has rate $\Omega(R_i^{ed})\bullet\lambda_i^{pcl}$ and is dependent upon the per-patient resource, R_i^{ed}, allocated to node EDC and the screening efficacy $\Omega(R_i^{ed})$. Of the traffic into node ETS, the ultimate removal of class pcl individuals depends upon the per-capita allocation of resources to ETS and the consequent return $\Theta(R_i^{pcl})$. Thus the fraction of latently ill individuals presented at node EDC and whom are cured after node ETS is given by the product $\Omega(R_i^{ed})\bullet\Theta(R_i^{pcl})$. Maximization of this function must be accomplished within the bounds of resources assigned to mass screening and early treatment. This is the task of the following non-linear mathematical program:

$$\text{(MSRA)} \quad \underset{R_i^{edc},R_i^{pcl}}{Max} \quad \Omega\left(R_i^{edc}\right)\bullet\Theta\left(R_i^{pcl}\right)$$

s.t.

$$(\lambda_i^h + \lambda_i^{pcl})R_i^{edc} + \Omega(R_i^{edc})\lambda_i^{pcl}R_i^{pcl} = R_i^{edts} \qquad \forall i \qquad (7.2.1)$$

$$0 \leq \Theta(R_i^{pcl}) \leq 1 \qquad \forall i \qquad (7.2.2)$$

$$0 \leq \Omega(R_i^{edc}) \leq 1 \qquad \forall i \qquad (7.2.3)$$

$$R_i^{edc}, R_i^{pcl}, R_i^{edts} \geq 0 \qquad \forall i \qquad (7.2.4)$$

The resource constraint is (7.2.1) and states that the total cost of testing both the healthy and latently ill arrivals plus the cost of treating the detected latently ill must equal the assigned resource amount, R_i^{edts}. Constraints (7.2.2) and (7.2.3) follow from the modelling of the resource-quality functions $\Omega\left(R_i^{edc}\right)$ and $\Theta\left(R_i^{ets}\right)$ as strictly concave non-decreasing functions bounded between 0 and 1. Constraint (7.2.4) is the non-negativity constraint.

The functional forms $\Omega(R_i^{ed})$ and $\Theta(R_i^{pcl})$ used in the solution of program MSRA are given as:

$$\Omega\left(R_i^{edc}\right)=\left(\frac{R_i^{edc}}{w^{edc}}\right)^{1/3} \text{ and } \Theta\left(R_i^{ets}\right)=\left(\frac{R_i^{pcl}}{w^{pcl}}\right)^{1/4} \qquad (7.2.5)\text{-}(7.2.6)$$

The solution is summarized in the following

Proposition 7.1. For a population with characteristics such that λ_i^h and λ_i^{pcl} are age group i class h and class pcl generated rates, and with an allocation to early detection and treatment of R_i^{edts}, the maximum proportion of latently ill detected and cured is given by:

$$\Xi\left(R_i^{edts}\right)=\left(\frac{1}{w^{edc}w^{pcl}}\right)^{1/4}\sqrt{\frac{R_i^{edts}}{2}}\left(\frac{1}{(\lambda_i^h+\lambda_i^{pcl})\lambda_i^{pcl}}\right)^{1/4} \qquad (7.2.7)$$

Further, the optimal per-capita allocation of resources between detection and treatment are given by:

$$R_i^{ed}=\left(\frac{R_i^{edts}}{2(\lambda_i^h+\lambda_i^{pcl})}\right) \qquad (7.2.8)$$

$$R_i^{pcl}=\left(\frac{w^{ed}}{4}\right)^{1/3}\left(\frac{(\lambda_i^h+\lambda_i^{pcl})^{1/3}}{\lambda_i^{pcl}}\right)(R_i^{edts})^{2/3} \qquad (7.2.9)$$

Proof: (Outline). Rearranging terms in the constraint (7.2.1), we obtain:

$$R_i^{pcl}=\left(\frac{R_i^{edts}-(\lambda_i^h+\lambda_i^{pcl})R_i^{ed}}{\lambda_i^{pcl}(R_i^{ed}/w^{ed})^{1/3}}\right) \qquad (7.2.1')$$

Substituting (7.2.1') into the objective function of MSRA, and using first order conditions, we obtain the results of the proposition.

Proposition 7.1 yields the following principal insights: The first is to give the functional dependence of the efficacy of an early detection and treatment on the relative ratios of latently ill to healthy persons in the population. In Figure 7.2 below we plot the return on resources allocated to mass screening and the effectiveness of a mass screening program for several ratios of latently ill to healthy. The figure displays that the efficacy of

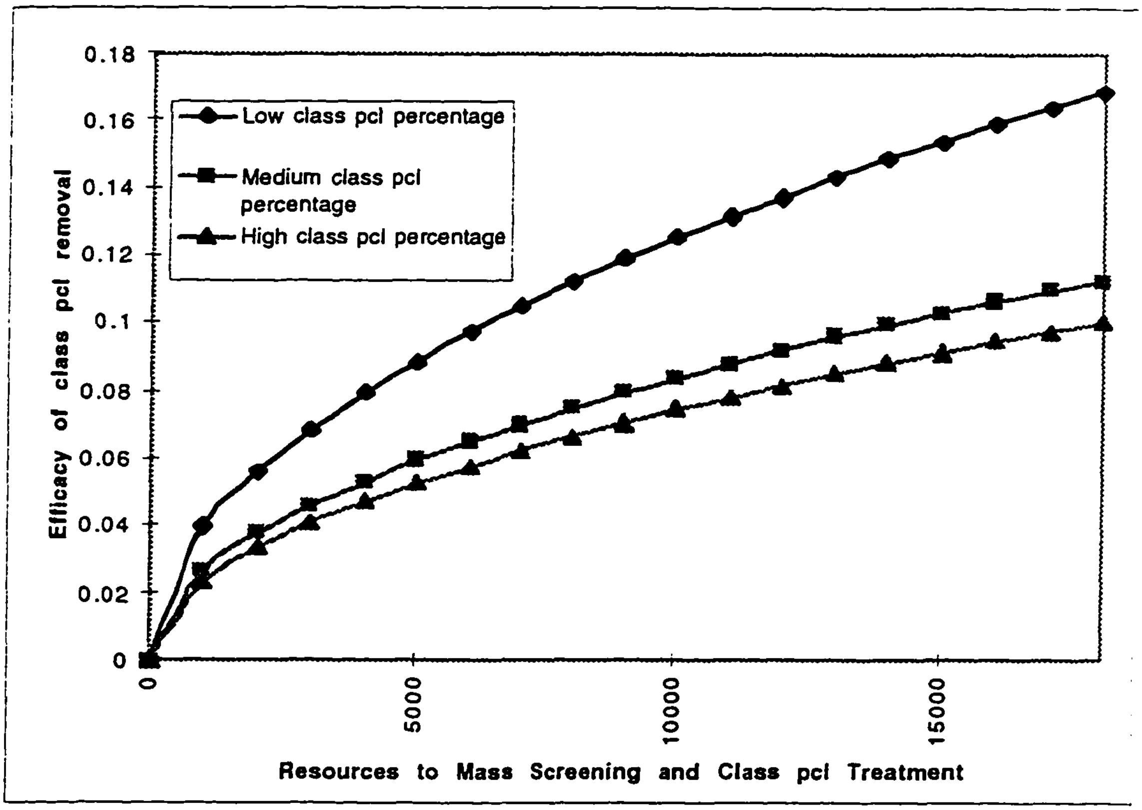

Figure 7.3. Efficacy of Early Detection and Treatment by Population Characteristics

detection and early treatment of breast cancer, as a function of resources expended, is inversely related to the population density of persons latently ill with breast cancer. This result is formalized in the following

Corrollary 7.2: For a fixed i th age group resource allocation, R_i^{edts}, the efficacy of detection and cure of pre-clinical breast cancer is subject to the following property: if $(\lambda_i^{pcl}, \Xi(R_i^{edts}))$, and $(\lambda_i{}^{*pcl}, \Xi^*(R_i^{edts}))$ are ordered pairs, and if $|\lambda_i^{pcl}| > |\lambda_i{}^{*pcl}|$, then $\Xi(R_i^{edts}) < \Xi^*(R_i^{edts})$.

Proof: Follows immediately from the functional form of $\Xi(R_i^{edts})$ given in Equation (7.2.7).

Figure 7.2 and Corrollary 7.2 yield that the efficacy of detection and cure is inversely proportional to the ratio of ith age group class *pcl* to class *h* individuals. The result follows by noting that $\Omega(R_i^{edc}) \bullet \lambda_i^{pcl} > \Omega(R_i^{edc}) \bullet \lambda_i^{pcl*}$ for $\lambda_i^{pcl} > \lambda_i^{pcl*}$. Consequently, with a greater number of arrivals at the early treatment center, ETS, the available per-patient resources and quality of treatment is decreased. In Corrollary 7.3, we show that even though the efficacy of detection and cure decreases with an increasing proportion of pre-clinical local stage individuals in the population, the absolute numbers of class pcl persons successfully detected and cured is an increasing function of the same.

Corrollary 7.3: While the efficacy of detection and cure is inversely proportional to the ratio of class *pcl* to class *h* individuals, $\lambda_i^{pcl}/(\lambda_i^h+\lambda_i^{pcl})$, the absolute numbers of class *pcl* individuals screened and cured is directly proportional to the ratio $\lambda_i^{pcl}/(\lambda_i^h+\lambda_i^{pcl})$.

Proof: For $\lambda_i^h+\lambda_i^{pcl} = \lambda_i^{h*}+\lambda_i^{pcl*}$, and $\lambda_i^{pcl} > \lambda_i^{pcl*}$, inspection of Equation (7.2.7) yields $\Xi(R_i^{edts})\bullet\lambda_i^{pcl} > \Xi(R_i^{edts})\bullet\lambda_i^{pcl*}$ ■

Figure 7.3 allows us to make the following observation: For two different decisions on the resources allocated to early detection and treatment in period i, we obtain simply

$$R_i^{edts} > R*_i^{edts} \Rightarrow \frac{\lambda_{i+1}^{pcl}}{\lambda_{i+1}^{h}} > \frac{\lambda*_{i+1}^{pcl}}{\lambda*_{i+1}^{h}} \qquad (7.2.10)$$

That is, a specific resource allocation to the combined node EDTS affects the ratio of latently ill to healthy in the next period with the property given in Equation (7.2.10). Further, a greater expenditure on early detection and treatment for a population at a certain age group will decrease the efficacy of detection and early treatment when the same population has aged to the next age group. Consequently, the efficacy of early detection-treatment as a function of allocated resources is strongly dependent upon decisions made in previous periods. It does not suffice to assume a fixed response curve when seeking to conduct a mass screening program through time.

The second insight provided by Proposition 7.1 is summarised in the following corrolary:

Corrolary 7.4. For an assigned budget R^{edts} to the screening and early treatment sectors, the optimal allocation of resources to each of the sectors, independent of the access level, is $R^{edts}/2$.

Proof: The budgets at each of the sectors is given by the per - patient expenditures multiplied by the traffic to the sectors. The proof is demonstrated if taking the ratio of the budgets at the screening and early treatment sectors yields that the ratio is equal to 1. Thereby,

$$\frac{(\lambda_i^h+\lambda_i^{pcl})\left(\dfrac{R_i^{edts}}{2(\lambda_i^h+\lambda_i^{pcl})}\right)}{\left(\dfrac{R_i^{edts}}{2(\lambda_i^h+\lambda_i^{pcl})}\right)^{1/3}\left(\dfrac{1}{w^{edc}}\right)^{1/3}\lambda_i^{pcl}\left(\dfrac{w^{ed}}{4}\right)^{1/3}\left(\dfrac{(\lambda_i^h+\lambda_i^{pcl})^{1/3}}{\lambda_i^{pcl}}\right)(R_i^{edts})^{2/3}}=\frac{2R^{edts}}{2R^{edts}}=1.$$

The managerial consequence of Corrollary 7.4 is that it simplifies the process of optimally allocating resources between screening and early treatment, i.e.simply divide the available resources in two and equally allocate between the two sectors. The allocation results primarily because independent of access level, and independent of the ratio of pre-clinical to healthy persons that are presented for screening, the allocation of resources to early screening is $R^{edts}/2$. Stated differently, regardless of the access level and of the ratio of pre clinical to healthy, the same amount is optimally spent on screening. The monetary compensation for a higher ratio of pre - clinical cases is made in their early treatment -- a higher concentration of pre clinical resources means that detected cases will be allocated fewer resources in treatment.

Equation (7.2.7) provides the functional response, $\Xi(R_i^{edts})$, of resources allocated to mass screening-early detection. This functional response is used in the next model we present to determine the resources, R_i^{edts}, that should be allocated to mass screening-early treatment, as well as the resources that should be allocated to the other components of the health care system.

7.3. Distribution of Resources Among Health Care Centers

Following the reverse hierarchical development of our models, in this section we assume an allocation, to a age group i, of a fixed resource amount, R_i^{tot}. Further, we assume that the health status of individuals entering from the previous period into period i is encapsulated by the vector $A_{i-1} = (\lambda_{i-1}^{h}, \lambda_{i-1}^{pcl}, \lambda_{i-1}^{csl}, \lambda_{i-1}^{r}, \lambda_{i-1}^{d}, \lambda_{i-1}^{po}, \lambda_{i-1}^{c})$. We ask how available resources should be allocated among the treatment, mass screening, and early detection centers in Figure 7.1 so that given $\mathbf{A}_{i-1}$, the national health quality of individuals in age group i, NHQ_i, is maximized. From Chapter 5 we note that the mean numbers of persons within a health class is given by

$$N_i^v = \lambda_i^v E[S_i^v]; \qquad v \in \mathbf{HC}, \qquad i \in \mathbf{I} \qquad (7.3.1)$$

With the constancy of $E[S_i^v]$, the maximization of N_i^v is equivalent to the maximization of λ_i^v, $v \in \mathbf{HC}$. Thus, the maximization of NHQ_i is equivalently obtained by the following mathematical program HCSRA:

$$HCSRA = \underset{R_i^x, x=edc,pcl,csl,r,d}{Max} \quad M_1\lambda_i^h + M_1\gamma_i^c + M_2\lambda_i^{pcl} + M_3\lambda_i^{csl} + M_4\lambda_i^r + M_5\lambda_i^d + M_6\lambda_i^{po}$$

s.t.

Equations (5.3.2) - (5.3.8) $\qquad \forall i \qquad (7.3.2)$

$$[\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl} R_i^{csl} +$$

$$[\alpha + (1-\alpha)\kappa^{r}]\lambda_{i-1}^{r} R_i^{r} + \lambda_{i-1}^{d} R_i^{d} + R_i^{edts} = R_i^{tot} \quad \forall i \qquad (7.3.3)$$

$$0 \le \Psi(\bullet) \le 1 \qquad (7.3.4)$$

$$0 \le \Phi(\bullet) \le 1 \qquad (7.3.5)$$

$$0 \le \xi(\bullet) \le 1 \qquad (7.3.6)$$

The objective in problem HCSRA is obtained as a weighted sum of components in the health status vector H_i. The weights $M_1, ..., M_6$ reveal the relative desirability of the different health classes, as outlined in Chapter 5. Constraints (7.3.2) capture the derivation of the vector $\mathbf{AR}_i$ as a function of vector $\mathbf{AR}_{i-1}$, and of a particular resource allocation instance $\mathbf{Q}_i = (R_i^{edts}, R_i^{csl}, R_i^{r}, R_i^{d})$ across treatment centers. Constraint (7.3.3) is the resource constraint. It gives that the sum of treatment center budgets must equal period i resource allocation, R_i^{tot}. Constraints (7.3.4)-(7.3.6) express the bounds on the resource quality functions.

The problem HCSRA is solved using functional form $\Xi\left(R_i^{edtc}\right)$ obtained from Proposition 7.1 and Equation 7.2.7. In addition, to provide analytical tractability, the resource quality functions for the treatment of class csl, r, and d patients are given as

$$\chi\left(R_i^{y}\right) = \sqrt{\frac{R_i^{y}}{w^{y}}}, \qquad (7.3.7) - (7.3.9)$$

where $\chi = \Psi, \Phi, \xi$, y = (csl, r, d) and the constants w^y ensure compliance with constraints (7.3.4) - (7.3.6). The solution is summarized in the following

Proposition 7.4. The optimal per-capita division of resources among the treatment centers is:

$$R_i^z = \frac{A_i^z \cdot R_i^{tot}}{u} \quad z=(\text{edts, csl, r,d}) \qquad (7.3.10)\text{-}(7.3.13)$$

In addition, the optimal age group specific budget expenditure rate (derived as the traffic to the treatment centers multiplied by the per-patient expenditures) are given as:

$$\text{Mass screening budget} = \alpha\varepsilon\left(\lambda_{i-1}^h + \lambda_{i-1}^{pcl}\right) \bullet R_i^{edc} \qquad (7.3.14)$$

$$\text{Early treatment or stage pcl budget} = \alpha\varepsilon\Omega\left(R_i^{edc}\right)\lambda_{i-1}^{pcl} \bullet R_i^{ets} \qquad (7.3.15)$$

$$\text{Stage csl (r) budget} = (\alpha + (1-\alpha)\kappa^{csl(r)})\lambda_{i-1}^{csl(r)} \bullet R_i^{csl(r)} \qquad (7.3.16)$$

$$\text{Stage d budget} = \lambda_{i-1}^d \bullet R_{i-1}^d \qquad (7.3.17)$$

Finally, the maximum return on national health quality for a single period expenditure of R_i^{tot} is given as:

$$\text{NHQ}_i\left(R_i^{tot}\right) = B_{i-1}^h\lambda_{i-1}^h + B_{i-1}^{pcl}\lambda_{i-1}^{pcl} + B_{i-1}^{csl}\lambda_{i-1}^{csl} + B_{i-1}^r\lambda_{i-1}^r + B_{i-1}^d\lambda_{i-1}^d + \left(4R_i^{tot}u\right)^{1/2} \qquad (7.3.18)$$

where

$$A_i^{edts} = \frac{\alpha^2\varepsilon^2}{8}\left(\frac{1}{w^{edc}w^{pcl}}\right)^{1/2}\left[M_1 - B_{i-1}^{pcl}\right]^2\left(\frac{(\lambda_{i-1}^{pcl})^3}{\lambda_{i-1}^n + \lambda_{i-1}^{pcl}}\right)^{1/2} \qquad (7.3.19)$$

$$A_{i-1}^y = \frac{1}{4w^y}\left[M_1 - B_{i-1}^y\right]^2 \quad y=(\text{csl,r,d}) \qquad (7.3.20)\text{-}(7.3.22)$$

Making the following identification: (h, pcl, csl, r, d, po) --> (1, 2, 3, 4, 5, 6) we obtain

$$B_{i-1}^{k} = M_k\left(1-\delta^k\right) + \sum_{l=k+1}^{6} M_l \delta^{k,l} \qquad (7.3.23)\text{-}(7.3.27)$$

for k = (1, 2, 3, 4, 5, 6). Note $\left(B_{i-1}^{h}, B_{i-1}^{pcl}, B_{i-1}^{csl}, B_{i-1}^{r}, B_{i-1}^{d}\right) = \left(B_{i-1}^{1}, B_{i-1}^{2}, B_{i-1}^{3}, B_{i-1}^{4}, B_{i-1}^{5}\right)$.

The variable u is obtained as

$$u = A_{i-1}^{edts} + A_{i-1}^{csl}\left(\alpha+(1-\alpha)\kappa^{csl}\right)\lambda_{i-1}^{csl} + A_{i-1}^{r}\left(\alpha+(1-\alpha)\kappa^{r}\right)\lambda_{i-1}^{r} + A_{i-1}^{d}\lambda_{i-1}^{d}. \qquad (7.3.28)$$

Proof: Since the objective function of Program HCSRA is strictly concave, and the resource constraint is linear, we use the Lagrange method (K-T conditions) to find the global maximum. Thereby, the Lagrangian is

$$\begin{aligned} L = C_1 &+ \alpha\varepsilon\Xi(R_{i-1}^{edts})[M_1 - B_{i-1}^{pcl}]\lambda_{i-1}^{pcl} + [\alpha+(1-\alpha)\kappa^{csl}]\Psi(R_{i-1}^{csl})[M_1 - B_{i-1}^{csl}]\lambda_{i-1}^{csl} + \\ &[\alpha+(1-\alpha)\kappa^{r}]\Phi(R_{i-1}^{r})[M_1 - B_{i-1}^{r}]\lambda_{i-1}^{r} + \xi(R_{i-1}^{d})[M_1 - B_{i-1}^{d}]\lambda_{i-1}^{d} + \\ &\mu\left[[\alpha+(1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl}R_i^{csl} + [\alpha+(1-\alpha)\kappa^{r}]\lambda_{i-1}^{r}R_i^{r} + \lambda_{i-1}^{d}R_i^{d} + R_i^{edts} - R_i^{tot}\right] \end{aligned} \qquad (7.3.29)$$

where $C_1 = B_{i-1}^{h}\lambda_{i-1}^{h} + B_{i-1}^{pcl}\lambda_{i-1}^{pcl} + B_{i-1}^{csl}\lambda_{i-1}^{csl} + B_{i-1}^{r}\lambda_{i-1}^{r} + B_{i-1}^{d}\lambda_{i-1}^{d}$

Taking the gradient of the Lagrangian with respect to R_i^{edts}, R_i^{csl}, R_i^{r}, R_i^{d}, and μ, and setting it to zero, we obtain

$$\nabla L = \begin{bmatrix} (1/2\sqrt{2})\alpha\varepsilon[M_1 - B_2]\lambda_{i-1}^{pcl}[w^{edc}w^{pcl}(\lambda_{i-1}^{h} + \lambda_{i-1}^{pcl})\lambda_{i-1}^{pcl}]^{-1/4}(R_{i-1}^{edts})^{-1/2} + \mu \\ [\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl}\left[(1/2)[M_1 - B_3](w^{csl}R_{i-1}^{csl})^{-1/2} + \mu\right] \\ [\alpha + (1-\alpha)\kappa^{r}]\lambda_{i-1}^{r}\left[(1/2)[M_1 - B_4](w^{r}R_{i-1}^{r})^{-1/2} + \mu\right] \\ \left[(1/2)[M_1 - B_5](w^{d}R_{i-1}^{d})^{-1/2} + \mu\right]\lambda_{i-1}^{d} \\ [\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl}R_i^{csl} + [\alpha + (1-\alpha)\kappa^{r}]\lambda_{i-1}^{r}R_i^{r} + \lambda_{i-1}^{d}R_i^{d} + R_i^{edts} - R_i^{tot} \end{bmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{pmatrix}$$

Simultaneous solution of Equations (7.3.30)-(7.3.34) and setting $1/\mu^2 = R_{i-1}^{tot}/u$, we obtain the results of the proposition.

In summary, this section provides the optimal age group i resource allocation (i.e. budget) policies if the objective is solely the societal goal of maximizing the health index of the population. The results of this section will be used, in Chapter 9, in the final hierarchy of decisions, i.e. the allocation of resources across age-groups.

Chapter 8

The Service Quality Resource Allocation Models

8.1. Introduction

This chapter develops optimization models that serve to maximize the service quality available in the health care system, subject to available resources. The models are developed similarly to the optimization models of Chapter 7 and consider the same hierachy of resource allocation decisions. That is, we first develop an i th age group service quality return on resources allocated to early screening and treatment. We shall refer to this functional as SQ_i^{edts}. Second, using SQ_i^{edts}, we develop the return on service quality across all health care treatment centers for a given age group resource allocation. We will refer to this service functional as SQ_i^{TC}.

In order to obtain specific resource allocations that maximize service quality, it behooves us first to define the appropriate service quality value funtion. In the models for CQA (Chapters 2 to 6) we captured service quality via the resource quality functions $\Omega(R_i^{edc})$ for the detection of pre-clinical breast cancer, and $\vartheta(R_i^x)$ where $\vartheta = (\Theta, \Psi, \Phi, \xi)$ and x=(*pcl*, *csl*, *r*, *d*) for the treatment of breast cancer. The fundamental property of the resource-quality functions of interest is the following: if $\left(R_i^x, \vartheta(R_i^x)\right)$ and $\left(R_i^{x*}, \vartheta(R_i^{x*})\right)$ are ordered pairs that define two instances of the resource quality functions, and if $\vartheta(R_i^x) > \vartheta(R_i^{x*})$, then $\vartheta(R_i^x) > \vartheta(R_i^{x*})$. In each of the maximization problems below we will define an appropriate service quality value function that will maintain the fundamental property of interest. However, the choice of the value function will be such as to render analytical solutions in each of the optimization problems.

8.2. The Maximization of Early Detection and Treatment Service Quality

The maximization of early detection and treatment service quality is subject to the resource constraint specified in Equation (7.2.1). We choose the product form $\Omega(S_i^{edc}) \bullet \Theta(S_i^{pcl})$ as the service quality value function. Consequently, we obtain:

$$\text{(MSQRA)} \quad \underset{S_i^{edc}, S_i^{pcl}}{Max} \; \Omega\left(S_i^{edc}\right) \bullet \Theta\left(S_i^{pcl}\right)$$

s.t.

$$(\lambda_i^h + \lambda_i^{pcl})S_i^{edc} + \Omega(S_i^{edc})\lambda_i^{pcl}S_i^{pcl} = S_i^{edts} \qquad \forall i \qquad (8.2.1)$$

$$0 \leq \Theta(S_i^{pcl}) \leq 1 \qquad \forall i \qquad (8.2.2)$$

$$0 \leq \Omega(S_i^{edc}) \leq 1 \qquad \forall i \qquad (8.2.3)$$

$$S_i^{edc}, S_i^{pcl}, S_i^{edts} \geq 0 \qquad \forall i \qquad (8.2.4)$$

Whence, we observe that problem MSQRA is the same as problem MSRA, leading to the following

Proposition 8.1. For a population with characteristics such that λ_i^h and λ_i^{pcl} are age group i class h and class pcl generated rates, and with an allocation to early detection and treatment of S_i^{edts}, the optimal edts service quality value function is of the form:

$$SQ_i^{edts}\left(R_i^{edts}\right) = \left(\frac{1}{w^{ed}w^{pcl}}\right)^{1/4} \sqrt{\frac{R_i^{edts}}{2}} \left(\frac{1}{(\lambda_i^h + \lambda_i^{pcl})\lambda_i^{pcl}}\right)^{1/4} \qquad (8.2.5)$$

Further, the maximization of edts service quality is obtained under the following division of resources between detection and treatment:

$$S_i^{edc} = \left(\frac{S_i^{edts}}{2(\lambda_i^h + \lambda_i^{pcl})} \right) \tag{8.2.6}$$

$$S_i^{pcl} = \left(\frac{w^{ed}}{4} \right)^{1/3} \left(\frac{(\lambda_i^h + \lambda_i^{pcl})^{1/3}}{\lambda_i^{pcl}} \right) (S_i^{edts})^{2/3} \tag{8.2.7}$$

Proof: See accompanying proof to Proposition 7.1.

The results of Proposition 8.1 mirror the results of Proposition 7.1, thereby we obtain the following corrollaries:

Corrollary 8.2: For a fixed i th age group resource allocation, S_i^{edts}, the efficacy of detection and cure of pre-clinical breast cancer is subject to the following property: if (λ_i^{pcl} , $SQ_i^{edts}(S_i^{edts})$), and (λ_i^{*pcl} , $SQ_i^{*edts}(S_i^{edts})$) are ordered pairs, and if | λ_i^{pcl} | > | λ_i^{*pcl} |, then $\Xi(R_i^{edts}) < \Xi^*(R_i^{edts})$.

Proof: See accompanying proof to Corrollary 7.2.

Corrollary 8.3: While the efficacy of detection and cure is inversely proportional to the ratio of class *pcl* to class *h* individuals, $\lambda_i^{pcl}/(\lambda_i^h+\lambda_i^{pcl})$, the absolute numbers of class *pcl* individuals screened and cured is directly proportional to the ratio $\lambda_i^{pcl}/(\lambda_i^h+\lambda_i^{pcl})$.

Proof: See accompanying proof to Corrollary 7.3.

Equation (8.2.6) provides the functional response, $SQ_i^{edts}\left(S_i^{edtc}\right)$, of resources allocated to mass screening-early detection. This functional response is used in the next model we present to determine the resources, S_i^{edtc}, that should be allocated to mass screening-early treatment, as well as the resources that should be allocated to the other components of the health care system.

8.3. Optimal Service Quality Distribution of Resources Among Health Care Centers

In this section we develop a mathematical program to allocate resources among the treatment, mass screening, and early detection centers such that a maximal return on service quality among all the treatment centers, SQ_i^{TC}, is obtained. In the derivation of SQ_i^{TC} we will assume that the health status of individuals entering from the previous period is encapsulated by the vector $\mathbf{H}_{i-1} = \left(\lambda_{i-1}^{h}, \lambda_{i-1}^{pcl}, \lambda_{i-1}^{csl}, \lambda_{i-1}^{r}, \lambda_{i-1}^{d}, \lambda_{i-1}^{p}, \gamma_{i-1}^{c}\right)$. We will obtain the service quality functional by assuming, as in Section 7.3, an i th age group allocation of a fixed resource amount, S_i^{tot}. Further, we will choose the summation of resource-quality functions to serves as the service quality value function. The maximization of SQ_i^{TC} is obtained by the following mathematical program MTCSQ:

$$MTCSQ = \underset{S_i^x, x=edts,pcl,csl,r,d}{Max} \quad SQ_i^{edts}(S_i^{edts}) + \Psi(S_i^{csl}) + \Phi(S_i^{r}) + \xi(S_i^{d})$$

s.t.

$$[\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl} S_i^{csl} +$$

$$[\alpha+(1-\alpha)\kappa^r]\lambda^r_{i-1}S^r_i+\lambda^d_{i-1}S^d_i+S^{edts}_i=S^{tot}_i \qquad \forall I \tag{8.3.1}$$

$$0\le SQ^{edts}_i(\bullet)\le 1 \tag{8.3.2}$$

$$0\le \Psi(\bullet)\le 1 \tag{8.3.3}$$

$$0\le \Phi(\bullet)\le 1 \tag{8.3.4}$$

$$0\le \xi(\bullet)\le 1 \tag{8.3.5}$$

The objective in problem MTCSQ is the combination of resource-quality functions across the treatment centers. Constraint (8.3.2) is the resource constraint. It gives that the sum of treatment center budgets must equal age group i resource allocation. The problem MTCSQ is solved using functional forms described in Equations (7.3.7)-(7.3.9). Further, the functional form, $SQ^{edts}_i\left(S^{edtc}_i\right)$, is obtained from Proposition 8.1 and Equation 8.2.5. The solution is summarized in the following

Proposition 8.4. To maximize the treatment center service quality value function, the optimal per-capita division of resources among the treatment centers is:

$$S^{edts}_i=\frac{1}{2}\left(\frac{1}{w^{edc}w^{pcl}}\right)^{1/2}\left(\frac{1}{(\lambda^h_{i-1}+\lambda^{pcl}_{i-1})\lambda^{pcl}_{i-1}}\right)^{1/2}\frac{S^{tot}_i}{z^2}; \tag{8.3.5}$$

$$S^{csl}_i=\frac{1}{w^{csl}}\left(\frac{1}{[\alpha+(1-\alpha)\kappa^{csl}]\lambda^{csl}_{i-1}}\right)^2\frac{S^{tot}_i}{z^2}; \tag{8.3.6}$$

$$S^r_i=\frac{1}{w^r}\left(\frac{1}{[\alpha+(1-\alpha)\kappa^r]\lambda^r_{i-1}}\right)^2\frac{S^{tot}_i}{z^2}; \tag{8.3.7}$$

$$S^d_i=\frac{1}{w^d}\left(\frac{1}{\lambda^d_{i-1}}\right)^2\frac{S^{tot}_i}{z^2} \tag{8.3.8}$$

Further the maximum return on service quality for a single period expenditure of S_i^{tot} is given as:

$$\mathrm{SQ}_i^{TC}\left(S_i^{tot}\right) = z \cdot (S_i^{tot})^{1/2} \tag{8.3.9}$$

where

$$z^2 = \left(\frac{1}{4w^{edc}w^{pcl}(\lambda_{i-1}^{h} + \lambda_{i-1}^{pcl})\lambda_{i-1}^{pcl}}\right)^{1/2} + \left(\frac{1}{w^{csl}[\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl}}\right) + \left(\frac{1}{w^{r}[\alpha + (1-\alpha)\kappa^{r}]\lambda_{i-1}^{r}}\right) + \left(\frac{1}{w^{d}\lambda_{i-1}^{d}}\right) \tag{8.3.10}$$

Proof: Since the objective function of Program HCSRA is strictly concave, and the resource constraint is linear, we use the Lagrange method (K-T conditions) to find the global maximum. With the multiplier v the Lagrangian is

$$L = \left(\frac{1}{w^{edc}w^{pcl}}\right)^{1/4}\left(\frac{1}{(\lambda_i^{h} + \lambda_i^{pcl})\lambda_i^{pcl}}\right)^{1/4}\sqrt{\frac{S_i^{edts}}{2}} + \sqrt{\frac{S_i^{csl}}{w^{csl}}} + \sqrt{\frac{S_i^{r}}{w^{r}}} + \sqrt{\frac{S_i^{d}}{w^{d}}} - v\left[S_i^{edts} + \lambda_{i-1}^{d} S_i^{d}\right]$$
$$- v\left[[\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl} \cdot S_i^{csl} + [\alpha + (1-\alpha)\kappa^{r}]\lambda_{i-1}^{r} \cdot S_i^{r} - S_i^{TOT}\right] \tag{8.3.11}$$

Taking the gradient of the Lagrangian with respect to S_i^{edts}, S_i^{csl}, S_i^{r}, S_i^{d}, and v, and setting it to zero, we obtain

$$\frac{1}{2\sqrt{2}}\left(\frac{1}{w^{edc}w^{pcl}}\right)^{1/4}\left(\frac{1}{(\lambda_{i-1}^{h} + \lambda_{i-1}^{pcl})\lambda_{i-1}^{pcl}}\right)^{1/4}\left(S_i^{edts}\right)^{-1/2} = v \tag{8.3.12}$$

$$\frac{1}{2\sqrt{w^{csl}}}\left(S_i^{csl}\right)^{-1/2} = v \cdot [\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl} \tag{8.3.13}$$

$$\frac{1}{2\sqrt{w^{r}}}\left(S_i^{r}\right)^{-1/2} = v \cdot [\alpha + (1-\alpha)\kappa^{r}]\lambda_{i-1}^{r} \tag{8.3.14}$$

$$\frac{1}{2\sqrt{w^{d}}}\left(S_i^{d}\right)^{-1/2} = v \cdot \lambda_{i-1}^{d} \tag{8.3.15}$$

$$S_i^{edis} + [\alpha + (1-\alpha)\kappa^{csl}]\lambda_{i-1}^{csl} \cdot S_i^{csl} + [\alpha + (1-\alpha)\kappa^{r}]\lambda_{i-1}^{r} \cdot S_i^{r} + \lambda_{i-1}^{d} \cdot S_i^{d} = S_i^{TOT} \qquad (8.3.16)$$

Simultaneous solution of Equations (8.3.12)-(8.3.16) yields the results of the propostion.

Having developed the optimal distribution of age-group i resources to maximize service quality, we note immediately that the national health, i.e. Equations (7.3.11) - (7.3.13), and service quality, i.e. Equations (8.3.6) - (8.3.8) optimal distributions differ. However, we are concerned with the relative effects of increased access upon the different resource allocations for the different types of quality, before managerial implications are discussed. The dynamics of the resource allocations with access is the subject of the next chapter.

Chapter 9

Resource Allocation Dynamics as a Function of Access

9.1. Introduction

Having established the optimal distribution of per-patient resources and the maximal return on national health quality, we explore the period i system dynamics as a function of the access level. This is the subject of the following two propositions:

Proposition 9.1. $NHQ(R_i^{tot})$, R_i^{edts}, and the budgets for early detection and treatment are concave functions of access level α; R_i^{csl}, R_i^{r}, and R_i^{d} and the budget for distant stage breast cancer are convex functions of access level.

Proof: It is sufficient to show $\delta NHQ/\delta\alpha > 0$, $\delta R_i^{edts}/\delta\alpha > 0$, $\delta R_i^{csl}/\delta\alpha < 0$, $\delta R_i^{r}/\delta\alpha < 0$ and $\delta R_i^{d}/\delta\alpha < 0$. We omit the details.

Corrollary 9.2. The treatment budgets for mass screening and early treatment, and for the treatment of class csl patients are concave functions of access, with treatment budgets for the treatment of class r and d patients convex functions of access.

Proof. We show $\delta B_i^{edts}/\delta\alpha > 0$, $\delta B_i^{csl}/\delta\alpha > 0$, $\delta B_i^{r}/\delta a < 0$, $dB_i^{d}/\delta\alpha < 0$. We omit the details.

From Proposition 9.1 we note that the optimal allocation of resources to maximize national health quality will increases with access. However, the cost will be decreased levels of quality of treatment available for class *csl*, and class *r* and class *d* individuals. Corrollary 9.2 yields that the decreased quality available for the disease classes is the result of a resource shift with increasing access to the early treatment centers. This resource shift is the result not just of increased traffic to the early treatment centers with increased access, but is also the result of greater per-patient expenditures on screening and early treatment. These results are displayed in Figure (9.1)

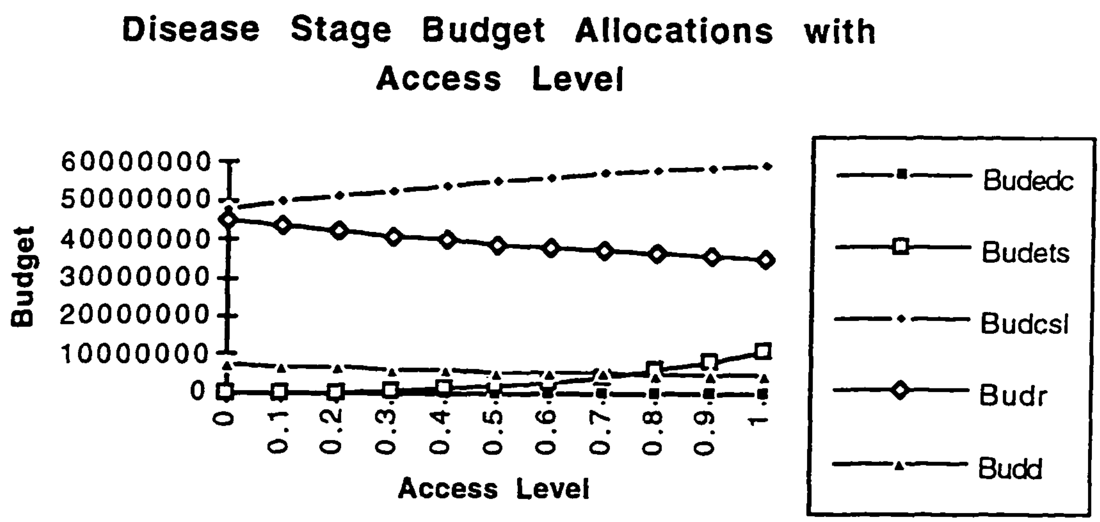

Figure 9.1: Disease Stage Budget Allocations with Access Level

Figure 9.1 indicates that change in resource requirements as the level of access is changed. Of note is that as the access level is increased, resources are shifted from the treatment of regional and distant stage patients to the detection of and treatment of pre-clinical breast cancer, and to the treatment of clinical breast cancer. Equation (7.2.14) explains that the budgetry rise in the resources devoted to early detection and cure of breast cancer is due not just to increased traffic to mass screening and early treatment, but also to an increased per-capita expenditure per patient. Equation (7.2.15)

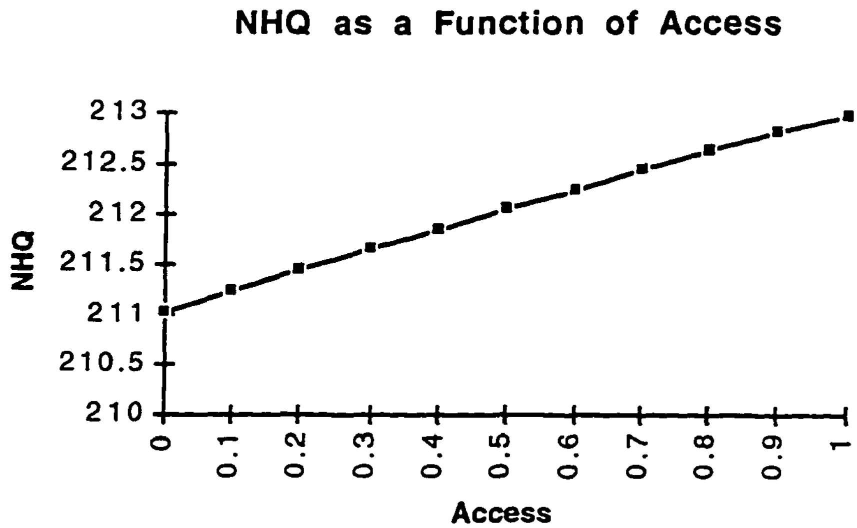

Figure 9.2: Societal Health as a Function of Access

indicates that the rise in resources with access for the treatment of class csl patients is due to the increased traffic more than compensating for the decreased per-patient resources. Similarly, the decline in resources for class r is due to the decline in allocated per-patient resources more than compensating for the increase in class r traffic. Finally, the decline in resources available to class d patients is solely due to the decreased per-patient resources, given that the distant stage traffic is unchanged.

Proposition 9.3: S_i^{edts} and S_i^d are concave functions of access level, α , whilst S_i^{csl}, S_i^r, and SQ_i^{TC} are convex functions of the same.

Proof. Taking first order derivatives with respect to α , we obtain $\delta S_i^{edts}/\delta\alpha > 0$, $\delta S_i^d/\delta\alpha > 0$ $\delta S_i^{csl}/\delta\alpha < 0$, $\delta S_i^r/\delta\alpha < 0$, and $\delta SQ_i^{TC}/\delta\alpha < 0$. Details are omitted.

The principal results of propositions 9.1 and 9.3 are that national health quality and service quality diverge with access within a given age group. The trade off mapping is given in Figure 9.3 where the convex combination $f(\alpha) = \gamma NHQ_i(\alpha) + (1-\gamma)SQ_i(\alpha)$ is plotted against access level.

Figure 9.3 indicates to decision makers the levels of quality available for a desired level of access. In particular, an increase of access from a particular level α* with an associated level of service quality SQ*, i.e. (α*, SQ*) may be increased to universal access, α = 1, and service quality service quality SQ*, i.e. (α = 1, SQ*) by

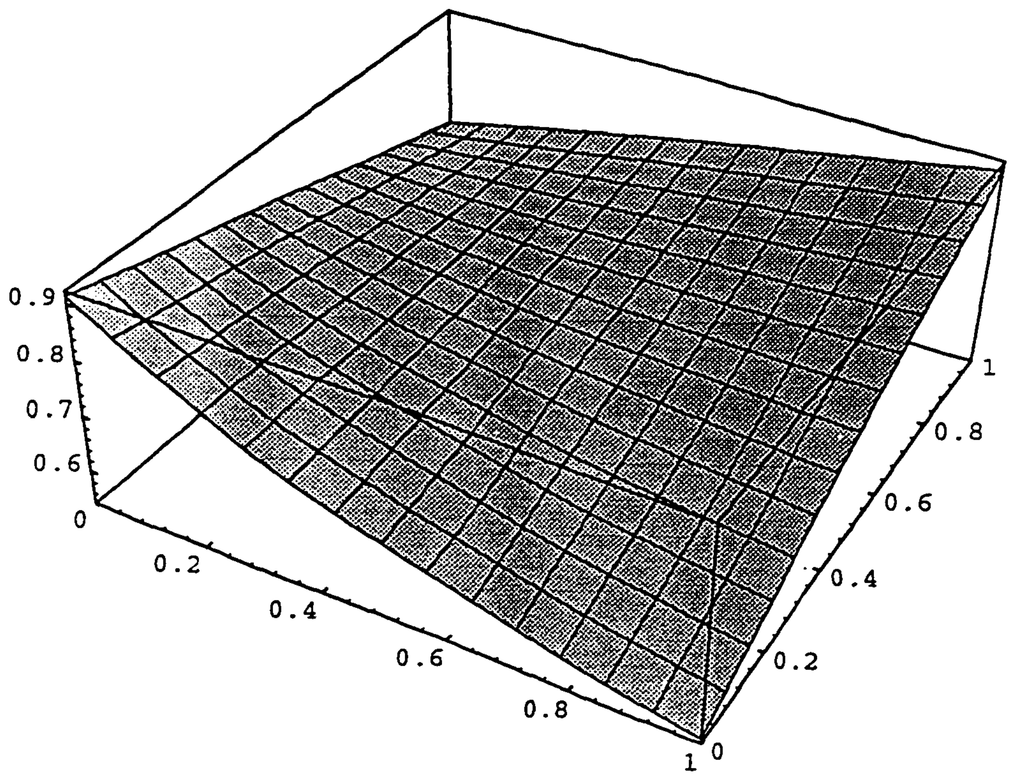

Figure 9.3: National Health and Service Quality Tradeoff as a Function of Access

reducing the level of national health quality appropriately. Resources may be allocated across the treatment centers to engineer the appropriate levels of national health and service quality to achieve full access.

9.2. The Multi-Period Allocation of Resources

The final task in the hierachy of decisions is the allocation of resources to each period, R_i^{tot}, for all periods such that the desired quantity of either national health quality, i.e. NHQ, or service quality, i.e. SQ, is maximized. In Chapters 7 and 8 we developed separate optimization models for NHQ and SQ. Our purpose was to understand the interplay between the optimal allocations as a function of access. In this section, however, we will obtain the optimal resource allocation for the convex combination, across age groups, of NHQ and SQ. Our reason is the following: Suppose we develop separate optimization models for the two types of quality. Suppose, further, that the optimal resource allocations for each age group and for each type of quality are different and are given by R_i^{NHQ} and R_i^{SQ}. The decision maker is asked, then, to choose a level of expenditure, R_i^{tot}, for each age group such that $R_i^{tot} \in (R_i^{NHQ}, R_i^{SQ})$. The possible resource allocation sample paths are infinite in number. The determination of the optimal resource allocation sample path may be determined by the following mathematical program:

$$(\text{DAR}) = \underset{R_i^{tot}}{Max} \left(\sum_{i=1}^{I} \gamma_i \text{NHQ}_i\left(R_i^{tot}\right)\right) + \left(\sum_{i=1}^{I} (1-\gamma_i) SQ_i^{TC}\left(R_i^{tot}\right)\right)$$

s.t.

Equations (5.3.2) - (5.3.8) (9.2.1) - (9.2.7)

$$\sum_{i=1}^{I} R_i^{tot} \leq R^{BC} \tag{9.2.8}$$

In problem (DAR), the objective function is the summation of the age-group specific convex combination of national health quality and service quality. Note that the cumulative measure of national health quality is the summation of the single period NHQ measures. Similarly, the value function we choose to represent the cumulative measure of service quality is the summation of the single period SQ measures. The return on national health quality for age group i, as a function of resources allocated to that age group, $NHQ_i(R_i^{tot})$, is obtained from Proposition 7.4 and Equation (7.3.18). Similarly, the age specific return on service quality as a function of allocated resources is obtained from Proposition 8.4 and Equation (8.3.9). Constraints (9.2.1) - (9.2.7) give the one-step evolution of the health status vector as a function of allocated resources. Constraint (9.2.8) recognizes budgetary restrictions on resources available to treat breast cancer.

Problem DAR can be solved as a two stage, one decision variable dynamic programming problem using a tree search. The breadth of the tree is specified by the number of stages s; the initial stage 0 node contains information on initial conditions at the begining of the time period of study. The algorithm for obtaining the resource allocations across stages is accomplished by the following tree search dynamic programming algorithm:

Algorithm: Optimal Resource Allocation by Successive Tree Search

1. Let the endowment be R. Choose s; obtain choice set of cardinality s+1. The choice of s determines the coarseness of the search. For computational purposes, a "grid" search may be initiated with succeedingly finer bounds on the solution.

2. List rewards in stage j+1 according to

$$f_{j+1}(H_j(i_{j+1} \cdot R / s, i_{j+1}), \forall i_{j+1} \in (0, R - \sum_{k=1}^{j} i_k \cdot R / s),$$

where $H_j(i_{j=1}R/s, i_{j+1})$ specifies the return on the health status vector given conditions H_{j-1} and the resource expenditure $i_{j-1} \bullet R/s$.

3. Obtain Max $\sum_{j=1}^{t} \sum_{i=1}^{j} f_{j+1}(H_j(i_{j+1} \cdot R / s, i_{j+1}), i)$

4. Trace the tree to obtain the resource allocations and health status vectors across stages and obtain the maximum expenditure across stages. Let the maximum expenditure be given by m.

5. Re-run the program for a maximum expenditure in each stage of m+(R/s). Include the priviso that the maximum expenditure across stages cannot exceed the endowment.

The algorithm will yield a successively finer search and may be terminated at the desired level of accuracy. We run the algorithm on the example data of Appendix Table 1 and the health status vector H_{init}=(96721, 1544, 796, 831, 107, 0, 0), and present the age group specific optimal allocation of resources in Table 6.2. It is to be noted that the dynamic programming algorithm is computationally expensive, with the number of calculations growing exponentially with the number of stages and with the coarseness of the search.

Age Group	1	2	3	4	5
Expenditure	$0.12*R^{BC}$	$0.2*R^{BC}$	$0.24*R^{BC}$	$0.2*R^{BC}$	$0.24*R^{BC}$

Table 9.1. Optimal Age-Group Specific Allocation of Resources.

In summary, this chapter provides decision makers with an expenditure schedule across age groups. Having obtained the age group specific resource allocations, the decision maker may appropriately allocate resources to treatment centers. This chapter completes the mathematical programming procedures to optimally allocate available single disease resources within a capitated budget environment.

Chapter 10

Summary and Conclusions

The backdrop to the research in this dissertation is constituted of two parts. The first is the continuing problems in the health care system, typified by i.) a rapid rise in health care costs, and ii.) a significant proportion of the population without access to medical care. The second part of the backdrop is constituted by the emerging trend of and possible future dominance of managed care as the primary form of health care delivery. In this dissertation, our purpose is to present a disease management approach to addressing the problems in the health care system, and to understanding how the evolution of the health care system can be best managed to meet specified goals.

The central tenet of this work is to note that three measures are generally accepted as defining the operational efficiency of a national health care system: cost, quality, and access. The foundation for the research in this dissertation has been to introduce a methodology, called CQA, that quantitatively relate these three measures in the health care system. Having developed CQA, we note that these relationships will form a quantitative foundation for the analysis and/or comparison of national policies in health care delivery.

We demonstrate the efficacy of CQA as a policy tool by showing, within the limitations of the model, that it is possible by a multiple of resource allocation policies in a policy set to achieve universal access without sacrificing quality, or increasing cost for diseases where early detection and treatment has benefit. Further, we provide a method to distinguish the supremum policy from within the policy set.

The methodology CQA is also used to consider the formulation of policies in managed care systems. The primary constraint under the paradigm of managed care is capitation. We ask, in this dissertation, how managerial policy should be formulated in the new environment. Specifically, we consider the attainment of societal goals, such as the maximization of a health index, when constrained by service quality considerations at the medical facility level. We formalize these goals by considering the societal goal as the maximization of the health index NHQ, and capture the service quality goals in the index SQ. Further, we are interested in understanding the dynamics of capitated systems with respect to access level, with our agenda being the provision of universal access without the sacrifice of quality.

With these objectives in mind, we consider a fundamental question in the formulation of managerial policy for managed care systems -- that of resource allocation. We develop a three tiered optimization system for resource allocation. At the highest level of the hierarchy, we consider the determination of the optimal resource allocation sample path across age groups to a convex combination of the objectives NHQ and SQ. Having determined the age group specific resources, further optimization models determine, for each of the objectives NHQ and SQ, the resource division to the treatment of the different disease classes. The last stage in the hierarchy of decisions is the resource division between mass screening and early treatment.

Exploring the dynamics of the resource allocations with access level provide the following results: (i.) The efficacy of early detection and treatment of pre-clinical stage breast cancer patients is inversely proportional to the density of class pcl persons in the population. (ii.) With increased access, both mass screening and early treatment will claim a greater share of the age group specific allocation of resources. Further, the lion's share of the resources will be directed towards the treatment of detected patients, with a lesser share directed towards mass screening. (iii.) The per-patient expenditure on clinically surfaced breast cancer patients will decrease in inverse proportion to access level, however, the resources allocated to the treatment of class csl patients will rise with increased access. (iv.) Both regional and distant stage per-patient expenditures and treatment budgets decline with increased access. The resource allocation dynamics with access require decision makers to shift equipment and personnel specialization's with societal access levels to medical care.

The real world application of the dissertation's methodology requires i.) data on transition rates between the different disease classes, and ii.) knowledge of the resource-quality functions. We envision that the data on mutation rates can be obtained through clinical studies, although there are significant ethical issues to consider. For example, obtaining the transition rate of clinical stage to regional stage breast cancer patients may require observing the transitions of *consciously* untreated patients. Thus, the methodologies presented in this dissertation must either rely on the best educated guesses of medical experts or wait to a time when science can provide the appropriate transition rates without need to observe untreated patients.

The resource quality functions have been approximated as having a root form. This choice provides analytical tractability. However, the root form functions suffer the drawback of providing non-zero marginal returns for all values of resource. The results presented in this dissertation may be strengthened by the use of resource quality functions that will eventually saturate as a function of resource. A possible method by which to obtain real world resource quality functions is to conduct a properly controlled cross-national study of breast cancer outcomes with resource expenditure. For both the mutation rates, and the resource quality functions, the necessary studies to provide the appropriate numbers and functional forms are beyond the scope of this dissertation.

Finally, avenues for future research include: i.) using CQA to test further policies, ii.) developing optimization models for the optimal allocation of resources in the health care system, given a set of objectives, and iii) extending the models and framework presented in this dissertation to the inclusion of multiple diseases. The study of multiple diseases will introduce additional complexities into the model. For example: The early detection and cure of a disease in a person may lead to the prolonging of the person's life beyond that which it would have been if the disease had not been treated in a timely manner. With the prolonging of the person's life, he/she may acquire a second disease that will be presented to the medical system. Consequently, a program of early detection and treatment may lead to additional demands placed on the medical system from other diseases. Thus, the study of multiple diseases must account for the interactions among the diseases. This will, however, lead to a more complete analysis of the health care system.

Appendix

Construction of a Representative Data Set

In this appendix we construct the representative data set used in the policy experiments of Chapter 6. The data set is constructed in the following order: (i) derivation of the initial health status vector $\mathbf{A}_0$, (ii) derivation of the service time matrix $\mathbf{m}$, (iii) presentation of the mean mutation probabilities, (iv) explanation of the assumed resource quality functions, (v) derivation of initial conditions, and (vi) presentation of the weight matrix $\mathbf{W}$.

I. Initial Health Status Vector

Age specific data for the incidence of breast cancer is reported in 5-year age groups by the National Cancer Institute (Ries, *et al.*, 1994). In this paper we consider the effects of the policies on patients aged between the ages of 40 and 65, leading to five age groups in our time window of study. 40-44; 45-49; 50-54; 55-59; and 60-64. The initial health status vector in this paper will be based upon the stage specific incidence rates reported by the NCI for women in 1989-1990 in the 40-45 age group. These numbers are reproduced here as (Harras, 1995):

Stage	Incidence/ 100 000
Localized	1194
Regional	831
Distant	107

Table A.1: Stage Specific Incidence Rates of Breast Cancer in Women Aged 40-44, 1989-1990.

Since the NCI reports only known cases of breast cancer, the cases of local stage breast cancer include both known pre-clinical (discovered by screening measures), as well as clinically surfaced local stage breast cancer. To estimate the numbers of pre-clinical breast cancer in this population, we note that the numbers of 40-44 year old women reported to have had screening mammograms in in 1990 was 31% (Breen and Kessler, 1994). Further, it was reported in a study (Walter and Day, 1983) that 40% of local stage breast cancer was determined by mammography. We assume this number and consequently estimate the numbers of pre-clinical cancer cases.

To derive the number of clinically surfaced and regional stage cases present in the population, we assume that 10% of clinically surfaced, and 5 % of regional stage breast cancer cases are not brought forward for treatment because of a lack of access. We will assume that all cases of distant stage breast cancer are presented for treatment. We arrive at the following *representative* initial health status vector:

$$\mathbf{H}_0 = \left(\lambda_0^{h}, \lambda_0^{pcl}, \lambda_0^{csl}, \lambda_0^{r}, \lambda_0^{d}, \lambda_0^{p}, \gamma_0^{c}\right) = (96721,\ 1545,\ 796,\ 831,\ 107,\ 0,\ 0). \qquad \text{(A.1)}$$

Note that the numbers of individuals in the classes of persons cured, i.e. c, and of persons who have passed on, i.e., po, are assigned the value of zero, as we assume that the initial vector represents the population state prior to entrance into the medical system.

II. Service Time Matrix

The diagonal elements of the service matrix **m** are obtained as follows: To obtain the service rates of class h individuals, $E[S_i^h]$, we note that the service rates can be obtained as the inverse of the median remaining life expectancy of individuals in the i th age group, i.e., $E[S_i^h] = T_i$, where T_i defines the median remaining life expectancy of the i th age group. The life expectancy of women in the United States in 1989 was 78.8 years

(Statistical Abstract, 1995). Consequently, for the five age groups in the time window of study, we obtain:

$$E[S_1^h] = 35.7; E[S_2^h] = 31.25; E[S_3^h] = 26.3 \quad \text{(A.2) - (A.4)}$$

and

$$E[S_4^h] = 21.28; \text{ and } E[S_5^h] = 16.4 \quad \text{(A.5)-(A.6)}$$

We assign the following service times for individuals in the disease classes: for pcl, 2 years; for csl, r, and d, 1 year. The service time of individuals in class po is infinite. Therefore, the diagonal elements of the service rate matrix are given as:

$$\mathbf{m}_i = \{E[S_i^h], E[S_i^{pcl}], E[S_i^{csl}], E[S_i^r], E[S_i^d], E[S_i^{po}], E[S_i^h]\} \quad \text{(A.7)}$$

where the elements of $\mathbf{m}_i$ have values given as $\mathbf{m}_i$ = {$E[S_i^h]$, 2, 1, 1, 1, 0, $E[S_i^h]$} where $E[S_i^h]$) is given by Equations (A2) - (A6).

III. Mutation Probabilities

In the following tables we present mean mutation probabilities. Table A.2 presents the age specific mean mutation probabilities from class h to more advanced disease classes.

Mutation from class h	Period 1	2	3	4	5
h to pcl	0.017	0.013	0.016	0.02	0.03
h to csl	0.006	0.0046	0.006	0.009	0.01
h to r	0.006	0.005	0.006	0.007	0.006
h to d	0.0004	0.0007	0.0006	0.00095	0.001
h to po	0	0	0	0	0

Table A.2. Age specific mean mutation probabilities from class h to classes pcl, csl, r, d, and po

	pcl_{i+1}	csl_{i+1}	r_{i+1}	d_{i+1}	po_{i+1}
pcl_i	0.1	0.5	0.3	0.1	0
csl_i	0	0.1	0.6	0.3	0
r_i	0	0	0	0.5	0.5
d_i	0	0	0	0	1
po_i	0	0	0	0	1

Table A.3. Mean mutation probabilities between the disease classes pcl, csl, r, and d.

IV. Resource-Quality Functions

We assume that the resource-quality functions have the following functional form:

$$\vartheta(R_i^z) = 1 - e^{-\beta^z R_i^z}, \tag{A.8}$$

where z= (edc, pcl, csl, r, d), R_i^z is defined per Modelling Assumption VII in Section 5.1, and β_i^z assume values given by:

$$\beta^{edc} = 2.3 \times 10^{-2}; \beta^{etc} = 1.15 \times 10^{-4}; \beta^{csl} = 2.4 \times 10^{-5}; \qquad \text{(A.9) - (A.11)}$$

$$\beta^{r} = 6.8 \times 10^{-6}, \beta^{d} = 1.05 \times 10^{-6} \qquad \text{(A.12) - (A.13)}$$

The non-linear resource-quality functional forms (Equation (A.8)) follows from the expectation in medical treatment that the probability of detection and/or cure of breast cancer will initially rise rapidly with resource expenditure, experience diminishing returns, and finally will eventually saturate. The values of the parameters β_i^z are obtained by noting the following: i) that with current technology the mammographic sensitivity is 90%, with an expense of $100 (ref), ii) the mean probabilities of cure for diseases in the pre-clinical, clinically surfaced, regional, and distant stages of breast are 0.9, 0.7, 0.4, and 0.1

respectively, iii) we assume mean expenditures to treat breast cancer as $20, 000, $50, 000, $75, 000, and $100, 000 for class pcl, csl, r, and d respectively.

V. Initial Conditions

To obtain the initial operating conditions of the health care system, we observe the evolution of the arrival rate vector A_0 through repeated application of Equation (5.2.9), under an access level, $\alpha = 0.8$, mean mutation probabilities specified in Tables A.2 and A.3, and functional forms and parameters specified by Equations (A.8)-(A.13). Consequently, for the five age groups in our time window of study, we obtain:

Health Class	Age Group				
	45-49	49-54	55-59	60-64	65-69
h	93911	91807	89253	85851	81973
pcl	1734	1385	1535	2017	2548
csl	1200	1178	1166	1459	1734
r	1165	1108	1103	1256	1350
d	353	417	436	473	566
po	506	892	921	936	1045
c	1132	1576	1478	1500	1841

Table A.4: Class Specific Incidence Rates Under Status Quo Operation.

The initial budgets are obtained from the arrival rates of Table A.4 and the expense parameters, R_i^z, specified in (IV) above.

Health Class	Age Group				
	40-44	45-49	50-54	55-59	60-65
detection	7.9	7.7	7.46	7.26	7.0
pcl	22.2	24.9	19.9	22.1	29
csl	35.8	54.	49.86	49.64	56.52
r	59.832	83.88	79.78	79.42	90.43
d	10.7	35.3	41.7	43.6	47.3

Table A.5: Initial Budgets Under Status Quo Operation ($millions)

VI. Weight Matrix

The weight matrix is obtained following the assumptions in Equations (5.3.6) and (5.3.7):

$$\mathbf{W} = \{M_1, M_2, M_3, M_4, M_5, M_6\}$$
$$= \{1, 0.1, 0.01, 0.001, 0.0001, 0\}. \qquad \text{(A.14)}$$

References

AYANIAN, J. Z., B. A. KOHLER, T. ABE, AND A. M. EPSTEIN, "The Relationship Between Health Insurance Coverage and Clinical Outcomes Among Women with Breast Cancer," *The New England Journal of Medicine*, 329, 5 (1993), 326 - 331.

BLENDON, R. J., BRODIE, M., AND BENSON, J., "What Happened to America's Support for the Clinton Health Plan," *Health Affairs*, Summer 1995, 7 - 23.

BREEN, N., KESSLER, L., "Changes in the Use of Screening Mammography: Evidence from the 1987 and 1990 National Health Interview Surveys," *American Journal of Public Health*, 84, 1, (1994), 62 - 67.

BUTLER, J. R. G., FURNIVAL, C. M., AND HART, R. F. G., "Estimating Treatment Cost Functions for Progressive Diseases: A Multiproduct Approach with an Application to Breast Cancer," *Journal of Health Economics*, 14, 1995, 361 - 385.

CHAFEE, MR., "Health Equity and Access Reform Today Act of 1993," *Bill No. S.1770*, 103rd Congress, 1st Session, United States Senate, Washington, D.C., 1993.

COOK, A. R., AND DRESSER, P. D., *Cancer Sourcebook for Women*, Omnigraphics, Inc., Detroit, MI, 1996.

CONGRESSIONAL RESEARCH SERVICE, "*Health Insurance and the Uninsured: Background Data and Analysis*," Library of Congress, U. S. Government Printing Office, Washington, D. C., 1988.

COUGHLIN, K. M., "After 20 years, HMO's are still challenged to deliver quality," *Business and Health*, 11, 1, 1993, 17 - 23.

DANZON, P., "Hidden Overhead Costs: Is Canada's System Really Less Expensive?" *Health Affairs*, (Spring 1992), 21 - 43.

DAVIS, K., ROWLAND, D., ALTMAN, D., COLLINS, K. S., MORRIS, C., "Health Insurance: The Size and Shape of the Problem," *Inquiry*, 32, Summer 1995, 196 - 203.

DEBROCK, L., AND ARNOULD, R. J., "Utilization Control in HMO's," *Quarterly Review of Economics and Finance*, 32, 3, Autumn 1992, 31 - 53.

EDDY, D. M., "A Mathematical Model for Timing Repeated Medical Tests," *Medical Decision Making*, 3 (1983), 45 - 62.

EYLES, J., AND BIRCH, S., "A Population Needs Based Approach to Health Care Resource Allocation and Planning in Ontario: A link Between Policy Goals and Practice?" *Canadian Journal of Public Health*, 84, 2, Mar - Apr 1993, 112 - 117.

FAHS, M. C., MANDELBLATT, J., SCHECHTER, C., AND MULLER, C. "Cost Effectiveness of Cervical Cancer Screening for the Elderly," *Annals of Internal Medicine*, 117, 6, 15 September 1992, 520 - 527.

FELDMAN, R., SLOAN, F., AND PARINGER, L, "Compensation Arrangements Between Hospitals and Physicians," *The Bell Journal of Economics*, 12, 1, Spring 1981, 155 - 170.

FOX S. A., P. J. MURATA, AND J. A. STEIN, "The Impact of Physician Compliance on Screening Mammography for Older Women," *Archives of Internal Medicine* 151 (1991), 50 - 56.

GILLIS, A. J., "Allocation of Healthcare Resources: The Case for Health Promotion," Nursing Forum, 27, 4, Oct - Dec 1992, 21 - 26.

GOOLD, S. D., HOFER, T., ZIMMERMAN, M., HAYWARD, R. A., "Measuring Physician Attitudes Toward Cost, Uncertainty, Malpractice, and Utilization Review," *Journal of General Internal Medicine*, 9, 1994, 544 - 549.

GROSS, D., AND HARRIS, C. M., *Fundamentals of Queueing Theory, 2nd Edition*, John Wiley & Sons, New York, NY, 1985.

HAFNER-EATON, C., "Physician Utilization Disparities Between the Uninsured and Insured," *Journal of the American Medical Association*, 269, 6 (1993), 787 - 792.

HARRAS, A., *Private Communication*, National Cancer Institute, Bethesda, MD, 1995

HECLO, H., "The Clinton Health Plan: Historical Perspective," *Health Affairs*, Spring 1995, 86 -98.

HENDERSON, I. C., J. R. HARRIS, D. W. KINNE, AND S. HELLMAN, "Cancer of the Breast," *Cancer, Principles & Practice of Oncology, 3rd Edition*, edited by V. T. DeVita, S. Hellman, and S. A. Rosenberg, J. B. Lippincott, Philadelphia, PA, 1989, 1197 - 1268.

HENNEY, J. E., AND V. T. DEVITA JR., "Breast Cancer," *Harrison's Principles of Internal Medicine*, McGraw-Hill, New York, NY, 1987, 1567- 1574.

HILLMAN, A. L., "Financial Incentives for Physicians in HMO's. Is There a Conflict of Interest?" *New England Journal of Medicine*, 317, 1987, 1743 - 8.

JONSSON, B., "Improving Patient Care: Consequences for Resource Allocation," *Cardiology*, 84, 6, 1994, 420 - 426.

LEITCH, A. M., AND R. F. GARVEY, "Breast Cancer in a County Hospital Population: Impact of Breast Screening on Stage of Presentation," Annals of Surgical Oncology, 1, 6 (1994), 516 - 520.

LEVIT, K. R., AND C. A. COWAN, "The Burden of Health Care Costs: Business, Households, and Governments," *Health Care Financing Review*, 12, 2 (Winter 1990), 127 - 137.

LEVIT, K. R., G. L. OLIN, AND S. W. LETSCH, "America's Health Insurance Coverage, 1980-91," *Health Care Financing Review*, 14, 1 (Fall 1992), 31 - 57.

LEVIT, K. R., A. L. SENSENIG, *et al.*, "National Health Expenditures, 1993," *Health Care Financing Review*, 16, 1 (Fall 1994), 247 - 294.

LEVIT, K. R., LAZENBY, H. C., AND SIVARAJAN, L., "Health Care Spending in 1994: Slowest in Decades," *Health Affairs*, 15, 2, 1996, 130 -144.

LOVE, S., AND K. LINDSEY, *Dr. Susan Love's Breast Book*, Addison-Wesley, Reading, MA, 1991.

MILLER, A. B., J. CHAMBERLAIN, N. E. DAY, M. HAKAMA, AND P. C. PROROK, *Cancer Screening*, Cambridge University Press, Cambridge, England, 1991.

MILLER, B.A., E. J. FEUER, AND B. F. HANKEY, "The Significance of the Rising Incidence of Breast Cancer in the United States," *Important Advances in Oncology 1994,* edited by V. T. DeVita, S. Hellman, and S. A. Rosenberg, J. B. Lippincott, Philadelphia, PA, 1994, 193 - 207.

MITCHELL, MR., "Health Security Act," *Bill No. S.2351*, 103rd Congress, 2nd Session, United States Senate, Washington, D.C., 1994.

MORREIM, E. H., "Whodunit? Causal Responsibility of Utilization Review for Physician's Decisions, Patients' Outcomes," *Law, Medicine & Health Care*, 20, 1-2, Spring-Summer 1992, 40 - 56.

MORRISON, A. S., *Screening in Chronic Disease, 2nd ed.*, Oxford University Press, Oxford, England, 1992.

MOYNIHAN, MR., "Health Security Act," *Bill No. S.2351*, 103rd Congress, 2nd Session, United States Senate, Washington, D.C., 1994.

MURRAY, C. J., KREUSER, J., AND WHANG, W., "Cost - Effectiveness Analysis and Policy Choices: Investing in Health Systems," Bulletin of the World Health Organization, 72, 4, 1994, 663 - 674.

NATHANSON, P., "Influencing Physician Practice Patterns," *Topics in Health Care Financing*, 20, 4, 1994, 16 - 25.

PONTES, M. C., "Agency Theory: A Framework for Analysing Physician Services," *Health Care Management Review*, 20, 4, 1995, 57 - 67.

REAGAN, M. D., *Curing the Crisis*, Westview Press, Boulder, Colorado, 1992.

REIS, L.A.G., MILLER, B.A., HANKEY, B. F., KOSARY, C. L., HARRAS, A., EDWARDS, B. K., (eds). *SEER Cancer Statistics Review, 1973-1991: Tables and Graphs, National Cancer Institute*. NIH Pub. No. 94-2789. Bethesda, MD, 1994.

ROBINSON, R., "Cost - Utility Analysis," *British Medical Journal*, 307, 6908, Oct 2 1993, 859 - 862.

ROCHAIX, L., "Information Assymetry and Search in the Market for Physicians' Services," *Journal of Health Economics*, 8, 1989, 53 - 84.

ROSENBERG, S. N., ALLEN, *et al.*, "Effect of Utilization Review in a Fee for Service Health Insurance Plan," *The New England Journal of Medicine*, 20, 333, 1995, 1326 - 1354.

ROWLAND, D., B. LYONS, A. SALGANICOFF, AND P. LONG, "A Profile of the Uninsured in America," *Health Affairs*, (Spring (II) 1994), 283 - 287.

SCHAPIRA, D. V., J. STUDNICKI, D. D. BRADHAM, P. WOLFF, AND A. JARRETT, "Intensive Care, Survival, and Expense of Treating Critically Ill Cancer Patients," *Journal of the American Medical Association*, 269, 6 (1993), 783 - 786.

SCHIEBER, G. J., J. P. POULLIER, AND L.M. GREENWALD, "Health Spending, Delivery, And Outcomes in OECD Countries," *Health Affairs*, (Summer 1993), 120 - 129.

SKOCPOL, T., "The Rise and Resounding Demise of the Clinton Plan," *Health Affairs*, Spring 1995, 66 - 85.

SONDIK, E. J., "Breast Cancer Trends," *Cancer* 74, 3 (1994), 995 - 999.

STARR, P., *The Logic of Health Care Reform*, Penguin Books USA, New York, NY, 1994.

STATISTICAL BULLETIN, "1992 Hospital Charges for Two Major Surgical Procedures for Cardiovascular Diseases. Part I: Coronary Artery Bypass Grafts," *Statistical Bulletin, Metropolitan Insurance Companies*, 75, 1 (1994), 12 - 20, 28.

VINCENZINO, J. V., "Developments in Medical Care Costs: An Update," *Statistical Bulletin*, Jan - Mar 1993, 29 - 35.

VINCENZINO, J. V., "Developments in Medical Care Costs: An Update," *Statistical Bulletin*, Jan - Mar 1994, 30 - 35.

VINCENZINO, J. V., "Who are the Medically Uninsured in the United States?" *Statistical Bulletin*, Apr - Jun 1994, 20 - 27.

VINCENZINO, J. V., "Health Care Costs: Market Forces and Reform," *Statistical Bulletin*, Jan - Mar 1995, 29 - 35.

WALTER S. D., AND N. E. DAY, "Estimation of the Duration of a Pre-Clinical Disease State Using Screening Data," *American Journal of Epidemiology*, 118, 6 (1983), 865 - 886.

THE WHITE HOUSE DOMESTIC POLICY COUNCIL, *The President's Health Security Plan*, Times Books, New York, NY, 1993.

WOOLHANDLER, S., AND D. U. HIMMELSTEIN, "The Deteriorating Administrative Efficiency of the U. S. Health Care System," *The New England Journal of Medicine* 324 (1991) 1253 - 1258.

ZAPKA, J. G., A. M. STODDARD, M. E. COSTANZA, AND H. L. GREENE, "Breast Cancer Screening by Mammography: Utilization and Associated Factors," *American Journal of Public Health* 79 (1989) 1499 - 1502.