ABSTRACT
The US Healthcare System: Finding a Model of Quality
and Cost-Effectiveness

Parker Dalglish

Director: Dr. Bill Neilson

The United States, while home to some of the best medical procedures and doctors in the world, is currently in a healthcare crisis. The entire system is not cost-effective at both the macro level of government spending and the micro level of clinical analysis. Even with the passing of the Affordable Care Act and a focus on containing costs, the United States spends more than twice the amount of money per capita compared to the second highest spending nation in the world. This thesis analyzes certain aspects of the U.S Healthcare system in order to identify how future reform can improve quality while maintaining cost-effectiveness. The history of healthcare in the United States is examined in depth to understand why comprehensive reform has been so difficult and how such reform may be achieved. This thesis argues that clinical epidemiology, comparative analysis, and cost-effectiveness analysis can all be tools that can be used to achieve this reform through the use of Accountable Care Organizations.

APPROVED BY DIRECTOR OF HONORS THESIS:

_____

Dr. Bill Neilson
Clinical Professor in the Honors Program

APPROVED BY THE HONORS PROGRAM:

_____

Dr. Elizabeth C. Corey
Honors Program Director

DATE: _____

THE US HEALTHCARE SYSTEM: FINDING A MODEL OF QUALITY AND COST-EFFECTIVENESS

A Thesis Submitted to the Faculty of

Baylor University

In Partial Fulfillment of the Requirements for the

Honors Program

By

Parker A. Dalglish

Waco, Texas

May, 2016

TABLE OF CONTENTS

Table of Contents ii

List of Figures iii

Acknowledgments iv

Introduction 1

Chapter One: History of Healthcare Costs 5

Chapter Two: Cost-Effectiveness & Clinical Epidemiology. 17

Chapter Three: Comparative Effectiveness & Cost-Effectiveness Analysis 27

Chapter Four: Accountable Care Organizations 33

Conclusion 45

Bibliography 48

LIST OF FIGURES

Graph 1: Healthcare Costs Per Capita vs. Life Expectancy 2

Table 1: Medical Effectiveness in Free Care vs. Cost Share Planes 10

Graph 2: Hospital Payment-to-Cost Ratios for Private Payers, Medicare, and Medicaid 12

Graph 3: Inputs of Health Care vs. Costs and Health Benefits 18

Table 2: MRI and CT Scans in the US vs OECD Countries 20

Table 3: 2x2 Table 23

Equation Table 1 24

Table 4 and Table 5: Example Problem and Application 25

Table 6: Social Value of QALY in the UK 28

Table 7: Agreement Percentages between Comparative Effectiveness and Cost-Effectiveness Analysis 30

Graph 4: Social Value Threshold vs Percent Agreement of Interventions 32

Table 8: Early Performance of Pioneer ACOs 37

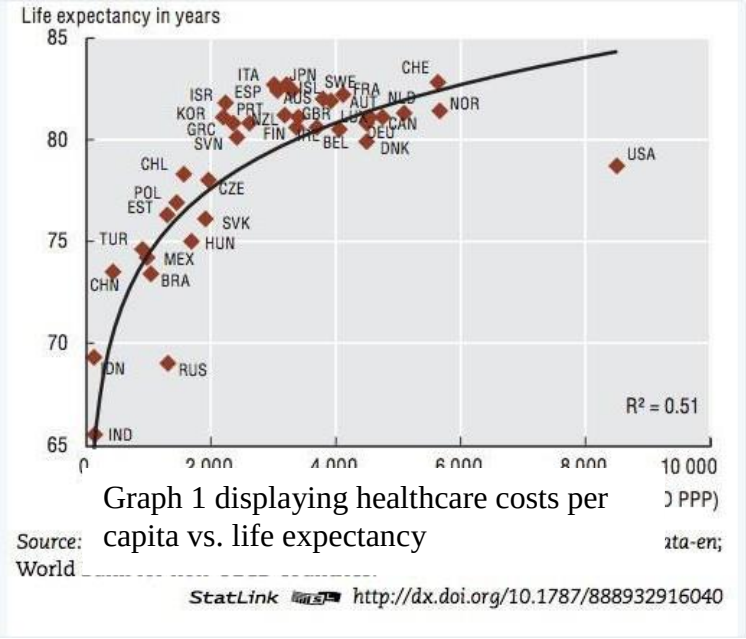
Table 9: Early Performance of MSSP 40

ACKNOWLEDGEMENTS

A special thanks to Dr. Bill Neilson for all of his guidance and knowledge, without which this thesis would not have been completed, and to Dr. Jennifer Good for her dedication and hard work. Thanks also to Dr. Al Beck for his help and patience in completing this thesis.

INTRODUCTION

The United States of America is home to some of the best healthcare options throughout the world, ranked first in 2014 by the Commonwealth fund in terms of providing the 'right care' for a given condition as defined by standard clinical guidelines and gave it especially high marks for preventive care. The United States leads the world in healthcare research and specialized care, with 5-year survival rates greater than other countries (Brownstein, 2014). In 2013, the United States spent $9,086 per person annually with a life expectancy of 78.8 years. The second highest spender, Sweden, spent $6,325 per person annually with a life expectancy of 82.9 years (Squires and Anderson, 2015).Yet; the increased amount of money spent on healthcare in the United States does not lead to greater life expectancy, which remains equal to, or lower than other comparable countries. Data published by the Organization for Economic Cooperation and Development (OECD) found that the United States fares far worse in life expectancy compared to other developed countries, as shown in Graph 1. The United States healthcare system is not cost-effective at the macro level (government programs and spending) or the micro level (clinical decision making by physicians in regards to prescribing tests and procedures), leading to increased spending without increasing benchmarks in quality. Comprehensive reform of the US Healthcare system is necessary in order for the United States to provide quality care at a sustainable cost.



Graph 1 displaying healthcare costs per capita vs. life expectancy

Several vital resources can be employed in order to understand the state and consequence of these issues. The major sources drawn upon for this thesis include Starr, *Health Affairs*, Shi and Singh, the *New England Journal of Medicine* (NEJM), and the *Journal of the American Medical Association* (JAMA). Paul Starr, a sociologist with a focus on healthcare, provides an extensive history and background for the US healthcare system, outlining the evolution of the system found today. Shi and Singh are leading public health experts and their work provides an excellent assessment of healthcare policy. Publications in *Health Affairs*, the NEJM, and JAMA provided research and data for assessment of quality and cost-effectiveness in spending. Many other individual articles were also used to support the arguments advanced in this thesis.

Chapter 1 introduces a brief history of healthcare in the US leading to the current system. Beginning in the 20th century, details of healthcare insurance and practice enter into the conversation, and it was at this point that healthcare costs began to rise to an extent that garnered national attention. Early efforts towards achieving healthcare reform

in the 1910s were directed on reform in the workforce by guaranteeing paid sick leave at 2/3 pay but ultimately failed to pass. The New Deal (1933-1938), focusing heavily on social welfare, sought to move healthcare reform forward with an introduction of universal healthcare.

Many factors caused this reform to fail and also brought about new policies that resulted in further increased medical costs. Increased government spending designed to improve healthcare coverage was part of reform efforts that followed World War II, with the two largest, and most controversial, reform efforts being Medicare and Medicaid. The 1960s, a period following implementation of these major changes resulted in small improvements to coverage and also increases in spending. In the 1990s, the Clinton plan focused on employer based healthcare insurance with government subsidies in an attempt for universal coverage. This plan will be examined at length because it was the closest the US came to comprehensive reform until the passing of the Affordable Care Act (ACA). Finally, the ACA is outlined and pros and cons of this law respective to treatment and costs discussed.

Chapter 2 surveys ways in which physicians and hospitals can work to improve healthcare quality and reduce spending, including a detailed discussion of these methods. The study of cost-effectiveness is introduced in this chapter, and its role in analyzing costs in the healthcare system examined. The use in modern medicine of technology, testing, treatment and how the overuse of these three can lead to large amounts of wasteful spending are discussed along the lines of cost-effectiveness.

This chapter argues that choices made by a practitioner around technology, testing, and treatment all play crucial roles in clinical decision-making. Deciding whether

or not to run a test, risks associated with running a test, and how to weigh the risks against the benefits are all part of this discussion. Prevalence, sensitivity, and specificity will be defined as means by which physicians can evaluate testing and treatment to determine whether or not a test or therapy is a viable course of action. An example will be given involving mammography, a common screening test in the United States.

Chapter 3 looks at the similar concept of comparative effectiveness and how it can be used to weigh how recommendations for treatments are made along with cost-effectiveness. While the two concepts seem to be contrasting in their emphasis, the two are often found in agreement with one another. This similar result demonstrates that the US healthcare system can improve upon quality, or keep it constant, while reducing the amount spent on healthcare.

The viability of Accountable Care Organizations (ACOs) as a future option for healthcare reform is the subject of Chapter 4. ACOs offer an alternative to the fee-for-service system currently used in the United States and, based on early results, have the potential to achieve comprehensive reform in how healthcare is delivered and paid for. The two main programs that discussed are the Medicare Pioneer Program and the Medicare Shared Savings Program (MSSP). Both give incentives to hospitals and physician groups to provide quality care at lower costs by providing bonuses if certain benchmarks are surpassed. These ACOs show early promise, however, there are also risks and challenges; the design of such ACOs and how the concerns regarding the risks might be managed are the subject of this chapter.

CHAPTER ONE

History of Healthcare Costs

The costs of healthcare and notion of health insurance in the United States first garnered national attention between 1915 and 1919. Prior to this period, healthcare was purchased like any other good or service and health insurance was nonexistent (Shi and Singh, 2015). In 1915, due to labor conditions and low pay rates, the American Association for Labor Legislation called for "sick pay at two-thirds wages for up to 26 weeks, payment of medical bills, a maternity benefit for both insured women and the wives of insured men, and a small benefit at death to cover burial costs" (Starr, p31). Many workers believed that "because Americans had been persuaded to adopt compulsory insurance against industrial accidents, they could also be persuaded to adopt compulsory insurance against sickness" (Shi and Singh, 2015).This plan was opposed by business representatives who believed it would promote "sickness" in their workforce. Physician groups were also against this type of health insurance, fearing loss of income. Insurance companies also opposed the plan, leading to its end in 1919 (Starr, p32). The Progressive Party was largely responsible for pushing for health insurance and, following their breakup in 1916 and poor election results for progressives in 1920, their campaign for healthcare was ended (Starr, p34). The crusade for health insurance remained in hiatus until the late 1930s when Franklin Delano Roosevelt advocated for it as part of the New Deal.

In the period of the New Deal and its aftermath, healthcare started down a road that led to repeated revisions, additions, and intricacies, creating the convoluted and confusing system that we currently have today. FDR pushed for national health insurance on three separate occasions but was stopped every time (the third attempt ending abruptly due to his death in 1945) (Starr, p40). Following his death, President Truman again pushed for national health insurance. "Initial public reaction to Truman's plan was positive…However, when a government-controlled medical plan was compared with privately obtained insurance, polls showed a drastic decline in public support" (Shi and Singh, 2015).

One key point that led to rising health costs later was that, while national health insurance was not passed, many programs that called for increased medical research and construction of new hospitals did pass (Starr, p41). By investing in the development of advanced technology, Americans received state of the art medical treatments but also had far larger costs than other countries throughout the world, a trend that continues to the present day. People became more willing to accept increased government intervention, contrary to historical record, since "medical costs had grown larger than wages lost to illness and now represented a problem not just for workers, but for many in the middle class as well" (Starr, p36).

A few of the reasons for rising costs beginning in the late 1930s may have been increased application in technology and a higher standard of education in medical professionals. The American Medical Association (AMA) spent two decades from 1910 to 1930 raising education requirements for physicians in an effort to increase effectiveness. In the end, "by requiring credentials that may or may not have had any

consumer benefit, a committee (AMA) convened to reduce costs almost certainly ended up raising them" (WPRI). By enacting these stricter laws and requirements, the supply of available physicians was drastically reduced while the need for their services was increasing (Starr, p36). Due to the principle of supply and demand, healthcare costs began to rise accordingly. As our knowledge of medicine broadened and the benefits of its practice became apparent, the number of hospitals also increased. Whereas in the 1800s and very early 1900s, hospitals were mainly used as hubs for the impoverished to receive charitable care, due to a lack of money or funds from family members (WPRI), by the 1920s, they "had transformed into medical workshop(s) for doctors treating patients from all classes" (Starr, p36). During this time, "Hospital costs rose from 7.6% of total family medical bills in 1918, to 13% in 1929" (WPRI).

Following World War II, healthcare once more became a pressing issue in politics and underwent major changes in an attempt to grant affordable health insurance amidst the rising costs of medical care. Under President Johnson in 1965, employer based insurance became the standard form of private health insurance in America with government funded programs of Medicare and Medicaid for the elderly and poor respectively. This led to the American healthcare system reaching a paradox as it did not cover everyone but provided enough coverage that those who received its benefits were resistant to change.

Medicare and Medicaid were created to extend health insurance coverage to those who may not otherwise be able to afford it. Medicare was seen as an addition to Social Security while Medicaid was an addition to welfare programs. While these two programs saw millions of Americans insured who otherwise would not be, they also contributed to

excessive spending from the government, causing the percent of national GDP spent on healthcare to rise at an "average annual rate of 7.6% between 1965 and 1970" (Shi and Singh, 2015) and continue to rise to the present day, with Medicare spending growing 5.5% and Medicaid growing 11.0% in 2014 ("NHE Fact Sheet", 2015). Through Medicare, seniors receive health insurance, with the consensus in America being that this healthcare has been "earned" by the elderly since people pay a payroll tax that supposedly covers their Medicare for when they get older, "even though seniors have never paid enough in payroll taxes to earn their insurance coverage" (Starr, p47). Medicare was drawing more money out of the system than taxpayers were putting in (Starr, p47). The opposite has been true of Medicaid. Since Medicaid was based on welfare, and welfare requirements differed state to state, coverage also varied amongst the poor. In order to qualify for health insurance, many people were "trapped" in a state of unemployment or low income jobs, diminishing the work force.

The most costly portion of these two programs was how hospitals and doctors were paid under Medicare. Medicare Part A paid hospitals according to their costs, with no regulation on rates. Medicare Part B "paid doctors their 'customary fees, assuming them to be…reasonable but the legislation set no standard for reasonableness" (Starr, p48). If costs were lowered, then providers were reimbursed less. This left little incentive for hospitals or doctors to reduce their costs. Immediately after Medicare passed, doctor and hospital fees increased.

Concurrently, Blue Cross was at the forefront of basing health insurance on employee groups and was successful at first. Initially a non-profit organization which provided the same coverage for all employee groups, Blue Cross was forced to change its

practices in order to compete with for-profit insurance companies who were "cherry-picking" employers with younger and healthier workers (Starr, p47). This left Blue Cross with high-cost high-risk groups and forced it to raise its rates (Starr, p47).

The combination of increasing costs for private plans with government programs for only a few select groups has resulted in a system that is fragmented and yet beneficial enough to resist change. Compromising, which had allowed Medicare and Medicaid to pass, now forced doctors and medical personnel to hire far more administrative personnel as the system had grown extremely complicated (Shi and Singh, 2015). These costs were added onto the cost of healthcare. While Medicare and Medicaid were successful in their goal of increasing coverage for Americans, they continue to be a source of increasing expenditures as the number of people qualifying for these programs grows and their treatment costs also rise.

Following the passage of Medicare and Medicaid, medical costs continued to rise. In the time span of 1965-1970, "state and federal expenditures rose at an annual rate of 20.8 percent" (Starr, p52). The cause for these increased costs stemmed from the problems created earlier by reimbursement of physicians and hospitals based on "costs" and the overuse of medical technology. Studies were conducted that "identified patterns of excessive hospitalization, unnecessary surgery, and other inefficiencies" (Starr, p52).

Medical Effectiveness Category	Predicted Percentage of Participants with at Least One Episode of Care in a Single Year, by Medical Effectiveness Category and Plan			
	Adults		Children	
	Free Care	Cost Sharing	Free Care	Cost Sharing
Highly effective				
Acute	28	19	32	23
Acute/chronic	17	13	19	16
Chronic	13	11	4	2
Quite effective	23	18	22	18
Less effective	30	19	13	10
Rarely effective	11	7	5	3
Rarely effective but equally effective with self-care or doctor	39	29	36	24

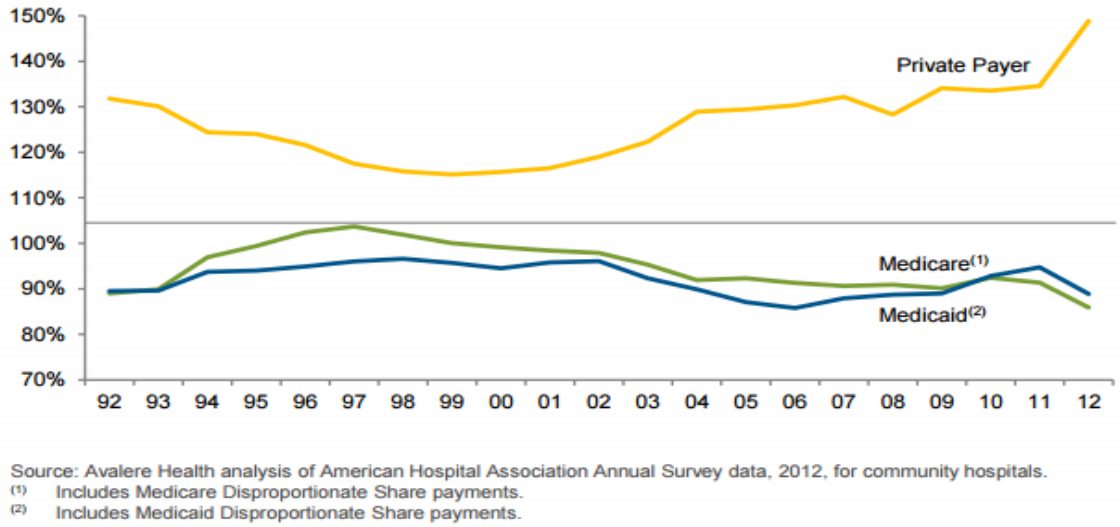
Cost Sharing Reduced the Use of Medical Services at all Levels of Effectiveness
Source: Lohr et al., 1986, p. 34.
Note: Percentages refer to participants who had at least one episode of care within the relevant diagnostic categories during a year of the experiment.

Table 1 demonstrating medical effectiveness in free care vs. cost share plans for varying qualities/types of healthcare.

Efforts to contain spending increased during the 1970s and 80s, culminating in the largest health policy study in U.S. history, the RAND Health Insurance Experiment. This experiment sought to explore the relationship between cost-sharing and the cost/quality of healthcare received. Participants were placed in one of five groups with one group being the "free" group (a Health Maintenance Organization (HMO) that requires minimal cost-sharing) and three plans with differing percentages of cost-sharing (25, 50, and 90%). From this study, taking place from 1971-1986, a number of conclusions were drawn. The most obvious was that people with a cost sharing plan who paid for a portion of their healthcare were less likely to utilize healthcare services than the free group. The second conclusion was that cost sharing "did not significantly affect the quality of care received by participants" (RAND). Cost sharing was found to reduce the use of both

effective and less-effective treatments. This averaged out to be roughly the same quality of care received by those in the free group, meaning cost-sharing plans were found to reduce spending while having no overall impact on quality (Table 1) with exceptions only arising in the most impoverished groups, (usually in areas of vision or dental health/hygiene) although these groups would be more likely to be covered by Medicaid and not need cost share plans. (RAND). This result led to the increases in use of cost sharing plans during the 1980s and 1990s as a means of reducing costs.

To further cut costs, in the mid -1980s, a new method of payment to hospitals was enacted based on diagnosis-related-groups (DRGs) which used prospective payment in contrast to Medicare which, at the time, reimbursed hospitals retrospectively on the basis of the costs they incurred (Starr, p64). This new method did not cover every single cost a hospital incurred and forced the hospital to accept some financial responsibility for the care they offered. Hospitals were paid a set fee and, if costs went above this limit, they lost money, and if they stayed under this limit, they turned a profit (this is similar to Accountable Care Organizations covered in Chapter 4). The initial problem was that payments were based on the costs that had been established in previous years, not leading to any reductions in spending. When these payments started to be cut, hospitals moved the costs onto privately insured patients in a process called cost-shifting.



Source: Avalere Health analysis of American Hospital Association Annual Survey data, 2012, for community hospitals.
(1) Includes Medicare Disproportionate Share payments.
(2) Includes Medicaid Disproportionate Share payments.

Graph 2 demonstrating aggregate hospital payment-to-cost ratios for Private Payers, Medicare and Medicaid with the horizontal line representing the percent for the hospital to make a profit. Years vs. Percent Payment of Costs.

Hospitals were losing money through Medicare and Medicaid patients and so, to make up for this, charged more when private insurance was involved in order to maintain a profit (see Graph 2).This in turn led to a shift of the healthcare industry towards more of a reliance on HMOs (Health Maintenance Organizations) and PPOs (Preferred Provider Organizations). These focused more on preventative care and reduced costs by helping people before they reached a more advanced stage of illness/disease.

Due to these issues, it became clear that the United States needed to undergo comprehensive reform in both public and private sectors. The complex healthcare system led to needless waste of money and technological advances combined with payment methods for Medicare and Medicaid drove costs to incredibly high amounts. On top of that, "public policy and the health-care financing of the system had promoted investments in hospitals and other costly facilities and encouraged the training of medical specialists rather than primary practitioners" (Starr, p72). This undermined the goal of reducing

costs for preventative care under HMOs as there was a shortage of primary practitioners who could provide it. Healthcare costs were growing at an unsustainable rate compared to the economy, with costs in the healthcare sector increasing an additional 1 percent every 34 months from 1980 to 1992 and private insurance premiums jumping 90% compared to a wage increase of only 28% between 1987 and 1993 (Starr, p72).

It was during this time of "healthcare crisis" that President Clinton came into office and became the first president to do so promising massive healthcare reform. The initial proposal was to have "managed competition with a budget." Employers would pay a certain set rate for the lowest health plan with the rest coming from employees. Any additional coverage could be chosen from a multitude of plans so that the individual could select the care they needed. Smaller businesses would be subsidized by the government in order to cover their employees and there would be an overall budget to cap spending should any of the other precautions fail (Starr, p97). When this plan failed to amass the support necessary for passage, compromise was sought between the Republican and Democratic parties but, instead, the plan eventually faced opposition from both. This amended plan was known as the Cooper plan and removed a section from the original plan that mandated employers to provide health insurance. The restriction on health insurance being mandatory for all Americans was also removed (Starr, p106). With the means of funding the plan through employers removed without an alternative solution, and with the threat of additional government intervention in American economics, Republicans did not support the plan. The plan did not garner much approval from Democrats either as the revised plan did not mandate universal coverage. The political battle waged between the two parties, and their inability to reach an agreement,

as well as the association of the Clinton name with the plan (which caused Americans to distrust the plan due to the Whitewater scandal), resulted in its failure. This was a key moment in American politics as it confirmed how difficult reform was going to be. Even when the call for reform was at its peak and with attempts to reach a compromise between parties, no agreement could be reached.

The most recent reform of healthcare comes in the form of the Affordable Care Act (ACA), known more publicly as "Obamacare." The main goal of this act is to ensure insurance coverage of all Americans and grant everyone access to healthcare. While an additional 6% of Americans received coverage one year following its passage compared to one year prior (12% uninsured in 2015 compared to 18% uninsured in 2013), it still has a long way to go in order to reach its goal (Shi and Singh, 2015). In fact, the ACA does not seem to provide a way to guarantee this 100% coverage with health policy researchers Nardin et al (2013) stating the following: "Our finding…runs counter to the common perception that the ACA will cover virtually all legal residents. The ACA will leave tens of millions uncovered. It will do little to alter racial disparity in coverage… The ACA, whatever its merits, will fall well short of its stated goal in providing affordable care for all Americans." Another issue arises in the cost of covering the additional people who have gained insurance through the ACA. The major source of this income currently is through increased taxes and fines placed on the individual and the healthcare industry (Kliff, 2012). The desire to use this money most effectively has led to the formation of Accountable Care Organizations (ACOs). ACOs attempt to save costs through integration. They integrate groups of providers and then contract with the government to deliver coordinated services to Medicare beneficiaries and share in any

resulting savings (Shi and Singh, 2015). ACOs are covered more in depth in Chapter 4. Additionally, while the government currently subsidizes the states to care for Medicaid patients, it may eventually start withdrawing support with no clear plan as of now to fill in the gap (Scott, 2012). As the ACA covers more people, there is an economic issue with paying for this coverage and, unless more revisions are made to the plan or the healthcare system as a whole, the ACA will become economically unviable several years down the line.

Healthcare reform has taken several forms over the course of American history, with small strides in coverage resulting in additional costs. While coverage has certainly expanded since talk of healthcare reform began in the United States, many people still remain uninsured or unable to pay for healthcare, as costs have exponentially increased at unsustainable rates. Healthcare reform at the macro level is incredibly difficult and intricate, as history has proven. While concessions and compromises have succeeded in making minor changes to the system, comprehensive reform has proven unsuccessful due to the inability to motivate enough parties/people (Republicans, Democrats, physicians, insurance companies, employers, seniors, impoverished, etc). Healthcare reform has become a risky enterprise for politicians as they bring high risks to presidential campaigns with few benefits in terms of public opinion. Attempts at reform have been unsuccessful, with many presidents only proposing reform to improve failing approval ratings. For these reasons, it is unlikely for comprehensive reform to take place in the coming years. Currently, one of the focal points in healthcare is the rationing and reduction of costs. While large scale reform is left to politicians, there are many options for hospitals and physicians to reduce costs while still providing high quality care by

using cost-effectiveness analysis. The next chapter will discuss what cost-effectiveness is and how it can be applied to not only cut down on excessive spending, but also to improve our healthcare system and raise the quality of care for patients. Chapter 3 will expand further on this with the addition of comparative effectiveness and Chapter 4 provides a vehicle in the form of ACOs for which this level of reform could be accomplished.

CHAPTER TWO

Cost-Effectiveness & Clinical Epidemiology

The main goal of cost-effectiveness analysis is to identify the relative costs and outcomes of several courses of action in order to find opportunities to perform tests or procedures that are relatively inexpensive or simple but provide substantial gains in the healthcare of a patient. Cost-effectiveness analysis can be explained as a scale that weighs the "gains in health relative to the costs of different health interventions" (ncbi). It "evaluates the additional (marginal) benefits derived in relation to the additional (marginal) costs incurred" (Shi and Singh, 2004). It is important to note that the "costs" calculated in cost-effectiveness analysis also include risks to the health of a patient and are not just monetary figures, thus ensuring high quality of care. The general idea is that "each unit of technology utilization is likely to provide benefits in excess of its costs" (Shi and Singh, 2004). This spending is very efficient early in treatment and can greatly increase the health of a patient for every dollar spent. Eventually though, a point is reached where benefits for a patient who is given additional care begin to flat line and approach zero. This is demonstrated in Graph 3. Any additional spending is wasteful and, in some cases, can actually harm the patient (such as overuse of CT scans).

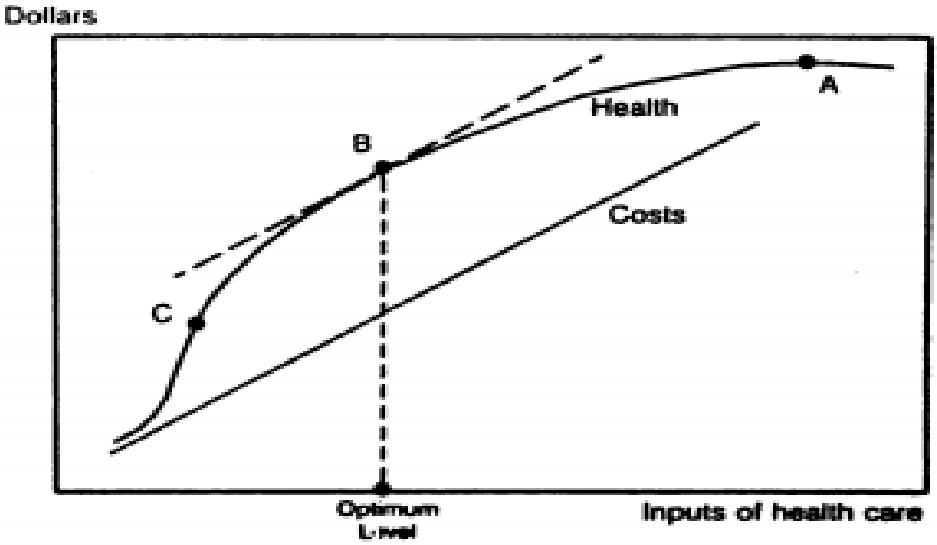


Fig. 1. "Flat of the curve" health care
B, socially optimum expenditures when marginal cost equals marginal benefit.
A, situation where marginal benefit becomes vanishingly small

Graph 3 illustrates Inputs of Health Care vs. Costs and Health Benefits (Massaro)

One can determine the cost-effectiveness ratio for a given treatment by dividing the cost in terms of currency by the change in health status, usually measured in years added to life expectancy. When calculating the average cost-effectiveness ratio for a given treatment, one adds up the total amount of money to treat all the patients that underwent said treatment and then divide by the total number of years added to life expectancy. It is common for the life years gained in these calculations to be replaced with quality- adjusted life years (QALYs) which allow "quality of life" for the patient to be taken into account during calculations. A year of perfect health is assigned the value of 1 with diminishing value as the amount and severity of disabilities increases. The value assigned to each of these disability weights is calculated by the World Health Organization. While these values carry a degree of subjectivity, they help with many of the ethical problems of assigning a monetary value to life in that they allow physicians to

use this science to make difficult decisions pertaining to patients' lives. In this manner, "cost-effectiveness analysis draws attention exclusively to health benefits, which are not monetized" (ncbi). These can then be used like an algorithm in order to find what course of action would be most beneficial to a patient.

An area of excessive spending in medicine, as well as an area that can cause potential harm to a patient, is in conducting tests. One flaw in the United States healthcare system is the abundant use of testing (most likely in relation to the development of technology as described in Chapter 1). The United States spends more on healthcare testing than any other country in the world. According to Dr. Steven Weinberger, the CEO of the American College of Physicians (PCP), the United States spends $200-$250 billion per year on excessive testing. He cites overuse of imaging techniques such as Computerized Tomography (CT) scans and Magnetic Resonance Imaging (MRI) scans as a significant reason for this waste. In many situations, it is far more effective to use a less expensive, alternative form of care if the results are nearly identical. There are several ways that scanning technology can be overused. One is by frequency of testing. Another is through unnecessary testing which can occur if additional tests are performed whose results will not affect what course of action the physician decides to take. The variation in the number of tests performed in the United States compared to other countries can be seen in Table 2. If the United States wishes to reduce the cost of healthcare, one route is to cut down on costs and expenditures through cost-effective analysis.

	United States	Rank compared with OECD countries	OECD average
MRI units	31.6 per million population	2nd	12.5 per million population
MRI exams	97.7 per 1 000 population	2nd	46.3 per 1 000 population
CT scanners	40.7 per million population	3rd	22.6 per million population
CT exams	265.0 per 1 000 population	3rd	123.8 per 1 000 population
Tonsillectomy	254.4 per 100 000 population	1st	130.1 per 100 000 population
Coronary bypass	79.0 per 100 000 population	3rd	47.3 per 100 000 population
Knee replacements	226.0 per 100 000 population	1st	121.6 per 100 000 population
Caesarean sections	32.9 per 100 live births	6th	26.1 per 100 live births

Source: OECD Health Data 2012.

Table 2 shows how the United States gives almost double the MRI and CT scans as the average found by the Organization for Economic Cooperation and Development.

There are several factors that can determine whether or not a physician should choose to run a specific test for a patient. These factors are prevalence, sensitivity, and specificity, which are all gathered from screened populations. The data from these screenings and values of these three factors can be used to determine the course of action a physician will take in regard to the single sample size of his one patient. The first of these, prevalence, "determines the total number of cases at a specific point in time, in a defined population" and is "useful in quantifying the magnitude of illnesses of a relatively long duration" (Shi and Singh, 2004). Sensitivity "is the proportion of truly diseased persons in the screened population who are identified as diseased by the screening test" ("Concepts", 2016). For example, if 1,000 people are screened for a disease and are known to have the disease, but the screening test results for these people say that only 850 people have the disease, the sensitivity of that test is 85%. Specificity is the proportion of people without the disease who the test will correctly determine to not have the disease. In the same manner as sensitivity, if 1,000 people are screened and are

known to not have the disease, but the screening test results say 100 of them do have the disease, the specificity is 90%. Sensitivity and specificity describe how accurate a test is at diagnosing whether a patient has a specific disease or not.

While no test can be entirely accurate, it is important to recognize tests that are subpar and avoid using them if possible. One such example relates to coronary artery disease (CAD). A study published in the Journal of the American Heart Association "found exercise echocardiography and myocardial perfusion imaging… considerably less sensitive and more specific" (DAIC) than previously believed. This was due to a referral bias of patients who had abnormal stress tests. Since these patients were more likely to be diagnosed correctly due to already having abnormal stress levels, their diagnoses contributed to the misconception of the tests having higher sensitivities than they actually did since they were high risk patients. Small findings such as this can reveal tests that are not as useful as previously believed and allow us to eliminate them in order to cut down on costs and to avoid incorrect diagnoses. An incorrect diagnosis can lead to a decrease in quality of life for a patient and alter a person's lifestyle. These incorrect diagnoses are called false positives and false negatives.

A test can provide one of four possible outcomes: true positive, true negative, false positive and false negative. True positive and true negative are if the test result is correct in its diagnosis. Of all the outcomes, true negatives are the ones physicians and their patients hope for and are the best result. True positives are bad in the sense that it means the person has the disease but are good in that they diagnose the problem and allow for a course of action to take place. If no treatment is available, running a test serves little purpose as the result does not change what can be done but can do potential

harm. False positives and false negatives are more complicated in the harm they cause. False positives occur when the test result says the patient has the disease when he[1] does not and can cause a patient to undergo unnecessary, and potentially damaging, treatments as well as cause the patient to alter his lifestyle. Since the patient believes he actually has the disease, he may change his lifestyle drastically or undergo operations that are not required, leading to a lower quality of life. False negatives occur when the test result declares that the patient does not have the disease when, in actuality, he does. This result is life threatening because the patient will not undergo vital treatments he needs in order to cure or slow the progression of said illness. Without treatment, the patient's health will deteriorate, occasionally resulting in death.

In the real world, it is impossible for a test result to be 100% accurate. However, by using prevalence, sensitivity, and specificity, it is possible to calculate the chance of each of the four above outcomes.

The most basic way to calculate the aforementioned values is through use of 2x2 tables. On the top of this grid, there should be two possible outcomes: "Has the disease", and "Does not have the disease". On the left of the grid, the two outcomes should be "Test Positive" and "Test Negative". The top left box will contain true positives and be labeled Box A, the top right box will contain false positives and be labeled Box B, the bottom left box will contain false negatives and be labeled Box C, and the bottom right box will contain true negatives and be labeled Box D. When completely labeled, the table should look similar to Table 3 below.

[1] In this paper, "he/she" and "his/her" have been shortened to "he" and "him" in order to be more cohesive, but imply "she" and "her" alongside it.

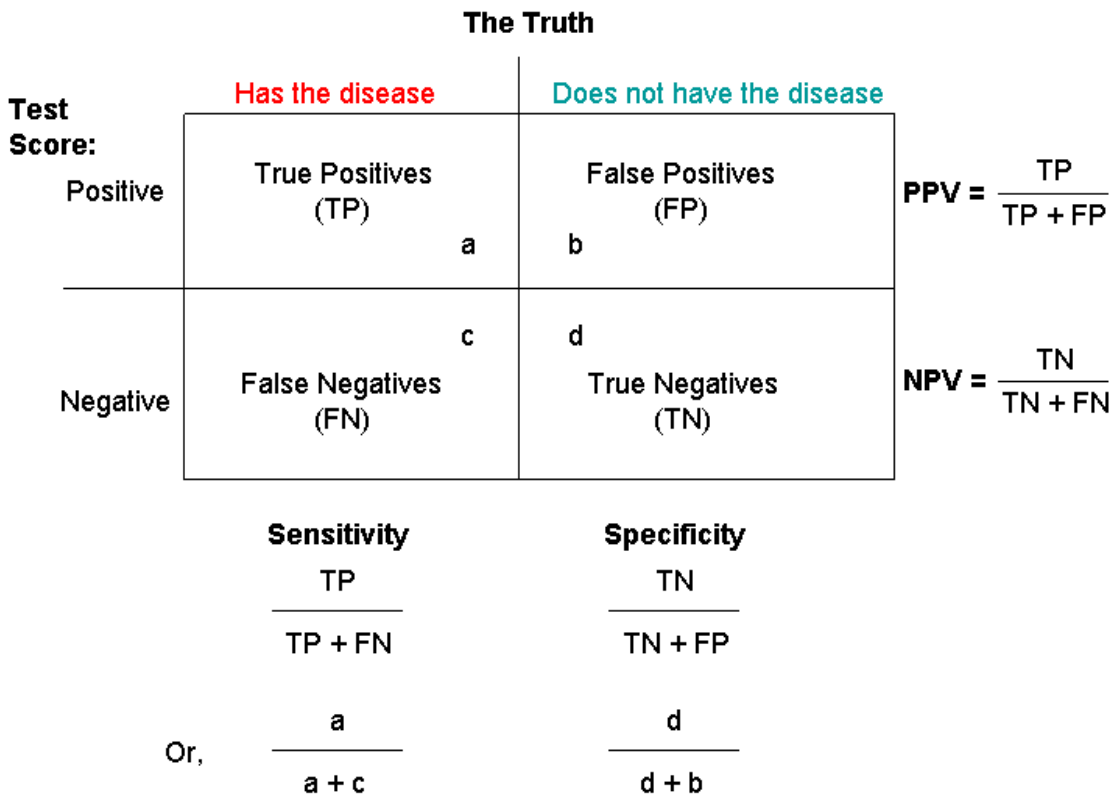


Table 3 showing how to properly set up the 2x2 table (Concepts: Sensitivity, Specificity, Predictive Values and Likelihood Ratios)

In order to calculate the percentage of true positives, one multiplies the prevalence by the sensitivity (prev*sens). Calculating the percentage of true negatives is similar except sensitivity is replaced with specificity and one multiplies it by 1 minus the prevalence ((1-prev)*spec). Percentage of false positives can be found by multiplying one minus the prevalence by one minus the specificity ((1-prev)*(1-spec)). Percentage of false negatives can be found by multiplying the prevalence by one minus the sensitivity (prev*(1-sens)). All of these equations, alongside equations for determining positive predictive value and negative predictive value can be found in the following equations

table.

What is attempting to be calculated	How to calculate it
Percentage of True Positives	(prev*sens)
Percentage of True Negatives	((1-prev)*spec)
Percentage of False Positives	((1-prev)*(1-spec))
Percentage of False Negatives	(prev*(1-sens))
Positive Predictive Value	(TP/(TP+FP))
Negative Predictive Value	(TN/(TN+FN))

Equation Table 1

The positive predictive value represents the probability that subjects with a positive screening test truly have the disease. The negative predictive value is the ratio of true negatives over the combined totals of true negatives and false negatives (TN/(TN+FN)). The negative predictive value represents the probability that subjects with a negative screening test truly do not have the disease. Whatever these values come out to be, they can be compared to our pretest probability as our posttest probabilities and used to decide a course of action.

The following example of mammography screening will be used to demonstrate the application of the above paragraphs. An article published on Medscape and updated on May 7th, 2015 stated that the sensitivity and specificity of a mammography screening test to be 67.8% and 75% respectively (Newton). A recent report of prevalence (October 23rd, 2015) reveals that, for a 40 year old woman, "the probability of developing invasive breast cancer in the next 10 years is 1.45%, or 1 in 69" ("Risk of Developing Breast Cancer", 2015). With these figures, it is possible to conduct a mock test to see what the results would tell us about the likelihood of a woman having breast cancer at the age of 40. The pretest probability would be equal to the prevalence (1.45%) as long as the sample female does not have any other risk factors. Using our prevalence, sensitivity,

specificity, and equations discussed earlier in the chapter, one can fill out a table to calculate True Positives, True Negatives, False Positives, and False Negatives.

	Has Breast Cancer	Does Not Have Breast Cancer	
Test Score Positive	0.00983 a	0.25 b	0.25983
Test Score Negative	0.00467 c	0.74 d	0.74467
	0.01450	0.99	1.0045 ≈ 1.0
Prevalence (0.0145)	Sens (0.678)	Spec (0.75)	

Table 4

In Table 4, Box a represents True Positives, Box b represents False Positives, Box c represents False Negatives, and Box d represents True Negatives. All the values should together add up to 1.0 (100%) which is demonstrated in the column to the right of "Does Not Have Breast Cancer" and in the row underneath "Test Score Negative". One can now use this data to calculate or PPV and NPV in Table 5.

	Has Breast Cancer	Does Not Have Breast Cancer	Predictive Values
Test Score Positive	0.00983 a	0.25 b	PPV = TP/(TP+FP) PPV ≈ 0.038
Test Score Negative	0.00467 c	0.74 d	NPV = TN/(TN+FN) NPV ≈ 0.99

Table 5

The Positive Predictive Value from the data comes out to be 0.038 (3.8%). This means that, should a random woman at the age of 40 with no preexisting risks for breast cancer undergo screening and the test is found to be positive, her risk of having breast cancer would only increase from 1.45% to 3.8%, even though the test says she has it. If

the test were to come back negative, her chances of not having breast cancer would increase around 0.45% from a 98.55% chance to approximately a 99% chance.

The above example demonstrates that, even though testing and treatment therapies are considered essential parts of healthcare, they must be used appropriately and in moderation to be effective. The following chapter will delve into comparative effectiveness and cost-effectiveness to explain how recommendations for specific tests are made and how selecting the option most beneficial to patients can actually result in decreased healthcare expenditures.

CHAPTER THREE

Comparative Effectiveness & Cost-Effectiveness Analysis

In the previous chapter, cost-effectiveness analysis was argued to be beneficial in reducing spending and helpful in clinical decision making by assigning a numerical value for treatments based on the ratio of costs of treatment to the number of quality-adjusted life years (QALY). However, as of 2010, the Affordable Care Act (ACA) has placed a greater emphasis on comparative effectiveness and research surrounding it, focusing more on quality of treatment and less on costs. Lawmakers and provisions in the ACA oppose giving the "newly established Patient-Centered Outcomes Research Institute, which will oversee federal comparative research efforts, the responsibility or even the authority to measure value" (Garber & Sox, p1806). Concerns of the government and the public toward healthcare decision making based on cost-effectiveness and not on absolute benefit to the patient have been the main factors in this opposition. Despite this, evidence suggests that comparative effectiveness and cost-effectiveness analysis are not mutually exclusive and that they are often found in agreement with one another.

Comparative effectiveness research is defined as "research evaluating and comparing health outcomes and clinical effectiveness, risks, and benefits of 2 or more medical treatments, services, and items" (Garber &Sox, p 1805). These services and items can range from protocols for treatment, care management, and delivery to medical devices and pharmaceuticals. By comparing all the options available for a specific category, it is possible to make a comparative effectiveness recommendation. This

recommendation is the "adoption of the therapy with the larger point estimate for QALYs from the denominator of the cost-effectiveness ratio" (the cost-effectiveness ratio being Cost/QALYs) (Glick et al., 2015). This recommendation is based solely on point estimate for QALYs and ignores the cost of each treatment. For example, if surgical treatment added more QALYs than medical treatment (drug therapy), the surgical treatment would be recommended as the course of action.

Cost-effectiveness recommendation, alternatively, is the "adoption of the therapy with the larger point estimate for QALYs if it dominates the alternative (that is, costs less and is more effective) or has a cost-effectiveness ratio that is less than or equal to the threshold social value of a QALY" (Glick et al., 2015). Cost-effectiveness analysis recommends therapy that is both high quality and low cost or high quality that costs more but still falls into socially accepted values for cost of QALYs (called social value of a QALY). The social value of a QALY is based on willingness-to-pay for a prevented fatality or a serious morbidity. While the United States does not currently use cost-per-QALY ratios, the United Kingdom does use the willingness-to-pay method. The accepted social values are shown in Table 6 and, converting pounds to USD, amount to approximately $100,000 for life-saving treatment, $50,000 for life-extending treatment, and $15,000 for quality-of-life-enhancing treatment (BMC Health Services Research).

Basic modelling approach	Value of a QALY (£)
Life-saving	£70,000
Life-extending	£35,000
Quality-of-life-enhancing	£10,000

Table 6: Social Value of QALY in the UK based on a survey taken by a cross-section of the population on how much the government should spend per QALY (BMC Health Services Research).

Additionally, comparative effectiveness research can be of direct aid to cost-effectiveness research. The denominator for cost-effectiveness ratios is often derived from comparative effectiveness research, "providing comparative health outcomes data of high, perhaps unprecedented, quality" (Garber & Sox, p1808). Comparative effectiveness analysis also includes patient preferences for different health outcomes in its research, contrary to most clinical studies, allowing opportunity to include patient opinions on varying side effects and benefits of treatment in the decision making processes. Cost-effectiveness can likewise aid in comparative effectiveness analyses since it allows comparison of treatments that may offer marginal benefits but vastly differ in cost. While federally sponsored research does not partake in cost-effectiveness analysis due to political and legal questionability, private parties such as health insurance companies, physician groups, and hospitals are encouraged to apply "cost-effectiveness analysis based on comparative effectiveness results in order to reduce spending as well as aid patients who may wish to take into account the financial implications of their health care decisions" (Garber & Sox, p1810).

In many cases, comparative effectiveness and cost-effectiveness are found to be in agreement with one another. The most beneficial treatment for a patient is often the most cost-effective. A recent study published in Health Affairs (Glick et al., 2015) conducted an analysis comparing the two recommendations coming from comparative effectiveness and cost-effectiveness in order to access the level of agreement between them, accounting for several social values for QALYs. In their sample of 2,027 cost-per-QALY ratios, "28 percent represented therapies that were more effective than the alternative and were less costly" (Glick et al., 2015). While a minority of cases, these therapies represent clinically

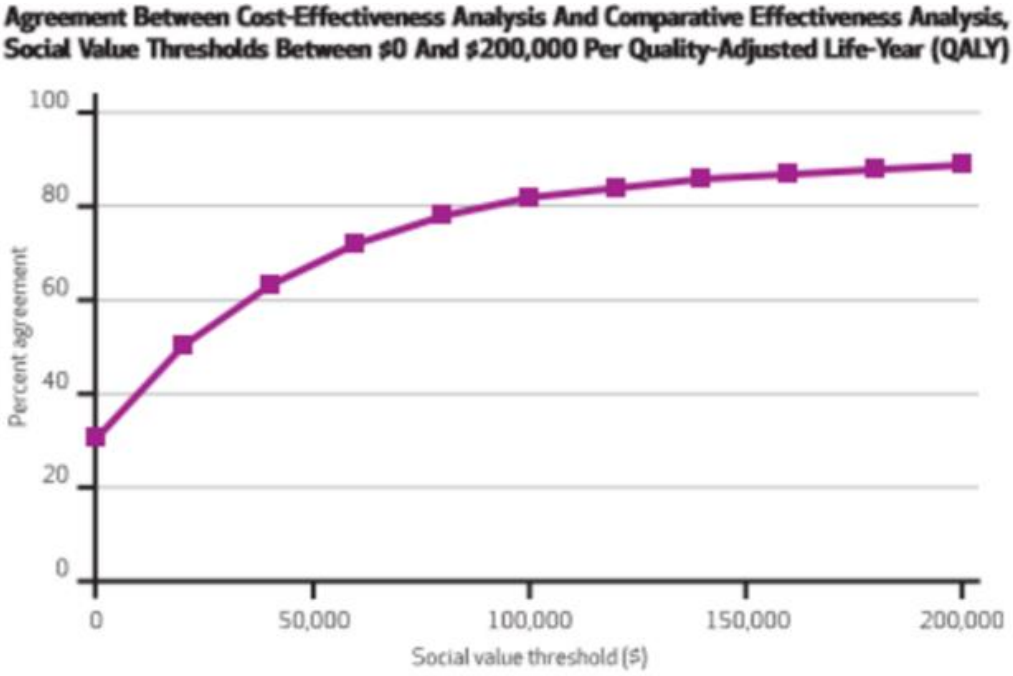
effective options that are preferred under any social value threshold and hold promise for large amounts of savings should physicians follow these recommendations. At the social value threshold of $100,000 per QALY, “81.1 percent of cases were ones in which comparative analysis and cost-effectiveness analysis agreed and led to the same choice as preferred therapy (Glick et al., 2015). This data is summarized in Table 7.

Agreement On Interventions Based On Cost-Effectiveness And Comparative Effectiveness Analyses, With Social Value Threshold Equal To $100,000

Category	Number of analyses	Agree	Disagree
Overall	2,027	81.1%	18.9%
TYPE OF INTERVENTION*			
Care delivery	227	83.7	16.3
Diagnostics	292	79.1	20.9
Health education and promotion	155	83.2	16.8
Immunization	93	86.0	14.0
Medical device	197	86.8	13.2
Medical procedures	306	82.7	17.3
Pharmaceuticals	929	80.2	19.8
Screening	342	75.4	24.6**
Surgical	330	86.1	13.9
Other	60	86.7	13.3
DISEASE CATEGORY*			
Cancer	348	78.2	21.8
Cardiovascular	482	85.1	14.9
Digestive	154	79.9	20.1
Endocrine	161	85.7	14.3
Environmental	92	83.7	16.3
Infection	571	81.3	18.7
Musculoskeletal	197	69.0	31.0***
Mental health	236	81.4	18.6
Maternal and perinatal	86	82.6	17.4
Respiratory	68	77.9	22.1
Sensory	96	85.4	14.6
Other	304	86.8	13.2
FUNDING SOURCE			
Pharmaceutical	393	91.6	8.4****
Other	1,634	78.6	21.4
PREVENTION STAGE			
Primary	376	77.4	22.6
Secondary	591	80.7	19.3
Tertiary	1,060	82.6	17.4
DATE OF STUDY			
Before 2005	623	80.6	19.4
2005	1,404	81.3	18.7

Table 7 displaying agreement percentages between comparative effectiveness and cost-effectiveness analysis (Glick et al., 2015)

Even by lowering the threshold for the QALYs by half from $100,000 to $50,000, the proportion of agreement between comparative effectiveness and cost-effectiveness analysis was still found to be 68 percent (Glick et al., 2015). Graph 4 illustrates the relationship between social value threshold for QALYs and the percent agreement found between comparative effectiveness and cost-effectiveness analysis. Early increases in spending amount to large increases in percent agreement while additional spending over $100,000 causes the curve to level off and provides no significant change in agreement. This demonstrates that the majority of interventions that provide the best outcomes for patients are also achieved with small increases in spending.



Graph 4 displaying the curve of Social Value Threshold ($) vs. Percent Agreement of Interventions (Glick et al., 2015)

Upon review of the societal values for QALYs that were found in the UK and their value in USD, it can be seen that the majority of American interventions based on comparative effectiveness and cost-effectiveness analysis fit the model of willingness-to-

pay (the amount the public believes the government should spend on a given intervention) for a QALY (up to $100,000 for life-saving treatment and up to $50,000 for life-extending treatment). This means that there is a large percentage of interventions where our healthcare system could deliver high quality care rates which are still cost-effective.

While 81.1 percent of treatments in cost-effectiveness studies were also found to be the most clinically effective, there remained 18.9 percent that were not in agreement. The cost that these interventions incur is significant to the degree of $412 billion, "or about 14 percent of the total annual US health care spending of $3 trillion" (Glick et al., 2015). Dissuading use of these therapies has the potential to save the US billions of dollars annually but comes at the cost of 1.21 million QALY (Glick et al., 2015). Currently, the US places a higher value on patient outcomes and prohibits the Patient-Centered Outcomes Research Institute from generating information related to increasing cost-effectiveness of treatments. Additionally, the government is making no efforts to convince physicians and patients to forgo these treatments for increased cost-effectiveness as well. With increasing expenditures under the ACA, it remains to be seen whether or not there will be a shift in the future towards these more cost-effective treatments. Nevertheless, there are still many areas where the US can improve quality of healthcare while decreasing spending. One method of achieving these goals is through implementation of Accountable Care Organizations (ACOs) which are designed to integrate the healthcare delivery system in the United States in an effort to reduce spending while maintaining, or improving, quality of healthcare. What ACOs are, what

they hope to achieve, and risks/challenges facing ACOs will be covered in the next chapter, as well as whether they have been successful thus far.

CHAPTER FOUR

Accountable Care Organizations

An Accountable Care Organization (ACO) is a "local, provider-led entity comprised of a wide range of collaborating providers ("Accountable Care Organizations (ACOs) - Health Cost Containment," 2016). They integrate care across numerous care settings (such as hospitals, physician practices, or clinics) and are held accountable to healthcare payers (Medicare, Medicaid, and private insurers) for providing quality care and limiting cost for a population. ACOs may come in a "variety of organizational forms, including integrated delivery systems, primary care or multispecialty medical groups, hospital-based systems, and even… networks of physicians" (Greaney, 2011) but are ultimately responsible for offering a vehicle for integrated healthcare delivery while addressing the shortcomings of our fee-for-service (FFS) system.

Fee-for-service, the traditional US method of paying for healthcare, pays physicians and hospitals based on the quantity of services provided and rewards high volume of care rather than low cost or high quality (Greaney, 2011). Advanced ACOs shift payments toward accountable care "through caps or reductions in some FFS rates and payments instead coming in the form of a per-member-per-month account that is tied to quality measures for the patient population" (McClellan, 2015). This system pays a certain amount for each patient a month and is comparable to payment through health maintenance organizations (HMOs) in this regard. This provides additional net revenues to healthcare organizations if costs are lower than this amount and quality is maintained.

Excess spending prominent under the FFS system is curbed due to healthcare organizations being at risk of taking a financial loss. Lack of quality care results in patients having to make repeat visits to hospitals. This increases spending, which places the healthcare organization at financial risk, and ensures that quality care is prioritized. This system encourages providers to give high quality care at a low cost (see comparative effectiveness and cost-effectiveness analysis in the previous chapter).

ACOs have seen increased use since the passing of the Affordable Care Act (ACA), especially in Medicare which has two major programs; Shared Savings ACOs and Pioneer ACOs. Shared Savings ACOs take a "middle ground" approach between FFS payment and the more "advanced" ACOs. Under the Shared Savings ACOs "clinicians still receive their usual FFS payments, but if the health care organization has spending lower than a benchmark, and if care results are as good or better than a set of measured quality benchmarks, then the ACO clinicians share in the savings" (McClellan, 2015). Participants in this system are not punished for going over this benchmark, which is set by the Center for Medicare and Medicaid Services (CMS) at the "average level of Medicare spending for patients served by an ACO during a baseline period before the start of the contract and is updated for each contract year on the basis of national Medicare spending growth since the baseline period" (McWilliams et al., 2016). Shared Savings ACOs are likely to see increased use in the future, with US Department of Health and Human Resources Secretary, Sylvia Mathews Burwell calling for "30% of Medicare payments to be based on non-fee-for-service (only) models by the end of 2016, and 50% to be so by the end of 2018" (Cutler, 2015). For comparison, "such payments did not

exist in 2011 and account for about 20% of Medicare payments (as of 2015)" (Cutler, 2015).

The Pioneer program, implemented in 2011, was similar to the Shared Savings plans for the first 2 years and adds a higher risk/reward system. Pioneer ACOs that "showed specified level of savings over the first two years (were) eligible to move a substantial portion of their payments to a population-based model" ("Pioneer ACO Model").

A population based model assesses the healthcare needs for a certain population and makes "health care decisions for the population as a whole rather than for individuals" ("What is Population-Focused Health Care"). This is addressed through measures such as eliminating distribution inequalities between population subgroups, educating a population about healthcare risks and what they can do to ensure healthiness, and a heavy emphasis on preventative care in an attempt to either prevent illness or catch it in early stages ("What is Population-Focused Health Care").

Thus far, Pioneer ACOs and Shared Savings ACOs have found reasonable levels of success and hold promise as an option for healthcare reform. In its first performance year of 2012, "the Pioneer model generated $147 million in total program savings, which exceeded the CMS' (Centers for Medicare & Medicaid Services) actuarial calculation of total savings ($87 million in the first year), and that the majority of ACOs outperformed rates in Medicare fee-for-service models for all 15 quality metrics for which comparable data were available" (Pham, Cohen, & Conway, 2014). In comparison to other Medicare plans, "Pioneer ACOs were associated with lower expenditure trends of approximately $427 in total Medicare Part A and Part B spending per beneficiary per year in the first

year and about $134 per beneficiary in the second year" (McClellan, 2015). Savings in the first year were largely attributed to few hospital readmissions and physician encounters while savings for the second year, after adjustment of the spending benchmark from the previous year, was due to preventable hospitalizations (McClellan, 2015). Unfortunately, not all the ACOs participating under the Pioneer model found the model a sustainable fit for their organization, leading 9 of the original 32 ACOs to transition to the Shared Savings Program. However, those that remained generated more total program savings in the second performance year than the first ($96 million vs $87 million) (Pham, Cohen, & Conway, 2014).Table 8 shows the combined results of the Pioneer Model ACOs for 2012 and 2013, including the ones that left the program after 2012.

Table. Early Performance in the Pioneer Accountable Care Organization (ACO) Model, 2012-2013[a]

	Year			
	2012[b]	2013[b]	2012[c]	2013[c]
Total program savings, millions, $[d]	87	96	128	96
Mean clinical quality score, %	70.8	84.0	73.0	84.0
Mean patient experience score, %	86.3	88.0	86.7	88.0

[a] Data from the Centers for Medicare & Medicaid Services financial reconciliation for the Pioneer ACO model in 2012 and 2013, based on Medicare claims for years 2009-2013, and clinical quality data submitted by Pioneer ACOs in 2013 and 2014.

[b] Results for 32 Pioneer ACOs participating in 2012 and 23 Pioneer ACOs participating in 2013.

[c] Results for the 23 Pioneer ACOs that participated in both 2012 and 2013.

[d] All savings minus all losses generated by the model.

Table 8 displays the total program savings, mean clinical quality score, and mean patient experience score for the Pioneer ACOs during 2012 and 2013 (Pham, Cohen, & Conway, 2014)

A separate article published in The New England Journal of Medicine, using a difference in differences approach[2] and researching the Medicare Shared Savings

[2] A difference in differences approach compares the effect of a treatment on a 'treatment group' versus a 'control group'. This is done by finding the change in the experimental group from before the independent variable was added to after it was added and then subtracting the difference found in the

Program (MSSP) found that the differential change in "total adjusted spending was -$144 per beneficiary in the 2012 ACO cohort as compared with the control group" (McWilliams et al., 2016) with the control group representing Medicare beneficiaries served by non-ACO providers. These results were not duplicated in 2013 with adjusted annual savings being "only -$3 per beneficiary in the 2013 ACO cohort as compared with the control group" (Williams et al., 2016). While still a decrease in spending, there remains a large decrease in the reduction found from 2012 to 2013. This decrease was mostly associated with entrants to the Shared Savings program in 2013 that did not participate in 2012, indicating these organizations may have struggled organizing their system in a manner conducive to ACOs, a risk discussed later.

Along with decreased spending, improvements in quality were also seen. McWilliams et al found that performance on "quality measures differentially improved in the ACO cohorts as compared with the control group for a few measures and did not differentially change for other measures." Referring back to Table 8, the Pioneer ACOs were found to have a "mean overall quality score of 84.0% in 2013 compared with 70.8% in 2012" (Pham, Cohen, & Conway, 2014). Additionally, "the mean performance score of all Pioneer ACOs improved in 28 of 33 quality measures" and their "mean performance score improved in 6 of 7 patient/caregiver experience ratings" (Pham, Cohen, & Conway, 2014).

Table 9 presents the differential changes in spending and quality from the McWilliams et al study between the 2012 cohort of MSSP ACOs compared to the control group. While decreased spending was found in almost all types of service and care

---

control group for the same period. This ensures that any variation found is due solely to the independent variable and no other outside factors.

settings, changes in quality were minimal. While not as promising as the data from the Pioneer ACOs, the MSSP ACOs were still able to find decreases in spending with no change in quality, indicating that reduction of services can be achieved under MSSP ACOs without adverse changes in healthcare This data illustrates how even with reductions in spending, ACOs show no decrease in overall quality and can actually be beneficial in many areas. Still, there are many risks and challenges facing ACOs, and these risks must be addressed if they are to become an integral part of our healthcare system in the future.

Measure	Mean in the ACO Group†	Difference between ACO Group and Control Group in Precontract Period	P Value‡	Differential Change in 2013 for ACO vs. Control Group	P Value‡
Total annual per-beneficiary spending ($)	9,942	133	0.22	−144	0.02
Annual per-beneficiary spending according to type of service and care setting ($)§					
Total acute inpatient care	3,321	42	0.28	−48	0.07
Facility¶	2,852	25	0.42	−43	0.05
Professional services	469	17	0.05	−6	0.22
Total outpatient care	3,250	−66	0.002	−4	0.81
Office	1,780	4	0.87	27	0.03
Hospital outpatient department	1,469	−70	0.006	−31	0.002
Total postacute care	1,133	76	0.23	−54	0.08
Facility	1,083	66	0.23	−49	0.07
Skilled nursing facility	855	56	0.31	−52	0.05
Rehabilitation facility	228	11	0.11	3	0.55
Professional services	50	10	0.27	−5	0.29
Home health care	623	30	0.13	−17	0.07
Durable medical equipment	305	−2	0.54	5	0.18
Hospice	208	24	0.18	−11	0.27
Annual quality measure‖					
30-Day readmissions (no.)	0.263	−0.002	0.46	0.004	0.15
Hospitalizations for ACSCs (no.)	0.056	−0.001	0.16	0.000	0.60
Congestive heart failure	0.016	0.000	0.06	0.001	0.11
COPD or asthma	0.012	0.000	0.50	−0.001	0.004
Cardiovascular disease or diabetes	0.024	−0.001	0.10	0.001	0.13
Screening mammography for women 65–69 yr of age (%)	56.9	2.0	<0.001	0.3	0.45
Preventive services for beneficiaries with diabetes (%)					
Glycated hemoglobin testing	75.0	1.5	<0.001	0	0.86
LDL cholesterol testing	79.0	0.9	0.12	1.1	0.05
Diabetic retinal examination	56.6	1.1	<0.001	0.3	0.17
All three	40.2	1.7	<0.001	0.6	0.06
Low-value services provided (no.)	0.437	−0.006	0.27	−0.003	0.25

* COPD denotes chronic obstructive pulmonary disease, and LDL low-density lipoprotein.
† We calculated values by adding the adjusted precontract difference between the ACO group and the control group to the unadjusted mean in 2013 for the control group to approximate the expected mean for the ACO group in 2013 in the absence of a differential change. Thus, dividing the differential change by this expected mean yields an estimate in relative terms.
‡ P values are for the comparison with zero.
§ Estimates for each category do not sum to total spending because we did not analyze lesser contributions separately from other categories of services or miscellaneous care settings. Professional services refer to physician and ancillary services appearing in the insurance-carrier claims file and reimbursed under Part B. Spending for an acute or postacute care facility was defined as the portion of acute or postacute care spending that was reimbursed under Part A.
¶ Inpatient-facility spending did not include capital payments, disproportionate-share hospital payments, or indirect medical-education payments. To adjust for between-hospital differences in Medicare payments for admissions in the same diagnosis-related group, we standardized inpatient-facility spending by calculating a national mean payment for each diagnosis-related group and summing mean payments across admissions for each beneficiary rather than using actual Medicare payments. Estimates were not appreciably changed by this standardization.
‖ Analysis of readmissions was limited to hospitalized beneficiaries. Hospitalizations for ambulatory care–sensitive conditions (ACSCs; conditions for which appropriate ambulatory care could potentially reduce the need for inpatient care) that were related to cardiovascular disease or diabetes included hospitalization for uncontrolled diabetes, short-term complications of diabetes, long-term complications of diabetes, amputation of the foot or leg, hypertension, angina without procedure, and congestive heart failure.[13] Estimates for all hospitalizations for ACSCs also included hospitalization for COPD or asthma and other ACSCs relevant to the Medicare population (dehydration, bacterial pneumonia, and urinary tract infection). Preventive services for diabetes were assessed among beneficiaries with a diagnosis of diabetes mellitus before the study year according to the CCW.

Table 9 displays the differential changes in spending and quality for the 2012 cohort of MSSP ACOs versus the Control Group (McWilliams et al., 2016)

ACO1s bear a striking resemblance to integrated delivery systems of the 1990s. Both models "create a care continuum and involve horizontal consolidation of hospitals; both may also create vertical integration of hospitals, physicians, and providers of post-acute care" (Burns and Pauly, 2012). Where the two differ is that integrated delivery systems are concerned solely with coordinated care whereas ACOs are also concerned with financial performance and clinical outcomes achieved under these systems ("What Are Accountable Care Organizations?", 2016). ACOs also "rely heavily on health information technology, data analytics, and decision support systems" such as comparative effectiveness and cost-effectiveness analysis which were discussed in chapter 3.The two are closely related to the point where integrated delivery systems have a place in the framework of ACOs. Therefore, it is important to recognize where integrated delivery systems failed in the past so those same mistakes are not repeated for ACOs in the future.

Integrated delivery networks did not achieve desired results for a variety of reasons. They "lacked the information technology, such as electronic health records… overpaid physicians for their practices…acquired hospitals without achieving economies of scale; and they failed to coordinate care for the population most in need, the chronically ill" (Burns and Pauly, 2012). Despite the label of "integrated delivery systems", "most did not really take a systems approach that involved integrated organizational planning from hiring and other personnel matters to physician culture" (Burns and Pauly, 2012). Fostering this level of integration between hospitals and physicians could be more difficult than anticipated as previous reviews "suggest that hospitals' efforts to partner with physicians using physician-hospital organizations or the

acquisition of physician practices have not promoted cooperation, improved quality, contained costs, or integrated clinical care" (Burns and Pauly, 2012). Integration of clinical decision support systems used to promote comparative effectiveness and cost-effectiveness analysis is also problematic as many physicians "often ignore the systems advice" (Burns and Pauly, 2012). Physicians value autonomy and are not easily swayed once they have become set in their practices, posing difficulties in adapting a revolutionary new system.

Another concern with the creation of ACOs and their development is the influence they may come to wield. Evidence suggests that "health care reform legislation has already prompted a number of mergers among health care providers" and economic evidence "indicates that market concentration has been a major factor spurring escalation in the cost of health insurance" (Greaney, 2011). As ACOs develop and come to take up more of the market concentration, they may be able to escalate cost of health insurance simply because they provide for so much of the population, similar to how monopolies can increase the prices of their product.

Finally, it is still uncertain whether ACOs will be successful on a widespread level and not solely in the experimental groups that have attempted them thus far. As mentioned earlier, nine of the original 32 Pioneer ACOs have dropped from the program, with two opting out of any form ACO model (Pham, Cohen, & Conway, 2014). For MSSP in 2012, the $238 million spending reduction "did not result in net savings to Medicare, because Medicare paid $244 million in bonuses without recouping losses from ACOs that had spending above benchmarks" (McWilliams et al., 2016). While this is still a decrease from the $252 million that was estimated to be spent ("NHE Fact Sheet",

2015), it carries the realization that not all ACOs will be able to keep spending below or at benchmarks set by the CMS. Proper application of comparative effectiveness, cost-effectiveness, and how the two interact needs to occur at a national level for ACOs to be successful. Furthermore, for a healthcare organization to transition to become an ACO in the first place, it must invest "heavily in electronic health records (EHRs) with installation costs; software licenses; hardware; paper records conversion; and office staff costs amounting to between $40,000 and $75,000 per physician" (Burt, 2012). This results in a large buy-in cost that many organizations may be unwilling to attempt unless potential bonuses from cost reduction and quality provision outweigh this risk.

Many events need to occur in order to bring about ACOs as a method of healthcare reform. First, and most importantly, the system must engage physicians and patients to encourage them to transition to an ACO system. Almost every ACO leader "highlights the importance of engaging physicians in the transition to this new model of care and payment" (Fisher et al., 2012). Patients may not initially understand the reduction in spending on healthcare and erroneously believe that such an endeavor will naturally lead to a decrease in quality. Physicians must take on a greater responsibility to educate their patients about treatment options and why one course of action is chosen over others, including the option of doing nothing at all, in order to keep patients involved in the program. Patient participation will directly affect physician participation as "providers' motivation to change current practices will increase with the proportion of their patient populations being cared for under ACO contracts" (Fisher et al., 2012). The incentive of increased compensation from bonuses given by providers, reduction of

spending, and increase of quality must be substantial enough to outweigh potential income from solely FFS models.

The benchmark set by the CMS should also accommodate ACOs based on the different subgroups under their care. ACOs should be held to a standard based on improvement for their population so that models serving disadvantaged[3] populations are still sustainable. This is because "ACOs taking care of a predominantly disadvantaged population would have a harder time meeting performance targets that reward a specific level of achievement…rather than the degree of improvement" (Fisher et al., 2012) and so would not be eligible for bonuses under the CMS. The subgroup of the population diagnosed with chronic conditions is of particular interest as chronic conditions make up the majority of Medicare spending ("NHE Fact Sheet", 2015).

Data collection will be of utmost importance to the success of ACOs. To achieve the goal of cost reduction while improving quality, intensive research must be pursued such as "tracking ACO formation; conducting qualitative research, such as case studies; conducting quantitative research, such as surveys; and aggregate performance data" (Fisher et al., 2012), much of which can be based on the comparative effectiveness and cost-effectiveness studies discussed in Chapter 3. Electronic medical records will play a vital part of this data collection process and will hopefully "maximize the opportunity to learn, minimize avoidable redundancy, and reduce burden imposed on payers and providers" (Fisher et al., 2012).

Ultimately, it is the "collaboration between evaluators, policy leaders, and other key stakeholders" that will be essential to the "successful implementation of a timely,

[3] The term "disadvantaged populations" refers to areas of increased poverty or of lower socioeconomic status where chronic illnesses, obesity, or other negative factors of health are more prevalent.

robust, and efficient approach to ACO evaluation" (Fisher et al., 2012). The transition to ACOs will require complete reform of the US healthcare system and getting people to accept this new ACO system will no doubt be the greatest hurdle. However, early results have proved promising in terms of spending reduction and quality improvement, giving policy leaders hope that this avenue may help alleviate many of the problems facing the US healthcare system.

CONCLUSION

The US healthcare system is currently undergoing reform but it is too soon to say if the changes being implemented will lead to comprehensive reform. Current reform and the application of an ACO model will impact the future of healthcare cost, accessibility, and quality in the US. The information from the previous chapters serves to outline a viable option for this comprehensive reform.

The history of US healthcare, the attempts at reform, and the reasons reform failed highlight mistakes of the past and how avoid those pitfalls in developing reform going forward. One lesson learned from that history is that incremental reform, small measures that affect only part of the healthcare system, lead to a fragmented and disjointed system. The United States has often found itself in a policy trap where millions are left without health insurance but enough people benefit from the system that it is resilient to change. Future reform efforts must be comprehensive in order to consider the needs of all groups and prevent this paradox.

Healthcare reform requires a certain environment in order for it to succeed. Values of the American public may be the most critical factor in the success of healthcare reform and it is important that any plan address the needs of the people to not only best satisfy the populace as a whole, but also to ensure the legislation matches the needs of the program in a timely manner. Moreover, physicians must properly apply clinical epidemiology and decision-making strategies to assess potential benefits and harms for their patients before determining the best course of action. A wide-ranging, non-partisan discussion on the needs and wants of the American public regarding healthcare would

help providers understand the realities of patients. Such a process would allow a clearer path to the policies and implementation that then could be legislated effectively.

To achieve high quality at a low cost, a physician calculates the probability of benefits and risks for an individual patient using prevalence, sensitivity, and specificity with an emphasis on providing high quality care at the most efficient and reasonable cost. Cost-effectiveness provides insight into how this clinical decision making and epidemiology can be used to reduce spending on unnecessary tests and treatments. Currently, physicians tend toward excessive utilization of technology, tests, and procedures that have no net positive effect on patients. Physicians need more access to that information and then can be tasked with properly choosing which test should be run, when, and addressing risks. Such a system will likely increase quality of healthcare, serve to improve cost-effectiveness, and increase patient and physician confidence in one another.

The opportunity for this program to improve quality and cost-effectiveness is increasingly being proven as comparative effectiveness, which focuses solely on absolute benefit for a patient, and cost-effectiveness analysis, which finds the ratio of costs over QALY, are often in agreement with one another. Small, early increases in spending show large increases in this agreement, implying that even options that are not the most cost-effective are still comparable in cost and are justified due to increased health benefits for patients. In tandem, comparative effectiveness and cost-effectiveness analysis should be researched and utilized jointly in order to best achieve high quality of care and low costs.

Accountable Care Organizations are the result of the early stages of those two mechanisms for looking at healthcare. They use integrated delivery systems as a means of

coordinating care, reducing the fragmentation usually observed in American healthcare providers. Clinical epidemiology is used when comparing tests and treatments to determine which one is superior in a given situation. Clinical epidemiology, applied to comparative effectiveness and cost-effectiveness analyses, must be utilized by ACOs. ACOs adopt these techniques in order to be successful and eventually may be applied at a national level. Medicare experimentation with ACOs has achieved promising results thus far, although there are still many challenges. The biggest of these is education, as patients and physicians must be convinced that this new system benefits them. Additional bonuses for achieving performance benchmarks can persuade physicians to adopt ACOs while increased quality of care and a more transparent cost structure will help patients to adapt to the new system as well. Reform toward such a system benefits patients, physicians, and the population as a whole. This system accounts for three tenets of healthcare: costs, access, and quality, and would serve to improve the health of our nation.

BIBLIOGRAPHY

"Accountable Care Organizations (ACOs) - Health Cost Containment."*National Conference of State Legislatures*. NCSL, n.d. Web. 24 Apr. 2016.

"BMC Health Services Research." The Social Value of a QALY: Raising the Bar or barring the Raise? N.p., n.d. Web. 24 Apr. 2016.

Brownstein, Ronald. The Atlantic. Atlantic Media Company, 13 Mar. 2014. Web. 26 Apr. 2016.

Burns, Lawton R. and Mark V. Pauly. "Accountable Care Organizations May Have Difficulty Avoiding The Failures Of Integrated Delivery Networks Of The 1990s" *Health Affairs* 31, no.11 (2012):2407-2416 doi: 10.1377/hlthaff.2011.0675

Burt, Jessica C. "ACO Value Metrics Series: A Cost-Benefit Analysis of ACO Formation." *Health Capital* 5.10 (2012): n. pag. Oct. 2012. Web.

"Concepts: Sensitivity, Specificity & Likelihood Ratios." Concepts: Sensitivity, Specificity & Likelihood Ratios. N.p., n.d. Web. 3 Jan. 2016.

Cutler, David M. "Payment Reform Is About to Become a Reality." *Jama*313.16 (2015): 1606-607. Web.

Davis, K. Stremikis, C. Schoen, and D. Squires, Mirror, Mirror on the Wall, 2014 Update: How the U.S. Health Care System Compares Internationally, The Commonwealth Fund, June 2014.

Fisher, Elliot S., Stephen M. Shortell, Sara A. Kreindler, Aricca D. Van Citters and Bridget K. Larson. "A Framework For Evaluating The Formation, Implementation, And Performance Of Accountable Care Organizations" *Health Affairs* 31, no.11 (2012):2368-2378 doi:10.1377/hlthaff.2012.0544

Garber, Alan M. and Harold C. Sox. "The Role Of Costs In Comparative Effectivenes Research" *Health Affairs* 29, no.10 (2010):1805-1811 doi: 10/1377/hlthaff.2010.0647

Glick, Henry A., Sean McElligott, Mark V. Pauly, Richard J. Wilke, henry Bergquist, Jalpa Doshi, Lee A. Fleisher, Bruce Kinosian, Eleanor Perfetto, Daniel E. Polsky and J. Sanford Schwartz. "Comparative Effectiveness And Cost-Effectiveness Analysis Frequently Agree On Value" *Health Affairs* 34, no.5 (2015):805-811 doi: 10.1377/hlthaff.2014.0552

Greaney, Thomas L. "Accountable Care Organizations - The Fork in the Road — NEJM." *New England Journal of Medicine*. NEJM, 6 Jan. 2011. Web. 24 Apr. 2016.

Kane, Jason. "Health Costs: How the U.S. Compares With Other Countries." PBS. PBS, 22 Oct. 2012. Web. 26 Oct. 2015.

Kliff, Sarah. "How Congress Paid for Obamacare (in Two Charts)."Washington Post. The Washington Post, 30 Aug. 2012. Web. 26 Apr. 2016.

Jamison, Dean T. Cost-Effectiveness Analysis. U.S. National Library of Medicine, n.d. Web. 26 Oct. 2015.

Massaro, Thomas A. Impact of New Technologies on Health-Care Costs and on the Nation's "Health" 36 (1990): n. pag. Web. 3 Jan. 2016.

McClellan, Mark. "Accountable Care Organizations and Evidence-Based Payment Reform." *Jama* 313.21 (2015): 2128-129. Web.

McWilliams, Michael, Laura A. Hatfield, Michael E. Chernew, Bruce E. Landon, and Aaron L. Schwartz. "Early Performance of Accountable Care Organizations in Medicare." *New England Journal of Medicine N Engl J Med* (2016): n. pag. Web.

Newton, Erin V., MD. "Breast Cancer Screening." : Practice Essentials, Overview, Guidelines. N.p., n.d. Web. 01 Feb. 2016.

"NHE Fact Sheet." - *Centers for Medicare & Medicaid Services*. CMS, n.d. Web. 25 Apr. 2016.

Pham, Hoangmai H., Melissa Cohen, and Patrick H. Conway. "The Pioneer Accountable Care Organization Model." *JAMA* 312.16 (2014): 1635-636. Web.

"Pioneer ACO Model." *Center for Medicare and Medicaid Services*. CMS, n.d. Web. 24 Apr. 2016.

"Risk of Developing Breast Cancer | Breastcancer.org." Breastcancer.org. N.p., 23 Oct. 2015. Web. 01 Feb. 2016.

Scott, Rick. US News. U.S.News & World Report, 24 July 2012. Web. 26 Apr. 2016.

Shi, Leiyu, and Douglas A. Singh. Delivering Health Care in America: A Systems Approach. Boston: Jones and Bartlett, 2004. Print

Shi, Leiyu, and Douglas A. Singh. Essentials of the U.S. Health Care System. N.p.: n.p., n.d. 2015 Print.

"Slide Rule." TRENDWATCH (n.d.): n. pag. American Health Association. Web. 31 Jan. 2016.

Starr, Paul. Remedy and Reaction: The Peculiar American Struggle over Health Care Reform. New Haven: Yale UP, 2011. Print.

"Study Reveals Sensitivity of Common Exercise Tests for Suspected CAD May Be Overestimated." DAIC. N.p., n.d. Web. 4 Jan. 2016.

""THE HISTORY OF HEALTH CARE COSTS AND HEALTH INSURANCE." Wisconsin Policy Research Institute 19.10 (2006): n. pag. Web. 4 Jan. 2016.

"The RAND Health Insurance Experiment: What You Need to Know." The Health Insurance Experiment: A Classic RAND Study Speaks to the Current Health Care Reform Debate. N.p., n.d. Web. 05 Jan. 2016.

"What Are Accountable Care Organizations?" *What Are Accountable Care Organizations?* American Academy of Pediatrics, n.d. Web. 25 Apr. 2016.

"What Is Population-Focused Health Care? Loyola Univ. Online." *What Is Population-Focused Health Care? Loyola Univ. Online*. Loyola University, n.d. Web. 24 Apr. 2016.