2018

A COMPARISON OF QUALITY INDICATORS BETWEEN MEDICARE ACCOUNTABLE CARE ORGANIZATIONS AND HEALTH MAINTENANCE ORGANIZATIONS USING PUBLICLY AVAILABLE DATA

William W. Campbell III
Virginia Commonwealth University

A Comparison of Quality Indicators Between Medicare Accountable Care Organizations and Health Maintenance Organizations Using Publicly Available Data

A dissertation submitted in partial fulfillment of the requirements for the Doctor of Philosophy in Health Related Sciences at Virginia Commonwealth University

by

W. Wesley Campbell, III
MHA, Virginia Commonwealth University, 1998
BSBA, University of Phoenix, 1996

Director: Carloyn Watts, PhD
Chair, Department of Health Administration

Virginia Commonwealth University
Richmond, Virginia
March, 2018

Acknowledgements

In 1984, as an undergraduate student, I flunked out of VCU. With the completion of this dissertation, the same institution will award me a PhD. The path I followed over thirty-four years between these events was circuitous and proof positive there is a right time for everything in one's life.

My heartfelt thanks to Dr. Carolyn Watts, who as my advisor and dissertation committee chair challenged and pushed me to explore more deeply while being a great advocate for my success. If not for a conversation with Dr. Watts in the spring of 2014, it is very likely I would never have enrolled in this program. I am grateful to my committee: Dr. Gloria Bazzoli, Dr. Askar Chukmaitov and Dr. Gina Engel, for their support and guidance.

I learned so much from my fellow students in the 2014 cohort, especially Dr. Tyler Corson and Dr. Jenny Inker. Their friendship and encouragement was invaluable to me and I will be forever appreciative. The faculty with whom I worked and from whom I learned were outstanding. Dr. Dianne Dodd-McCue is without doubt the finest teacher I ever encountered. Dr. Kupstas and the entire staff of the Dean's office were wonderfully accommodating and made a daunting endeavor seem much more manageable.

My children, Ella and Will, reminded me every day that I could do this. Their belief in me gives me strength. They always made it possible for me to be on-campus or working in the evening. I hope as they grow they will look back and remember that even with a full-time job and being a PhD student/candidate their dad was always there for them.

The one person to whom I owe the most is Sara, my wife, partner in life and best friend. When I have a tendency to get down on myself and worry whether I can do anything, she lifts me. She is my rock. If not for her, I doubt I would ever have found the confidence to embark on this journey. Thank you, Sara. I love you and admire you. This achievement is as much my family's as it is mine.

Table of Contents

List of Tables

List of Figures

Abstract

A COMPARISON OF QUALITY INDICATORS BETWEEN MEDICARE ACCOUNTABLE CARE ORGANIZATIONS AND HEALTH MAINTENANCE ORGANIZATIONS USING PUBLICLY AVAILABLE DATA

By W. Wesley Campbell, III, PhD

A dissertation submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Health Related Sciences at Virginia Commonwealth University

Virginia Commonwealth University, 2018

Major Director: Carolyn Watts, PhD, Chair, Department of Health Administration

The purpose of this study is to explore differences in quality between Medicare Accountable Care Organizations (ACO) and Health Maintenance Organizations (HMO). Three outcomes measures reported by these plans use different methodologies but possess enough alignment to permit comparison: percent of diabetic patients with last HbA1c > 9.0%, colon cancer screening rate and ER visits per 1,000. These outcomes are the dependent variables (DV). A secondary purpose is to explore differences in quality based on the size of the beneficiary population served, using the same measures.

As the Medicare program faces threats to its solvency in coming decades, with 10,000 baby boomers becoming eligible every day, and the ongoing national conversation about healthcare more generally, approaches to Value-Based Purchasing (VBP) are becoming more common. Organizations seeking to identify the types of VBP arrangements in which they should enter have precious little information on the comparative performance of VBP approaches relative to outcomes measures. Different structures create different incentives through the plan design and risk/reward. The convergence or dissipation of the plan incentives at the level of the provider, particularly in primary care, may be a source of variance.

This study is retrospective, non-experimental, and uses publicly available data on the performance of Medicare ACO and HMO plans in calendar year 2015, for the identified measures. Using the Donabedian Structure-Process-Outcome framework, this study explores the impact of structure by type of plan and size of population served, relative to the outcomes. Race, average Hierarchical Condition Category (HCC) risk score and duration of operations are control variables. The analysis uses multiple hierarchical regression to better understand the relationship between the independent variables (IV) and DVs, after the impact of the control variables (CV).

After controls, the IVs did offer some explanation of variation in outcomes. The ACO plans fared better on HbA1c control, while HMO plans had fewer ER visits per 1,000. No discernable difference existed between the HMO and ACO plans with regard to colon cancer screening rate. Serving larger populations led to better performance on all three measures. In general performance was worse on each measure in both models when the percent of not-White patients or average HCC risk score increased. A longer duration of operations also associated to better performance on the outcome measures.

Keywords: Accountable care organization, health maintenance organization, HEDIS, NQF, quality, SPO

Chapter 1: Introduction

The purpose of this study is to examine the role of structure relative to outcome using the Donabedian quality framework. The focus is on Medicare Health Maintenance Organizations (HMOs) and Accountable Care Organizations (ACOs). The outcomes of interest describe clinical, prevention and utilization measures based on the percent of population in each model meeting certain criteria. Because no existing similar studies appear to exist, this effort is timely and relevant to the "big" questions concerning quality and coverage ongoing in the national debate about healthcare reform.

Background

Questions about approaches to healthcare coverage, quality and ever increasing costs are not new. Theodore Roosevelt included a campaign plank for national healthcare in his "Bull Moose" Presidential run in 1912 (Hoffman, 2008). In the World War II era, the organized labor movement began to make inroads into employer-sponsored insurance through collective bargaining negotiations (Hoffman, 2008). In 1965, the Medicare program came into existence to contend with coverage and cost, based on policy discussions beginning in the Truman administration. As the US national healthcare expenditure (NHE) reaches 17% of gross domestic product (GDP), Medicare a/k/a the Centers for Medicare and Medicaid Services (CMS), consumes 20% of the NHE (Centers for Medicare & Medicaid Services (CMS), HHS, August, 2016).

Speculation about insolvency of the Medicare Trust Fund projects funding through perhaps 2030 (Dickson, 2016). CMS in more recent years began to develop novel approaches to confronting increasing costs first through the implementation of the Inpatient Prospect Payment System (1982), commonly known as "DRGs" and later through the Resource-Based Relative Value Scale methodology and development of the Medicare Fee Schedule (MFS) for Part B services (1992). As managed care became de rigueur, particularly in the 1980s, CMS developed Part C, Medicare Managed Care or what we know today as "Medicare Advantage" as an alternative to fee-for-service. In 2010, through the passage of the Patient Protection and Affordable Care Act (PPACA), CMS began to create innovative approaches by altering incentives for providers, particularly in primary care (PPACA, 2010). One of these new approaches is the Accountable Care Organization (ACO).

Medicare managed care plans and ACOs enhance the focus on primary care services, directly and indirectly, using different tools and financial relationships (Blackstone & Fuhr, 2016; Bolch, 2013). These relationships are, at some level, simply variations on a theme with strong resemblance to Value-Based Purchasing (VBP) arrangements, which continue to proliferate in commercial and government insurance offerings (Rappleye, 2015). Similarities between Medicare Advantage and ACOs exist, congruent with VBP, through a focus on quality and cost.

Despite the similarities in approaches, significant differences also exist between Medicare Advantage and ACOs. Viewed from the perspective of structure in the Donabedian "Structure-Process-Outcome" (SPO) framework, understanding these differences may better serve robust evaluation of the quality in these programs (A. Donabedian, 1966). This is particularly relevant when it comes to Health Maintenance Organizations (HMOs) in Medicare Advantage as they comprise 64% of the total available plans nationwide (Kaiser Family Foundation, May, 2016).

No published works exist documenting the differences in the structure of these programs relative to the SPO framework. Contrasting these programs through the evaluation of results increases the existing knowledge of their applicability and success in contending with increasing cost and the lack of apparent improvement in ambulatory quality of care (Levine, Linder, & Landon, 2016). The approach outlined herein appears to be novel and is an early step towards filling the literature gap regarding performance of these important programs.

The following describes the programmatic approaches and history of the Medicare Shared Savings Program (MSSP) ACO and HMO programs in Medicare. The conceptual framework discussion will integrate the SPO through two perspectives permitting a comparison of outcomes between ACOs and HMOs and by size of the beneficiary population served.

Medicare Shared Savings Program ACOs

CMS created the MSSP through PPACA in response to long-standing and increasing concerns over quality, cost and access. The ACO is a group of providers who voluntarily collaborate to deliver high quality, coordinated care to Medicare beneficiaries. These providers can include physicians, hospitals and other providers, e.g., long-term care, skilled nursing, DME, therapists (Centers for Medicare & Medicaid Services (CMS), HHS, 2015a).

Three “tracks” exist in the MSSP, each with differences in upside financial benefit and downside financial risk. All ACOs must meet certain quality benchmarks to access upside benefit. The level of performance relative to quality benchmarks determines the portion of the maximum “shared savings” available. ACOs must exceed a savings threshold to qualify for shared savings. Shared savings is the difference between medical expense budget for a population of beneficiaries, based on historical benchmarks, and actual expenses. Table 1 highlights differences in risk/shared savings between the tracks:

Table 1: *MSSP Share Savings and Risk by Track*

	Track 1	Track 2	Track 3
Shared savings	+	+	+
Shared risk	-	+	+
Shared savings cap	50%	60%	70%
Risk sharing cap	N/A	5%	7.5%

In track 1, actual shared savings may not exceed 10% of the ACO's updated medical expense benchmark. In tracks 2 and 3, actual shared savings may not exceed 15% and 20%, respectively, of the medical expense benchmark. Track 2 and 3 ACOs have the option to specify symmetrical medical savings rates (MSR) and medical loss rates (MLR), between 0% and 2% in .5% increments. Starting in calendar year 2017, a track 1+ option exists as an "Advanced Alternative Payment Model" (APM) for purposes of the Medicare Incentive Payment System (MIPS).

Calendar year 2015, was the third "performance year" for the MSSP ACOs, although not all operated for all three years. The vast majority of ACOs are non-risk bearing. Table 2 shows the number of ACOs by year, and distribution by track for 2015.

Table 2: *Count of MSSP ACOs*

Year	Count of MSSP ACOs
2012	114
2013	220
2014	333
2015	439
Track 1	401
Track 2	3
Track 3	35

CMS publishes publicly available performance data on all MSSP ACOs. Table 3 provides data on savings for the first three performance years, including all ACO tracks, reported by CMS.

Table 3: *MSSP ACO Shared Savings 2013-2015*

	2013	2014	2015
Total savings	$705M	$806M	$429M
Total shared savings	$315M	$341M	Not reported
ACOs earning shared savings	89	92	119
ACOs generating savings with no sharing	60	89	83

The "Not Reported" amount of total shared savings in 2015, performance year 3, was perhaps intentional on the part of CMS. Analysis of the performance data indicates 189 ACOs overspent their budgets by a total of $1,138,967,553. Another 203 ACOs generated savings of $1,568,222,249. The difference is the $429M total savings reported by CMS. However, summing the shared savings amount reported in the performance data shows $645,543,866 returned to a subset of the 202 ACOs (119) generating savings. This means CMS returned more dollars in shared savings than the total MSSP ACO savings – over $200M more. In other words, the MSSP program *increased* total costs by more than $200M in 2015 (Centers for Medicare & Medicaid Services (CMS), HHS, 2014; Centers for Medicare & Medicaid Services (CMS), HHS, 2015b; Centers for Medicare & Medicaid Services (CMS), HHS, 2016b).

Most ACOS in 2015, and earlier, exist through Independent Practice Associations (IPA) or group practices and nearly one-half include an acute care hospital (Centers for Medicare & Medicaid Services (CMS), HHS, 2017b). Nearly two thirds of these ACOs, 62%, operated in population dense metropolitan areas with just over 6M beneficiaries attributed to ACOs with HSS headquarters in: Boston, New York, Philadelphia, Atlanta and Chicago (Centers for Medicare & Medicaid Services (CMS), HHS, 2017b). Overlaying maps of ACO locations with physicians per population density and higher per capita beneficiary expense reveals what appears to be a high degree of correlation between the highest cost areas, more physicians per capita and presence of ACOs.

Medicare Advantage

Medicare managed care began in the 1980s, became "Part C" in the 1990s and was renamed "Medicare Advantage" in 2003 as Medicare began to more actively push the program. In 1992, total enrollment was 6.9M. Enrollment in 2016 was 17.6M and is projected to increase to 22M by 2020 (Neuman & Jacobsen, 2014). Plans offered through delivery systems, called "Provider-Sponsored Organizations" (PSO) are increasing in number, but in 2016 insurance-based plans accounted for 70% of all plans nationally. In fact, companies like: Anthem, UnitedHealthcare, Cigna and Humana offer 21% of all plans (Kaiser Family Foundation, May, 2016). The HMO plans are 64% of the total, with UnitedHealthcare and Humana offering 39% of them (Jacobsen, Casillas, Damico, Neuman, & Gold, 2016).

Plans bid contracts to CMS to provide care for a defined population based on geography at no smaller than a county level. The plans take risk from CMS in the form of a per member per year (PMPY) capitation and cover all Part A and B services. Many plans offer Part D benefits as well. ACOs do not offer Part D benefits. The plans enroll a provider network to provide care, and in the case of HMO plans pass risk onto primary care providers (PCP) through a capitation rate. Beneficiaries must have a PCP in HMO plans. Other plan types offer more flexibility. Most provider networks are limited to very narrow, particularly in the HMO plans (Jacobsen, Trilling, Neuman, Damico, & Gold, 2016). Plans have great flexibility to define premiums and cost sharing with beneficiaries.

Total Medicare Advantage HMO spending was slightly less than Medicare FFS spending by $378B, in 2016 (Biles, Casillas, & Guterman, 2016). In the top 25 counties with the greatest difference between Medicare Advantage and Medicare FFS expenses, HMOs underspent Medicare FFS by $5.27B. Everywhere else, the HMO plans overspent Medicare FFS by $1.0B while all other Medicare Advantage plans types overspent Medicare FFS by $3.8B (Biles et al., 2016). The net savings to Medicare was approximately $427M for the entire Medicare Advantage program.

The list of states with plan offerings covering two thirds of Medicare Advantage enrolled beneficiaries includes, in descending order of enrollment: California, Florida, New York, Texas, Pennsylvania, Ohio, Michigan, North Carolina, Minnesota, Georgia, Arizona, Tennessee and Wisconsin. This list certainly shares some geographic overlap with ACOs while covering a much larger geographic range including less population dense states.

Conclusions

Medicare clearly recognizes the need to develop better approaches to resolving economic instability in the program, while improving quality. The HMO program is older than the ACO program, but programmatic intervention by Medicare is not a new phenomenon. The programs share some common attributes while simultaneously possessing significant differences, explored more fully in the chapter, "Conceptual Framework." The SPO framework is useful to elucidate contrasts of structure in assessing the achievement of these programs on quality measures.

The structural differences between ACOs and HMOs originate with CMS through the design of the programs. The examples of these differences, discussed in the Conceptual Framework, are structural responses to the incentives running through these models. These incentives exist in layers first through the organization, then through the provider and lastly to the beneficiaries. The Conceptual Framework connects the SPO to these programs by examining the structural differences which represent points where the incentives may diverge.

Chapter 2: Conceptual Framework

Much of the focus in today's healthcare environment is on outcomes, which some consider to be indicative of quality. Many organizations, including the National Committee on Quality Assurance (NCQA) and the National Quality Foundation (NQF), promulgate quality measurements. Measurements could be end-point oriented, e.g., blood sugar control in a diabetic population. Measurements could also be cost or utilization oriented, e.g., ER utilization. Other measures could be process-related activities that represent an end-point or address a gap in care, but may not be a direct indicator of current health status, e.g., colon cancer screening. Value-based purchasing approaches place a great deal of emphasis on outcome measures.

"Quality" remains as fungible today as it was when Donabedian attempted to put some parameters around its meaning (A. Donabedian, 1966). Despite the efforts of the Institute of Medicine and others to cement a common understanding of quality, various methodologies for measurement and evaluation compete for attention often with different levels of measurement or foci (Institute of Medicine, 2000).

From a research perspective, a singular focus on the results of care is inadequate in understanding quality. The Donabedian SPO framework relies on assessing Structure, Process and Outcome in evaluating the quality of healthcare services (A. Donabedian, 1966). Applying the SPO framework is a time honored and frequent approach in outcomes research (Ayanian & Markel, 2016). The directional SPO elements provide the closest perspective to a 360˚ view of the delivery-side inputs. In the SPO, structure is foundational as it exists prior to any process and facilitates process. Outcome results directly from process. The SPO framework permits evaluation of quality as something more than simply an outcome. This framework enables the researcher to not only ascertain whether outcomes may be different but to explore hypotheses as to why the differences may exist.

The lack of agreement between the various organizations developing quality measures is in and of itself something of a contradiction to Donabedian's approach – something he might characterize as the absence of a "firm foundation of prior agreement" before attempting to assess quality (A. Donabedian, 1988). On balance, however, while specifics of individual measures may differ based on methodology some agreement exists between the various schemes as to the types of measures. Most include some health status measurements, HbA1c level or blood pressure, for example. Most include process measures geared towards prevention or early detection, like colon or breast cancer screening. Measures also exist to gauge the appropriateness of services, e.g., antibiotic use in upper respiratory infections or otitis media. Utilization measures are common as well, for example, ER visits per 1,000 population. So, while firm foundations may not exist regarding the specifics of measurements, like in the case of HbA1c level applicability to Type I or Type II diabetics, the idea that blood sugar control is important in a diabetic population supports the measurement methodology.

Structure

Donabedian referred to structure in terms of resources from the perspectives of allocation, location, funding and organization (A. Donabedian, 1966; R. C. Donabedian A., Wheeler, & Wyszewianski, 1982). In the SPO, structure describes attributes of the care setting (A. Donabedian, 1988). The attributes and resources are human and capital, organizational and financing. The Donabedian framework suggests that improvements in structure would enable process improvements leading to improvements in outcome. Differences in structure may be more easily observed than differences in process, and as a result structure is often an independent variable in outcomes research. The chapter entitled, "Literature Review," identifies several examples of contrasts in outcomes based on differences in structure.

The size and composition of a provider panel enrolled in a managed care contract or employed in an organization would be examples of structure. In this research, the size of the beneficiary population served by an ACO or HMO is an element of structure for three important reasons. First, in order to serve a greater number of beneficiaries, the plan or program would need more provider, locations, staff and equipment. Second, the size of the beneficiary population determines total medical expense. While cost is an outcome relative to the SPO, budget is an organizational attribute and as such is an element of structure. Third, greater mass may provide better leverage to optimize the response to incentives. Risk spread over a larger population might be easier to manage. In this study, the size of the beneficiary population is an independent variable and represents differences in structure.

In today's environment we could add the quality measurement as a structure variable because different systems of measurement exist. In using structure as a basis for comparison, the differences in what quality means relative to each model presents unique challenges in evaluating outcomes. It is for these reason, a differentiation of reporting mechanisms is part of the delineation of HMO and ACO attributes that follows.

The following section discusses several key structure differences between HMOs and ACOs. On balance, the incentives through the ACO program are more diffuse than incentives existing in the HMO program. This is in no small part because HMO plans, providers and patients share some portion of the same financial risk. In the delineation of structure differences these incentives become apparent through approaches to enrollment, provide participation, and payment/reporting.

Differences in Structure Between Medicare HMOs and ACOs

With a good understanding of HMO and ACO programs from the Introduction, further delineation of structure differences is possible. The following is not the complete menu of differences, but in terms of the SPO framework serves as the basis of quality comparison relative to the structure of these programs. If over time more ACOs move to risk-bearing models as CMS certainly desires, these differences may or may not remain significant. It is important to note, also, that while differences between Medicare Advantage and FFS are well chronicled, the recent implementation of MIPS may obscure future differences.

Difference: beneficiary enrollment versus attribution. Beneficiary enrollment in Medicare Advantage is active and voluntary. The beneficiary must choose to enroll. Premium payments go to the plan, not Medicare. In ACOs, beneficiaries become attributed based on where they receive the "preponderance of care" if their provider affiliates with an ACO. This attribution occurs at the end of a performance period as opposed to Medicare Advantage enrollment, which occurs during open enrollment.

When a beneficiary in an HMO plan presents for primary care with a network provider, the beneficiary has a plan ID card. When a beneficiary presents for primary care to an ACO provider, the beneficiary has a Part B card with no identification as to attribution status. The practical implications of this are that the HMO providers know when an HMO subscriber is receiving care at the time of care but the ACO providers will not know if that beneficiary will become attributed until the end of the calendar year. Beneficiaries enrolled in a Medicare Advantage plan cannot be attributed to an ACO.

Difference: provider participation in HMOs versus ACOs. A provider may only participate with one ACO. In Medicare Advantage, depending on the plan's network, providers could contract with and receive capitation from multiple HMOs. ACO beneficiaries can see any Part B participating provider. HMO subscribers must use in-network providers. An ACO beneficiary could receive primary care services from different providers in different organizations or systems during the course of a year. An HMO beneficiary must choose a PCP from the list of in-network providers. In order to see a different PCP, the beneficiary must notify the HMO of their desire to change.

ACO beneficiaries could see an ACO PCP while getting specialty care from providers not part of that ACO, or even in a different ACO. HMO beneficiaries can obtain specialty care from in-network providers, usually only with a referral authorized by the PCP. The provider "network" available to the ACO beneficiary is any Part B participating provider. The same is true of acute care providers. The HMO provider network is much more limited, including which hospitals the beneficiary could use. Of course, there are exceptions for emergencies and travel, etc. The Medicare Advantage Preferred Provider (PPO) networks are typically more open than HMO networks, but still not as open to beneficiaries as ACOs.

The practical implications relative to incentives are evident through the "shaping" of the provider networks and restrictions on accessing out-of-network providers. The ACO providers have limited ability to restrict beneficiaries from accessing care whenever and wherever they wish. Yet, the ACO is still responsible for the total cost of care and quality for Parts A and B for any attributed beneficiary who they do not know will become attributed until after the episodes of care. CMS rules permit Medicare Advantage plans to tell beneficiaries which providers they may see for care. ACOs cannot impose these kinds of restrictions.

Difference: payment and reporting in HMOs Versus ACOs. In Medicare Advantage, the link between payment and reporting is very different from Medicare ACOs. Generally speaking, and specifically for primary care services, reporting in Medicare HMOs does not drive payment. In ACOs, reporting, i.e., billing, does drive payment and it serves a secondary purpose regarding quality measurement.

Payment. Payment mechanisms for episodes of care are very different in HMOs and ACOs. Any provider rendering services to an ACO attributed beneficiary bills Medicare just as they would for any Medicare FFS beneficiary. This is true whether the provider is a PCP, specialist, hospital, diagnostic facility or anything else. Part B beneficiaries have a 20% cost sharing for most services covered under Medicare FFS. This applies to services provided through ACOs, which are by definition Medicare FFS. Medicare HMOs have the ability to change Medicare benefits particularly with respect to beneficiary cost-sharing for covered services. ACO attributed beneficiaries are FFS beneficiaries. Medicare HMO enrolled beneficiaries are not FFS beneficiaries.

In HMOs the PCPs are almost universally capitated at least for the "bucket" of primary care services covered through the plan. The PCP submits an encounter level report of services provided, with numeric codes mapped to the Healthcare Effectiveness Data and Information Set (HEDIS) measures promulgated by the NCQA, not for payment but rather for reporting. These PCPs also transmit information to the HMO plans to identify services rendered allowing the plans to adjudicate beneficiary cost-sharing. An in-network HMO specialist might take capitation for certain patients or programs, e.g., Special Needs Programs, but for many of these beneficiaries will submit a claim for payment to the HMO. The payment is then based on the HMO's fee schedule that the provider agrees to as part of their network participation contract.

The HMO fee schedule could be the same as the MFS or it could be different and the HMO beneficiary could have the same or different cost-sharing. The Medicare Advantage plans have the ability to alter Medicare benefits, particularly beneficiary cost-sharing. ACOs do not have this ability.

Reporting. Primary care providers in HMOs report the same Current Procedural Terminology (CPT) identifying services as any other Part B provider. The difference is how these codes determine payment, as discussed above. The PCPs also report additional data at the encounter level using "Category II" CPT codes. These are the numeric codes mapping to the HEDIS measures. For instance, if during a PCP encounter the HMO beneficiary has a blood pressure of 140/90, the PCP would report codes 3077F for most recent systolic pressure ≥140, and code 3080F for most recent diastolic pressure ≥90. If the visit was for routine follow-up, the PCP would also report an Evaluation and Management CPT code to identify the service. Correct and complete diagnosis coding is essential in Medicare Advantage, sometimes for payment purposes but more to identify patient risk and as a result HMO providers typically report every applicable diagnosis on at least an annual basis. Diagnosis codes and other factors map into Hierarchical Condition Category (HCC) risk scores (Centers for Medicare & Medicaid Services (CMS), HHS, 2016a). These risk scores determine the capitation rate plans get from Medicare. They also influence capitation rates PCPs receive from HMO plans.

Encounter level reporting in ACOs functions primarily for payment. It is no different from claims non-ACO providers submit for services. While Medicare encourages any part B provider to report Category II codes it is entirely optional outside Medicare Advantage. Other mechanisms for reporting also exist, e.g., Physician Quality Reporting System (PQRS), which are retrospective in nature, typically. In fact, it is through mechanisms like PQRS that ACOs sometimes report quality data to Medicare. In general, however, Medicare directs the ACO to sample assigned beneficiaries and extract data from the medical record on a patient-by-patient basis for ACO-specific quality measures (RTI International, 2015).

Similarities Between Medicare HMOs and Medicare ACOs

Despite all the areas of differentiation outlined above, one element of striking similarity exists between Medicare HMOs and ACOs. Namely, when it comes to financial performance the more beneficiaries in the pool, the better.

Figure 1 indicates in calendar year 2015, and preceding years, the ACO minimum savings rate, i.e., the threshold, decreases as the number of beneficiaries increases.

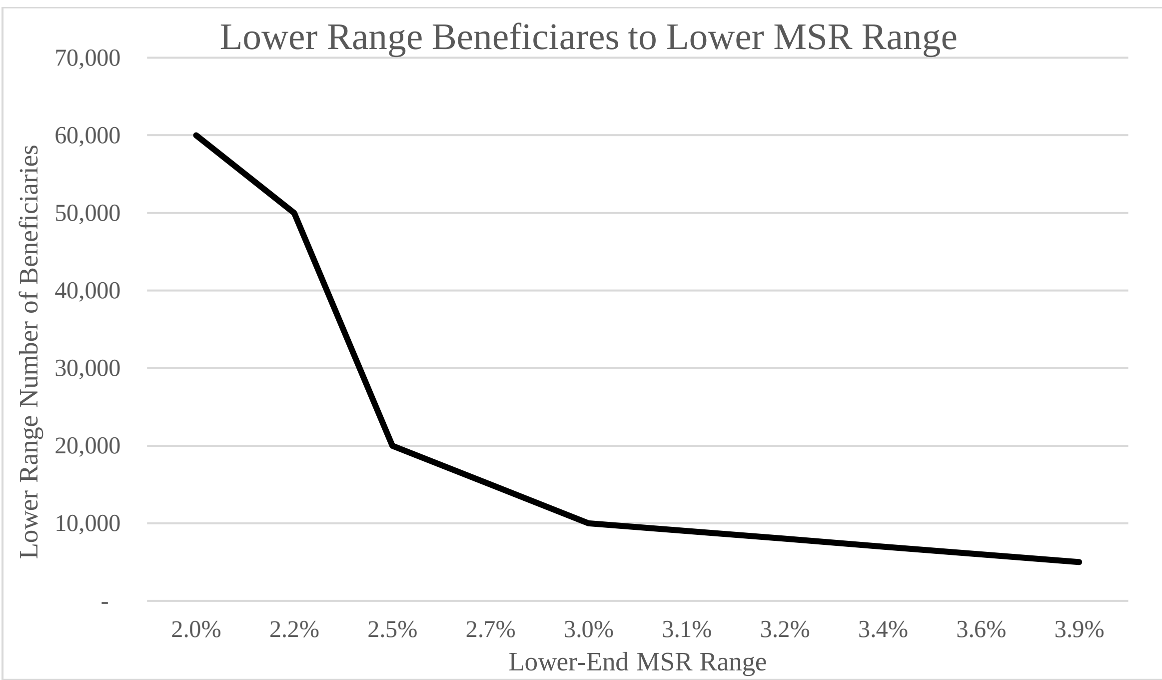

Figure 1: ACO Beneficiaries and MSR

The greater the number of attributed beneficiaries, and correspondingly the larger the total benchmark expenditures, the lower the percentage savings threshold to earn shared savings. An ACO with 10,000 attributed beneficiaries would have to nearly double the savings rate

(3.9%) of an ACO with 60,000 beneficiaries (2.0%) to earn shared savings. If the benchmark per capita expense is $10,000, the larger ACO would need to exceed $12M in savings to earn a share at a 2.0% MSR. The smaller ACO would have to exceed $3.9M in savings to earn a share at a 3.9% MSR. In this example on total benchmark expenses of $600M and $100M, this structure clearly poses a challenge to smaller ACOs. ACOs with 10,000 beneficiaries, 17% of the larger ACO with 60,000, have a dollar savings threshold of 33% of the larger ACO's. Larger ACOs need to reduce per capita expenses by less than smaller ACOs to earn shared savings.

In Medicare HMOs, the PMPY received from Medicare would exhibit the same tendency. Larger HMOs, will take in more revenue but need to save a smaller percentage per capita on a greater number of enrollees to benefit economically from its risk-taking. This applies as well to the PCPs taking risk from the HMOs in terms of the number of patients in their panels. The HCC risk scores impact this and existing evidence indicates Medicare Advantage plans with more enrollees falling into HCCs with greater margin potential generate more profit (Newhouse et al., 2013).

Regarding the research questions based on the size of the beneficiary population as a difference in structure, this incentive to create mass is important. If mass creates financial incentives for the HMO or ACO, or HMO PCP, does it similarly enable achievement on quality measures?

Process

The SPO describes process essentially as "what" in terms of activities or inputs to care. Where structure refers to attributes, process is the application of those attributes at the level of the patient (A. Donabedian, 1988). Process incorporates both the technical and interpersonal skills evident in a system of care (A. Donabedian, 1966). It might describe history taking, accurate diagnosis, appropriate follow-up, or the openness of the communication in the provider-patient relationship. Process is also often an independent variable in outcomes research frequently in conjunction with structure.

Process can also describe patient involvement (A. Donabedian, 1988). While it is not a patient-centric perspective, like the Anderson Behavioral Model of Health Services Utilization, it does incorporate elements of patient activation or care seeking (Andersen, 1995; A. Donabedian, 1988). Differentiating structure and process can be difficult as the boundaries are not always evident. The existence of a diabetes self-management program with diabetic educators and nutritionists would be a structural attribute, for example. At the same time, patient participation in a diabetes self-management program, recommendations made and adherence to them would fall under process.

Another example of the often confusing boundary between structure and process would be the need for PCP referral to access specialty care in an HMO. The existence of this requirement relates to payment and is therefore an element of structure. However, the determination of the need for specialty referral relates directly to technical process in the system of care.

The purpose of this study is not to assess the impact of variation in process on outcome. At the same time, process is not completely divorced from structure. All thing being equal, differences in the technical or interpersonal processes relative to the SPO should exhibit no more variation in Medicare HMOs than they do in ACOs. Board eligibility or certification would be a prerequisite to participation in either and ACO or HMO, for instance, in the vast majority of cases. The US Preventive Task Force (USPTF) recommendations for diabetic self-care, frequency of HbA1c testing and screening for nephropathy or retinopathy, are often the basis for treatment plans used by PCPs nationally. In fact, PPACA mandated coverage of USPTF recommendations with A and B evidence ratings with no patient cost-sharing (Shuval et al., 2017). The same is true for recommendations regarding cancer screening or other prevention-oriented services. Despite the difference in measurement methodology in some cases between HEDIS and NQF, the congruence in the domains of measurement supports the assumption that variation in technical process would not be significantly different between HMOs and ACOs.

This is not to say differences in process do not exist in these programs. Rather an assumption is being made that process variation exists in both programs to the same degree. In other words, differences in the "quality" of care provided through these programs may be attributable to structure more so than process (Muhlestein, Croshaw, Merrill, Pena, & James, 2013). This is in no small part because the different structures create different incentives for the participating organizations or providers. While the creation of these incentives based on structure is out of scope for this research, the nature of these incentives deserves some discussion. Since the incentives are mostly economic in nature, and relate to the cost of care or capitation rate at the ACO or HMO level, an obvious approach would be to focus on reducing unnecessary use of higher cost services in the ER or acute care setting.

ACO incentives. A larger ACO needs to save a lesser percentage of medical expense to generate shared savings than a smaller ACO. One-half of the MSSP ACOs include a hospital. Because ACO beneficiaries are FFS beneficiaries, they are not obligated to use either ACO providers or the ACO hospital. And, because these beneficiaries are FFS beneficiaries, provider revenue is dependent on providing a service. In the ACO model, reducing ER or acute care utilization at the ACO hospital means a reduction in revenue. Since much of the operating cost of these hospitals is fixed or has fixed-cost attributes, reducing this utilization may well mean a reduction in operating income. If the ACO is able to exceed its savings threshold, the return then is up to 50% of that savings, in the MSSP Track 1 and that amount may not exceed 10% of the total benchmark medical expense.

The ACO has an incentive to reduce costs. It may benefit from cost savings if it meets the ACO quality measures. The ACO may share none, some or all of the cost savings with the providers. The providers continue to receive FFS payments. In order to reduce cost these providers might focus on utilization. This may mean the providers receive less FFS payment. If there are 1,000 providers, each might share 1/1,000 of the 50% of the total earned shared savings. This amount would quite likely be less than the FFS revenues generated had those patients been seen. The beneficiary has no real incentive to use ACO providers, at least so far as the program incentives exist. Since the beneficiary can see any Part B participating provider, and has the same co-insurance responsibility, the decision about from whom to obtain services may not be economic.

HMO incentives. The HMO takes risk from Medicare. Said differently, the HMO gets paid in advance for care, based on actuarial projections using the HCC risk score and other factors. Reducing utilization for the HMO provides a direct economic benefit to the HMO. Because the HMO shares financial risk with providers, particularly primary care providers, the incentive to reduce utilization also exists at the provider level. Quality measures are important in these HMOs as attainment at higher levels results in higher Star Ratings and potentially additional financial benefit. Beneficiaries have restrictions about which providers they may see. Going outside the provider network may increase beneficiary cost sharing. Because the HMO can modify benefits, common services like presenting in the ER may increase beneficiary cost sharing if the visit is not pre-authorized by the PCP.

At the ACO or HMO organization level, the different bottom-line impacts related to FFS payment versus pre-payment via capitation and risk, means a reduction in ER or acute care services creates different financial implications. Digging deeper to focus on the PCP incentives relative to this focus on utilization reveals again disparity.

Viewing structure through study variables. The outcome measures of this study related to blood sugar control, colon cancer screening and ER utilization are representative of three important primary care functions: chronic disease management, prevention, and utilization management. Understanding why these activities matter, and how the incentives through the ACO and HMO may be different for each, is important to putting structure in the right context.

Diabetes management. Diabetes management almost universally involves management of blood sugar, frequently measured by the glycosylated hemoglobin, or HbA1c, which measures average blood sugar concentration over a three months period. Diabetes is a dangerous disease because the excess blood sugar damages blood vessels and nerves over time leading to macrovascular and microvascular complications. One of the most concerning microvascular complications is the onset of chronic kidney disease. Diabetes, particularly Type II, disproportionately affects certain populations. Race/ethnicity and age are risk factors for diabetes and chronic kidney disease as shown in Table 4.

Table 4: *Prevalence of Diabetes and CKD by Race/Ethnicity 2015*

Race/Ethnicity	All Medicare		Medicare Aged <65		Medicare Aged >65	
	Diabetes	CKD	Diabetes	CKD	Diabetes	CKD
Non-Hispanic White	24.4%	17.4%	22.6%	12.7%	24.7%	18.2%
Black/African American	36.6%	24.9%	30.6%	21.3%	39.8%	26.7%
Hispanic	36.5%	19.5%	32.5%	18.7%	37.9%	19.8%
Asian/Pacific Islanders	34.0%	17.9%	27.8%	18.8%	34.9%	17.7%
American Indians/Alaska Natives	38.0%	20.8%	33.0%	18.1%	40.4%	22.1%

In 2015, nationally for all beneficiaries, Medicare spent $15,920 per capita on individuals with diabetes. In the same year, the per capita spending on individuals with chronic kidney disease was $25,668. In 2015, the prevalence of the chronic condition dyad diabetes and chronic kidney disease was 14.7%, with a per capita cost of $24,481. Because gender is also a risk factor for many chronic conditions, it is interesting to note the prevalence of this dyad in the male population was 17.0% versus 12.9% in the female population. A diabetic patient with chronic kidney disease requires on average nearly $10,000 a year more to treat than a diabetic patient without chronic kidney disease. If the chronic kidney disease progresses to End-Stage Renal Disease (ESRD) which requires dialysis, the per capita spending increases to nearly $90,000.

While some debate exists as to the relevance of HbA1c in Type II diabetes, and consequences do exist should blood sugar levels get below a certain threshold, management and lowering of HbA1c is still an important part of diabetic care (Giugliano et al., 2015; Wilson & Perry, 2009).

The relevance of these data points to this research becomes obvious. Controlling for age/gender and risk combinations through the HCC risk score and race/ethnicity, if one structure associates to a lower percent of the diabetic population out-of-control for HbA1c, the historical cost data suggest a significant savings opportunity.

In primary care practice, approximately one-third of office visits are for diabetic care (Centers for Disease Control and Prevention, 2010). A typical approach for the practitioner would be to obtain the HbA1c every three months, screen for kidney disease annually, counsel the patient on self-management and use pharmacotherapy based on the level of blood sugar control.

Colon cancer screening. Screening and detection of colon cancer likewise translates directly into Medicare program savings. The National Colon Cancer Roundtable (2007) identified a $15 billion savings opportunity through increasing the rate of colonoscopy in the Medicare population aged 50-64. The savings opportunity associated to increasing stool blood testing may be close to $13 billion (National Colorectal Cancer Roundtable, 2007).

Processes geared to chronic disease management or prevention present a fundamental issue. Namely, while savings associated to prevention may accrue to Medicare they may not accrue to the ACO or HMO because the investment in prevention may avert future spending by which time the beneficiary may no longer be attributed to the ACO or enrolled in the HMO. This paradox exists in VBP arrangements in terms of identifying the recipients of the "value" created which may not include either patients or payors (Tanenbaum, 2016).

At a macro level, the incentives created by Medicare in both programmatic approaches through the ACO and HMO are cost-oriented with the attainment of quality measures being a "ticket to play." In other words, ACO or HMO realization of the micro-level savings and the temporal relationship to the resources expended in the prevention of future disease states may be disconnected or unavailable.

ER utilization. On the other hand, the financial benefit to ACOs and HMOs in preventing unnecessary ER utilization seems more apparent. Of the 130 million ER visits nationwide in 2013, persons aged 65 and older accounted for 20.7 million. The semi-urgent and non-urgent visits numbered 16.3 million in the same age group (CDC, 2013). These visits frequently result in imaging, procedures or prescriptions (Honigman, Wiler, Rooks, & Ginde, 2013). Nationally, in 2015 for all Medicare beneficiaries ER visits per thousand for individuals with diabetes were 1,010, CKD 1,494 and cancer 1,018. This means on average individuals with one of these conditions will present in the ER at least once each year. Over the three performance years of the MSSP ACO program (2013-2015) average ER visits per 1,000 actually increased from 323 to 731. Over the same time in all Medicare Advantage plans ER visits per 1,000 changed from 479 to 497. In the example of the ACO and HMO PCPs whose patient needs immediate care but does not need to visit the ER, this may provide some evidence of the physician-level incentives to intervene and prevent an unnecessary ER visit in the HMO and how at the individual practitioner level the same incentive may not exist in the ACO. The HMO creates this incentive by sharing financial risk with the PCP. The ACO lacks this ability because ACO attributed beneficiaries are by definition FFS beneficiaries and may present in the ER based on their own personal decision. Viewing these incentives through the SPO, it would be appropriate to say the incentives created by CMS through the HMO program result in a structural response to alter provider incentives. The providers respond through care processes to take advantage of these incentives.

The incentives created through policy and payment mechanism in the ACO or HMO structure are an area for additional inquiry. In the SPO, these incentives exist in the structures and create the differences in provider participation, patient choice, and payment methods used by the ACO/HMO. These differences then determine processes, which in turn produce outcomes.

Overall, Medicare clearly has a goal of reducing costs. The ACO and HMO programs are only two of the models currently in use, but collectively cover a very large part of the Medicare population. These ACOs and HMOs may invest in disease management or prevention programs designed to further attain quality measures and reduce costs. However, the benefit of the cost reduction may occur in the future perhaps after the beneficiary who would use resources is no longer part of the ACO or HMO. Understanding better the extent to which one approach may outperform another may be important to setting payment policy designed to steer organizations into that preferred model.

Outcome

In the SPO, outcome is the effect. It is where evidence of differences or changes in structure and/or process become evident. Assessing outcome requires knowledge of the relationships between structure and process (A. Donabedian, 1988). In outcomes research, this is almost by definition the dependent variable.

In approaches like the one used in this study, relying on publicly available data, structure may be easily observed but with less definition as to quantification of differences. Process may be more obscure making differences or changes more difficult to detect. Outcome may be the most easily measured, even if differences in the levels of measurement exist. Multiple systems of outcome measurement exist and form the basis for most VPB arrangements, like in the case of HEDIS and NQF measures used by HMOs and ACOs, respectively.

Outcome is a result. It might describe a snapshot of a moment in time, like a blood sugar measurement. Outcome might also describe what quality measures typically refer to as a process measure, but would not be process relative to the SPO, like a cancer screening. Not unlike the grey area sometimes existing between structure and process, some obscurity exists between process and outcome. In the case of a cancer screening, the completion of that screening would be an outcome whereas the recommendation to screen and discussion concerning need would be part of the technical process in the SPO.

Cost and utilization could be debated as to whether they represent outcomes. Donabedian seems to have evolved his thinking over time on this subject particularly regarding the inclusion of cost in determining efficacy. (A. Donabedian, 2003; R. C. Donabedian A. et al., 1982). Since many, if not most or all, VBP arrangements include cost/utilization measures they are close to being *de facto* outcome measures.

Judgements as to the benefit of outcomes is often an area for debate. In the areas of chronic disease management, like diabetes, typically lower HbA1c would be better. But, this is not an absolute and there are disagreements as to the role of HbA1c in the case of Type II diabetes particularly regarding serious complications of diabetes (Wilson & Perry, 2009). Generally, cancer screening is thought to be beneficial but again this is not an absolute as in the case of Prostate Specific Antigen testing to screen for prostate cancer (Cohn et al., 2014). Cost/utilization may be particularly difficult to assess in terms of benefit. Who is to say less cost is necessarily better? Some argue, at the macroeconomic level, the issue of the US NHE at 17% of the GDP is the value extracted as a result of the spend rather than the amount of that spending in the aggregate (Fuchs, 2010).

Conclusions

The data on ACOs overall is nascent. The MSSP plans completed the third performance year in calendar 2015 for early adopters. Some ACOs have only one or two years of performance data through 2016. Many Medicare Advantage HMOs have much longer tenure with plan start dates in the mid-1980s through the 1990s.

In the SPO framework, these organizations represent distinct structures and as such are a basis for quality comparison. Donabedian referred to structure as “blunt instruments” of quality assessment which should be a “major preoccupation” in system design (A. Donabedian, 1988). Figure 2 depicts the SPO and highlights the relationship of the different structures between ACOs and HMOs within the framework.

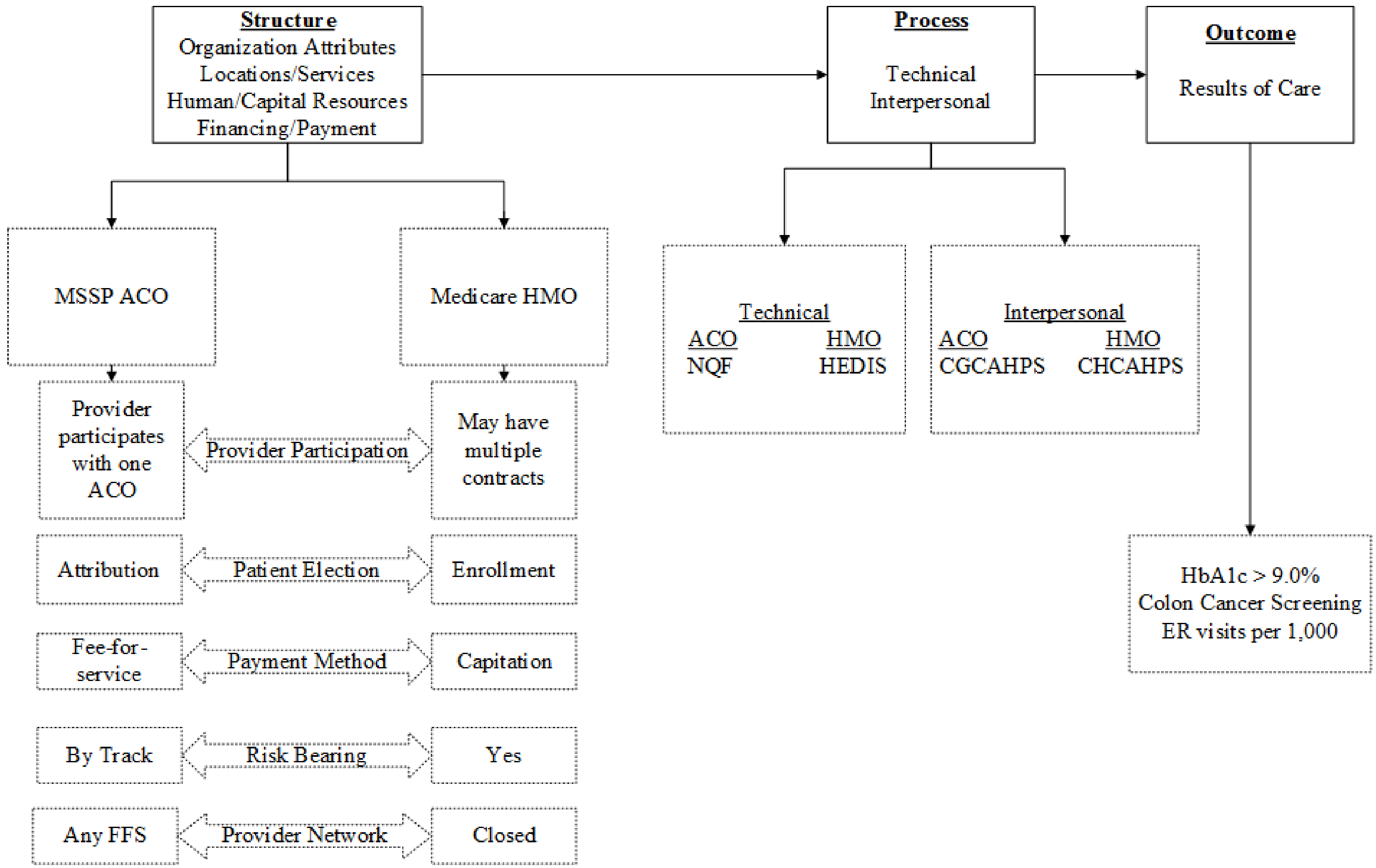

Figure 2: The SPO Framework

The incentives created by and through these programs exist in structure, becoming activated through process. Methodologically, comparing quality between these organization structures poses some challenges as will be discussed. The lack of a quality comparison between Medicare ACOs and HMOs is a gap in the existing literature. This study seeks to close this gap, at least in part, by using several dependent variables from HEDIS and NQF measures to analyze possible differences in quality of care through the use of a small sub-set of these measures where congruence in the population measured and level of measurement exists.

Shifting the perspective of structure slightly to view it through the number of beneficiaries served reveals a potential difference in the ability of these organizations to generate a return on investment (ROI). Because the number of beneficiaries relates to the number of providers, locations for service deliver, etc., this view of structure is appropriate in the SPO framework. The size of the population served and implications for ROI then raises questions concerning the ability to manage quality measures more effectively in larger populations because ROI is ultimately a function of medical expenses savings and quality measure attainment.

The two “slices” through structure inform the research questions of this study. First, across the various quality measures comparable between HMOs and ACOs, does one model perform better than the other? Second, do differences exist in the attainment of these plan comparable quality measures based on the size of the beneficiary population served?

In creating these structures with attendant incentives, CMS may be trying to obtain something very difficult to get: changes in provider behavior that are inconsistent with the provider best interest. In the risk-sharing model of the HMO, the provider will want to reduce expenses in the same time-period of the pre-payment of services. Expending resources to lower a patient’s blood glucose now with an associated cost savings in the future creates misalignment.

In the ACO model, expending resources now to lower a patient's blood sugar may exhibit the same lack of temporal relationship between investment and return. In addition, the ACO provider may have an incentive to increase the frequency of contact with that patient, and generate FFS payments, which does not necessarily align with the ACO incentive to reduce costs.

In both structures, the early detection and treatment of colon cancer can reduce cost. Investing resources that will lower expenditures in the contract period with a positive ROI is something CMS can almost universally count on providers to do. The contingency would be detecting a treatable cancer. Because the screening interval is ten years, many cancers will develop after a negative screening. Consuming resources to not find treatable cancers again hinders access to financial incentives for the providers. The availability of services may also be different because ACO beneficiaries could receive screening through any Part B provider whereas HMO beneficiaries must use in-network providers. Because ACOs must meet benchmarks on quality measures to be eligible for shared savings, if savings occur in excess of the threshold, the incentive for ACOs is different from HMOs. The HMOs may receive a quality bonus based on the Stars Rating, but the cost savings associated to detecting early stage disease creates a direct financial incentive. If the ACO design is indeed based on quality and coordination of care, then it is reasonable to expect they have a higher colon cancer screening rate.

The prevention of unnecessary ER utilization creates diverging incentives between these structures. In the HMO, the provider may have financial risk and the beneficiary may as well, when beneficiaries seek care in the ER for non or semi-urgent services. In the former, if the PCP authorizes a non-urgent ER visit they may have to "pay" for the service through their capitation.

In the latter, the beneficiary may have to pay out of pocket the entire cost of a non-urgent ER service not pre-authorized by the PCP. The ACO beneficiary would have the same cost sharing resulting from an ER services regardless of acuity or need. The PCP may not even be aware until after the fact as the FFS beneficiary can chose to seek services in the ER at any time.

As a result of these incentives, the hypotheses are the ACO beneficiaries are in better blood sugar control and a higher proportion receive colon cancer screening. The HMO beneficiaries should exhibit lower ER utilization. The hypotheses based on the size of the population served are that larger organizations have more patients in blood sugar control, a higher proportion screened for colon cancer and lower ER utilization. This is because the issue of ROI is much easier to leverage across a larger population. This research will examine the publicly available data to see if these hypotheses hold up based on the two views of structure.

Chapter 3: Literature Review

Introduction

This literature review constitutes a comprehensive examination of the peer-reviewed literature concerning comparisons of cost, quality and/or access between payment systems, within payment systems and between delivery models and payment systems. It includes non-peer reviewed materials in the form of dissertations and reports by not-for-profit or industry organizations. Despite the lack of peer review, the information obtained from not-for-profit or industry reports summarized mostly publicly reported data from a variety of sources, e.g., The Kaiser Family Foundation report on total HMO enrollment.

I performed a systematic search of peer-review articles in August 2016, using PubMed and Google Scholar. This literature review focuses primarily on Medicare because the sampling frame of this study is Medicare Advantage HMOs and Medicare ACOs. Follow-up references in retrieved articles supplement this systematic search. Search terms included: Medicare advantage, accountable care organization, healthcare cost, healthcare access, healthcare quality, managed care, HMO, fee-for-service, diabetes mellitus, hypertension, National Quality Foundation (NQF) and HEDIS. Results in each search iteration included variations of terms, e.g., healthcare cost became "analysis, cost." These literature searches identified 1,723 peer-reviewed publications. Sorting these publications by the PubMed Identification number indicates 233 references returned for more than one of the above searches.

Studies selected for further review followed these inclusion criteria:

1. Reflected an attempt to compare quality, or cost or outcomes between payment models, within payment models or payment models to delivery models.
2. Used HEDIS or NQF measures as outcomes variables.
3. Identified Medicare Advantage as independent variable.

Based on these inclusion criteria, 78 studies merited additional review. Of these, only 15 previous investigations were directly relevant to the research questions of this study. The following is a discussion and evaluation of these studies by theme.

Quality Measurement in Medicare Advantage and ACOs

Medicare Advantage plans report quality measures using the HEDIS measures. Medicare ACOs use quality measures from the NQF. As we will see in the methods chapter, these dimensions of quality measurement do not line up exactly. This disconnection in the domains and levels of quality measurement no doubt speaks to the lack of a "template" approach. The body of evidence regarding head-to-head comparisons between financing and delivery mechanisms is scant generally, and specifically within Medicare programs. Interplan comparisons within Medicare Advantage however are routine and actually support the Star Ratings System Medicare uses to evaluate these plans (Gold, 2009; Herman, 2015).

Survey of Existing Literature

The literature reveals mixed findings in comparing between payment models, within payment models and payment models to delivery models with regard to quality and cost. In general, HMOs when compared to FFS arrangements tend to be lower cost with some indications of lower quality, particularly for at-risk populations, e.g., the elderly, minorities or those with chronic conditions (Davidson, 1997; Hellinger, 1998; Miller & Luft, 1997). In Medicare specifically, the results from Medicare Advantage to FFS comparisons tend to follow the same pattern: Medicare Advantage may demonstrate less cost but the trade-off may be lower quality particularly for at-risk populations and other demographics. Because the ACO is a relatively new delivery/payment model, direct comparisons of ACO performance to other delivery or payment models appear not to exist.

Comparisons between models. Several important studies from the mid-to-late 1990s help identify important differences in results based on model. These studies used dependent variables essentially similar to HEDIS or NQF, but not actual HEDIS or NQF measures, or used patient self-assessment via the 36-Item Short Form Survey. All were observational. The research questions posed were all variations on the "which one is better?" theme. Six studies were head-to-head comparisons between a managed care product, HMO or Medicare Advantage, and a FFS product. The unit of analysis in each case was patient-level.

All but one study generated data from administrative/claims records. Davidson, et al (1997) audited medical charts to derive measurements for dependent variables. If differences existed in the dependent variables based on structure, the independent variable in each of these investigations, they were minor and typically favored FFS models for clinical measure improvement/change and HMO/Medicare Advantage models for utilization/cost measures. Some findings demonstrated selection bias in the plan design resulting in less favorable outcomes for at-risk populations (Retchin et al., 1992; Ware, Bayliss, Rogers, Kosinki, & Tarlov, 1996).

Methods, sample sizes, power and risk adjustment. Alpha levels were uniformly set at .05. Power calculation discussions were uniformly absent. The methodological approach in virtually all of these investigations was a detection of group differences through t-Test. Where not specifically discussed, the tabular data presented and discussion of results clearly indicated a test of mean differences likely using a t-Test. Using G*Power 3.1.9.2 to calculate the necessary sample size at α=.05 with β=.80 for a 2-tailed t-Test indicates group samples of 394, each, are necessary to detect small effects. Sample sizes ranged from 353 charts to over 5 million observations. Insufficient power may be problematic for Davidson (1997), and Cohen, et al (2012). Only one study used a risk adjustment methodology (Brennan & Shepard, 2010). The lack of risk adjustment is a limitation generally in the literature of this approach reviewed for this study.

Summary of results. Table 5 illustrates the purpose, methods/variables, design/sample and conclusions from each of these approaches.

Table 5: *Summary of Literature Findings*

Author(s)	Year	Purpose/Research Question	Methods/Variables	Design/Sample	Conclusions
Silcox et al	2003	Determine differences in quality of care between FFS and HMO for CHF post-hospitalization	t-Test (2 sample), data extracted by chart review, Odds Ratio to Relative Risk by insurance plan and outcome for: mortality, 30-day readmission, ACE inhibitor use, 2-week follow-up visit, discharge instructions	Observational, case control. 154 subjects, ages 19-98 stratified by attending MD, chart review.	Slightly shorter ALOS for HMO. No differences in mortality OR. HMO with higher readmission. No significant differences by attending specialty. No differences in follow-up or discharge instructions
Ware et al	1996	Compare outcomes of chronically ill between HMO and FFS	t-Test (2 tail) SF-36	Observational, 4-year. 2,235 patients with NIDDM, HTN, Acute MI, HF, Depression, 1986-1990	No average differences. Elderly showed more decline in HMO for physical function but mental function better. Lower poverty levels did better in FFS, higher income levels did better in HMO

Table 5: *Continued*

Author(s)	Year	Purpose/Research Question	Methods/Variables	Design/Sample	Conclusions
Davidson	1997	Which model is better for DM care, MA or FFS?	HbA1c level plus 5 process measures: eye exam, kidney screening, HbA1c screening, foot exam, lipid panel. Includes number of PCP visits and specialty referrals, likely t-Test	353 audited charts	ADA standards not met, generally. FFS slightly better.
Brennan & Shepard	2010	Compare clinical quality of care between MA and FFS	11 HEDIS process measures, mean differences	Sociodemographic adjustment	MA better on 8 of 11, FFS better on 3, newer measures
Cohen et al	2012	Does an MA Special Needs Plan (SNP) improve care for patients with NIDDM	Annual utilization rates: admissions and LOS, hospital outpatient visits, physician office visits	12 months of beneficiary level data for FFS and Care Improvement Plus (SNP)	SNP had lower admissions, shorter LOS, less hospital outpatient visits, greater physician office visits compared to FFS
Landon et al	2012	Compare MA and FFS utilization patterns	HEDIS utilization: med/surg hospitalization, outpatient visits, ambulatory surgery, ER visits plus 12 specific surgical procedures, t-Test (2 tailed)	Matched samples: 3.1M MA in 2003 to 5.7M in 2009. 5.0M in FFS 2003 to 5.6M 2009	MA HMO patients used fewer services
Landon et al	2015	Compare utilization MA v. FFS for DM and CV disease	HEDIS plus RRU (created for FFS with administrative data). IP utilization, E&M services, surgeries, ED visits. Quality measures for LDL and HbA1c screening plus eye exam, likely t-test	Matched FFS samples. MA sample: 680,000 w/DM, 270,000 w/CV disease. 20% national FFS sample	MA had lower RRU and ED visits. More established MA plans had lower RRU with higher quality compared to newer plans

Comparisons within models. The results of comparisons based on HEDIS measures drive the "Star Rating" methodology CMS uses to assess Medicare Advantage plan performance. The Star Ratings directly affect plan incentive payments. One commonality between Medicare Advantage plans appears to exist: favorable selection (Newhouse, Price, McWilliams, Hsu, & McGuire, 2014). This phenomenon is well known, understood and quantifiable (Newhouse et al., 2014). In recent years, however, the implementation of the HCC risk scoring methodology may be moderating the effect (McWilliams, Hsu, & Newhouse, 2012).

In the comparison of Medicare Advantage plans one interesting finding was the better performance of not-for-profit plans (McBean, Jung, & Virnig, 2005). This is consistent with findings regarding HMO-to-HMO performance in previous efforts (Himmelstein, Woolhandler, Hellander, & Wolfe, 1999; B. Landon & Epstein, 2001). Typically not-for-profit Medicare Advantage plans demonstrate better quality (Xu, Burgess, Cabral, Soria-Saucedo, & Kazis, 2015). To the extent quality of care overall was roughly equal in these comparisons, some subgroups did experience poorer outcomes based on plan design (McBean et al., 2005). Variation in plan performance clearly exists as evidenced by the variation in Star Ratings. In 2012, the total percent of Medicare Advantage plans with a 4 or 5 Star Rating was 20%, and this increased to 40% in 2016 (Kaiser Family Foundation, 2015).

Sources of variation. The literature seems consistent as to the source of these variations including: favorable or unfavorable selection, not-for-profit or for profit status, age of the plan and plan experience with specific measures. One interesting phenomenon in Medicare Advantage plans appears to be the willingness of beneficiaries to change plan if the Star Rating is one level higher (Reid, Deb, Howell, & Shrank, 2013). Medicare permits beneficiaries to quit their current plan and enroll in a five-star plan at any time, even off-cycle from open enrollment periods. Generally speaking, the literature in this domain is sparse, which may mostly be attributable to the existence of the Star Ratings methodology derived largely from HEDIS measures.

One of the most relevant approaches, examining variations in diabetic care based on plan case mix, used multivariate methods (Abraham, Marmor, Knutson, Zeglin, & Virnig, 2012). In this study, case-mix associated to quality with a modest effect size. In other words, more favorable case-mix would result in higher quality.

Summary. Variation exists between Medicare Advantage plans as to quality. Some of this is attributable to case-mix or selection. Some variance may occur based on not-for-profit or for-profit status with not-for-profit plans generally outperforming for-profit plans. Because HEDIS measures serve as the basis for Star Ratings, the overall rating for Medicare Advantage plans, much of the research effort in comparing Medicare Advantage plans focused on chronic conditions like DM.

Comparisons of integrated delivery systems to medicare advantage. While no standard definition of an Integrated Delivery System (IDS) exists, the term suggests information sharing and coordination of care through established relationships between providers (Hwang, Chang, Laclair, & Paz, 2013). This theme exists also in Medicare's description of ACOs: groups of physicians and other providers who come together voluntarily to offer high quality, coordinated, care to beneficiaries.

Comparisons of IDSs to Medicare Advantage are meaningful because many IDSs are financing models as well. For instance the VA Health System is a self-contained financing and delivery structure. The ACO is similar in this regard with what is a FFS payment structure existing, prior to MIPS, between a delivery model and a secondary payment model via the shared savings opportunity.

Many IDSs have, or are developing, provider-sponsored Medicare Advantage plans as well. A trend towards having multiple track MSSP ACOs and a Medicare Advantage plan seems to be developing as a strategy to gain more experience with managing population risk (Gabriel, 2015).

In the literature search one study comparing Medicare Advantage to the VA appeared consistently. It compared elderly patients receiving care in the Veteran's Administration (VA) health system directly with Medicare Advantage on several HEDIS outcomes measures (Trivedi & Grebla, 2011). The independent variable in this study is structure, namely whether the individuals received care in the VA or Medicare Advantage. The dependent variables using HEDIS measures were:

1. HbA1c control
2. LDL control
3. Diabetic annual eye exam performed
4. Blood pressure control
5. Beta blocker use in hypertension
6. Mammography screening
7. Colorectal cancer screening

In this study, a retrospective cross-sectional design following patients between 2000 and 2007, the VA outperformed Medicare Advantage with the performance gap widening over the years (Trivedi & Grebla, 2011). The sample included 293,554 patients in 142 VA medical centers and 5,768,573 enrolled in 305 Medicare Advantage plans. Significant covariates included age group, gender, Census division, proportion of persons age 65 and older in the enrollee's ZIP code with income less than the federal poverty level, and proportion with at least some college education. The authors applied a risk adjustment method based on individual demographic characteristics, income and education. Because the VA does not use HEDIS measures, the authors re-stated internal VA measures, called the "External Peer Review Program" (EPRP), relative to the HEDIS definitions. This study is remarkably similar in terms of content and approach. It is a direct comparison between Medicare Advantage and another structure, the VA Health Sytsem, and applied the General Linear Model (GLM) to analyze the data. However, the main author subsequently identified a significant weakness when examining the same individuals receiving care in the VA and Medicare Advantage (Trivedi et al., 2016). The issue relates to the structural differences between Medicare Advantage and VA. The conclusion was the Medicare Advantage plans simply under-report performance as they take data from their contracted providers. At the same time, because ACOs also self-report based on the sampling guidelines, this inherent weakness in reporting method exists in both structures.

Administrative Data in Healthcare Outcomes Research

As more and more data become available for researchers, the convenience and cost savings associated to using administrative data is making its use more common. Administrative data has limitations, perhaps prime among these is the fact that its primary use is not for research purposes (Fung, Brand, Newhouse, & Hsu, 2011; Hashimoto, Brodt, Skelly, & Dettori, 2014). Research using patient-level data can be difficult to complete and poses its own challenges in terms of validity. A reasonable approach to comparing quality differences in Medicare HMOs and ACOs would be to compare patient-level outcomes data. This would be possible for Medicare HMOs, as CMS maintains beneficiary-level data including HEDIS-linked Category II CPT Codes. No such database exists for Medicare FFS beneficiaries. So, the next best alternative is to take the data that is publicly available and make comparisons where congruence exists between outcomes measures used in both models. The research questions and hypotheses flow from the nature of the data available.

Research Questions and Hypotheses

The purpose of this study is to examine differences quality relative to enrollment in a Medicare Advantage HMO or attribution to an ACO. The results of previous studies of different models demonstrate varying conclusions. This may be especially true when viewing outcomes in the aggregate, i.e., at the population level, without detailed examination of the impact to sub-populations, e.g., the frail, elderly or economically disadvantaged.

The research questions relate directly to the differences in structure between Medicare Advantage HMOs and ACOs and explore two themes: the role of structure, HMO or ACO, on chronic disease management, cancer prevention screening and utilization; the impact of the size of the population served on chronic disease management, cancer prevention screening and utilization. Both HEDIS and NQF define other measures of all these dimensions of care. Unfortunately, very little alignment exists between HEDIS and NQF except for the three outcome measures used in this study.

Conclusions

The dearth of literature studying head-to-head measurements of quality based on structure clearly demonstrates the need for more study. Previous works explored differences within and between models, e.g., Medicare Advantage to Medicare Advantage, Medicare Advantage to FFS, with more emphasis on comparing Medicare Advantage to Medicare FFS. Much of the early focus in these efforts was cost-based or based on self-perception of health status. Medicare Advantage plans are the only Medicare products consistently using HEDIS measures. Although both the ACO and Medicare Advantage HMOs use similar but not identical outcomes measures, no studies appear in the current literature comparing these models. Track 1 MSSP ACOs are only in the fourth performance year, in 2016, for the oldest of these organizations so the lack of comparison is understandable. The number of Medicare beneficiaries in an ACO or Medicare Advantage plan is, however, quite large and projected to increase over time. With the repeal of the sustainable growth rate (SGR) methodology to annual Medicare Fee Schedule (MFS) updates and the new pressures pushing providers toward Advanced APM, the financial risk in providing care to a Medicare population will increase (Civic Impulse, 2016).

Financial pressure on the Medicare Trust Fund requires relief through some model able to not only control costs but improve population health status, at least insofar as defined through the current set of quality measures. CMS clearly has an interest in identifying this ideal model as it continues to promulgate new rules and programs designed to elicit these responses. This study is an early contribution to the examination of quality differences in Medicare HMOs and ACOs. While narrow in focus it appears unlike anything currently published.

Chapter 4: Methods

Introduction

This study is an examination of performance on several outcome indicators for Medicare beneficiaries receiving care through the Medicare Track 1 MSSP ACOs and Medicare Advantage HMO plans. These indicators, the dependent variables, are: percent of population with diabetes mellitus and most recent HbA1c > 9.0%, percent population screened for colon cancer based on age-group, and Emergency Room (ER) visits per 1,000 beneficiaries not resulting in hospital admission. These outcomes describe clinical, prevention and utilization outcomes. The purpose of this chapter is to elucidate the rationale for the design and methods, describe the analytic procedure and explain the process of interpreting results.

Research Questions and Hypotheses

The research questions are as follows:

1. Do Medicare ACOs have a greater percent of population with diabetes mellitus and most recent HbA1c > 9.0% than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?
2. Do Medicare ACOs have a greater percent of population screened for colon cancer than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?

3. Do Medicare ACOs have lower ER utilization per 1,000 not resulting in hospital admission than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?
4. Do ACOs and HMOs serving larger beneficiary populations have fewer patients with diabetes mellitus and most recent HbA1c>9.0%, controlling for race/ethnicity, average HCC risk score and plan years of operations?
5. Do ACOs and HMOs serving larger beneficiary populations have more patients screened for colon cancer, controlling for race/ethnicity, average HCC risk score and plan years of operations?
6. Do ACOs and HMOs serving larger beneficiary populations have fewer ER visits per 1,000 not resulting in hospital admission, controlling for race/ethnicity, average HCC risk score and plan years of operations?

The hypotheses of this study are as follows:

- Hypothesis 1: Medicare ACOs have fewer diabetic patients with last HbA1c > 9.0%
- Hypothesis 2: Medicare ACOs screen more age-appropriate patients for colon cancer.
- Hypothesis 3: Medicare HMOs have fewer ER visits per 1,000 not resulting in hospital admission.
- Hypothesis 4: Organizations serving larger populations will have fewer diabetic patients with last HbA1c > 9.0%
- Hypothesis 5: Organizations serving larger populations will have a higher rate of colon cancer screening.
- Hypothesis 6: Organizations serving larger populations will have fewer ER visits per 1,000 not resulting in hospital admission.

Design and Approach

This study is non-experimental and retrospective. The data come from CMS publicly available data for Medicare HMOs and ACOs through various CMS sources. Multiple databases containing information on each model require concatenation to align data elements for analysis. Data sources are used only if it is possible to link them through an identification number consistently used between the files.

The analysis will use multiple hierarchical regression and input variables into the equation through a series of steps. The first variables will be the control variables, input in three discrete steps. The independent variables will be group (HMO or ACO) and total beneficiary enrollment. The dependent variables will be the quality measures. The model will return the unstandardized beta weights and significance at each step allowing for interpretation of the impact of the control or independent variable on the outcome.

Threats to validity. Internal validity is a concern because HMO and ACO providers self-report quality data. The providers in Medicare Advantage plans self-report HEDIS measures to CMS through the claims adjudication process and if not the HMO plans audit patient-level encounter data to extract the information. ACOs report attainment on quality measures based on a patient sampling methodology identified by Medicare (RTI International, 2015). Because the design is retrospective, utilizing secondary data, there are no procedures available to test equipment and verify lab controls, training protocols or data entry accuracy.

The primary threat to external validity is simply that the main purpose of the data used is not research. It is administrative data. In the cases of Medicare Advantage HMOs and ACOs these data support decisions regarding program incentives and payment which are outcome-oriented. Another threat to validity is selection bias. Although the amount of favorable selection in Medicare Advantage is perhaps less today than in previous years, it may possibly remain (Mello, Stearns, Norton, & Ricketts, 2003). Some selection bias may also be occurring in ACOs. For instance the percent of population with diabetes mellitus and most recent HbA1c > 9.0%, in year one of operations, was lower in 2015 compared to 2013 for plans starting in those years. Why this may be is not discernable but it could be a result of "design" relative to populations managed through ACOs in some organizations particularly when those organizations participate in other population-based models. As age is a risk factor for virtually all chronic conditions, and because so many baby boomers are aging into Medicare, it could also be the average beneficiary age in the newly starting ACOs is lower. However, because the average HCC risk score, and beneficiary race (White or not-White) and duration of operations are control variables in this study the issue of selection bias is less concerning. It would be better to match individual beneficiaries in cohorts based on HCC risk score, and other factors, to further protect external validity but this is not possible with the publicly available data.

Subjects. This is a non-experimental, and due to the aggregate level of the publicly available data it is not a human subjects study (Department of Health and Human Services, 2009). No patient identifiable information exists in the publicly available CMS data. All data available through CMS is at the ACO or HMO level. Institutional Review Board (IRB) approval is not necessary (Department of Health and Human Services, 2009).

Sample. In the first set of research questions, the outcomes relate to how many beneficiaries in either an ACO or HMO meet certain criteria. As such, the sample power derives from the number of beneficiaries included. In the second set of research questions, the outcomes relate to the size of the beneficiary population meeting certain criteria receive care through the ACOs and HMOs. Again, the sample power derives from the number of beneficiaries included rather than the number of ACOs and HMOs. This is consistent with other similar approaches identified in the literature review (Trivedi & Grebla, 2011).

In total, ACO attribution and Medicare Advantage enrollment is slightly less than half of the 55M Medicare beneficiaries in total in 2016. Including all Track 1 MSSP ACOs and Medicare Advantage HMOs operating in 2015, in the sampling frame yields 609 total plans serving 17,898,330 beneficiaries. There are 226 HMOs with an average of 47,941 enrolled beneficiaries and 383 ACOs with an average of 18,708 attributed beneficiaries.

Sample size calculations used the software package G*Power 3.1.9.2 specifying $\alpha=.05$ and $\beta=.80$, and a small (.10) effect size. The calculation for R^2 change in multiple regression using two independent predictors indicates a sample size of 100 is the minimum necessary to achieve the desired power.

Risk adjustment. Risk adjustment methodologies are imperfect and can never completely account for differences in populations (Lane-Fall & Neuman, 2013b). Adjusting risk is particularly important, however, in the use of administrative data to analyze outcomes (Sernyak & Rosenheck, 2003). Sufficient detail does not exist in the CMS publicly reported data to use it for risk adjustment between the organization types, e.g., geographical dispersion of plans. Different data are reported for HMOs and ACOs. However, the publicly available data does report average HCC risk score at the plan level for Medicare Advantage. The ACO publicly available performance data reports average HCC risk score for several populations in the ACO: the disabled, end-stage renal disease, aged non-dual eligibility and aged dual eligibility. Weighting these average HCC risk scores by the number of beneficiaries in those categories within the population of each ACO returns the weighted average HCC risk score for the entire ACO population. Using the plan-specific average HCC risk score provides a solid basis for adjustment. The HCC risk score factors chronic conditions, age/gender, and disability/ESRD status in the calculation methodology. Also, if a beneficiary became Medicaid eligible during the calculation year, the HCC risk score factors this in the final score for that beneficiary. This provides some level of socioeconomic risk adjustment because Medicaid eligibility implicates lower economic status.

Based on the distribution of ACO and HMO beneficiaries, further risk adjustment to incorporate variation in socioeconomic factors based on geography would not be meaningful. The distribution of these beneficiaries reveals significant overlap as shown on the density map in Figure 3.

Data Sources and Collection

The secondary outcome data come from two sources: CMS publicly available data for HMOs and ACOs. The performance and demographic data for both exist in separate datasets and require concatenation to be useful. Concatenation is possible using unique identification codes common to the datasets for each. These sets for HMOs and ACOs do not appear with consistent record layouts. Constructing a dataset for analysis requires manual effort to populate data elements using a taxonomy and record layout useful for the analysis.

Medicare HMO plans may operate in multiple geographies each representing a different case in the data sets. Each contract with CMS has a unique identifier as do the plans under each contract. For instance, an insurance company might bid for multiple plans in different geographies. Each plan award has an identification code with all plans under the same company using the same contract identification code. Summary data by plan code is obtainable by concatenating these data sets at the level of the contract identification code

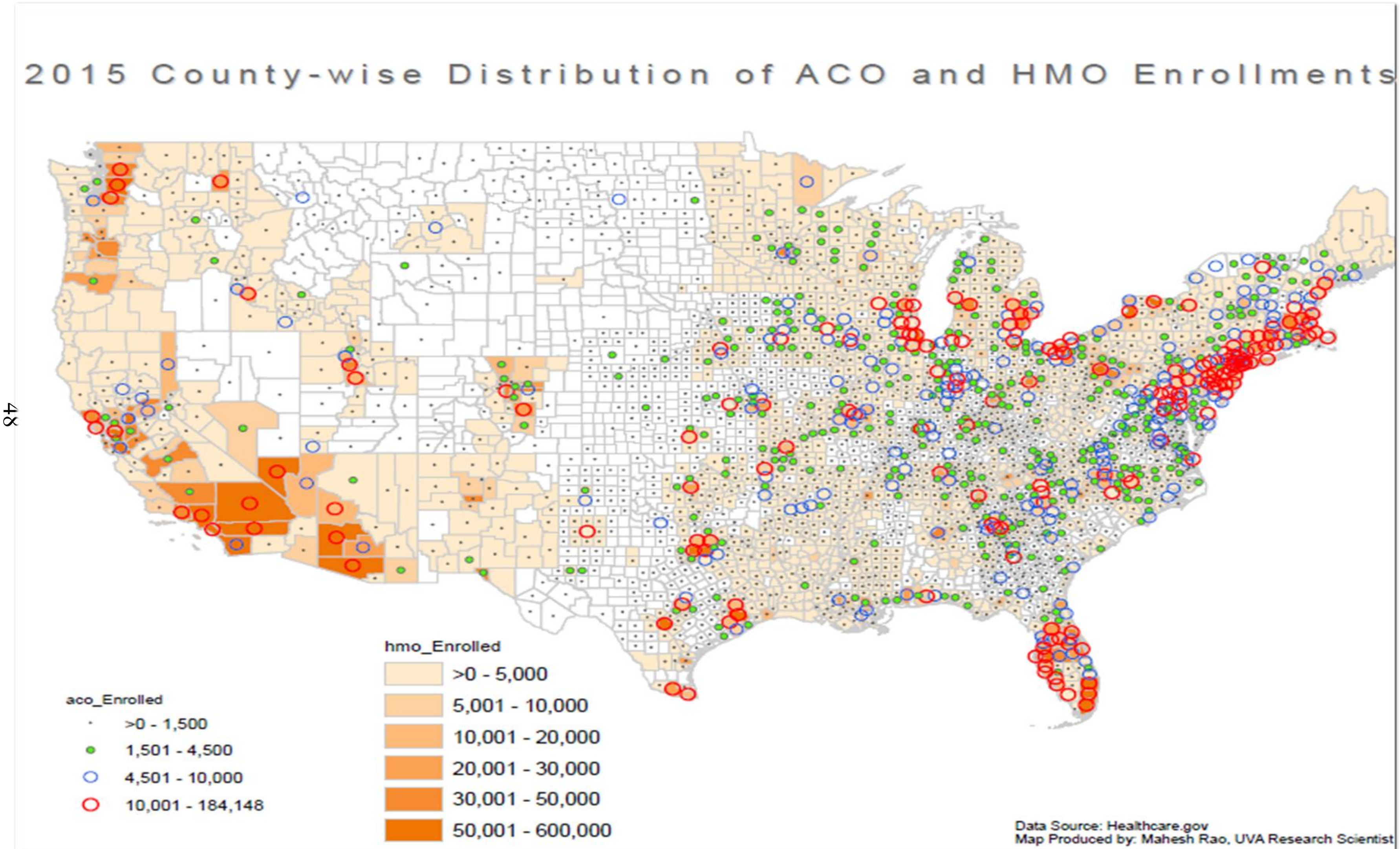

Figure 3: Map of HMO and ACO Beneficiaries Calendar Year 2015

Data Exploration and Cleaning

All data exploration, cleaning and analysis will occur in SPSS v. 24. Using the unique code identifier for each plan, a new dataset will be constructed with the following elements:

1. ID Number: ACO number of HMO contract ID code.
2. Cohort: 0 for HMOs, 1 for ACOs.
3. Plan start date.
4. Years of operations, based on an end date of December 31, 2015.
5. Total beneficiaries in calendar year 2015.
6. HbA1c: percent population with most recent HbA1c > 9.0%.
7. ColonCA: percent population screened for colon cancer.
8. ERVisits: ER visits per 1,000 population not resulting in hospital admission.
9. Average HCC Risk score: published for HMOs, calculated weighted average for ACOs.
10. White: percent of population identified as Caucasian.
11. Non-White: percent of population in all other race/ethnic categories.

Data cleaning. After producing descriptive statistics, identification of non-normal distributions will occur through visualization of histograms. Identification of univariate outliers will occur by calculating the Z-score associated to each case and variable. A multiple regression procedure using a dummy DV with all CVs and DVs specified above entered as IVs will facilitate the identification of multivariate outliers based on the Mahalanobis distance. Decisions about case deletion, missing data replacement or data transformations will depend on the cause of the outlier. The preferred approach, while preserving minimum cell sizes, is to delete cases. If this is not possible, transforming the data based on the nature of the variable will be the next option. Data replacement would next follow with continuous variables replaced by means and categorical variables replaced by medians.

Data transformation. In the case of skew or kurtosis associated to any DV, the natural log can normalize these distributions. Homoscedasticity will be verified through visualization of a residual scatterplot of observed versus predicted values for each DV. Linearity and homogeneity of variance observations will use normal P-P plots of standardized residuals and histograms for each DV. The Mahalanobis distance Chi-Square will guide decisions concerning deletion of outlier cases. As the number of cases exceeds the minimum sample size, based on the number of plans, case deletion can occur without sacrificing unnecessarily degrees of freedom.

Data substitution. The SPSS Missing Value Analysis (MVA) will depict the pattern of missing data and locations of the missing data. The zero-order correlation matrix will reveal possible collinearity between DVs. Data substitution will only occur in the case of greater than 5% missing values and then only if the case cannot be deleted.

Instrumentation

Because these data are secondary and include a variety of variable types, some explanation as to the nature of these variables is appropriate. It is somewhat difficult to place clinical variables in the appropriate context without an understanding of their relevance.

Clinical outcome. Measurement of blood sugar frequently occurs in the physician office using a "point-of-care" laboratory test called "glycosylated hemoglobin" frequently abbreviated HbA1c. This test essentially uses hemoglobin in the blood to estimate average blood sugar level over the previous three months, expressed as a percentage, indicating the concentration of sugar in the blood. Individuals with diabetes mellitus have HbA1c >6.5% at diagnosis although after treatment and management may be below 6.5%.

The HEDIS measures used by Medicare HMOs report several measures related to diabetes mellitus, one of these is the Category II CPT code indicator for HbA1c > 9.0%. The population measures are individuals with Type I and Type II diabetes mellitus. These are distinct clinical entities. The NQF measures used by Medicare ACOs also report the percent of population with most recent HbA1c > 9.0%. The population measured for NQF is individuals with Type II diabetes mellitus. Because the prevalence of Type I diabetes is less than 5% of all diabetes mellitus, the difference between HEDIS and NQF poses no serious threats to validity (Centers for Disease Control and Prevention, 2014).

Prevention outcome. Both HMOs and ACOs offer patients opportunities to undergo screening for colon cancer. The first screening is usually recommended at age 50, with a repeat interval of ten years absent a clinical need for a shorter interval. Screening can take several forms including colonoscopy, occult fecal testing or the more recently develop stool DNA analysis. The method of screening is likely less important than getting screened but colonoscopy is still considered the “gold standard” by most gastroenterologists. Both HEDIS and NQF have a measure for colon cancer screening rate. The measures do not differentiate based on type of screening. The population for the HEDIS measure is beneficiaries aged 51-75. The population for the NQF measure is beneficiaries aged 50-74. While certainly a limitation, the minor difference in the age grouping of the screening population for reporting does not invalidate the comparison on this measure.

Utilization outcome. This measure is identical between HEDIS and NQF. It is the number of ER visits per 1,000 attributed or enrolled that did not result in a hospital admission.

Data Analysis

After appropriate data cleaning, preparation and transformation, three separate multiple hierarchical regression models will detect the association between the IVs, HMO or ACO, and size of population served, relative to the three DVs: percent population with diabetes and last HbA1c > 9.0%, age-appropriate colon cancer screening rate, and ER visits per 1,000 not resulting in hospital admission.

The research questions and design of this study indicate multiple regression is the appropriate test with DVs being continuous, one categorical IV and one continuous IV. The hierarchical model is useful to detect improvements in the model by entering the IVs in a series of steps after entering the control variables.

Control variables. Identifying appropriate control variables is essentially to understand the impact of the IVs at each step in the models. The data for these variables are part of the HMO and ACO datasets discussed earlier. Issues of normality likely exist, particularly positive skew and leptokurtosis. The log transformation of these variables will make the distributions near normal. Differences in publicly available data for HMOs and ACOs to some degree limits the effectiveness of these control variables.

Race/ethnicity. Percent of population in certain race/ethnicity categories in the public data do not directly line up between HMOs and ACOs. It is for this reason, the control variable regarding race/ethnicity will be a dummy variable defined as 0=Caucasian, 1=non-Caucasian. Obviously, this would not permit detection of model changes based on sub-categories. As race/ethnicity are risk factors in the case of all three outcomes, this is a weakness of the study design but not avoidable.

Duration of operations. As the literature review indicates, the level of experience with quality measures may matter when it comes to reporting attainment of those measures. The influence of tenure of the organization on HEDIS and NQF measures is an area for future exploration but out of scope for this study. Tenure of patients in these organizations would also be of interest, but out of scope for this study.

Duration of operations is a control variable because it will remove the effect of longevity and experience with quality measures so a better understanding of the role of the IVs is possible. The duration of operations is expressed in years and tenths, based on the initial start date of the HMO or ACO reported in the publicly available data through 12/31/2015.

Hierarchical condition category risk score. The HCC risk score is a relatively new metric, implemented in 2003, used by Medicare for a variety of purposes. It consists of several elements. The methodology produces a numeric value that identifies the relative risk associated to the factors in the model, i.e., it is an actuarial determination of future utilization. Its primary purpose is payment particularly in Medicare Advantage plans relative to the PMPY amount the plans receive to manage care for a beneficiary population.

The HCC methodology maps the roughly 9,000 existing ICD-10 diagnosis codes in a one-to-many relationship to 79 HCCs. This mapping occurs at the individual beneficiary level based on data providers report to Medicare. Some diagnosis codes have no relevance in the HCC risk score. For instance, the set of ICD-10 diagnosis codes describing disorders of metabolism, like diabetes mellitus, may map into multiple HCCs. If an individual with diabetes suffers a fractured finger, the diagnosis code associated to the fracture would not change the individual HCC risk score for that beneficiary.

Age group, and then gender are the next elements in the HCC methodology. Age groups are <65 years, 65-74 years, 75-79 years, 80-84 years and ≥ 85 years. Each age group exists by gender in the HCC methodology.

The HCC methodology also identifies institutional or community residential status by beneficiary as well as whether the beneficiary is disabled or in end-stage renal disease status. The methodology also includes economic status by changing the HCC risk score if the beneficiary was Medicaid eligible during the calculation period.

While many criticism of the HCC risk scoring methodology exist and it is not perfect, it is very useful as a risk-adjustment tool in this analysis. It considers elements frequently recommended in developing a risk-adjustment methodology, namely, age, gender, and socioeconomic status.

CMS reports data on the average HCC risk score at the Medicare Advantage contract level. CMS reports average HCC risk score by beneficiary demographic category for ACOs. The Medicare Advantage average is the arithmetic mean. Because this mean derives from individual beneficiary HCC risk scores, and the ACO data reports the arithmetic mean by demographic category, it is necessary to weight the ACO HCC average risk scores by demographic group. Using the “mean of the means” in the ACO data would assume the contribution of each demographic category average HCC risk score should be the same. Since beneficiary level HCC risk scores are not available in the ACO data, weighting the average HCC risk score by ACO beneficiary category is the only way to establish some equivalence between the HMO and ACO data.

Limitations

Several limitations are important to note. First, these data are secondary. Improper laboratory controls, inaccurate Category CPT II code assignment and inaccurate ICD-10 diagnosis coding may exist. Second, non-reporting for any beneficiary regarding HbA1c or colon cancer screening places that beneficiary in the numerator of the HEDIS or NQF measure calculation. For example, a PCP may have a very well controlled diabetic patient in his/her panel. This patient's HbA1c may be <8.0%. However, if on the last patient encounter during the period of measurement the patient did not have an HbA1c level done, but should have, that patient will be considered out of control. In other words, missing data for a given beneficiary is the same as "fail" on these quality measures. Third, data on race/ethnicity are patient self-reported although the data may populate the patient's record. Race or ethnicity is not a medical issue, per se, and as such virtually all providers simply accept what the patient reports and enter that data in the record. Medicare maintains this data separately as well through its enrollment processes. An individual from Spain, for instance, could identify as Caucasian or Hispanic.

Limitations exist also regarding the HEDIS and NQF measures as they do not directly correlate between the measurement systems. The HbA1c measure in Medicare HMO includes beneficiaries with Type I and Type II diabetes mellitus. The NQF measure does not include Type I diabetes mellitus. Because the prevalence of Type I diabetes mellitus is 5% or less its existence in the HEDIS data for Medicare HMOs should not skew results in a meaningful way.

The age groups for colon cancer screening are slightly different. The HEDIS measure includes people aged 50-75. The NQF measure includes people aged 51-75. This is a small difference and does not fundamentally alter the applicability of this analysis. The importance of this measure is the screening process much more so than the age range.

Finally, some evidence suggests the structural factors, HMO or ACO, could lead to under-reporting of quality measures based on administrative claims data when compared to a delivery system using a single electronic health record (Trivedi et al., 2016). No way exists to assess the impact of this dynamic in this research.

Delimitations

This effort is an examination of several dependent variables for MSSP ACOs and Medicare Advantage HMOs using publicly available data. Analysis of outcomes based on subpopulations in either model are not part of this study. Medicare Advantage PPO plans are likewise not part of this study. This study focuses on outcomes relative to structure as viewed through the SPO framework. The specific incentives created by the plan type or design are not part of this study except by virtue of the fact that incentives are part of structure. Further study into the relationship between incentives, behaviors and outcomes would be meaningful.

Alternate Methodologies

With regard to multivariate analysis, discriminant analysis determines which variables discriminate between at least two naturally occurring groups. It is similar to Multiple analysis of the variance (MANOVA). In this case, the research questions are the inverse of a discriminant analysis as this study seeks to understand, based on group, differences in quality (Tabachnick, 2013). The purpose is not to predict group membership, rather, it is to understand the differences between the groups.

Factor analysis is similar to discriminant analysis in the prediction of group membership. However, it also involves identifying variance within the set of DVs. Using factor analysis would help understand the impact of latent and unobserved variables.

Multiple analysis of the covariance (MANCOVA) would be very applicable to this study with slightly different research questions. Ample evidence exists regarding the relationships between the DVs and age, gender, race/ethnicity, socioeconomic status and education. The purpose of this research is not to further quantify those relationships relative to structure. The MANCOVA approach also would not further clarify the role of structure on the DVs. MANOVA would likely be most appropriate in a comparison of a structure using patient-level data, where individual covariates would need control.

Conclusions

This is a retrospective, non-experimental cohort study seeking to understand differences between ACOs and HMOs regarding several outcome measures used similarly between these structures. The analysis will use multiple hierarchical regression with several control variables entering the model first. The IVs of interest will then entire the model in successive steps. The change in beta weights at each step of the model will reveal the contribution of the IV to the model.

The analysis will view structure as the IV. It will be operationalized through group assignment, HMO or ACO, and by size of the beneficiary population served. This will permit detection of how the model and the size of population result in model refinement.

Chapter 5: Results

Based on the enumerated differences in structure between Medicare ACOs and HMOs, would quality differ between the organization types? Would the number of beneficiaries served impact quality, regardless of plan type? Of all the measures available between HEDIS and NQF for these models, only three measures are similar enough for meaningful comparison: percent of population with diabetes and HbA1c > 9.0%, colon cancer screening rate, and ER visits per 1,000. Data analysis occurred using SPSS 24. This chapter reports the results of the analysis.

Data Collection

Data were collected from CMS Public Use Files (PUF) for ACOs and HMOs. CMS publishes these data at different web sites. The process of creating files with the necessary identifiers and variables was different for each organization type due to the format of the data published by CMS.

ACO data. The web site containing links to the ACO PUFs is: https://data.cms.gov/Special-Programs-Initiatives-Medicare-Shared-Savin/Medicare-Shared-Savings-Program-Accountable-Care-O/x8va-z7cu. These data were last updated on August 25, 2016, the date of publication. The data are available in a database which can be filtered to include only certain fields and then exported in a Microsoft Excel format. Data fields included in the dataset and used for this analysis were:

ACO ID number

Initial start date

Track_1

Total assigned beneficiaries and total by category in the performance year:

End-stage renal disease (ESRD)

Disabled

Aged/dual eligible

Aged/non-dual eligible

Number of beneficiaries by race/ethnic category:

White

Black

Asian

Hispanic

Native

Other

HCC risk score for 4 categories of beneficiaries in the performance year:

ESRD

Disabled

Aged/dual eligible

Aged/non-dual eligible

ER Visits not resulting in inpatient stay

ACO-19, colorectal cancer screening

ACO-27, percent of beneficiaries whose HbA1c is in poor control (>9.0%)

HMO data. CMS publishes Medicare Advantage data at the web site: https://www.cms.gov/ Research- Statistics -Data-and- Systems/ Statistics-Trends-and-Reports/ MCRAdvPartDEnrolData/index.html?redirect=/mcradvpartdenroldata. Several discrete files are needed to build a database containing the relevant information for this analysis:

1. Monthly enrollment by plan
2. MA HEDIS public use files
3. MA plan directory

A different web site contains the plan payment data where HCC risk score by plan is published: https://www.cms.gov/Medicare/Medicare-Advantage/Plan-Payment/Plan-Payment-Data.html

All the above files were downloaded and then opened in Microsoft Excel as the CMS Medicare Advantage data web sites are not interactive like the ACO web site. Using a series of lookups based on the Contract ID number, a new dataset was created containing the following information:

1. Contracts identified with an ID number HXXXX, and plan type = HMO, indicating the contract is an HMO
2. Contract enrollment, including Part D, as of December 2015. This was verified using the MA plan payment data using the plan type for HMO.
3. Average Part C risk score, which is the contract level HCC risk score.
4. HEDIS measure results:
    a. EOC020-0040, percent of population with HbA1c > 9.0%
    b. UOS528-0020, ER visits per 1,000 with no inpatient admission
    c. Race/Ethnicity data:

i. PD1808-1160, percent of white beneficiaries

ii. PD1808-1210, percent of black beneficiaries

iii. PD1808-1310, percent of Asian beneficiaries

iv. PD1808-1360, percent of native Hawaiian or other Pacific Islanders

v. PD1808-1570, percent of Hispanic beneficiaries

vi. PD1808-1410, percent of some other race

vii. PD1808-1230, percent of Native beneficiaries

d. EOC040-0010, reported rate of colorectal cancer screening

Data Concatenation

Using Microsoft Excel, the following specifies the steps needed to concatenate the data sets to create a research database suitable for use in SPSS 24:

1. Create a dummy variable for cohort and assigned ACO=0, HMO=1
2. Calculate the weighted average HCC risk score in each ACO by computing the percent of total assigned beneficiaries in each category in the performance year and multiplying by the HCC risk score for that category of beneficiaries. The sum of these products is the weighted average HCC risk score for the ACO.
3. Calculate the percent of ACO and HMO beneficiaries who are Caucasian and the percent non-Caucasian to create two new dummy variables, percent White and percent not-White.
4. Using the ACO and HMO Contract start dates, calculate operating years to the nearest tenth using an end date of December 31, 2015.

The resulting database includes the following fields:

1. ID: the ACO ID or HMO Contract ID numbers
2. Cohort: HMO=0, ACO=1

3. Population
4. HbA1c: beneficiaries with diabetes per the NQF or HEDIS measure, with last HbA1c>9.0%.
5. ColonCA: percent of beneficiaries screened for colon cancer per the NQF and HEDIS measures.
6. ER Visits: ER visits per 1,000 beneficiaries not resulting in inpatient admission.
7. HCC Score: average HCC risk score.
8. Operating Years: years and tenths of years calculated using the start date of the plan and ending December 31, 2015.

Data Cleaning and Preparation

In order to use the research dataset, missing values and univariate outliers needed identification as the existence of either can distort the analysis (Tabachnick, 2013). Missing values were missing at random, indicated by Little's MCAR ($\chi2 = 14.142$, df=25, p=.439).

Identification of univariate outliers occurred by examining Z scores for values > 3.29. A case is an outlier if is deviates from the mean by +/- three standard deviations (Tabachnick, 2013). Of the 619 cases in the data, 28 cases were outliers on one or more of the variables leaving cases in the dataset for analysis. Because deleting outliers did not compromise power, these 28 cases were deleted.

Three cases had missing values for at least one data element. Deleting these cases, without replacement, further reduces the total sample to 588 cases with 15,777,561 beneficiaries. Because deleting these cases does not materially reduce power, they were also deleted.

Study Variable Intercorrelations

Bivariate correlation analysis identified relationships between variables. None of the correlations were above tolerance of the Pearson product-moment correlation ($r > .70$) although several correlations were significant at $p<.01$ in the two-tailed analysis. Most interesting were the negative correlations between operating years and ER visits per 1,000 (-.557) and between colon cancer screening and HbA1c > 9.0% (-.533). Table 6 presents these results.

Table 6: *Correlations Between Study Variables*

	% Not White	Avg. HCC	HbA1c	ColonCA	ER Visits	Population	Operating Years
% not-White		.207**	.182**	.020	-.075	.072	.293**
Avg. HCC			.157**	-.001	.148**	.125**	.343**
HbA1c				-.533**	.283**	-.139*	-.036
ColonCA					-.461**	.272**	.338**
ER Visits						-.305**	-.557**
Population							.452**
Operating Years							

*p<.05, **p<.01.

Descriptive Statistics

Table 7 contains the descriptive statistics for each variable.

Table 7: *Descriptive Statistics*

	HbA1c	ColonCA	ER Visits	HCC Score	not-White	Population	Years
Range	.96	.71	1,171.44	1.03	.98	326,257.00	29.61
Minimum	.04	.21	192.45	.68	.02	513.00	1.00
Maximum	1.00	.92	1,364.00	1.70	1.00	326,770.00	30.60
Mean	.20	.64	639.45	1.07	.26	27,109.21	7.20
Std. Dev.	.09	.12	222.16	.15	.27	31,613.23	7.90
Skewness	2.95	-.54	.36	1.07	1.61	3.46	1.51
Kurtosis	18.27	.14	.05	1.82	1.69	19.13	1.19

In sample sizes greater than 300 the standard errors of skewness or kurtosis may cause Type 1 error for the Z-test of the null hypothesis (Hae-Young Kim, 2013). The Z-scores for the kurtosis statistic exceeded 1.96 except for colon cancer screening rate and ER visits per 1,000.

The Z-scores associated to skewness exceed +/- 1.96 for each variable. Visualization of the histograms and the absolute values of skewness, for values >2, and kurtosis, for values >7, should guide decisions about variable transformation in large sample sizes (Hae-Young Kim, 2013).

A case could be reasonably made to use these variables without transformation. However, the decision should take the assumptions of multiple regression into account. If the natural state of these variables should cause heterogeneity of residual variance or heteroscedasticity, then transformation would be appropriate.

Assumptions of Multiple Regression

Multiple regression relies on several important assumptions. The ratio of cases to IVs exceeds the minimum recommendation of twenty (Tabachnick, 2013). The data should demonstrate linearity, with homogeneity of variance and homoscedasticity. These are assessed through visualization of histograms, P-P plots and scatterplots of residual versus predicted values. While multiple regression is robust to violations of normality, assessment of the data and residuals as normal or near-normal strengthens the interpretation of results (Tabachnick, 2013). Because this analysis tests the relationship between organization type and population size on three separate DVs, assessment and reporting of these assumptions will occur in the discussion of each model.

Data transformation. In each model, however, the assessment of violations of these assumptions revealed obvious deviations from normality using un-transformed variables. As a result, the multiple regression procedure occurred using the natural log transformation of all continuous variables. The unstandardized beta weights of the un-transformed variables reflected the same relationships as those found using the transformed variables reported in the data analysis.

Multivariate outliers. Multivariate outliers were assessed as part of a separate regression procedure. In this procedure the dummy variable for cohort served as the DV with all the study CVs and DVs, along with the IV for population, input in one step. The Mahalanobis distance statistics indicated the presence of multivariate outliers. However, the maximum Cook's Distance values were less than 0.5, indicating these values were low leverage and distance, and removing them from the models would not materially alter results. The existence of multivariate outliers with low leverage and distance characteristics suggest caution interpreting results but does not invalidate the use of those variables (Tabachnick, 2013). To remove uncertainty, however, Mahalanobis distance values with $p<.01$ were identified. Six cases had $p<.01$ and were deleted resulting in 582 cases in the dataset for analysis.

Multivariate collinearity was assessed as part of the regression procedure. Eigenvalues ranged from .001 to .900. Variance Inflation Factors (VIF) were only >2 for two variables and those values were less <3 (ER visits and operating years). The condition index was <30 for all variables except colon cancer screening at 36.21.

While multivariate outliers were present and some collinearity exists, these conditions are not surprising based on the nature of the data. The assessment of linearity, homogeneity of variance and homoscedasticity using transformed variables, showed in each model normal or near-normal conditions such that continuation of the analysis was appropriate (Tabachnick, 2013).

Data Analysis

The following describes the results of the multiple hierarchical regression models testing the previously stated hypothesis concerning HbA1c, colon cancer screening and ER utilization based on organization type and population size. To avoid confusion in the interpretation of the sign of beta weights for each DV based on the IVs, specific identification of these relationships occurs coincident with the reported results on each DV.

Model one: organization type. Multiple hierarchical regression was used to determine whether a relationship exists between the type of organization, ACO or HMO, and the DVs previously described. The dummy variable created for cohort identified HMO=0 and ACO=1. The IVs and CVs were transformed using the natural log as previously discussed.

Percent population with HbA1c > 9.0%. The research question is: do Medicare ACOs have a greater percent of population with diabetes mellitus and most recent HbA1c > 9.0% than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?

The model was significant, $F(1,571) = 24.524$, p=.000, adjusted $R^2 = .141$. The R^2 change at the fourth step in the model was .015. Table 8 reports the regression coefficients and standard errors:

Table 8: *Multiple Hierarchical Regression of HbA1c Based on Cohort*

	B	Standard Error	*t*	Significance
Log not-White	.109	.018	6.069	.000
Log HCC	.559	.118	4.728	.000
Log Years	-.149	.037	-4.891	.000
Cohort	-.215	.061	-3.208	.001

These results indicate that a higher percentage of not-White patients would have a higher percentage of diabetic patients with the most recent HbA1c > 9.0%. A higher average HCC risk score would indicate the same finding. A longer duration of operations would associate to a lower percentage of diabetic patients with last HbA1c > 9.0%. After accounting for the variation on these CVs, being in the ACO cohort would associate to a lower percentage of patients with last HbA1c > 9.0%.

The histogram reflects some kurtosis but appears near-normal. The P-P plot of standardized residuals show small deviation from the diagonal line of identity.

The scatterplot of residual v. predicted values demonstrates sufficient homoscedasticity.

Figure 4 is the histogram of the log transformed HbA1c > 9.0%

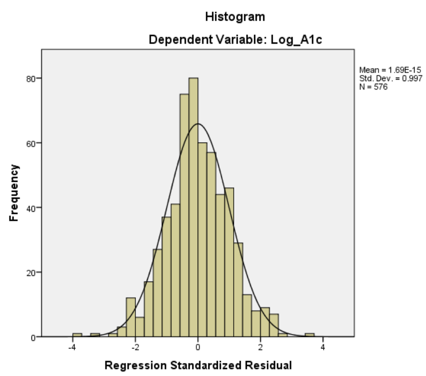

Figure 4: Histogram of Log Transformed HbA1c > 9.0%

Figure 5 is the P-P plot of standardized residuals.

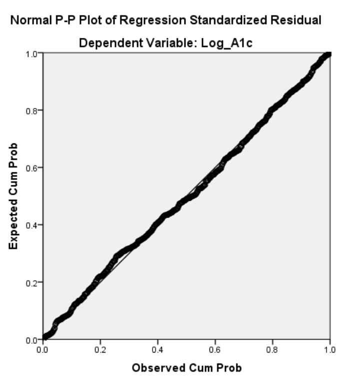

Figure 5: P-P Plot of Standardized Residuals

Figure 6 is the scatterplot of standardized residual v. predicted values.

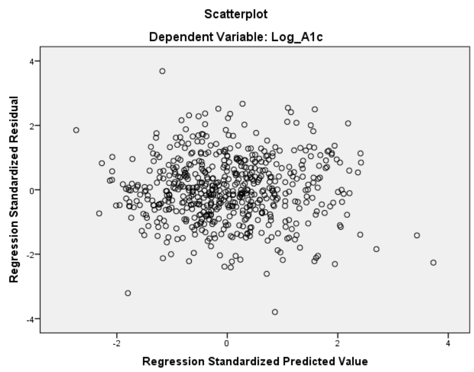

Figure 6: Scatterplot of Residual v. Predicted Values

Percent population screened for colon cancer. The research question is: do Medicare ACOs have a greater percent of population screened for colon cancer than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?

The model was significant, $F(1,571)$ = 38.95, p=.000, adjusted R^2=.206. The R^2 change at the fourth step in the model was .000. Table 9 reports the regression coefficients and standard errors.

Table 9: *Multiple Hierarchical Regression of Colon Cancer Screening Based on Cohort*

Variable	B	Standard Error	*t*	Significance
Log not-White	-.038	.009	-4.201	.000
Log HCC Risk Score	-.220	.059	-3.699	.000
Log Operating Years	.102	.015	6.642	.000
Cohort	-.003	.034	-.083	.934

A larger percentage of not-White beneficiaries and a higher average HCC risk score would associate to a lower colon cancer screening rate. A longer duration of operations would also associate to a higher screening rate. However, after the CVs, the cohort identification did not return a significant result and as such is an equivocal finding.

The histogram continues to reflect some kurtosis and perhaps negative skew but appears near-normal. The P-P plot of standardized residuals show small deviation from the diagonal line of identity but are near-normal. The scatterplot of residual v. predicted values demonstrates sufficient homoscedasticity, although there is some visual evidence of multicollinearity. Figure 7 is the histogram of the log transformed colon cancer screening rate.

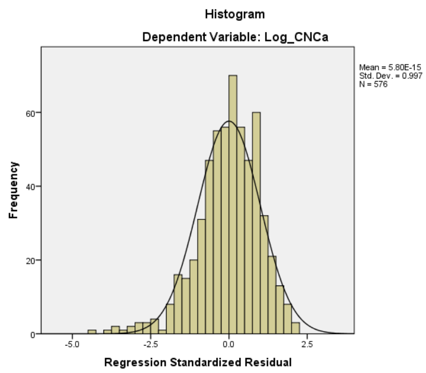

Figure 7: Histogram of Log Transformed Colon Cancer Screening Rate

Figure 8 is the P-P plot of the standardized residuals.

A higher percentage of not-White beneficiaries and a higher HCC risk score would indicate a greater number of ER visits per 1,000

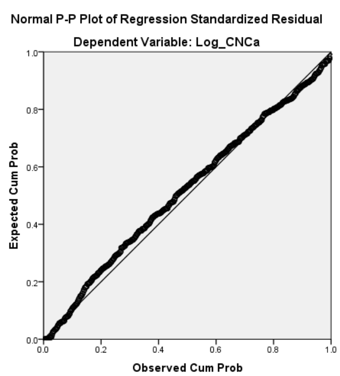

Figure 8: P-P Plot of Standardized Residual

Figure 9 is the scatterplot of standardized residual v. predicted values.

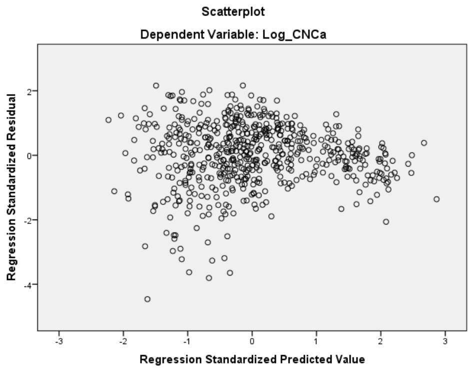

Figure 9: Scatterplot of Residual v. Predicted Values

ER utilization based on cohort. The model was significant, *F(1,571)* = 11.376, p=.000, adjusted R^2=.559. The R^2 change at the fourth step in the model was .044. Table 10 reports the regression coefficients and standard errors.

Table 10: *Multiple Hierarchical Regression of ER Visits per 1,000Bbased on Cohort*

Variable	B	Standard Error	*t*	Significance
Log Not-White	.052	.012	4.186	.000
Log HCC Risk Score	1.037	.082	12.702	.000
Log Operating Years	-.145	.021	-6.893	.000
Cohort	.350	.046	7.591	.000

A longer duration of operations would indicate lower ER utilization. After the CVs, the identification of cohort with a positive beta weight indicates higher ER utilization in the ACOs.

The histogram continues to reflect some kurtosis but appears near-normal. The P-P plot of standardized residuals show small deviation from the diagonal line of identity but are near-normal. The scatterplot of residual v. predicted values demonstrates sufficient homoscedasticity. Figure 10 is the histogram of the log transformed ER visits per 1,000.

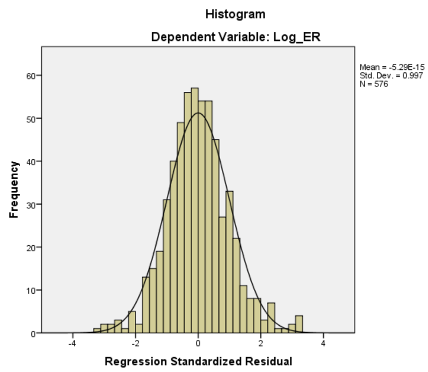

Figure 10: Histogram of Log Transformed ER Visits per 1,000

Figure 11 is the P-P plot of standardized residuals.

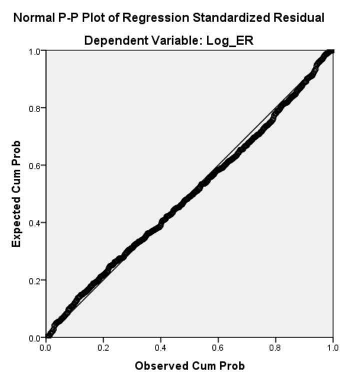

Figure 11: P-P Plot of Standardized Residual

Figure 12 is the scatterplot of standardized residual v. predicted values.

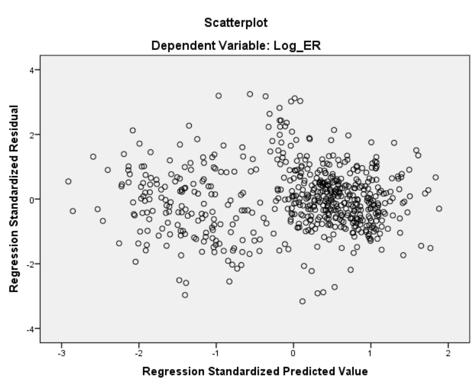

Figure 12: Scatterplot of Residual v. Predicted Values

Model two: population size. This model views the IV associated to structure through the perspective of the number of beneficiaries served by the plans to examine the effect on the DVs.

Population with HbA1c > 9.0%. The research question is: do ACOs and HMOs serving larger beneficiary populations have fewer patients with diabetes mellitus and most recent HbA1c>9.0%, controlling for race/ethnicity, average HCC risk score and plan years of operations?

The model was significant, $F(1,571) = 7.271$, p=.000, adjusted R^2=.15. The R^2 change at the fourth step in the model was .029. Table 11 reports regression coefficients and standard errors.

Table 11: *Multiple Hierarchical Regression of HbA1c Based on Population*

Variable	B	Standard Error	*t*	Significance
Log Not-White	.123	.017	7.301	.000
Log HCC Risk Score	.505	.117	4.301	.000
Log Operating Years	-.035	.018	-1.944	.052
Population	-.083	.019	-4.454	.000

A higher percentage of not-White patients and higher HCC risk score would indicate a higher percentage of diabetic patients with most recent HbA1c > 9.0%. This is consistent with the results of the cohort analysis. The negative beta weight associated to duration of operations indicates newer plans have a greater proportion of diabetic patients with last HbA1c > 9.0%. This result approached significance but was not conclusive. After the CVs are taken into account, this model indicates a larger population of beneficiaries would have a lower percentage of diabetic patients with last HbA1c > 9.0%, regardless of plan type.

The histogram continues to reflect some kurtosis but appears near-normal. The P-P plot of standardized residuals show small deviation from the diagonal line of identity but are near-normal. The scatterplot of residual v. predicted values demonstrates sufficient homoscedasticity.

Figure 13 is the histogram of the log transformed HbA1c > 9.0%.

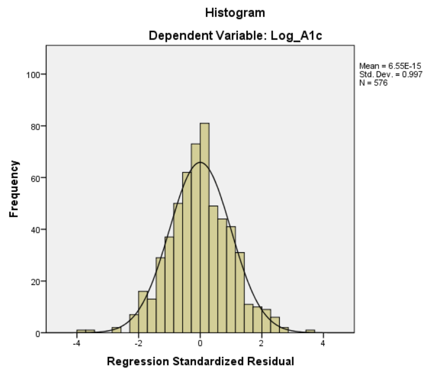

Figure 13: Histogram of Log Transformed HbA1c > 9.0%

Figure 14 is the P-P plot of standardized residuals.

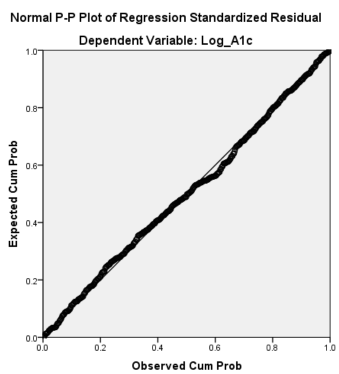

Figure 14: P-P Plot of Standardized Residual

Figure 15 is the scatterplot of standardized residual v. predicted values.

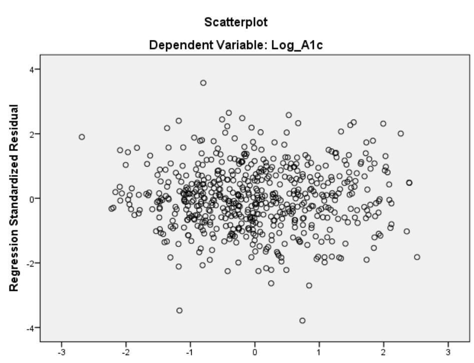

Figure 15: Scatterplot of Residual v. Predicted Values

Percent population screened for colon cancer. The research question is: Do ACOs and HMOs serving larger beneficiary populations have more patients screened for colon cancer, controlling for race/ethnicity, average HCC risk score and plan years of operations?

The model was significant, $F(1,571) = 43.174$, p=.000, adjusted R^2=.227. The R^2 change at the fourth step in the model was .020. Table 12 reports regression coefficients and standard errors.

Table 12: *Multiple Hierarchical Regression of Colon Cancer Screening Based on Population*

Variable	B	Standard Error	*t*	Significance
Log Not-White	-.036	.008	-4.279	.000
Log HCC Risk Score	-.206	.059	-3.512	.000
Log Operating Years	.089	.009	9.907	.000
Population	.036	.009	3.886	.000

Having more not-White beneficiaries and a higher HCC risk score would indicate a lower colon cancer screening rate in the population. A longer duration of operations associates to a higher screening rate. This is consistent with the findings in the cohort model. The IV for population size, indicates a larger population would associate to a higher screening rate after controls. Where the finding on in the cohort analysis for this DV was not significant, this finding is significant.

The histogram continues to reflect some kurtosis and perhaps negative skew but appears near-normal. The P-P plot of standardized residuals show small deviation from the diagonal line of identity but are near-normal. The scatterplot of residual v. predicted values demonstrates sufficient homoscedasticity, although there is some visual evidence of multicollinearity. Figure 16 is the histogram of the log transformed colon cancer screening rate.

ER utilization per 1,000. The model was significant, *F(1,571)* = 158.526, p=.000, adjusted R^2=.523. The R^2 change at the fourth step in the model was .009. Table 13 reports regression coefficient and standard errors.

Table 13*: Multiple hierarchical regression of ER visits per 1,000 based on population*

Variable	B	Standard Error	*t*	Significance
Log Not-White	.021	.012	1.694	.091
Log HCC Risk Score	1.056	.085	12.445	.000
Log Operating Years	-.263	.013	-20.225	.000
Population	-.044	.013	-3.245	.001

A higher percentage of not-White beneficiaries and a higher HCC risk score would indicate more ER visits per 1,000. A longer duration of operations would indicate lower ER utilization. These results are consistent with the cohort model. This result also indicates that a larger population of beneficiaries served would have fewer ER visits per 1,000, after controls.

The histogram continues to reflect some kurtosis but appears near-normal. The P-P plot of standardized residuals show small deviation from the diagonal line of identity but are near-normal. The scatterplot of residual v. predicted values demonstrates sufficient homoscedasticity. Figure 16 is the histogram of the log transformed colon cancer screening rate.

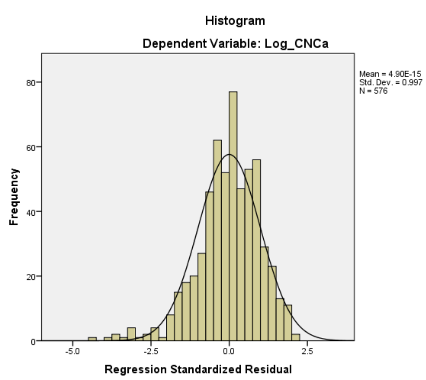

Figure 16: Histogram of Log Transformed Colon Cancer Screening Rate

Figure 17 is the P-P plot of standardized residuals.

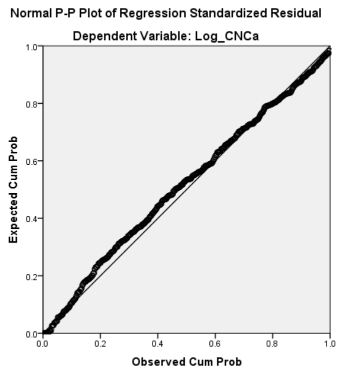

Figure 17: P-P plot of standardized residual

Figure 18 is the scatterplot of standardized residual v. predicted values.

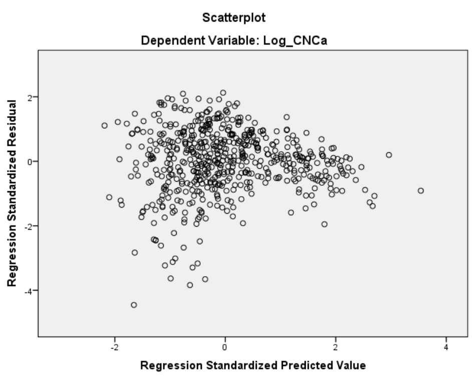

Figure 18: Scatterplot of residual v. predicted values

Figure 19 is the histogram of the log transformed ER visits per 1,000.

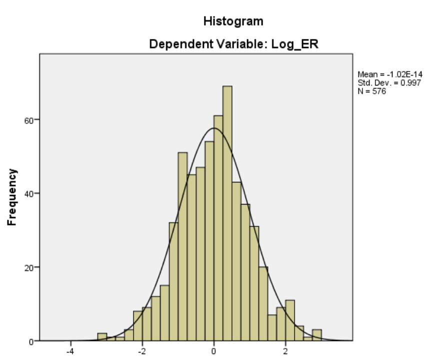

Figure 19: Histogram of Log Transformed ER Visits per 1,000

Figure 20 is the P-P plot of standardized residuals.

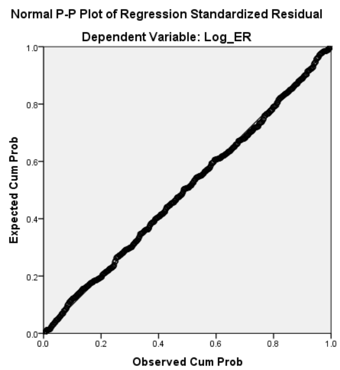

Figure 20: P-P Plot of Standardized Residuals

Figure 21 is the scatterplot of standardized residual v. predicted values.

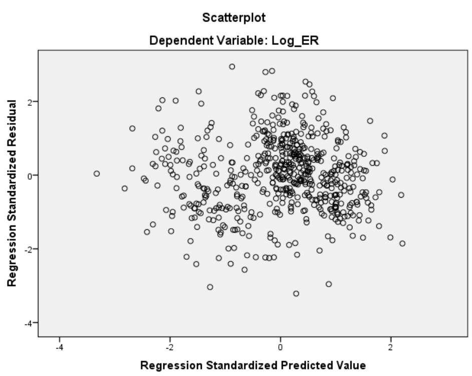

Figure 21: Scatterplot of Residual v. Predicted Values

Summary of Findings

Findings associated to each hypothesis are summarized in Table 14.

Table 14: *Summary of Findings*

Hypothesis	Supported
H1: Medicare ACOs have fewer beneficiaries with last HbA1c > 9.0%	Yes
H2: Medicare ACOs screen more age-appropriate patients for colon cancer	No
H3: Medicare HMOs have fewer ER visits per 1,000 not resulting in hospital admission	Yes
H4: Organizations serving larger populations will have fewer diabetic patients with last HbA1c > 9.0%	Yes

Table 14: *Continued*

Hypothesis	Supported
H5: Organizations serving larger populations will have a higher rate of colon cancer screening	Yes
H6: Organizations serving larger populations will have fewer ER visits per 1,000 not resulting in hospital admission	Yes

- Hypothesis 1. Findings indicate ACOs have a smaller population of diabetic patients with last HbA1c > 9.0% (β = -.215, $p<.001$). Cohort accounted for 1.5% of the regression model. These results supported hypothesis one.
- Hypothesis 2. Findings are equivocal as to whether ACOs screen more age-appropriate patients for colon cancer (β=-.003, p=.934). Cohort accounted for 0% of the regression model. The results did not support hypothesis two.
- Hypothesis 3: Findings indicate HMOs have lower ER visits per 1,000 not resulting in hospital admission (β=.350, $p<.001$). Cohort accounted for 4.4% of the regression model. The results support hypothesis three.
- Hypothesis 4: Findings indicate the larger the population served, the lower the percentage of diabetic population with last HbA1c > 9.0% (β=-.083, $p<.01$). Population size accounted for 2.9% of the regression model. The results support hypothesis four.
- Hypothesis 5. Findings indicate serving a larger population associated to a higher age-appropriate colon cancer screening rate (β=.036, $p<.01$). Population accounted for 2.0% of the model. The results support hypothesis five.
- Hypothesis 6: Findings indicate serving a larger population associated to lower ER visits per 1,000 not resulting in hospital admission (β=-.044, p=.001). Population accounted for .09% of the regression model. The results support hypothesis six.

Conclusions

On balance, the control variables explained a significant amount of variation in both models for each DV. The R^2 change in each model where the IV entered the equations were uniformly small, the largest being .350 when analyzing ER visits based on cohort. Only one finding, for hypothesis two, was not significant. The unstandardized beta weights associated to the transformed values of the variables require caution in their interpretation as the percentage change in a log value will not equal the percentage change on the variable prior to transformation. However, the consistency of the direction of the beta weights for the control variables gives some indication that the demographics of the population served are meaningful in terms of quality measures.

Chapter 6: Discussion

This research examined the role of structure relative to outcome using the Donabedian quality framework. It was retrospective, non-experimental, and used publicly available data to evaluate quality based on two operational definitions of structure.

This chapter summarizes the introduction, conceptual framework, literature review and methods. It reviews results and offers interpretation of their implications. Thoughts regarding additional lines of inquiry are also identified as well as limitations and delimitations.

Introduction

The introduction provided background regarding the ongoing national conversation and debate about healthcare reform. Despite the per capita spend in the US being twice as high as other first-world countries, the health status of the US population is arguably worse than some third-world countries. The US has shorter life expectancy, more infant mortality, more obesity and a greater chronic disease burden than other countries spending far less. The inexorable year-over-year increase in the NHE, now estimated at $3.3 trillion each year, and increasing, might be justified if the return matched the investment (Centers for Medicare & Medicaid Services (CMS), HHS, August, 2016; Fuchs, 2010). At nearly 17% of the GDP, however, the focus on quality in conjunction with decreasing at least the rate of NHE increase has profound implications on the US economy.

These conversations about reform, cost, access and even quality are not new. The idea of creating a social health insurance program no doubt came up before the Truman administration but it was during that time the idea began to gain traction. It culminated with the creation of Medicare Part A in 1965. Cost pressures in the Medicare program are building as trust fund insolvency looms in the not-too-distant future while 10,000 baby boomers each day age into eligibility. Medicare has ample experience developing programmatic approaches to improving quality and managing cost, e.g., DRGs, Part B fee schedule. Today, Medicare spending consumes 20% of the NHE.

One of the most significant programs developed was Medicare managed care. Initially called "Part C" and now called "Medicare Advantage," today nearly 20 million beneficiaries receive care through the program. Nearly 64% of Medicare Advantage plans are HMOs, most of them offered by health insurance companies like: United Healthcare, Anthem, Humana and Cigna. Roughly 30% of Medicare Advantage plans are "provider sponsored" meaning the plan is part of an umbrella organization that includes providers. The majority of plans use contracted networks to deliver care. These plans bid risk contracts to Medicare in specific geographies. The HMOs in turn pass financial risk to their provider networks, which are typically narrowly constituted.

With the implementation of PPACA, Medicare created a novel structure, the ACO. The most common form is the MSSP Track 1 ACO. In 2016, 91% of MSSP ACOs were in Track 1, which do not bear direct financial risk. Medicare reported the composition of MSSP ACOs to be 56% organized as networks of individual practices (IPAs), 36% as group practices and 38% as hospital/professional partnerships (Centers for Medicare & Medicaid Services (CMS), HHS, 2017a).

While similarities exist between ACOs and HMOs, there are profound differences. These differences are structural responses to incentives created by Medicare. The ACOs and HMOs then create incentives for the providers. The alignment or diffusion of these incentives varies with structure.

Conceptual Framework

Donabedian offered the Structure-Process-Outcome (SPO) model to aid in a more complete understanding of quality, coincidentally right after Medicare came into existence (A. Donabedian, 1966). The SPO framework identifies structure as resources, resource allocation, location, funding and organization. Process describes technical and interpersonal skill. The framework is hierarchical building from structure through process to outcome. But, as a common definition of "quality" is still elusive, the SPO framework suggests that the result achieved is only part of "quality."

Taken together, the differences in beneficiary participation, provider participation, the ability of HMOs to modify benefits, and the differences in payment and reporting, support an operational definition of HMOs and ACOs as different structures. The differences in funding and organization yield varying incentives which may shift priority between quality and cost based on structure.

Because structure also refers to the quantity, allocation and location of resources, a second operational definition of structure measures by the size of the beneficiary population served is also tested. As the size of the beneficiary population increases the number of resources needed to provide care must also increase.

Literature Review

The literature is not abundantly populated with comparisons of quality between structures. However, the SPO is often used in quality research (Ayanian & Markel, 2016). Previous research explored differences between Medicare FFS and HMOs using variables like the SF-36. Other studies examined differences between for-profit and not-for-profit Medicare HMOs. One approach involved comparing quality in the VA Health System and Medicare Advantage, essentially using the Health Effectiveness Data and Information Set (HEDIS) measures (Trivedi & Grebla, 2011). This study is important because it identified differences in quality and because the composition of ACOs more closely aligns to integrated delivery systems (IDS), like the VA Health System, than does the ownership of HMOs.

Methods

The research questions explored the impact of structure on outcome. The research questions were:

1. Do Medicare ACOs have a greater percent of population with diabetes mellitus and most recent HbA1c > 9.0% than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?
2. Do Medicare ACOs have a greater percent of population screened for colon cancer than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?
3. Do Medicare ACOs have lower ER utilization per 1,000 not resulting in hospital admission than HMOs, controlling for race/ethnicity, average HCC risk score and plan years of operations?
4. Do ACOs and HMOs serving larger beneficiary populations have fewer patients with diabetes mellitus and most recent HbA1c>9.0%, controlling for race/ethnicity, average HCC risk score and plan years of operations?
5. Do ACOs and HMOs serving larger beneficiary populations have more patients screened for colon cancer, controlling for race/ethnicity, average HCC risk score and plan years of operations?

6. Do ACOs and HMOs serving larger beneficiary populations have fewer ER visits per 1,000 not resulting in hospital admission, controlling for race/ethnicity, average HCC risk score and plan years of operations?

The data were contained in Public Use Files (PUF) and existed in multiple disparate databases. The HMO plans use HEDIS to report quality performance. The ACO plans use the National Quality Foundation (NQF) measures. The three measures have minor differences in the level of measurement but were essentially congruent. These measures were the dependent variables. They were: the percent of population with diabetes mellitus and last HbA1c > 9.0%, colon cancer screening rate and ER visits per 1,000 not resulting in hospital admission.

The hypotheses were as follows:

- Hypothesis 1: Medicare ACOs have fewer diabetic patients with last HbA1c > 9.0%
- Hypothesis 2: Medicare ACOs screen more age-appropriate patients for colon cancer.
- Hypothesis 3: Medicare HMOs have fewer ER visits per 1,000 not resulting in hospital admission.
- Hypothesis 4: Organizations serving larger populations will have fewer diabetic patients with last HbA1c > 9.0%
- Hypothesis 5: Organizations serving larger populations will have a higher rate of colon cancer screening.
- Hypothesis 6: Organizations serving larger populations will have fewer ER visits per 1,000 not resulting in hospital admission.

The sample power derived from the number of beneficiaries included. The sampling frame included 609 HMOs and ACOs serving 17,898,330 beneficiaries. The analysis used multiple hierarchical regression. As the data were publicly available with no beneficiary identifiers, this was not a human subjects study.

Risk adjustment occurred through the use of several control variables (CV). The CVs were the percent of population not-White, plan level average Hierarchical Condition Category (HCC) risk score and duration of operations through December 31, 2015.

The data were explored and cleaned. Univariate outliers were identified using Z-scores. Multivariate outliers were identified by the Mahalanobis distance. This resulted in the deletion of 28 cases. Three additional cases were deleted due to missing data which was determined to be missing at random using Little's MCAR. Positive skew and kurtosis were corrected using the natural log transformation of continuous variables.

The cohort model used a dummy variable, with HMO = 0 and ACO = 1 as the independent variable (IV). The CVs entered the equation in three steps in the order of the percent not-White, HCC risk score and duration of operations. The DVs were the percent with last HbA1c > 9.0%, colon cancer screening rate and ER visits per 1,000.

The population model used a continuous independent variable which was the size of the beneficiary population for each plan. The CVs were entered in three steps as above. The DVs were the same as the cohort model

Results

The beta weights associated to the DVs in both models are summarized in Table 15. The results of the cohort model perhaps illuminate how aligned the incentives created in each program really are. The ACOs had fewer beneficiaries with diabetes and elevated blood sugar, after controls. There was no difference between ACOs and HMOs in the rate of colon cancer

Table 15: *Summary of Beta Weights for Each Variable*

	Cohort			Population		
	HbA1c	ColonCA	ER Visits	HbA1c	ColonCA	ER Visits
Not-White	.109	-.038	.052	.123	-.036	.021
HCC Score	.559	-.220	1.037	.505	-.206	1.056
Operating Years	-.149	.102	-.145	-.145	.089	-.263
Cohort	-.215	-.003*	.350			
Population				-.083	.036	-.044

*p>.05

screening, after controls. HMOs had fewer ER visits per 1,000 than ACOs, after controls. The results support the hypothesized relationships except for the colon cancer screening rate.

The results of the population model support all three hypotheses. As the organization increases in size to serve more beneficiaries, performance on the quality measures improves. These findings are in keeping with a common approach in value-based purchasing (VBP) of enrolling subscribers at some scale. The idea of spreading the risk of fewer high-risk patients over a larger population of lower risk patients is a commonly employed strategy.

Discussion

Cohort model. Consider two individuals with Type II diabetes. One beneficiary has an HbA1c of 8.9%, the other has HbA1c at 9.1%. In terms of how well controlled a physician might consider their disease, this difference is negligible. We might prefer the lower number, but in reality, a practitioner might not view these numbers very differently. In terms of utilization, there is not likely to be significant variance based on the minor difference in blood sugar level. Because any HbA1c reading is a “snapshot” of three months, we might counsel the person at 9.1% to make some minor changes and re-check in three months. In fact, that is likely to be the exact approach for the individual at 8.9%.

If the beneficiary with their most recent HbA1c at 9.1% is in an ACO, then the ACO fails that measure for that beneficiary. If enough of these exist in the attributed population, then the

ACO does not qualify to share savings no matter how much it might actually save. If this is an HMO enrollee, then an opportunity for a quality incentive through the Stars Rating might not exist. But, all things being equal, the HMO is taking financial risk and passing that on to its provider network so no financial penalty results from this last HbA1c level. If the ACO alters its utilization pattern in managing this beneficiary to reduce cost but has this HbA1c result, the reduced utilization does not yield a shared savings and the ACO providers forego the FFS payment for services they might otherwise receive. If the HMO changes its utilization pattern, and gets the same HbA1c result, then it presumably generates profit but may not get the Stars Rating incentive. The priority of meeting quality measures and financial implications results in a decidedly mixed incentive in comparing these structures.

The results on colon cancer screening did not differentiate based on the IV for structure in cohort model. The incentives to screen and detect are difficult to identify. Prevention activities have a cost. The temporal relationship between the cost and realizing the financial benefit of preventing that disease exists over a period of years. During that interval, the beneficiary may change plans meaning the benefit of the investment in prevention might accrue to another ACO or HMO. Colonoscopy, still the gold standard, is at best difficult for patients and more expensive than other tests. Patients prefer other less invasive and lower cost methods of screening while physicians prefer colonoscopy (Hawley, Lillie, Cooper, & Elston Lafata, 2014). Even though individuals aged 66 and older get screened at a higher rate than individuals aged 50-65, the screening rate in the 66 plus age group was only 66% in 2015 (American Cancer Society, 2017). Chronic disease and utilization measures are more “now,” where prevention is more “later.” Said differently, if a patient does not accept a recommendation for screening colonoscopy this year, the same recommendation will likely be offered next year. This may be contributing to performance as the priority for this screening may take lower priority than others (Coronado, Petrik, Bartelmann, Coyner, & Coury, 2015).

The preventable deaths and possible savings associated to better screening and detection are compelling. As less invasive, lower cost and higher sensitivity test options become available, e.g., Cologuard, the incentives for ACOs and HMOs, physicians and patients might align better to increase colon cancer screening rates. Other forms of cancer screening are easier to accomplish, e.g., mammography, and additional research may find better success on other screening measures in both structures.

The findings on ER visits per 1,000 are in keeping with prior research identifying the ability of HMOs to manage utilization. When primary care practitioners take financial risk, preventing avoidable ER visits results in financial benefit. In the ACO program, as the proportion of evaluation and management services rendered by primary care physicians increases, so does the ER utilization (Herrel, Ayanian, Hawken, & Miller, 2017). Medicare Advantage plans have lower ER utilization than fee-for-service Medicare (B. E. Landon et al., 2012). This dynamic is a result of the financial incentives created through these programs.

Population model. Dr. Atul Gawande, a well-known physician and author, opined as to the benefits of scale, i.e., structure, with regard to process, saying, “As that [moving more risk to providers] emerges…it seems to me that bigger organizations are in a better position to take on that risk” (Berenson, 2014). Dr. Gawande’s point directly address the idea of spreading the fewer numbers of high risk patients across a larger population of low risk patients. As more value-based arrangements come into existence, consolidation will continue to occur particularly for provider organizations. This is already occurring with pace as it seems VBP is in many ways a euphemism for shifting financial risk from insurers to providers (Lineen, 2014).

The results of population model may also indicate the leveraging of structure in replicating processes at scale. Dr. Gawande referred to this as “big medicine” (Gawande, 2012). As the size of the beneficiary population increases, the structure to support it may provide greater opportunity to reduce variation through the use of informatics and systems with a heightened focus on execution (Berenson, 2014). The medical expense budgets increase with the size of the population, meaning more dollars are at-risk when the ACO or HMO is larger. The relationship between mass and risk relative to quality is an important area for future research.

Control variables. The control variables held great importance in both models, yielding consistent results across models for all three DVs. This indicates a high degree of stability and strengthens the interpretation of these results. More not-White beneficiaries and higher average HCC risk scores meant poorer performance on the quality measures. Also in keeping with the literature, the findings on duration of operations indicates the longer the plans exist the better their performance on these quality measures.

Limitations and Delimitations

The primary limitation of this research is publicly available data are designed for administrative purposes, not research. These data are highly diverse in taxonomy, format and content. This diversity may be preventing robust analysis of this kind. This is an issue going all these programs would make the analysis of results much easier.

The purpose of this study did not include an assessment of costs. Drawing inferences about cost based on the ER utilization measure is reasonable but far from certain or definitive. Incorporating cost into future research would be valuable. The specific incentives created through the structures embodied by HMOs and ACOs are not part of this effort. Medicare PPO plans were not included. Analysis at the sub-population would be an important future area of research but was not part of this study.

Conclusion

The results of this research suggest logical steps for future research. A comparison of outcomes over time at the beneficiary level between a small group of HMOs and ACOs would be meaningful. Exploration of whether favorable selection is occurring in ACOs, and perhaps occurring less in HMOs, would likewise be informative. Replicating these results with calendar year 2016 data would add value to our understanding of the quality of care in these programs. In any of these approaches, more refined risk adjustment would increase applicability. A study further exploring the economic incentives created by Medicare, actualized by the plans and realized by the providers would also be an important next step. Further research into the relationship between the mass of the plan or organization, in terms of covered lives, and quality, cost or access results would add to our understanding of VBP programs and their effectiveness.

This research fills an important role in developing a better understanding of the quality, measured by HEDIS or NQF, of the ACO and HMO programs. The relevance of this research is in demonstrating that differences do exist between ACOs and HMOs with regard to performance on quality measures. The results by size of the population served are also relevant in that they may be added confirmation of a common approach in VBP driving provider consolidations and mergers.

Healthcare financing models create incentives through the payment and quality programs they design. Healthcare provider systems develop structures to maximize their incentive opportunities relative to payment or quality. These organizational structures create incentives to elicit provider-level responses driving positive outcomes for the organization. The interaction of these incentives and responses is key in realizing positive health and economic outcomes for patients, providers and healthcare systems.

References

Abraham, J. M., Marmor, S., Knutson, D., Zeglin, J., & Virnig, B. (2012). Variation in diabetes care quality among medicare advantage plans. *American Journal of Medical Quality, 27*(5), 377-382. 10.1177/1062860611428529

American Cancer Society. (2017). *Colorectal cancer facts & figures 2017-2019.* ().American Cancer Society.

Andersen, R. M. (1995). Revisiting the behavioral model and access to medical care: Does it matter? *Journal of Health and Social Behavior, 36*(1), 1-10. Retrieved from http://www.jstor.org/stable/2137284

Ayanian, J. Z., & Markel, H. (2016). Donabedian's lasting framework for health care quality. *The New England Journal of Medicine, 375*(3), 205-207. 10.1056/NEJMp1605101

Berenson, R. A. (2014). Is bigger better? the implications of health care provider consolidation.

Biles, B., Casillas, G., & Guterman, S. (2016). Does medicare advantage cost less than traditional medicare? *Issue Brief (Commonwealth Fund), 2*, 1-9.

Blackstone, E. A., & Fuhr, J. P. (2016). The economics of medicare accountable care organizations. *American Health & Drug Benefits, 9*(1), 11.

Bolch, M. (2013). PCPs central to star ratings. Retrieved from http://managedhealthcareexecutive.modernmedicine.com/managed-healthcare-executive/news/managed-healthcare-executive/provider-networks/pcps-central-star-r

Brennan, N., & Shepard, M. (2010). Comparing quality of care in the medicare program. *The American Journal of Managed Care, 16*(11), 841.

Centers for Disease Control and Prevention. (2010). *National ambulatory medical care survey*

Centers for Disease Control and Prevention. (2014). Diabetes public health resource. Retrieved from http://www.cdc.gov/diabetes/statistics/a1c/a1c_dist.htm

Centers for Medicare & Medicaid Services (CMS), HHS. (2014). Fact sheets: medicare ACOs continue to succeed in improving care, lowering cost growth. Retrieved from https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2014-Fact-sheets-items/2014-09-16.html

Centers for Medicare & Medicaid Services (CMS), HHS. (2015a). Accountable care organizations (ACO). Retrieved from https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/ACO/index.html?redirect=/aco

Centers for Medicare & Medicaid Services (CMS), HHS. (2015b). Medicare ACOs provide improved care while slowing cost growth in 2014. Retrieved from https://www.cms.gov/newsroom/mediareleasedatabase/fact-sheets/2015-fact-sheets-items/2015-08-25.html

Centers for Medicare & Medicaid Services (CMS), HHS. (2016a). March 31, 2016, hhs-operated risk adjustment methodology meeting discussion paper. Retrieved from https://www.cms.gov/CCIIO/Resources/Forms-Reports-and-Other-Resources/Downloads/RA-March-31-White-Paper-032416.pdf

Centers for Medicare & Medicaid Services (CMS), HHS. (2016b). Medicare accountable care organizations 2015 performance year quality and financial results. Retrieved from https://www.cms.gov/newsroom/mediareleasedatabase/fact-sheets/2016-fact-sheets-items/2016-08-25.html

Centers for Medicare & Medicaid Services (CMS), HHS. (2017a). Fast facts: All medicare shared savings program (shared savings program) accountable care organizations (ACOs). Retrieved from https://www.cms.gov/Medicare/Medicare-Fee-for-Service-Payment/sharedsavingsprogram/Downloads/All-Starts-MSSP-ACO.pdf

Centers for Medicare & Medicaid Services (CMS), HHS. (2017b). Fast facts: All medicare shared savings program accountable care organizations. Retrieved from https://www.cms.gov/medicare/medicare-fee-for-service-payment/sharedsavingsprogram/downloads/all-starts-mssp-aco.pdf

Centers for Medicare & Medicaid Services (CMS), HHS. (August, 2016). NHE fact sheet. Retrieved from https://www.cms.gov/research-statistics-data-and-systems/statistics-trends-and-reports/nationalhealthexpenddata/nhe-fact-sheet.html

Cohn, J. A., Wang, C. E., Lakeman, J. C., Silverstein, J. C., Brendler, C. B., Novakovic, K. R., . . . Helfand, B. T. (2014). Primary care physician PSA screening practices before and after the final U.S. preventive services task force recommendation. *Urologic Oncology: Seminars and Original Investigations, 32*(1), 41.e23-41.e30. 10.1016/j.urolonc.2013.04.013

Coronado, G. D., Petrik, A. F., Bartelmann, S. E., Coyner, L. A., & Coury, J. (2015). Health policy to promote colorectal cancer screening: Improving access and aligning federal and state incentives. *Clinical Researcher (Alexandria, Va.), 29*(1), 50-55. 10.14524/CR-14-0044

Davidson, M. (1997). Diabetes care in health maintenance organisation and fee-for- service settings. *Disease Management & Health Outcomes, 2*(4), 189-197. 10.2165/00115677-199702040-00003 45 cfr 46, (2009).

Dickson, V. (2016). Trustees' report says medicare will be insolvent by 2028. Retrieved from http://www.modernhealthcare.com/article/20160622/NEWS/160629973

Donabedian, A. (1966). Evaluating the quality of medical care. *The Milbank Memorial Fund Quarterly, 44*(3), Suppl:166.

Donabedian, A. (1988). The quality of care. how can it be assessed? *JAMA : The Journal of the American Medical Association, 260*(12), 1743.

Donabedian, A. (2003). *An introduction to quality assurance in health care*. Oxford ; New York: Oxford ; New York : Oxford University Press.

Donabedian, R. C., A., Wheeler, R. C., J., & Wyszewianski, R. C., L. (1982). Quality, cost, and health: An integrative model. *Medical Care, 20*(10), 975-992.

Fuchs, V. R. (2010). Health care is different-- that's why expenditures matter. *Jama, 303*(18), 1859. 10.1001/jama.2010.593

Fung, V., Brand, R. J., Newhouse, J. P., & Hsu, J. (2011). Using medicare data for comparative effectiveness research: Opportunities and challenges. *The American Journal of Managed Care, 17*(7), 488.

Gawande, A. (2012). Big med. *The New Yorker, 20*, 53-63.

Giugliano, D., Maiorino, M., Bellastella, G., Petrizzo, M., Ceriello, A., Genovese, S., & Esposito, K. (2015). Setting the hemoglobin A1c target in type 2 diabetes: A priori, a posteriori, or neither? *Endocrine; International Journal of Basic and Clinical Endocrinology, 50*(1), 56-60. 10.1007/s12020-015-0549-2

Gold, M. (2009). Medicare's private plans: A report card on medicare advantage. *Health Affairs (Project Hope), 28*(1), w41. 10.1377/hlthaff.28.1.w41

Hae-Young Kim. (2013). Statistical notes for clinical researchers: Assessing normal distribution (2) using skewness and kurtosis. *Restorative Dentistry & Endodontics, 38*(1), 52-54. 10.5395/rde.2013.38.1.52

Hashimoto, R., Brodt, E., Skelly, A., & Dettori, J. (2014). Administrative database studies: Goldmine or goose chase?*05*(02), 074-076. 10.1055/s-0034-1390027

Hawley, S., Lillie, S., Cooper, G., & Elston Lafata, J. (2014). Managed care patients' preferences, physician recommendations, and colon cancer screening. *The American Journal of Managed Care, 20*(7), 555.

Hellinger, F. J. (1998). The effect of managed care on quality: A review of recent evidence. *Archives of Internal Medicine, 158*(8), 833.

Herman, B. (2015). *Medicare advantage star ratings to be released*. Chicago:

Herrel, L., Ayanian, J., Hawken, S., & Miller, D. (2017). Primary care focus and utilization in the medicare shared savings program accountable care organizations. *BMC Health Services Research, 17*10.1186/s12913-017-2092-8

Himmelstein, D. U., Woolhandler, S., Hellander, I., & Wolfe, S. M. (1999). Quality of care in investor-owned vs not-for-profit HMOs. *Jama, 282*(2), 159.

Hoffman, B. (2008). Health care reform and social movements in the united states. *American Journal of Public Health, 98*, S69-S79. 10.2105/AJPH.98.Supplement_1.S69

Honigman, L. S., Wiler, J. L., Rooks, S., & Ginde, A. A. (2013). National study of non- urgent emergency department visits and associated resource utilization. *The Western Journal of Emergency Medicine, 14*(6), 609. 10.5811/westjem.2013.5.16112

Hwang, W., Chang, J., Laclair, M., & Paz, H. (2013). Effects of integrated delivery system on cost and quality. *The American Journal of Managed Care, 19*(5), e175.

Institute of Medicine. (2000). *To err is human : Building a safer health system*. Washington, D.C.: Washington, D.C. : National Academy Press.

Jacobsen, G., Casillas, G., Damico, A., Neuman, T. & Gold, M. (2016). Medicare advantage 2016 spotlight: Enrollment market update. Retrieved from http://kff.org/medicare/issue-brief/medicare-advantage-2016-spotlight-enrollment-market-update/

Jacobsen, G., Trilling, A., Neuman, T., Damico, A. & Gold, M. (2016). Medicare advantage hospital networks: how much do they vary? Retrieved from http://kff.org/report-section/medicare-advantage-hospital-networks-how-much-do-they-vary-results/

Kaiser Family Foundation. (2015). Medicare advantage 2016 data spotlight: Overview of plan changes, quality ratings. Retrieved from http://kff.org/report-section/medicare-advantage-2016-data-spotlight-overview-of-plan-changes-quality-ratings/

Kaiser Family Foundation. (May, 2016). Medicare advantage. Retrieved from http://kff.org/medicare/fact-sheet/medicare-advantage/

Landon, B. E., Zaslavsky, A. M., Saunders, R. C., Pawlson, L. G., Newhouse, J. P., & Ayanian, J. Z. (2012). Utilization of services in medicare advantage versus traditional medicare since the passage of the medicare modernization act. *Health Affairs (Project Hope), 31*(12), 2609-2617. 10.1377/hlthaff.2012.0179

Landon, B., & Epstein, A. (2001). For- profit and not-for- profit health plans participating in medicaid. *Health Affairs, 20*(3), 162-71.

Levine, D. M., Linder, J. A., & Landon, B. E. (2016). THe quality of outpatient care delivered to adults in the united states, 2002 to 2013. *JAMA Internal Medicine,* Retrieved from http://dx.doi.org/10.1001/jamainternmed.2016.6217

Lineen, J. (2014). Hospital consolidation: " Safety in numbers" strategy prevails in preparation for a value- based marketplace. *Journal of Healthcare Management / American College of Healthcare Executives, 59*(5), 315.

McBean, A. M., Jung, K., & Virnig, B. A. (2005). Improved care and outcomes among elderly medicare managed care beneficiaries with diabetes. *The American Journal of Managed Care, 11*(4), 213-222. 2827 [pii]

McWilliams, J. M., Hsu, J., & Newhouse, J. P. (2012). New risk-adjustment system was associated with reduced favorable selection in medicare advantage. *Health Affairs (Project Hope), 31*(12), 2630-2640. 10.1377/hlthaff.2011.1344 [doi]

Mello, M. M., Stearns, S. C., Norton, E. C., & Ricketts, T. C. (2003). Understanding biased selection in medicare HMOs. *Health Services Research, 38*(3), 961-992. 10.1111/1475-6773.00156

Miller, R. H., & Luft, H. S. (1997). Does managed care lead to better or worse quality of care? *Health Affairs (Project Hope), 16*(5), 7.

Muhlestein, D., Croshaw, A., Merrill, T., Pena, C., & James, B. (2013). *The accountable care paradigm: More than just managed care 2.0* Center For Accountable Care Intelligence.

National Colorectal Cancer Roundtable. (2007). *Increasing colorectal cancer screening - saving lives and saving dollars: Screening 50-60 year olds reduces cancer costs to medicare*

Neuman, T., & Jacobsen, G. (2014). Medicare advantage: Take another look. Retrieved from http://kff.org/medicare/perspective/medicare-advantage-take-another-look/

Newhouse, J. P., Price, M., McWilliams, J. M., Hsu, J., & McGuire, T. (2014). How much favorable selection is left in medicare advantage? *National Bureau of Economic Research Working Paper Series, 20021*(March 2014)http://www.nber.org/papers/w20021

Newhouse, J. P., Mcwilliams, J. M., Price, M., Huang, J., Fireman, B., & Hsu, J. (2013). Do medicare advantage plans select enrollees in higher margin clinical categories? *Journal of Health Economics, 32*(6), 1278-1288. 10.1016/j.jhealeco.2013.09.003

Patient protection and affordable care act, 42 U.S.C. § 18001, (2010).

Rappleye, E. (2015, January 28, 2015). 20 major health systems, payers pledge to convert 75% of business to value-based arrangements by 2020. *Becker's Hospital CFO,*

Reid, R. O., Deb, P., Howell, B. L., & Shrank, W. H. (2013). Association between medicare advantage plan star ratings and enrollment. *Jama, 309*(3), 267. 10.1001/jama.2012.173925

Retchin, S. M., Clement, D. G., Rossiter, L. F., Brown, B., Brown, R., & Nelson, L. (1992). How the elderly fare in HMOs: Outcomes from the medicare competition demonstrations. *Health Services Research, 27*(5), 651.

RTI International. (2015). Web interface sampling methodology for the physician quality reporting system, the medicare shared savings program, and the pioneer aco model. Retrieved from https://www.cms.gov/Medicare/Quality-Initiatives-Patient-Assessment-Instruments/PQRS/Downloads/2015_WebInterfaceSamplingDoc.pdf

Shuval, K., Leonard, T., Drope, J., Katz, D. L., Patel, A. V., Maitin-Shepard, M., . . . Grinstein, A. (2017). *Physical activity counseling in primary care: Insights from public health and behavioral economics*10.3322/caac.21394

Tabachnick, B. G.,1936- author. (2013). *Using multivariate statistics* (6th ed.. ed.). Boston: Boston : Pearson Education.

Tanenbaum, S. (2016). What is the value of value- based purchasing? *Journal of Health Politics, Policy and Law, 41*(5), 1033. 10.1215/03616878-3632254

Trivedi, A. N., & Grebla, C. R. (2011). Quality and equity of care in the veterans affairs health-care system and in medicare advantage health plans. *Medical Care, 49*(6), 560-568. 10.1097/MLR.0b013e31820fb0f6

Ware, J. E., Bayliss, M. S., Rogers, M. H., Kosinki, M., & Tarlov, A. R. (1996). Differences in 4-year health outcomes for elderly and poor, chronically III patients treated in HMO and fee-for- service systems: Results from the medical outcomes study. *Jama, 276*, 1039. 10.1016/S1062-1458(97)83314-4

Wilson, G., & Perry, T. (2009). Is tight glycemic control in type 2 diabetes really worthwhile? no. *Canadian Family Physician Médecin De Famille Canadien, 55*(6), 581, 583, 585, 587, 588.

Xu, P., Burgess, J. F., Cabral, H., Soria-Saucedo, R., & Kazis, L. E. (2015). Relationships between medicare advantage contract characteristics and quality-of-care ratings: An observational analysis of medicare advantage star ratings. *Annals of Internal Medicine, 162*(5), 353. 10.7326/M14-0332

Vita

W. Wesley Campbell, III, was born on September 21, 1965, in Macon, Georgia and is an American citizen. He graduated from Clover Hill High School in 1983. He received an Associate of Applied Science in Legal Assisting from J. Sargeant Reynolds Community College in 1993. He received a Bachelor of Science in Business Administration from the University of Phoenix in 1996, and a Master of Health Services Administration degree from Virginia Commonwealth University in 1998. Wes is the Chief of Clinical Operations for the University of Virginia Physicians Group, a position he has held since 2010, and a Fellow in the American College of Healthcare Executives.