Value-Based Contracting Innovated Medicare Advantage Healthcare Delivery and Improved Survival

Aloke K. Mandal, MD, PhD; Gene K. Tagomori, BSc; Randell V. Felix, BSc; and Scott C. Howell, DO, MPH&TM

ABSTRACT

OBJECTIVES: In Medicare Advantage (MA) with its CMS Hierarchical Condition Categories (CMS-HCC) payment model, CMS reimburses private plans (Medicare Advantage Organizations [MAOs]) with prospective, monthly, health-based or risk-adjusted, capitated payments. The effect of this payment methodology on healthcare delivery remains debatable. How value-based contracting generates cost efficiencies and improves clinical outcomes in MA is studied.

STUDY DESIGN: A difference in contracting arrangements between an MAO and 2 provider groups facilitated an intervention-control, preintervention–postintervention, difference-in-differences approach among statistically similar, elderly, community-dwelling MA enrollees within one metropolitan statistical area.

METHODS: Starting in 2009, for intervention-group MA enrollees, the MAO and a provider group agreed to full-risk capitation combined with a revenue gainshare. The gainshare was based on increases in the Risk Adjustment Factor (RAF), which modified the CMS-HCC payments. For the control group, the MAO continued to reimburse another provider group through fee-for-service. RAF, utilization, and survival were followed until December 31, 2012.

RESULTS: The intervention group's mean RAF increased significantly (P <.001), estimating $2,519,544 per 1000 members of additional revenue. The intervention increased office-based visits (P <.001). Emergency department visits (P <.001) and inpatient hospital admissions (P = .002) decreased. This change in utilization saved $2,071,293 per 1000 enrollees. By intensifying office-based care for these MA enrollees with multiple comorbidities, a 6% survival benefit with a 32.8% lower hazard of death (P <.001) was achieved.

CONCLUSIONS: Value-based contracting can drive utilization patterns and improve clinical outcomes among chronically ill, elderly MA members.

Am J Manag Care. 2017;23(2):e41-e49

In caring for traditional Medicare beneficiaries, primary care physicians will need to transform their clinical practice and assume more fiduciary risk. CMS recently published "new policies to address and incentivize participation in alternative payment models" (APMs).[1] An APM broadly can be defined as any reimbursement model other than strict fee-for-service (FFS). Currently, 30% of traditional FFS Medicare is reimbursed through APMs, with the goal of 50% involvement by 2018.[2,3] Clinicians who want to become qualifying APM participants can expect to "bear more than a nominal amount of risk for monetary losses."[1] Through APMs and increasing risk assumption, these new policies aspire to promulgate high-value healthcare, as defined by "better care, smarter spending, and healthier people."[1,2] Whether FFS divestiture in favor of APMs and increasing capitation can generate cost efficiencies and also improve clinical outcomes remains debatable.[4,5]

Medicare Advantage (MA) provides an alternative to traditional FFS Medicare. It has been a commercial success, accounting for 17.5 million (30.6%) of all Medicare enrollees and $204.7 billion (28.9%) of Medicare's 2017 gross spending budget.[6,7] Because it is regulated by its own federal statutes,[8] MA is classified as an "Other Payer APM" and excluded from CMS' Proposed Rule for Medicare FFS.[1] For over a decade, MA has used the CMS Hierarchical Condition Categories (CMS-HCC) payment model to reimburse private plans (Medicare Advantage Organizations [MAOs]) with prospective, monthly, risk-adjusted or health-based capitated payments for the care of MA enrollees. The value of subsidizing MA often has been challenged.[9-14] Consequently, on October 3, 2016, CMS' Innovation Center (CMMI) announced its Medicare Advantage Value-Based Insurance Design (VBID) model to test whether new initiatives can "improve health outcomes and lower expenditures for Medicare Advantage enrollees."[15]

CMS adopted the CMS-HCC payment model with the concept that MAO recompense should reflect the disease and related cost burdens of the pertinent population and, thus, fundamentally changed how MAOs are reimbursed.[16-19] In return for providing healthcare benefits to MA enrollees during the calendar year (CY),

TAKE-AWAY POINTS

This study tested the hypothesis that payer-provider risk contracting promotes high-value care.

- In the future, clinicians increasingly will have to bear the monetary risks associated with healthcare utilization.
- The Medicare Advantage program provides a unique milieu for investigating provider groups that have either risk-bearing or fee-for-service contracts with private health plans.
- Full-risk capitation combined with a revenue gainshare agreement promulgated a clinical practice transformation at the provider group level, associated with increased office-based care and decreased hospital-based services.
- The clinical practice transformation resulted in a 6% survival benefit and lowered the hazard of death by 32.8%.
- Value-based contracting benefits all stakeholders of the Medicare Advantage program.

MAOs receive risk-adjusted payments during the following payment year (PY), modified by a Risk Adjustment Factor (RAF). Along with demographic data, specific diagnosis codes grouped into HCCs and associated with increased future expenditures impact the RAF. Certain HCC-HCC and demographic-HCC interactions have further additive effects on the RAF. Recalibration of the model occurs every 2 years so that the "typical" FFS Medicare beneficiary's total RAF is 1.00.[20] Therefore, when compared with an FFS Medicare beneficiary, a healthier MA enrollee's total RAF would be less than 1.00 and a sicker one's would be greater than 1.00.

The capitated revenue for MAOs is dependent on how its contracted providers document and code. For example, based on calculations derived from CMS' most recent policy update and data on both MA enrollment and risk-adjusted benchmark rates,[6,21,22] each 0.1 of the RAF has a nationally averaged valuation of $74.85 per member per month (PMPM), or $898.20 per member per year (PMPY). If providers document and code for a specific diabetic complication in CY2016, an MAO could anticipate a risk-adjusted revenue in 2017 (PY2016) that is 3-fold greater ($238.02 PMPM or $2856.28 PMPY) than if diabetes were coded without any specified complication ($77.84 PMPM or $934.13 PMPY). Assuming diagnostic coding accuracy, an MAO would be reimbursed at a higher rate during the PY for enrolling sicker MA members during the CY.

The CMS-HCC payment model thus should provide incentives to MAOs that reward continuous, high-value care from its contracted providers.[19,20] Of all the different types of MA plans available, coordinated care plans are the majority, of which health maintenance organizations (HMOs) account for 52% of all MA plans.[6] In applying for an HMO-type MA plan, the MAO must include "a network of providers that are under contract or arrangement with the organization to deliver the benefit package approved by CMS."[8] How MAOs reimburse providers has gone without much input from CMS, and most providers prefer FFS reimbursement.[8,23,24] For example, in Oregon, the HMO-type MA plan denoted by CMS contract #H3805 has capitated agreements with only a quarter of its contracted providers, and the rest are reimbursed on an FFS schedule (**eAppendix** [eAppendices available at **www.ajmc.com**]). If the MAO's capitated reimbursement from CMS is greater than its FFS disbursement to providers, there might be no incentive to encourage high-value care.[9,10,13,14]

The primary aim of this study tested the following hypothesis: value-based contracts augment the CMS-HCC payment model's ability to generate cost efficiencies and improve clinical outcomes. Because local provider groups remain the core business unit for both MAO revenue generation and overall cost-of-care management, a secondary aim was to delineate the specific clinical practice transformations implemented by a provider group pursuant to such contracting changes.

METHODS

Study Population, Study Design, and Provider Groups

The study population consisted of community-dwelling MA members 65 years or older, enrolled in 1 HMO-type MA plan (CMS contract #H3805),[6] and ascribed to 1 of 2 provider groups for their primary care in the Portland, Oregon, metropolitan statistical area (MSA). These members had to be enrolled entirely through CY2008, which was the preintervention period (eAppendix). The intervention occurred in CY2009 when the MAO's contract with Provider Group A differed from that with Provider Group B (**Table 1**). Effective CY2009, Provider Group A became financially responsible for most healthcare services (full-risk capitation) for intervention-group members. In turn, starting in 2010 (PY2009), the provider group would receive most of the CMS-HCC risk-adjusted monthly capitated payment (RAF gainshare). In response, Provider Group A appointed an "HCC Physician Champion" (HPC) in CY2009. The HPC's mission was to improve HCC documentation and coding, as well as to develop cost-effective approaches to primary care delivery. The contracting change and HPC assignation were considered a single intervention—all initiated in CY2009. In the control group, the MAO continued FFS reimbursement for Provider Group B. The postintervention period began on January 1, 2009, and ended on December 31, 2012, after which a revised CMS-HCC model (V22) was introduced and would have produced nonuniform changes in the RAF.[25,26]

Based on the timing of this intervention, an intervention–control, preintervention–postintervention, difference-in-differences (DID) approach among matching cohorts sought to determine the intervention's effect on the RAF, utilization, and survival. This approach within 1 MSA inherently controlled for variations in care because of geography, time of enrollment, MAO plan administration, provider groups, and enrollee characteristics.[27-29]

TABLE 1. Comparison of MA Enrollees, According to Matching Group

Characteristics	Intervention Group (n = 1230)	Control Group (n = 1230)
MAO and provider details		
MA contract number (type)	H3805 (HMO)	H3805 (HMO)
Provider group	Provider Group A	Provider Group B
RAF gainshare arrangement	Yes	No
Full-risk capitation arrangement	Yes	No
HCC physician champion	Yes	No
Mean age ± SEM (range)	80.18 ± 0.20 (65.00-102.50)	80.05 ± 0.19 (65.00-99.4)
Female (%)	763 (62.0)	758 (61.6)
Race (%)		
White	1177 (95.7)	1195 (97.2)
Asian	22 (1.8)	5 (0.4)
Black	7 (0.6)	7 (0.6)
Hispanic	3 (0.2)	3 (0.2)
Native American	2 (0.2)	1 (0.1)
Unknown or other	19 (1.5)	19 (1.5)
Original reason for Medicare entitlement (%)		
Age	1191 (96.8)	1192 (96.9)
Disability	39 (3.2)	38 (3.1)
End-stage renal disease	0 (0)	0 (0)
Disease burden		
Mean Charlson Comorbidity Index ± SEM	3.21 ± 0.07	3.16 ± 0.07
Mean number of chronic conditions ± SEM	3.84 ± 0.07	3.92 ± 0.07

HCC indicates Hierarchical Conditions Category; HMO, health maintenance organization; MA, Medicare Advantage; MAO, Medicare Advantage Organization; RAF, Risk Adjustment Factor; SEM, standard error of the mean.

A smaller, well-controlled study at the provider-group level with greater internal validity was chosen over one that compiled and summarized large datasets, which, although may provide greater external validity, would obfuscate exploring the specific clinical practice transformations implemented at the provider-group level. Provider Group A had 7 clinic locations, and Provider Group B had 5 locations, all within 2 to 18 miles from one another. Provider Group A had 25 primary care specialists (16 in internal medicine, 9 in family medicine) and Provider Group B had 19 (5 in internal medicine and 14 in family medicine). By the end of the study period, both provider groups were Oregon Health Authority–certified Patient-Centered Primary Care Homes (PCPCHs). Six of Provider Group A's 7 clinic locations were PCPCH-accredited in December 2012, with the last one certified in 2014. All of Provider Group B's clinic locations were accredited slightly earlier, in June 2012.

Propensity Score Model and Nearest-Neighbor Matching

Plan enrollment and provider-group election are by enrollee choice and subject to selection bias. Covariates known to affect healthcare utilization and expenditure include age, sex, ethnicity, original reason for Medicare entitlement, and disease burden.[30] By not applying the appropriate preprocessing methodologies to ensure comparison of statistically similar groups,[31] previous conclusions on MA healthcare delivery vis-à-vis traditional FFS Medicare may have been compromised.[27,28,32-35] For example, past studies relied on self-reported survey data for health status and were subject to recall bias.[27,28,32,34,35] This study's proprietary access to full encounter claims data for MA enrollees facilitated 2 objective measures of disease burden: the Charlson Comorbidity Index (CCI) and the number of CMS Chronic Conditions Warehouse (CCW) categories (eAppendix). Using the aforementioned covariates, the logistic regression model created propensity scores for each member.[31] Nearest-neighbor matching based on propensity scores created 2 well-balanced, statistically similar groups for subsequent analysis (**eAppendix Table 1** and **eAppendix Figure 2**), as delineated in **Table 1**.

Statistical and Economic Analyses

Primary outcomes included: 1) RAF data at the member level obtained from CMS' Monthly Membership Report, 2) utilization based on full-encounter claims data, and 3) survival data obtained from CMS' Daily Transactional Reply Report (eAppendix). A linear regression model evaluated the RAF as a continuous variable. Poisson regression models examined utilization as count data. DID analyses compared postintervention (CY2009-CY2012) with preintervention (CY2008) data. For survival, the log-rank test—using either time-on-study or age of the enrollees on the time scale—evaluated Kaplan-Meier survival curves for the 2 groups.[36,37] Permutation testing (randomization inference) validated the DID and survival analyses (eAppendix).[38] Statistical analyses were performed within the R statistical computing environment (R Foundation for Statistical Computing, Vienna, Austria).

RAF-based revenue was estimated using CMS' published rates for Multnomah County for each year (eAppendix).[22] FFS Medicare expenditures extracted from the Medical Expenditure Panel Survey data files provided cost-of-care data according to place of service (eAppendix). Since the last CMS-HCC-derived capitated payment was disbursed in 2013 (PY2012), all amounts were indexed to 2013 dollar values using the appropriate component of the Personal Healthcare Expenditure Index (eAppendix).

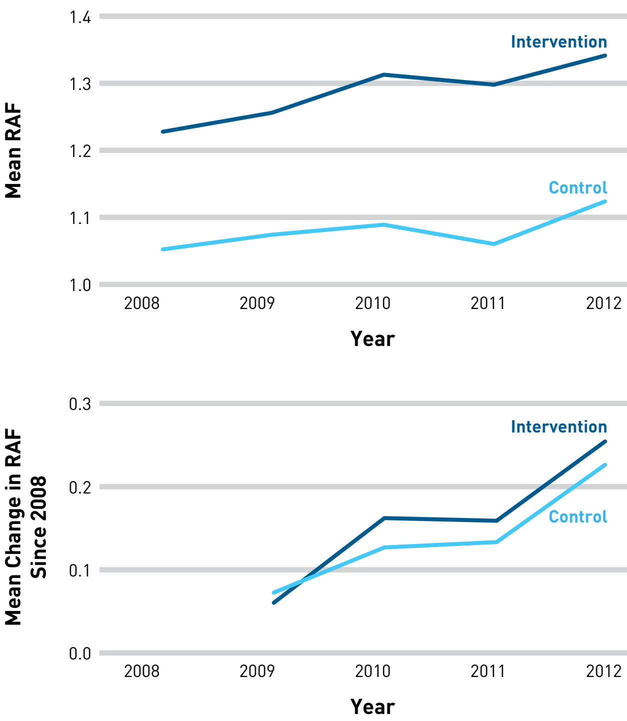
RESULTS

Effect on RAF

The HPC's sequential endeavors included: 1) in the beginning of 2009, focusing on coding specificity for diabetes and its complications; 2) toward the end of 2009, auditing all charts for disparities between clinical documentation and HCC coding; and 3) in the beginning of 2010, 1-on-1 mentoring of colleagues whose panel of intervention-group members had a mean RAF in the lower 50th percentile. **Figure 1** summarizes the mean RAF for both groups over the 5-year study period (top panel) and the changes from preintervention (CY2008) values (bottom panel).

Based on the DID analysis, in 2009, the intervention group increased the RAF by 4.1% compared with the control group ($P = .031$). In other words, a control-group member whose RAF was 1.000 in 2009 would have increased its documented RAF to 1.041 if in the intervention group. Subsequent yearly analysis demonstrated the following multiplicative increases for the intervention group, relative to control: 1.072 from 2009 to 2010 ($P < .001$); 1.071 from 2010 to 2011 ($P = .002$); and 1.069 from 2011 to 2012 ($P = .010$). Randomization inference confirmed that the intervention group increased the RAF at a greater rate than control ($P < .001$). Over the postintervention period, the increase in the RAF generated an additional $2,519,544 per 1000 intervention-group members (**eAppendix Table 2**), averaging $629.89 PMPY.

FIGURE 1. Changes in Risk Adjustment Factor During the Study Period[a,b]



RAF indicates Risk Adjustment Factor.

[a]Mean RAF values were obtained from CMS's final Monthly Membership Report for each corresponding year (top panel).

[b]The preintervention period was January 1, 2008, to December 31, 2008. The intervention began on January 1, 2009, with the postintervention period ending on December 31, 2012. The change from 2008 mean total RAF is also provided (bottom panel). Based on the difference-in-differences linear regression analysis, the intervention significantly increased mean RAF in the intervention group, relative to the control group. By exponentiating the coefficient for the intervention group, the average yearly risk ratio for increasing RAF was 1.061 ($P < .001$) compared with the control group. In other words, the intervention group averaged a 6.1% increase in RAF during each of the postintervention years compared with the control group.

Effect on Utilization

During the preintervention period, the intervention group had lower office-based utilization (6713 visits/1000 members) than the control group (7220 visits/1000 members). In 2009, the intervention group increased its office visits (7642/1000 members), comparable to the control group (7637/1000 members). Subsequently, relative to the control group and adjusted to 1000 members, the intervention group had 346 more office-based visits in 2010 ($P = .016$), 417 more in 2011 ($P = .008$), and 415 more in 2012 ($P = .026$). DID analysis yielded an exponentiated coefficient of 1.112 ($P = .005$), and randomization inference confirmed this statistical significance ($P < .001$). If reimbursed on a Medicare FFS schedule, additional expenditures for the increased office-based utilization for the intervention group would have been $258,334 per 1000 members (**eAppendix Table 3**), which Provider Group A assumed because of full-risk capitation. Furthermore, Provider Group A used this optimized RAF data to manage the increased workload. An index analysis validated Provider Group A's risk-stratification process (**Table 2**); by clustering intervention-group members into 3 groups based on 2010 office-based visits,

TABLE 2. Difference in RAF, CCI, and Subsequent Office Visits (2011-2012) Among Intervention-Group Members Grouped by 2010 Office Visits

Intervention-Group Terciles (Number of 2010 Office Visits)	RAF (2010)	CCI (2010)	Office Visits (2011)	Office Visits (2012)
Top tercile (10-60)	1.78 ± 0.06[a]	5.08 ± 0.15[a]	12.29 ± 0.45[a]	11.65 ± 0.46[a]
Middle tercile (5-9)	1.28 ± 0.04[b]	3.77 ± 0.12[b]	7.01 ± 0.24[b]	7.08 ± 0.30[b]
Bottom tercile (0-4)	0.99 ± 0.03	2.74 ± 0.11	3.96 ± 0.18	4.18 ± 0.23

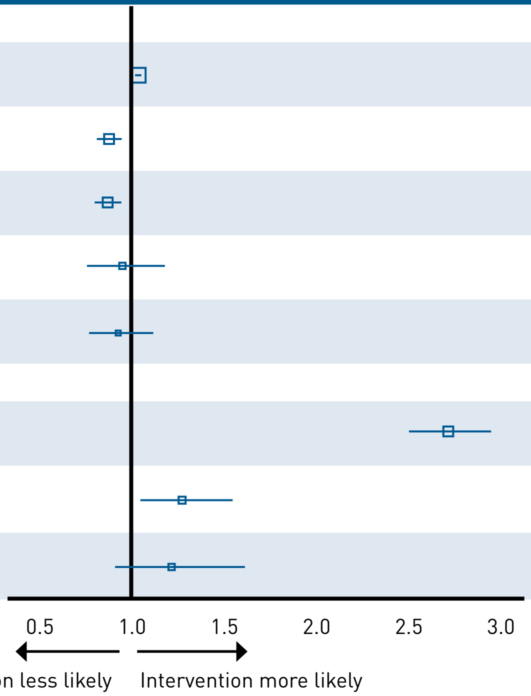
CCI indicates Charlson Comorbidity Index; RAF, Risk Adjustment Factor.

[a]$P < .001$, comparing top tercile with middle tercile, using analysis of variance (ANOVA).

[b]$P < .001$, comparing middle tercile with bottom tercile, using ANOVA.

FIGURE 2. Forest Plot of Place of Service and Preventive Care Utilization for Matching Cohorts of Intervention and Control Group Medicare Advantage Enrollees During the Postintervention Period[a]

Category	Control: Number of events / total (rate)	Intervention: Number of events / total (rate)	Incidence Rate Ratio (95% CI)	*P* Value for Intervention
Place of service utilization				
Office[b]	31,467/50,945 (0.618)	33,573/52,396 (0.641)	1.038 (1.022-1.054)	<.001
ED, hospital[b]	1528/50,945 (0.030)	1399/52,396 (0.027)	0.888 (0.826-0.955)	.001
Inpatient, hospital[b]	1246/50,945 (0.024)	1130/52,396 (0.022)	0.881 (0.813-0.954)	.002
30-day readmission[c]	168/1060 (0.158)	147/979 (0.150)	0.959 (0.769-1.195)	.707
60-day readmission[c]	237/971 (0.244)	204/907 (0.225)	0.933 (0.774-1.124)	.464
Preventive care utilization				
Preventive visits[d]	793/50,945 (0.016)	2223/52,396 (0.042)	2.734 (2.522-2.965)	<.001
Screening mammography[e]	187/373 (0.501)	235/366 (0.642)	1.280 (1.056-1.551)	.012
Screening colonoscopy[f]	85/719 (0.118)	108/748 (0.144)	1.221 (0.919-1.622)	.168



CI indicates confidence interval; ED, emergency department.

[a]A difference-in-differences approach under the framework of Poisson regression could account for not only utilization as count data but also members enrolled for different lengths of time. Incidence rate ratios demonstrated the difference in utilization between the intervention and control groups. The box sizes are proportional to the precision of the estimated outcomes, with larger boxes representing a greater degree of precision. The inverse of the standard error of the estimate determined the precision with which each outcome could be ascertained. The values then were scaled to add up to 100%. Because office utilization had much greater precision, a natural log of the values was used to adjust the box sizes. Error bars represent 95% CIs on a linear scale.

[b]For each place of service utilization category, rates were defined as the number of visits per member-months.

[c]Readmission was defined as hospital inpatient stay within 30 days or 60 days, following the incident hospital inpatient stay. Both the numerators and denominators were censored for death within 30 days or 60 days after the incident admission.

[d]The rate of preventive visits was per member-months.

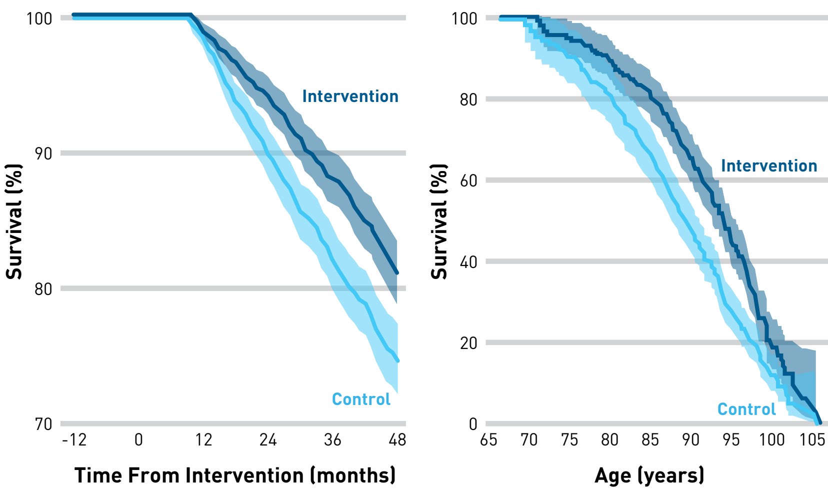
[e]The rate of screening mammography was restricted to women 74 or younger and adjusted to member-years.

[f]The rate of screening colonoscopy was restricted to those 75 years or younger and adjusted to member-years.

statistically different mean RAF, CCI, and subsequent office-based utilization counts were observed for all terciles (Table 2).

Next, the effect of increasing office-based visits on hospital-based services was analyzed. To this end, incident rate ratios (IRRs) for each place of service utilization were normalized to member-months. As with the previous analysis adjusted to 1000 members and as demonstrated in **Figure 2** when adjusted to member-months, intervention-group members were more likely to have increased office-based utilization (*P* <.001). The intervention concomitantly reduced emergency department (ED) visits and inpatient hospital (IP) admissions by 11.2% (*P* = .001) and 11.9% (*P* = .002), respectively (Figure 2). Randomization inference confirmed that the reduction in IP admissions was statistically significant (*P* = .037), whereas the reduction in ED visits approached statistical significance (*P* = .086). These 154 fewer ED visits per 1000 members yielded $100,915 in estimated savings, and 143 fewer IP admissions per 1000 members saved $1,756,869 (eAppendix Table 3). Because the intervention reduced both readmissions (numerator) and death-censored incident admissions (denominator), the actual readmission rate was not statistically different compared with the control group. The intervention also influenced certain preventive care measures. As evidenced in Figure 2, intervention-group members were almost 3 times more likely to undergo preventive care visits (*P* <.001), and intervention-group women 74 years or younger were 28% more likely

FIGURE 3. Kaplan-Meier Curves for Survival for Matching Cohorts of Medicare Advantage Enrollees During the Study Period[a,b]



[a]Shading denotes 95% confidence intervals.

[b]In the left panel, Kaplan-Meier survival curves were constructed with time from the intervention on the time scale. Differences in survival first were observed at 16 months after the intervention and continued throughout the remainder of the study. The right panel comprises Kaplan-Meier curves constructed with age on the time scale. A survival benefit was noticeable for intervention-group members aged 82 to 96 years. Based on the right panel data and a Cox Proportional Hazards model, the intervention-group members had a 32.8% lower risk of dying compared with control-group members.

to undergo screening mammography (*P* = .012). Randomization inference confirmed these findings for preventive care visits (*P* <.001) and screening mammography (*P* = .005). Although the intervention group had higher rates of screening colonoscopy (for those 75 years or younger), the difference was not statistically significant.

Effect on Survival

The intervention group's overall survival rate was 82%, and the control group's was 76%. This 6% survival benefit first became apparent at 16 months after the intervention (**Figure 3**, left panel), coinciding with the first year that the intervention group had higher office-based utilization than the control group. Age provided a natural time-scale for calculating the hazard of death for this elderly population with multiple comorbidities and a higher risk of all-cause mortality.[36,37] Intervention-group members had a 32.8% lower hazard of dying (*P* <.001). The survival benefit was more apparent among those aged 82 to 96 years (Figure 3, right panel). Randomization inference confirmed these survival data, whether time (*P* <.001) or age (*P* <.001) was the time scale.

DISCUSSION

In this study of MA enrollees, a contracting change between the MAO and a provider group leveraged the core initiatives proposed for reforming the Medicare FFS program, namely APM participation and increasing risk assumption. Key findings included: 1) MAO–provider collaboration optimized the RAF, 2) RAF optimization supported a risk-stratification process in the effective triage of office-based care, 3) intensive office-based care concomitantly reduced hospital-based services, and 4) this shift in healthcare delivery improved survival. The survival benefit first became apparent in CY2010, coinciding with the completion of the HPC's mentoring program and the first year of significantly increased office-based utilization. This sequence of events leads to the assertion that improved survival is related to and attributable to enhanced CMS-HCC data and value-based contracting, which then transform primary care delivery.

In redesigning primary care practices, a CMMI-funded study noted that the focus of most clinicians is to identify "new funding streams for their program activities."[39] The CMS-HCC payment model attempts to account for anticipated future costs through an assessment of disease burden by coding HCCs during face-to-face clinical encounters. HCC coding is an onerous task for physicians, especially with each revision of the CMS-HCC model.[26] For example, in V12 of the CMS-HCC model that was in effect during the study period, 2935 of the 14,567 (20.2%) unique *International Classification of Diseases, Ninth Revision, Clinical Modification* (*ICD-9-CM*) codes were aggregated into 70 HCCs that impact the RAF. In V22 that was phased in CY2013, 8667 of the 69,823 (12.4%) *ICD-10-CM* codes were grouped into 79 HCCs.

During the gradual implementation phase of the CMS-HCC payment model (2004-2007), Provider Group A began advancing its HCC documentation and coding practices, in contrast to most of the MAO's other contracted provider groups (as exemplified by Provider Group B). Although the 2 groups had similar disease burdens based on CCI and CCW categories, the RAF scores always were different because of Provider Group A's eagerness to learn about HCC during CMS-HCC's implementation phase (Figure 1, top panel).

With recognition of this potential new funding stream and the goodwill established from CMS-HCC revenue previously generated for the MAO, Provider Group A effectively argued for an RAF gainshare arrangement. The intervention further increased RAF scores relative to the control group (bottom panel, Figure 1), thus optimizing CMS-HCC's value in assessing the intervention group's disease burden and prospective healthcare expenditures. Such gainshare arrangements also benefit MAOs because federal statutes constrain MA plan profitability since the Medical Loss Ratio must be 85% or greater.[40] By including a full-risk capitation clause to the RAF gainshare, the MAO locked into a fixed rate of return and thereby limited both its upside and downside risk, analogous to the financial instrument of interest rate swaps.

MAOs receive monthly capitated prospective payments from CMS and become financially responsible for Part A and Part B benefits for their MA members.[8] Cost-of-care management agreements between an MAO and its provider networks are not mandatory. Instead, the Code of Federal Regulations meekly states that MA plans "*may* [emphasis added] include mechanisms to control utilization, such as referrals from a gatekeeper for an enrollee to receive services within the plan, and financial arrangements that offer incentives to providers to furnish high-quality and cost-effective care."[8] When assuming fiduciary risk, MAOs reflexively create their own programs for utilization management, care coordination, and disease prevention. Alternatively, risk-contracting with clinicians can reduce the MAO's clinical footprint. By enabling "incentives to providers to furnish high-quality and cost-effective care,"[8] there is a statutory ease of administration for MAOs to risk-contract with providers, based on percentages of the CMS-HCC capitated revenue and member premium. These percentages can range from 30% to over 80%, depending on the magnitude of monetary risk providers agree to accept.

Prior to 2008, many of Provider Group A's other contracts required it to reimburse all external specialists and for the plans to pay all hospital claims. The difference then was split equally between the provider group and the plans (50/50–shared risk). With this past experience, Provider Group A agreed to the challenges of taking on full monetary risk for healthcare expenditures—including facility charges—with only a few exceptions or "carve-outs" (eg, out-of-area, vision, hearing aid, most transplantation services). In contrast, Provider Group B remained on a strict FFS schedule, with no risk assumption of other services.

Despite previous risk-contracting, Provider Group A still had a steep learning curve to potentiate high-value care. Both provider groups were PCPCH-certified, and—similar to previous reports on patient-centered medical homes—this certification did not spontaneously transform clinical practice.[41,42] Because risk-stratification can cultivate novel care coordination protocols and workplace efficiency, CMMI has funded studies on targeting high-risk patients.[43,44] A novel finding of the current study was that, after optimizing HCC coding, intervention-group enrollees at greatest risk for hospital-based services could be identified and then managed with intensive office-based care. Provider Group A standardized care with: 1) a triage system so that frail, complex patients (as determined by higher RAF scores) had immediate access to their primary care physicians or nurse practitioners (Table 2); 2) improved care coordination that booked clinic visits soon after hospitalization, as the appointment dates proposed by hospital-based physicians often were overdue; and 3) scheduling members with specific HCCs (ie, heart failure [HF], chronic obstructive pulmonary disease [COPD], and diabetes) at regular intervals because elective office visits by members with these HCCs often were prodromal of imminent hospital-based services. The current study thus provides preliminary evidence that the RAF can direct primary care. In parallel, MAOs that participate in CMMI's upcoming VBID demonstration project can formulate different plan benefits for enrollees with the following HCCs: HF, COPD, diabetes, cerebrovascular disease, hypertension, coronary artery disease, and mood disorders.[15] Unlike the VBID demonstration project, in this study, the providers who best knew the enrollees designed these value-based protocols and not a third-party payer. Provider Group A's clinical practice transformation increased office-based utilization and reduced hospital-based services, with an estimated cost benefit of $2,116,118 per 1000 members, or $529.03 PMPY.

One clinical vignette exemplified the type of care intervention-group members received. The HPC cared for an 80-year-old man with ischemic cardiomyopathy, whose HF worsened and resulted in acute renal failure. Given the poor prognosis, the patient's cardiologist recommended hospice care. After the patient beseechingly asked about alternatives, the HPC embarked on an intensive schedule of weekly to fortnightly office visits with continual medication readjustment. The patient lived for another 3.5 years with good quality of life, only 1 hospital admission, and no admissions for HF. Evidence-based medicine has determined that exacerbation of HF is an "ambulatory care–sensitive condition"; nevertheless, ED visits for HF are commonplace and a significant drain on Medicare's budget.[45,46] Provider Group A shifted its practice standards after full-risk contracting. Respected clinicians within a provider group (eg, the HPC) can advocate for unique critical pathways, such as linking the CMS-HCC fiduciary model to innovative healthcare delivery.

By comparing disparate populations of MA and traditional Medicare beneficiaries and analyzing incomplete utilization data, previous reports have attempted to demonstrate that MA promulgates high-value care.[9,27,28,31] By comparing statistically similar groups of MA enrollees and analyzing full encounter claims data, the current study adds further credible evidence that value-based contracting generates cost efficiencies and improves health. Following this shift in utilization patterns, intervention-group members experienced a 6% survival benefit. Because a DID approach can underestimate confidence intervals and lead to erroneous conclusions about healthcare policy reform measures,[38,47] randomization inference was employed to validate that the intervention's effects were statistically significant with adequate statistical power to support the current conclusions (eAppendix). There were no deaths for the first postintervention year (Figure 3, left panel) in either group. Study inclusion required both a minimum age of 65 and survival during the entire preintervention period. Although the average life expectancy in the United States is 78.8 years, those aged 65 years can expect to live another 19.3 years, and those aged 75, another 6.6 years.[48] The study requirement of living the entire preintervention period most likely preselected for subsequent survivability.

Because of the observational nature of this study, survival differences might be attributable to factors not examined, such as environmental or social ones. Focusing the study within a single MSA and matching cohorts with propensity scores attempted to control for such factors. In a randomized trial, the propensity score is a known function. On the other hand, in an observational study, it is always unknown because differences in unobserved baseline characteristics may influence observed outcomes. Confirming the survival analysis with randomization inference rejoined many theoretical issues of relying solely on propensity score matching in this observational quasi-experiment. Of course, the most notable limitation is whether these findings can be applied to the broader MA population. Future studies must investigate how MAOs and their differing arrangements with providers interact with ethnic and geographic variations in healthcare delivery.[28,29]

CONCLUSIONS

Value-based contracting between MAOs and providers generate cost efficiencies and improve clinical outcomes in MA, which is the ultimate aim of the current initiatives for Medicare FFS reform. Empiric economic analyses—ignorant of differing MAO–provider contracting arrangements—have challenged the value of subsidizing the MA program and calculated a rather poor return on the taxpayers' investment.[9,10,13,14] CMMI's VBID model thus has been aptly timed, beginning on January 1, 2017, and running for 5 years.[15] In the interim and with minimal effort, MAOs and provider groups can alter contracting arrangements. In turn, providers organically develop innovative primary care strategies that are cost-efficient and improve clinical outcomes to the benefit of all MA program stakeholders. ■

Acknowledgments

The reporting and interpretation of these data are the sole responsibility of the authors. The authors wish to thank: 1) Joanna Martson, Director of Market Consultation (OR, WA, AK, HI) at Optum, for her additional insights on the Oregon MA market; 2) Larry Kunz, Director of Research, and his staff at SavvySherpa, Inc for their independent audit of our data files, assistance with our econometric analysis (Philip R. Vande Kamp, PhD), and fastidious statistical analytical support (Daniel J. Halterman, MS); and 3) Eric M. Peterson, Deputy General Counsel at Optum, for his input during the editing of this paper. Dr Howell's current affiliation is with the Heritage Provider Network.

Author Affiliations: Optum, UnitedHealth Group (AKM, GKT, RVF, SCH), Santa Ana, CA.

Source of Funding: None.

Author Disclosures: The authors report no relationship or financial interest with any entity that would pose a conflict of interest with the subject matter of this article.

Authorship Information: Concept and design (AKM); acquisition of data (AKM, GKT, RVF); analysis and interpretation of data (AKM, SCH); drafting of the manuscript (AKM, SCH); critical revision of the manuscript for important intellectual content (AKM, GKT, RVF, SCH); statistical analysis (AKM, SCH); provision of patients or study materials (RVF, GKT); obtaining funding (AKM); administrative, technical, or logistic support (AKM, GKT, RVF, SCH); and supervision (AKM).

Address Correspondence to: Aloke K. Mandal, MD, PhD, Optum, Mail Code CA152-0243, 3110 W Lake Center Dr, Santa Ana, CA 92704-6917. E-mail: aloke.mandal@optum.com.

REFERENCES

1. Proposed rule: Medicare program; Merit-Based Incentive Payment System (MIPS) and Alternative Payment Model (APM) incentive under the physician fee schedule, and criteria for physician-focused payment models. *Federal Register* website. http://federalregister.gov/a/2016-10032. Published May 9, 2016. Accessed December 31, 2016.
2. Better care. smarter spending. healthier people: improving quality and paying for what works. CMS website. https://www.cms.gov/Newsroom/MediaReleaseDatabase/Fact-sheets/2016-Fact-sheets-items/2016-03-03-2.html. Published March 3, 2016. Accessed December 31, 2016.
3. Burwell SM. Setting value-based payment goals—HHS efforts to improve U.S. health care. *N Engl J Med*. 2015;372(10):897-899. doi: 10.1056/NEJMp1500445.
4. Gosden T, Forland F, Kristiansen I, et al. Capitation, salary, fee-for-service and mixed systems of payment: effects on the behaviour of primary care physicians. *Cochrane Database Syst Rev*. 2000;(3):CD002215. doi: 10.1002/14651858.CD002215.
5. Flodgren G, Eccles MP, Shepperd S, Scott A, Parmelli E, Beyer FR. An overview of reviews evaluating the effectiveness of financial incentives in changing healthcare professional behaviours and patient outcomes. *Cochrane Database Syst Rev*. 2011;(7):CD009255. doi: 10.1002/14651858.CD009255.
6. Medicare Advantage/Part D contract and enrollment data. CMS website. https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/. Updated December 6, 2016. Accessed December 31, 2016.
7. HHS FY 2017 budget in brief—CMS—Medicare. HHS website. http://www.hhs.gov/about/budget/fy2017/budget-in-brief/cms/medicare/index.html. Updated February 12, 2016. Accessed December 31, 2016.
8. Part 422—Medicare Advantage Program. US Government Publishing Office website. http://www.ecfr.gov/cgi-bin/retrieveECFR?gp=&SID=5942b7e3c726955e6601f4af927598b7&mc=true&r=PART&n=pt42.3.422. Updated December 29, 2016. Accessed December 31, 2016.
9. Pizer SD, Frakt AB, Feldman R. Nothing for something? estimating costs and value for beneficiaries from recent Medicare spending increases on HMO benefits and drug benefits. *Int J Health Care Finance Econ*. 2009;9(1):59-81. doi: 10.1007/s10754-008-9047-x.
10. Song Z, Landrum MB, Chernew ME. Competitive bidding in Medicare: who benefits from competition? *Am J Manag Care*. 2012;18(9):546-552.
11. Guram JS, Moffett RE. The Medicare Advantage success story—looking beyond the cost difference. *N Engl J Med*. 2012;366(13):1177-1179. doi: 10.1056/NEJMp1114019.
12. Baicker K, Chernew ME, Robbins JA. The spillover effect of Medicare managed care: Medicare Advantage and hospital utilization. *J Health Econ*. 2013;32(6):1289-1300. doi: 10.1016/j.jhealeco.2013.09.005.
13. Duggan M, Starc A, Vabson B. Who benefits when the government pays more? pass-through in the Medicare Advantage program [NBER working paper No. 19989]. The National Bureau of Economic Research website. http://www.nber.org/papers/w19989. Published March 2014. Accessed December 31, 2016.
14. Cabral M, Geruso M, Mahoney N. Does privatized health insurance benefit patients or producers? evidence from Medicare Advantage [NBER working paper No. 20470]. The National Bureau of Economic Research website. http://www.nber.org/papers/w20470. Published September 2014. Accessed December 31, 2016.
15. Medicare Advantage Value-Based Insurance Design Model. CMS website. https://innovation.cms.gov/initiatives/vbid/. Updated December 30, 2016. Accessed December 31, 2016.
16. Brown RS, Clement DG, Hill JW, Retchin SM, Bergeron JW. Do health maintenance organizations work for Medicare? *Health Care Financ Rev*. 1993;15(1):7-23.
17. Riley G, Tudor C, Chiang YP, Ingber M. Health status of Medicare enrollees in HMOs and fee-for-service in 1994. *Health Care Financ Rev*. 1996;17(4):65-76.
18. Mello MM, Stearns SC, Norton EC, Ricketts TC 3rd. Understanding biased selection in Medicare HMOs. *Health Serv Res*. 2003;38(3):961-992.
19. Pope GC, Kautter J, Ellis RP, et al. Risk adjustment of Medicare capitation payments using the CMS-HCC model. *Health Care Financ Rev*. 2004;25(4):119-141.
20. Pope GC, Kautter J, Ingber MJ, Freeman S, Sekar R, Newhart C. Evaluation of the CMS-HCC risk adjustment model: final report. CMS website. https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/downloads/evaluation_risk_adj_model_2011.pdf. Published March 2011. Accessed December 31, 2016.
21. Announcement of calendar year (CY) 2017 Medicare Advantage capitation rates and Medicare Advantage and Part D payment policies and final call letter. CMS website. https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtg-SpecRateStats/Downloads/Announcement2017.pdf. Published April 4, 2016. Accessed December 31, 2016.
22. Ratebooks & supporting data. CMS website. https://www.cms.gov/medicare/health-plans/medicareadvtg-specratestats/Ratebooks-and-Supporting-Data.html. Updated April 25, 2016. Accessed December 31, 2016.
23. Bodenheimer T, Pham HH. Primary care: current problems and proposed solutions. *Health Aff (Millwood)*. 2010;29(5):799-805. doi: 10.1377/hlthaff.2010.0026.
24. Tilburt JC, Wynia MK, Sheeler RD, et al. Views of US physicians about controlling health care costs. *JAMA*. 2013;310(4):380-388. doi: 10.1001/jama.2013.8278.
25. Announcement of calendar year (CY) 2013 Medicare Advantage capitation rates and Medicare Advantage and Part D payment policies and final call letter. CMS website https://www.cms.gov/Medicare/Prescription-Drug-Coverage/PrescriptionDrugCovContra/Downloads/RateNoticeA.pdf. Published April 2, 2012. Accessed December 31, 2016.
26. Comments on proposed changes to the CMS-HCC risk adjustment model for payment year 2017. CMS website. https://www.cms.gov/Medicare/Health-Plans/MedicareAdvtgSpecRateStats/Downloads/RiskAdj2017ProposedChangesComments.pdf. Published December 18, 2015. Accessed December 31, 2016.
27. Landon BE, Zaslavsky AM, Saunders RC, Pawlson LG, Newhouse JP, Ayanian JZ. Analysis of Medicare Advantage HMOs compared with traditional Medicare shows lower use of many services during 2003-09. *Health Aff (Millwood)*. 2012;31(12):2609-2617. doi: 10.1377/hlthaff.2012.0179.

28. Matlock DD, Groeneveld PW, Sidney S, et al. Geographic variation in cardiovascular procedure among Medicare fee-for-service vs Medicare Advantage beneficiaries. *JAMA*. 2013;310(2):155-162. doi: 10.1001/jama.2013.7837.
29. Ayanian JZ, Landon BE, Newhouse JP, Zaslavsky AM. Racial and ethnic disparities among enrollees in Medicare Advantage plans. *N Engl J Med*. 2014;371(24):2288-2297. doi: 10.1056/NEJMsa1407273.
30. National Center for Health Statistics. Analytic issues in using the Medicare enrollment and claims data linked to NCHS surveys. CDC website. http://www.cdc.gov/nchs/data/datalinkage/cms_medicare_analytic_issues_final.pdf. Published December 2012. Accessed December 31, 2016.
31. Haukoos JS, Lewis RJ. The propensity score. *JAMA*. 2015;314(15):1637-1638. doi: 10.1001/jama.2015.13480.
32. Lemieux J, Sennett C, Wang R, Mulligan T, Bumbaugh J. Hospital readmission rates in Medicare Advantage plans. *Am J Manag Care*. 2012;18(2):96-104.
33. Wong ES, Hebert PL, Maciejewski ML, et al. Does favorable selection among Medicare Advantage enrollees affect measurement of hospital readmission rates? *Med Care Res Rev.* 2014;71:367-383. doi: 10.1177/1077558714533823.
34. McWilliams JM, Hsu J, Newhouse JP. New risk-adjustment system was associated with reduced favorable selection in Medicare Advantage. *Health Aff (Millwood).* 2012;31(12):2630-2640. doi: 10.1377/hlthaff.2011.1344.
35. Rahman M, Keohane L, Trivedi AN, Mor V. High-cost patients had substantial rates of leaving Medicare Advantage and joining traditional Medicare. *Health Aff (Millwood).* 2015;34(10):1675-1681. doi: 10.1377/hlthaff.2015.0272.
36. Lamarca R, Alonso J, Gomez G, Munoz A. Left-truncated data with age as time-scale: an alternative for survival analysis in the elderly population. *J Gerontol A Biol Sci Med Sci*. 1998;53(5):M337-M343.
37. Cho H, Mariotto AB, Mann BS, Klabunde CN, Feuer EJ. Assessing non-cancer-related health status of US cancer patients: other-cause survival and comorbidity prevalence. *Am J Epidemiol*. 2013;178(3):339-349. doi: 10.1093/aje/kws580.
38. Kaestner R. Did Massachusetts health care reform lower mortality? no according to randomization inference. *Stat Public Policy*. 2016;3(1):1-6. doi: 10.1080/2330443X.2015.1102667.
39. Mathematica Policy Research. Evaluation of Health Care Innovation Awards (HCIA): primary care redesign programs—second annual report. CMS website. https://downloads.cms.gov/files/cmmi/hcia-primarycareredesignprog-secondannualrpt.pdf. Published March 2016. Accessed December 31, 2016.
40. Medicare program; medical loss ratio requirements for the Medicare Advantage and the Medicare Prescription Drug Benefit Programs; final rule. US Government Printing Office website. https://www.gpo.gov/fdsys/pkg/FR-2013-05-23/pdf/2013-12156.pdf. Published May 23, 2013. Accessed December 31, 2016.
41. Jaén CR, Ferrer RL, Miller WL, et al. Patient outcomes at 26 months in the patient-centered Medical Home National Demonstration Project. *Ann Fam Med*. 2010;8(suppl 1):S57-S67; S92. doi: 10.1370/afm.1121.
42. Jackson GL, Powers BJ, Chatterjee R, et al. Improving patient care. the patient centered medical home. a systematic review. *Ann Intern Med*. 2013;158(3):169-178.
43. NORC at the University of Chicago. First annual report: HCIA complex/high-risk patient targeting. CMS website. https://innovation.cms.gov/Files/reports/HCIA-CHSPT-FirstEvalRpt.pdf. Published November 7, 2014. Accessed December 31, 2016.
44. NORC at the University of Chicago. Second annual report: HCIA complex/high-risk patient targeting. CMS website. https://downloads.cms.gov/files/cmmi/hcia-complexhighriskpattargeting-secondevalrpt.pdf. Published March 2016. Accessed December 31, 2016.
45. Berwick DM, Nolan TW, Whittington J. The triple aim: care, health, and cost. *Health Aff (Millwood)*. 2008;27(3):759-769. doi: 10.1377/hlthaff.27.3.759.
46. Dharmarajan K, Hsieh AF, Lin Z, et al. Diagnoses and timing of 30-day readmissions after hospitalization for heart failure, acute myocardial infarction, or pneumonia. *JAMA*. 2013;309(4):355-363. doi: 10.1001/jama.2012.216476.
47. Sommers BD, Long SK, Baicker K. Changes in mortality after Massachusetts health care reform: a quasi-experimental study. *Ann Intern Med*. 2014;160(9):585-593. doi: 10.7326/M13-2275.
48. Table 15: life expectancy at birth, at age 65, and at age 75, by sex, race, and Hispanic origin: United States, selected years 1900–2014. CDC website. http://www.cdc.gov/nchs/data/hus/2015/015.pdf. Updated April 21, 2016. Accessed December 31, 2016.

eAppendix for:

Value-Based Contracting Innovated Medicare Advantage Healthcare Delivery and Improved Survival Aloke K. Mandal, M.D., Ph.D.; Gene K. Tagomori, B.Sc.; Randell V. Felix, B.Sc.; and Scott C. Howell, D.O., M.P.H. & T.M.

Contents

Background and Overall Methodological Approach:

Barriers to Research on the Medicare Advantage (MA) Program:

MA enrollment has doubled over the past 10 years, currently comprising 17.5 million members or close to one-third of all Medicare beneficiaries.[1] For FY2017, MA shall account for $204.7 billion (28.9%) of Medicare's gross spending.[2] In light of the commercial success of the MA program and the federal largesse that subsidizes it, a more detailed analysis is warranted.

Unfortunately, data on MA enrollees available from the Centers for Medicare and Medicaid Services (CMS) to researchers is not as complete as those for Medicare FFS beneficiaries. Prior to 2012, MAOs were not required to report full encounter claims data, replete with the associated bill, revenue, diagnostic, and procedural coding information.[3] Claims data has been limited to: (1) those specific for the mandatory reporting of Healthcare Effectiveness Data and Information Set (HEDIS) measures and (2)—in specific states that separately identify MA enrollees—inpatient (IP) claims data from the Healthcare Cost and Utilization Project (HCUP). Instead, much of the utilization and health status data are derived from self-reported survey data, which may be subject to considerable recall bias by elderly Medicare beneficiaries with multiple chronic conditions and age-related cognitive impairment. Moreover, CMS-HCC Risk-Adjustment Factor (RAF) scores at the MA member level are not readily available.[4,5] Finally, details on the payer-provider arrangements to ensure the delivery of CMS-approved benefit packages are not in the public domain. Accordingly, because of this incomplete data, conflicting conclusions on the MA program has emerged, based on differing methodologies and empiric modeling approaches.[6,7,8]

---

[1] Monthly Contract and Enrollment Summary Report Items. Centers for Medicare and Medicaid Services, Baltimore, MD. November 7, 2016. (Accessed on December 15, 2016, at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/Monthly-Contract-and-Enrollment-Summary-Report-Items/Contract-Summary-2016-04.html?DLPage=1&DLEntries=10&DLSort=1&DLSortDir=descending.)

[2] HHS FY2017 Budget in Brief-CMS-Medicare. Department of Health & Human Services. Washington, DC. February 12, 2016. Accessed on December 15, 2016, at http://www.hhs.gov/about/budget/fy2017/budget-in-brief/cms/medicare/index.html.)

[3] Medicare Program; Revisions to Payment Policies Under the Physician Fee Schedule, Clinical Laboratory Fee Schedule, Access to Identifiable Data for the Center for Medicare and Medicaid Innovation Models & Other Revisions to Part B for CY 2015; Final Rule. Department of Health & Human Services. Washington, DC. November 13, 2014. Federal Register 79:67547-68092. (Accessed on December 15, 2016, at https://www.federalregister.gov/articles/2014/11/13/2014-26183/medicare-program-revisions-to-payment-policies-under-the-physician-fee-schedule-clinical-laboratory.)

[4] Kronick, R, Welch WP. Measuring coding intensity in the Medicare Advantage Program. Medicare Medicaid Res Rev 2014; 4: E1-E19.

[5] Rahman M, Keohane L, Trivedi AN, Mor V. High-cost patients had substantial rates of leaving Medicare Advantage joining traditional Medicare. Health Aff 2015; 34: 1675-1681.

[6] Guram JS, Moffett RE. The Medicare Advantage Success Story — Looking beyond the Cost Difference. N Engl J Med 2012; 366(13):1177-1179.

[7] Baicker K, Chernew ME, Robbins JA. The spillover effect of Medicare managed care: Medicare Advantage and hospital utilization. J Health Econ 2013; 32: 1289-1300.

Optum assists 27 private health plan clients on over 600 MA plans. It houses demographic, full encounter claims, and CMS administrative data on 9.5 million MA members (52% of all MA enrollees). Through monitoring risk-adjustment and healthcare quality activities of over 12,000 provider groups contracted with these MAO clients, knowledge of different MAO-provider arrangements also exists. Analyzing this proprietary database should circumvent the aforementioned barriers to MA research. The primary aim of this study tested the following hypothesis: value-based contracts augment CMS-HCC payment model's ability to generate cost-efficiencies and improve clinical outcomes.

Study Design:

Under Contract H3805 listed in CMS' website,[9] UnitedHealthcare of Oregon is the MAO for a HMO-type Local Coordinated Care Plan. Optum assists this MAO in its risk-adjustment activities. About three-quarters of the providers contracted through this plan are reimbursed on a FFS basis and only a quarter through some capitated arrangement.[10] With this knowledge, rather than comparing a group of MA enrollees with a disparate group of Medicare FFS beneficiaries (which has been a study design flaw of past MA research since none of these studies appropriately identified statistically similar groups for comparison or used objective criteria for health status[5,7,11,12,13]), it was possible to identify enrollees in the same MA plan, with one group treated under a FFS model and another through a capitated one. Portland, Oregon was chosen as the Metropolitan Statistical Area (MSA), since it has a relatively homogenous population and a high degree of MA penetration (>50%).[14,15] By using this MSA, it was possible to control for racial and market variations, especially since areas with high MA penetration have more uniform utilization rates.[7,13] Since the date of enrollment also affects

---

[8] Duggan M, Starc A, Vabson B. Who benefits when the government pays more? Pass-through in the Medicare Advantage program. March 2014. NBER Working Paper No. 19989. (Accessed on December 15, 2016, at http://www.nber.org/papers/w19989.)

[9] Plan Directory. Medicare Advantage/Part D Contract and Enrollment Data. Research, Statistics, Data and Systems. Baltimore, MD: Centers for Medicare and Medicaid Services. (Accessed on December 15, 2016, at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/MA-Plan-Directory.html.)

[10] This estimate was obtained from Joanna Martson, Director of Market Consultation (OR, WA, AK, HI) at Optum. In addition, from the authors' own experience, the FFS reimbursement of providers from this capitated payment model persists throughout the nation, with very few MAO's having risk-bearing agreements with their provider networks.

[11] Landon BE, Zaslavsky AM, Saunder RC, et al. Analysis of Medicare Advantage HMOs compared with traditional Medicare shows lower use of many services during 2003-2009. Health Affairs 2012; 31: 2609-2617.

[12] Matlock DD, Groenveld PW, Sidney S, et al. Geographic Variation in Cardiovascular Procedure among Medicare Fee-for-Service vs Medicare Advantage Beneficiaries. JAMA 2013; 310(2): 155-62.

[13] Cabral M, Geruso M, Mahoney N. Does privatized health insurance benefit patients or producers? Evidence from Medicare Advantage. September 2014. NBER Working Paper No. 20470. (Accessed on December 15, 2016, at http://www.nber.org/papers/w20470).

[14] Quick Facts: Multnomah County, Oregon. March 31, 2015. Washington, D.C. U.S.Census Bureau. (Accessed on December 15, 2016, at http://quickfacts.census.gov/qfd/states/41/41051.html.)

[15] MA State/County Penetration. Medicare Advantage/Part D Contract and Enrollment Data. Research, Statistics, Data and Systems. Baltimore, MD: Centers for Medicare and Medicaid Services. (Accessed on December 15, 2016, at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/MCRAdvPartDEnrolData/MA-State-County-Penetration.html.)

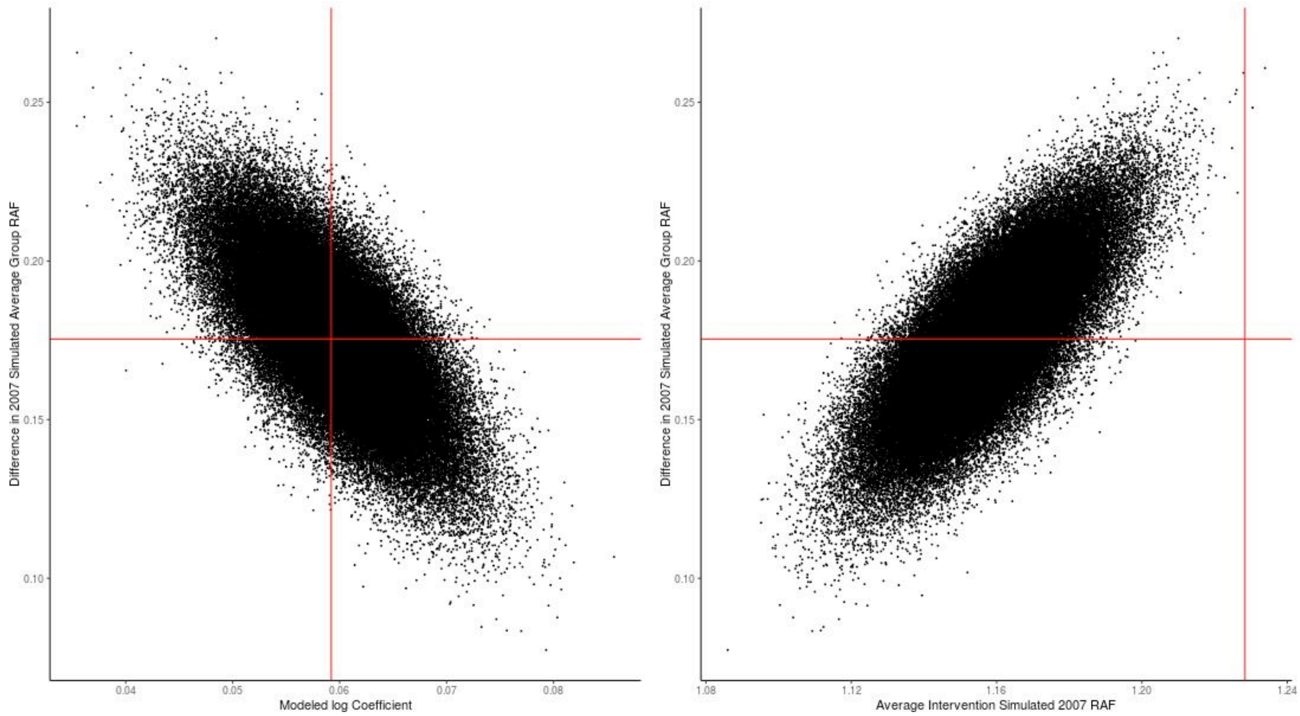
utilization among MA enrollees,[11] this study was confined to a discrete “same-store” group of MA enrollees with well-defined preintervention and postintervention periods. A difference-in-differences (DID) approach was chosen in order to account for other factors which could influence outcomes over time.

Sensitivity Analysis—Comparing a 1-Year (2008) with a 2-Year (2007-2008) Preintervention Period:

Simulations for 2007 RAF: The CMS-HCC payment model gradually was phased in, beginning in 2004 with full implementation in 2007. It wasn’t until 2007 when most MAOs fully concerned themselves with CMS-HCC risk-adjustment reimbursement, with most provider groups’ delaying their understanding of the payment model for even longer. Optum’s client services include provider training on HCC coding. Based on this experience, it was noted that only a few provider groups began to be interested in Optum’s outreach training program on or around 2008. Even currently, many provider groups contracting with our MAO clients lack a fundamental understanding of the CMS-HCC payment model, as evidenced by lower-than-expected RAF scores based on case-mix. Therefore, Optum’s RAF data prior to 2008 was disjointed among many provider groups. Accordingly, we were confident only of RAF data housed at Optum, starting in 2008. Thus, only one-year of preintervention data for RAF was available.

In order to determine the adequacy of a one-year preintervention period when compared with a two-year preintervention period, the following simulation was programmed. Using 2008 RAF and 2009 RAF for the members in the study, this relationship predicted what each member’s 2007 RAF could have been. For each member, a predicted distribution was created (left panel, *eAppendix Figure 1*). A predicted RAF for each member in 2007 was extrapolated, averaged, and then combined with the actual 2008 RAF for the preintervention RAF score in order to develop models for a two-year preintervention period (right panel, *eAppendix Figure 1*). By running models that estimated the average change in RAF over the postintervention period for each group compared to the preintervention RAF score; the coefficient, the p-value of this coefficient, and the average RAF scores for each group for each simulation were obtained. After 100,000 simulations, the p-values always were less than 0.0015. Although these results are dependent on an ability to predict members’ 2007 RAF score, this modeling nevertheless delineates how using 1 year of preintervention RAF scores did not compromise the DID analysis when compared with using putatively 2 years of RAF for the preintervention period.

eAppendix Figure 1: Simulations of 2007 RAF data for 2-Year Preintervention Data.



Left Panel demonstrated the difference in the Average 2007 simulated RAF for each group. After simulating 2007 RAF and taking the averages, the difference in those values were calculated. The modeled log coefficient is on the horizontal axis. All values are above 0.03, suggesting the intervention group always would have had a higher change in RAF in the difference-in-differences approach. The simulation provided a range of what the coefficient would be, dependent on reasonable 2007 RAF values. The red lines depict the observed difference in 2008 and the observed log coefficient when only using 2008 as preintervention. *Right Panel* demonstrates the relationship between how the intervention Group's average simulated 2007 RAF increases and how the differential increases. Even when that difference is large and the Intervention Group's 2007 RAF is large, a significant p-value is obtained. The red lines in this figure illustrate what the average 2008 RAF actually was for the intervention group and the difference in average RAF in 2008.

Using 2-year (2007 and 2008) claims data for the preintervention period, compared with using 1-year (2008) of data: 2007 full encounter claims data were available from the UnitedHealth Group Analytics Platform (UGAP).[16] The results of using a one-year preintervention were compared with that of a two-year preintervention period. With regard to office based visits, the sensitivity of the p-value and coefficient when using a one year and a two year preintervention period were analyzed. When using one year of data, the incident rate ratio (IRR) was estimated at 1.112 with a p-value of 0.005. When using a two year period, the IRR was estimated at 1.068 with a p-value of 0.025. Although the two year period did lessen the estimated IRR, nevertheless, the preintervention-postintervention DID in office based visits remained statistically significant between these two groups.

[16] Information on the UnitedHealth Group Analytics Platform (UGAP) can be found at http://dmo.optum.com/products/ugap.html.

Conclusions: Based on the above simulations for 2007 RAF data and use of actual 2007 claims data, it appears that using only one year of data for the preintervention period would not compromise the DID analysis for changes in either RAF or utilization.

Data Sources:

The study population was enrolled in H3805 and ascribed to one of the two provider groups. The enrollees initially were identified by their HIC numbers and the provider groups by their Taxpayer ID numbers for each clinic locations. After determining the feasibility and obtaining approval for the study; enrollees, medical groups, and facilities were anonymized. Appropriate administrative, technical, procedural, and physical safeguards were in full force to protect the confidentiality of the data and prevent unauthorized access. Next, Optum's Division of Analytics and Oversight provided the following data.

- Data derived from CMS' Monthly Membership Report (MMR) provided information on age, sex, address (to confirm that enrollee's primary residence was the Portland metropolitan service area and was consistent with plan's local network), race, original reason for Medicare entitlement, and RAF;
- Data extracted from claims data provided information on place of service, dates and duration of service, diagnoses codes, and procedures codes;
- Data derived from CMS' Daily Transaction Reply Report (DTRR) provided information on survival.

SavvySherpa (https://savvysherpa.com/), with independent access to UGAP[16], then performed a third-party audit in order to assure integrity and reproducibility of the database used for this study.

Propensity Score Model and Nearest-Neighbor Matching:

Health plan enrollment and choice of providers are self-selecting and certainly non-random processes. In this observational study, threats to validity of the findings because of self-selection into the Intervention Group were mitigated by matching. Matching is a non-parametric pre-processing step that mimics the assignment stage in a randomized experiment and is separate from the analysis stage. Matching ensured that the Intervention and Control Groups overlapped on key covariates and thus were comparable.

Selection of Covariates:

Five preintervention (2008 baseline) characteristics of the study population included: age, sex, race, original reason for Medicare entitlement, and disease burden. The original reason for Medicare entitlement was included since aged beneficiaries who formerly were disabled have cost and utilization profiles different from other aged beneficiaries, especially between the ages of 65 and 74.[17] With the availability of diagnosis codes from claims data, it was possible to develop objective parameters for disease burden (health status) in two manners. Firstly, from the available claims data, a Charlson Comorbidity Index (CCI) was ascertained, using an "Enhanced ICD-9-CM Coding Algorithm."[18] Secondly, the number of chronic conditions for each member was determined using the algorithm to identify the original twenty-one chronic condition categories within CMS' Chronic Conditions Data Warehouse.[19]

Meaningful differences were considered those with a standardized mean difference (SMD) less than -0.1 or greater than 0.1. As demonstrated in *eAppendix Table 1*, prior to matching, meaningful differences were noted for race (Asian, Black), original reason for Medicare entitlement (Age, Disabled), and marginally for CCI. In terms of counts, the biggest differences were in white, female, black, and age as the original reason for Medicare entitlement.

eAppendix Table 1: Comparison of MA Enrollees in 2008 before and after Nearest-Neighbor Matching.

	Before Matching			*After Matching*		
Characteristics	***Intervention Group (n=1,230)***	***Control Group (n=1,530)***	***SMD***	***Intervention Group (n=1,230)***	***Control Group (n=1,230)***	***SMD***
Age ± S.E.M.	80.18±0.20	80.24±0.18	-0.009	80.18±0.20	80.05±0.19	0.018
Female	763	962	-0.025	763	758	0.008
Race						
-White	1,177	1,460	-0.015	1,177	1,195	0.079
-Asian	22	6	0.134	22	5	0.133
-Black	7	31	-0.130	7	7	0.000
-Hispanic	3	3	0.010	3	3	0.000
-Native American	2	1	0.029	2	1	0.023
-Other	19	20	0.019	19	19	0.000
Original Reason for Medicare Entitlement						
-Age	1,191	1,422	0.156	1,191	1,192	-0.005
-Disability	39	99	-0.156	39	38	0.005
-End-Stage Renal Disease	0	0	0.00	0	0	0.000
Burden of Disease						
-Charlson Comorbidity Index[17]	3.21±0.07	2.98±0.06	0.098	3.21±0.07	3.16±0.07	0.023
-No. of Chronic Conditions[18]	3.84±0.07	3.82±0.06	0.006	3.84±0.07	3.92±0.07	-0.037

[17] National Center for Health Statistics, Office of Analysis and Epidemiology. Analytic Issues in Using the Medicare Enrollment and Claims Data Linked to NCHS Surveys. December 2012. Hyattsville, Maryland. (Accessed on December 15, 2016, at http://www.cdc.gov/nchs/data/datalinkage/cms_medicare_analytic_issues_final.pdf.)

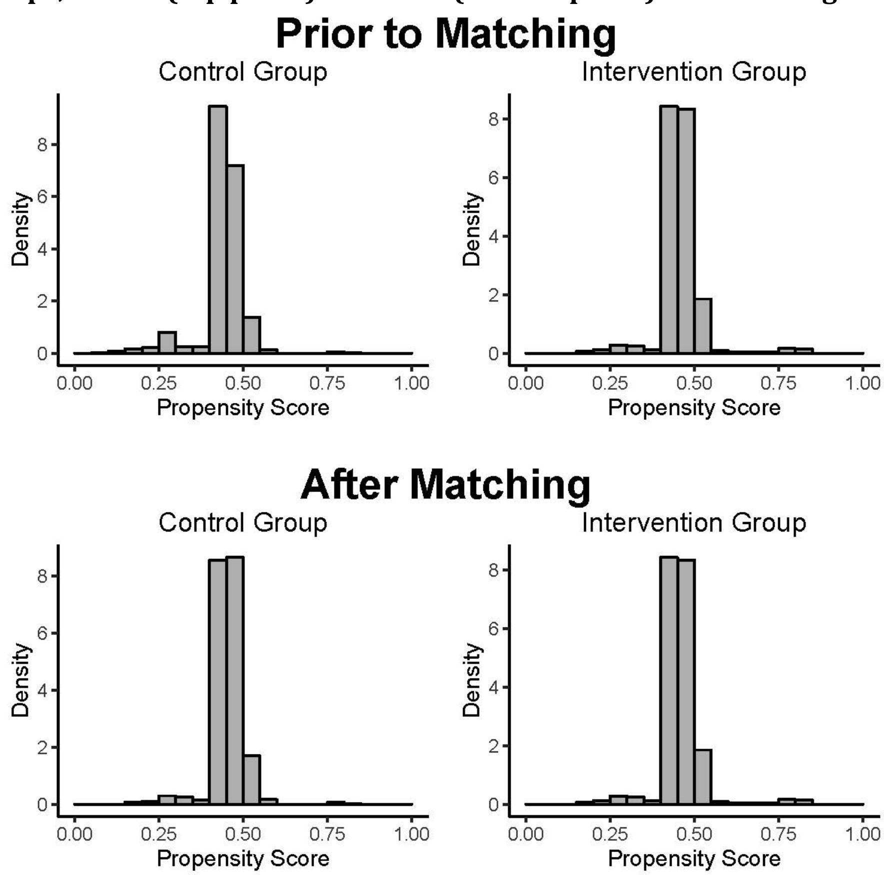
[18] Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. Med Care 2005; 43: 1130–1139.

[19] CMS Chronic Condition Data Warehouse Condition Categories. Centers for Medicare and Medicaid Services. 2016. (Accessed on December 15, 2016, at https://www.ccwdata.org/web/guest/condition-categories.)

Nearest-Neighbor Matching:

A logistic regression model was used to create propensity scores for each member, using the baseline covariates. Pairs of enrollees in the Intervention and Control groups were matched on the logit of the propensity score. All of the enrollees in the Intervention Group were matched to enrollees in the Control Group with a similar propensity score, using "greedy nearest-neighbor matching."[20,21] Unmatched enrollees were removed, providing a set of statistically similar Intervention and Control cohorts. From the Control Group, the matching process trimmed 265 White, 230 with age as the original reason for Medicare entitlement, 204 Female, and 24 Black enrollees (*eAppendix Table 1*). After matching, only Asian race persisted as a meaningful difference with a SMD of 0.133. As illustrated in *eAppendix Figure 2*, matching created well-balanced and statistically similar groups for subsequent analysis.

eAppendix Figure 2: Histogram of the Distribution of Propensity Scores for Intervention and Control Groups, before (top panel) and after (bottom panel) Nearest-Neighbor Matching.



---

[20] Ho DE, Imai K, King G, Stuart EA: MatchIt: nonparametric preprocessing for parametric causal inference. Journal of Statistical Software 2011; 42(8): 1-28.

[21] Austin PC. The use of propensity score methods with survival or time-to-event outcomes: reporting measures of effect similar to those used in randomized experiments. Statist Med 2014; 33(7): 1242-1258.

Study Variables and Rationale for Statistical Methodologies:

Power Calculations:

Based on the data available and the sample size, a power analysis was performed in order to determine which study variables could be analyzed. Variables proposed for the study included RAF (obtained from end-of-year MMR), utilization aggregated by place of service (obtained from full encounter claims data), preventive care utilization (full encounter claims data), death (obtained from the DTRR), and cause of death (full encounter claims data). Based on a power analyses, cause of death was excluded as a proposed outcome variable. Specifically, the percentages for different causes of death were obtained from Centers for Disease Control and Prevention data.[22] Sample sizes per group necessary to detect a 1%-5% difference in cause of death with 80% power were calculated for when the intervention group was less than the control. Based on the prevalence of heart disease as a cause of death,[22] over 50,000 members in each group would have been needed to detect any difference with 80% power.

RAF and Estimates on Risk-Adjusted Revenue:[23]

RAF was obtained from data extracted from the end-of-year MMR for each CY. The DID analysis determined the percent increase in RAF for the Intervention Group, compared with the Control Group. The Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (MMA) established the basic elements of the current MA payment scheme.[24] The formula for MAO payment from CMS is:

MAO Payment = (Risk Score x 100% Bid) – Max(100%Bid-100%Benchmark) + Rebate

Since the MAO's actual bid price is not within the public domain, CMS publicly reports various blended benchmarks with two components: (1) a statutory component of the weighted average of the county capitation rates and (2) a "plan-bid" component that is the weighted average of all standardized FFS bids for MA plans in the region.[25] Based on the above formula, it therefore is possible to estimate the risk-based revenue by multiplying RAF and the County-

---

[22] Table 20: Leading causes of death and numbers of deaths, by age: United States, 1980 and 2014. National Center for Health Statistics. Health, United States, 2015. (Accessed on December 15, 2016, at http://www.cdc.gov/nchs/data/hus/2015/020.pdf.)

[23] The authors would like to acknowledge Philip R. Vande Kamp, Ph.D., an economist at SavvySherpa for his assistance in developing the models for the revenue and cost estimates.

[24] Medicare Advantage Program. 42 Code of Federal Regulations §422. October 1, 2011. (Accessed on December 15, 2016, at https://www.gpo.gov/fdsys/granule/CFR-2011-title42-vol3/CFR-2011-title42-vol3-part422.)

[25] Ratebooks and Supporting Data Items. MedicareAdvantage Rates and Statistics. Centers for Medicare and Medicaid Services, Baltimore, MD. April 25, 2016. (Accessed on December 15, 2016, at https://www.cms.gov/medicare/health-plans/medicareadvtgspecratestats/Ratebooks-and-Supporting-Data.html.)

Level Benchmark rate.[26] One of these blended benchmarks is the county-level "Risk Rate."[25] In order to estimate the additional revenue benefit from the increases in RAF for CY2009 to CY2011, the county-level "Risk Rate" for Multnomah County was accessed for these three years.[25] As mandated by the Patient Protection and Affordable Care Act of 2010 (ACA) and starting with CY2012, bonuses were calculated into the county rate based on a plan's performance on healthcare quality measures (HEDIS/Stars). Because this study only was concerned with revenue generated from RAF, the CY2012 rate for plans with ≤2.5 Stars was used to calculate the PY2012 revenue.[25] For the sake of completion, it should be noted that moving forward in 2015, more factors have been included in ACA-mandated payment reforms, with increased emphasis on healthcare quality ratings, medical cost trends, administrative efficiency, greater offsets for coding intensity, and other regulatory changes.

Because the PY2012 payments occurred in 2013, all values were indexed to 2013 dollar values. The total Personal Health Care Expenditure (PHCE) Index found in Table 23 in CMS' National Health Expenditure Data website was used for the indexing RAF-based revenue.[27,28] The calculations are summarized in *eAppendix Table 2*.

eAppendix Table 2: Estimated Revenue Benefit from RAF Increase for Intervention Group, Compared with Control Group, Following Intervention and Adjusted to 2013 Dollars.

PY	*RAF (% Increase)*	*Monthly County Rate*	*PCHE Index for 2013**	*PCHE Index for each PY**	*Months per Year*	*Members*	*Total (2013 Dollars)*
2009	4.1	818.77	108.4	102.7	12	1,000	$425,193
2010	7.2	815.16	108.4	105.1	12	1,000	$726,412
2011	7.1	815.16	108.4	106.8	12	1,000	$704,921
2012	7.1	778.19	108.4	108.4	12	1,000	$663,018
						Total to MAO:	**$2,519,544**

* ***In order to index to 2013 US Dollars, the 2013 Total PCHE Index was divided by the Total PCHE Index for each of the following years, 2010 to 2013.***

Place of Service Utilization and Estimated Cost Benefits for the Intervention Group:

Place of Service Utilization: Office-based visits were identified by the appropriate Common Procedural Terminology (CPT) codes. Combinations of Type of Bill and Hospital

[26] Indeed, in reviewing the data with Provider Group A's Chief Financial Officer, the RAF Gain Share estimated in this study approximated the actual revenue within 3-5%.

[27] Table 23 in NHE Tables. National Health Expenditure Data. Research, Statistics, Data, and Systems. The Centers for Medicare and Medicaid Services, Baltimore, MD. November 24, 2015. (Accessed on December 15, 2016, at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/NationalHealthAccountsHistorical.html.)

[28] National Health Expenditure Accounts: Methodology Paper, 2013. Definitions, Sources, and Methods. The Centers for Medicare and Medicaid Services. (Accessed on December 15, 2016, at https://www.cms.gov/Research-Statistics-Data-and-Systems/Statistics-Trends-and-Reports/NationalHealthExpendData/Downloads/dsm-13.pdf.)

Revenue codes[29] determined the two types of hospital-based service: emergency room (ER) visits not accompanied by an IP stay and hospital IP stays. In addition, rates were calculated on a per member-month basis in order to censor for the month when enrollees died or disenrolled and to develop incidence rate ratios from a Poisson regression analysis.[30]

- **Office Based Visits:** Claims data were searched for CPT codes: 99201 to 99205, 99211 to 99215, 99241 to 99245; 99381 to 99387, 99391 to 99397, 99401 to 99404, 99411, 99412, 99420, and 99429.
- **ER Visits:** Claims data were filtered for "Type of Bill" codes that denoted Hospital Outpatient. Then, hospital revenue codes between 450 and 459 (excluding 456 for Urgent Care) as well as revenue code 981 were included in order to identify ER visits. All ER visits that occurred with an admission or discharge date of an IP stay were removed.
- **IP Hospital Admissions:** Claims data were filtered for "Type of Bill" codes beginning with "1" (denoting that the facility was a hospital) and then queried for hospital revenue codes between 0100 and 0219 but excluding those that denoted a skilled nursing facility, hospice, rehabilitation center, or sub-acute care facility. If "117" (the code for a replacement bill) appeared, then the previously submitted corresponding bill was excluded in order to avoid duplication. All claims that met the above criteria with an admission up to the day before the date of disenrollment were included because the date of death generally is one day before disenrollment.

Cost Estimates Based on Place of Service Utilization:[19] From the Medical Expense Panel Survey (MEPS) Household Component for each corresponding year, the average Medicare FFS expenditure for each type of service was obtained.[31] These data are summarized in *eAppendix Table 3*, using the following methodology:

- From the MEPS website,[31] SAS transport files of MEPS Household Component Public Use Files for Office-Based Medical Provider Visits (HC-126G, HC-135G, HC-144G, HC-152G), Emergency Room Visits (HC-126E, HC-135E, HC-144E, HC-155E), and Hospital Inpatient Stay Files (HC-126D, HC-135D, HC-144D, HC-152D) were downloaded.

---

[29] Medicare Claims Processing Manual, Chapter 25 - Completing and Processing the Form CMS-1450 Data Set. The Centers for Medicare and Medicaid Services, Baltimore,MD. Rev. 3435, December 31, 2015. (Accessed on December 15, 2016, at https://www.cms.gov/Regulations-and-Guidance/Guidance/Manuals/downloads/clm104c25.pdf.)

[30] Frome EL, Checkoway H. Use of Poisson regression models in estimating incidence rates and ratios. Am. J. Epidemiol 1985; 121: 309-323.

[31] Medical Expenditure Panel Survey (MEPS) Public Use Data Files were obtained as SAS Transport Files from the website (https://meps.ahrq.gov/mepsweb/), managed by Agency for Healthcare Research and Quality, Department of Health & Human Services , Rockville, MD.

- Per encounter costs were obtained by programming only for those Medicare beneficiaries, aged 65 years or older. Mean costs then were calculated.
- "Physician and Clinical Services" component of the PCHE Index was used to index office-based visits expenditures to 2013 dollar values.[27,28] The "Hospital Care" component of the PCHE Index adjusted ER visits and IP expenditures to 2013 dollars.[27,28] A previously published MEPS data report estimated care for the elderly, regardless of type insurance.[32] In that study, while office-based expenditures were similar between Medicare FFS and private health insurance enrollees, hospital services for the elderly were reimbursed by private health insurance carriers at higher rates than by Medicare.[32,33] The estimates in *eAppendix Table 3* thus seem reasonable for the purposes of this study.

All-Cause 30-day and 60-day Hospital Readmissions: From the previously obtained data on hospital IP Stays, anyone who died within 30-days or 60-days was excluded from both the numerator and denominator. Then, the previously described algorithm for determining IP stays was used and determined to be a 30-day or 60-day hospital readmission if an IP stay was within 30 or 60 days of a previous IP stay, respectively.

Preventive Care Utilization:

- Preventive Care Visits were obtained from claims data, filtering for the following CPT codes: 99381 to 99387, 99391 to 99397, 99401 to 99404, 99411, 99412, 99420, and 99429. Member-months were calculated for all members, censoring for the month when enrollees died or disenrolled.
- Screening Mammography: Female MA enrollees whose age was 74 years or younger were identified, and member-years were calculated. From claims data, screening mammography was based on the CPT codes: 77055 to 77057. Rates were based on number of screenings per member-years.
- Screening Colonoscopy: Enrollees aged 75 years or younger were identified, and member-years were calculated for this particular subpopulation. From claims data, screening colonoscopy was based on CPT codes: 44388 to 44394, 44397, 45355, 45378 to 45387, 45391, and 45392. Rates were adjusted to member-years.

[32] Mirel, L.B., Carper, K. Trends in Health Care Expenditures for the Elderly, Age 65 and Over: 2001, 2006, and 2011. Statistical Brief #429. January 2014. Agency for Healthcare Research and Quality, Rockville, MD. (Accessed on December 15, 2016 , at http://www.meps.ahrq.gov/mepsweb/data_files/publications/st429/stat429.pdf.)

[33] Hamavid H, Birger M, Bulchis AG, et al. Assessing the Complex and Evolving Relationship between Charges and Payments in US Hospitals:1996 – 2012. PLoS ONE 2016; 11(7): e0157912. (Accessed on December 15, 2016, at http://journals.plos.org/plosone/article/asset?id=10.1371/journal.pone.0157912.PDF.)

eAppendix Table 3: Cost Benefits from Utilization Changes for the Intervention Group, Compared with Control Group, Following the Intervention and Adjusted to 2013 Dollars.

Office-Based Visit Cost Difference per 1,000 Members					
CY	***Difference in Office-Based Visits per 1,000 Members***	***Office-Based Visit Cost***	***PCHE Component Index for 2013***	***PCHE Component Index, relative to CY2009***	***Total (2013 Dollars)***
2009	5	214.19	105	100.0	1,032
2010	346	226.72	105	102.3	80,529
2011	417	227.47	105	103.7	96,028
2012	415	194.22	105	104.9	80,744
				Cost Benefit:	***258,334***
Previously reported MEPS data[32] on elderly (regardless of type of insurance) estimated average cost of $228 (2011 Dollars), which is similar to amounts calculated for the current study. If the previously reported cost was used, the total benefit in additional office visit care would have been $273,087.					
ER Visit Cost Difference per 1,000 Members					
CY	***Difference in ER Visits per 1,000 Members***	***ER Visit Cost***	***PCHE Component Index for 2013***	***PCHE Component Index, relative to CY2009***	***Total (2013 Dollars)***
2009	-48	740.22	110.2	100.0	(39,546)
2010	-48	547.88	110.2	103.0	(27,867)
2011	-50	555.94	110.2	105.2	(29,182)
2012	-8	548.80	110.2	107.8	(4,320)
				Cost Savings:	***(100,915)***
Previously reported MEPS data[32] on elderly (regardless of type of insurance) estimated average cost of an ER visit at $884 (2011 dollars), which is considerably higher than amounts calculated for the current study derived from MEPS data on just Medicare beneficiaries. If the previously reported amounts were used, the cost savings would have increased to $142,449 (2013 Dollars).					
Hospital Inpatient Cost Difference per 1,000 Members					
CY	***Difference in Inpatient Hospital Stays per 1,000 Members***	***Inpatient Hospital Cost***	***PCHE Component Index for 2013***	***PCHE Component Index, relative to CY2009***	***Total (2013 Dollars)***
2009	-17	10,809.58	110.2	100.0	(207,867)
2010	-52	11,981.32	110.2	103.0	(660,299)
2011	-29	11,066.11	110.2	105.2	(337,445)
2012	-45	11,854.31	110.2	107.8	(551,258)
				Cost Savings:	***(1,756,869)***
Previously reported MEPS data[32] on elderly (regardless of type of insurance) estimated the average per diem cost of an inpatient stay at $3,199 with an average stay of about 5 days (or $15,995 in 2011 dollars), which is considerably higher than amounts calculated for the current study derived from MEPS data on just Medicare beneficiaries. If the previously reported amounts were used, the cost savings would have increased to $2,347,364 (2013 Dollars).					
Total Cost Benefit Based on MEPS Data Used for This Analysis (in 2013 Dollar Values):					**$2,071,293**
Total Cost Benefit Based on Previous MEPS Data[32] (in 2013 Dollar Values):					**$2,754,425**

Survival:

The date of death was obtained from CMS' DTRR. Kaplan-Meier Survival Curves were created with two different time-scales: (1) time-on-the-study and (2) age of enrollees. Using age of enrollees on the time scale provided a natural time-scale for calculating the hazard of death for this elderly population with multiple co-morbidities.[34,35] SEER reportedly uses age on the time scale for assessing survival wherein left-truncated data apply.[35]

Permutation Tests (Randomization Inference):

Randomization Inference on Difference-in-Differences Analysis of RAF and Utilization:

In this study, difference-in-differences models were used longitudinally over a 5-year period to analyze serially correlated outcomes (i.e. change in RAF over time, change in utilization over time), in which case standard errors may be underestimated.[36,37] A permutation test is performed by randomly sampling a variable of interest (in this case, the Intervention/Control group indicator) from the data. By randomly sampling this variable, the distribution of test statistics of interest (in this case, the preintervention-postintervention differences) under the null hypothesis that the permutation variable has no effect, given the data observed, can be explored. If the observed test statistic (effect size) without permutation is not likely under the resampled permutation distribution, the p-value under permutation will be small, providing evidence that the permuted variable does have a meaningful relationship to the outcome of interest. On the other hand, if under the permutation distribution the observed effect is likely, then the permuted variable does not have an effect on the outcome. Because the cohorts in the Intervention and Control groups were selected through a propensity score match, the permutation was performed in a way that maintained this matching of group members.[38]

Randomization Inference on Survival Analysis:

Randomization inference also was used to confirm the survival analysis in order to ensure that the survival difference was not because of a regression to the mean. By permuting

[34]Lamarca R, Alonso J, Gomez G, Munoz A. Left-truncated data with age as time-scale: an alternative for survival analysis in the elderly population. J.Gerontol Med Sci 1998; 53A:M337-M343.
[35] Cho H, Mariotto AB, Mann BS, Klabunde CN, Feuer EJ. Assessing non-cancer-related health status of US cancer patients: other-cause survival and comorbidity prevalence. Am J Epidemiol 2013; 178: 339-349.
[36] Sommers BD, Long SK, Baicker K. Changes in mortality after Massachusetts health care reform: a quasi-experimental study. Ann Int Med 2014; 160, 585–593.
[37] Kaestner R. Did Massachusetts health care reform lower mortality? No according to randomization inference. Statistics and Public Policy. 2016; 3(1): 1-6. (Accessed on December 15, 2016, at http://dx.doi.org/10.1080/2330443X.2015.1102667.)
[38]Rosenbaum PR (1984). Conditional permutation tests and the propensity score in observational studies. J Am Stat Assoc 1984; 79: 565-574.

the groups that a member was in many times and estimating the hazard ratio, a distribution of hazard ratios was obtained. If the hazard ratio observed is not likely under this permuted distribution, this would have to be interpreted as evidence that the group to which an individual was actually assigned is meaningful in determining the hazard. If there were regression to the mean, under the permuted distribution, the meaningful difference would happen often and therefore the permutation test p-value would be larger, indicating that the group assignation does not have an effect on the overall hazard rate.