United States Government Accountability Office

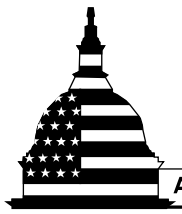
GAO

Report to Congressional Requesters

January 2012

MEDICARE ADVANTAGE

CMS Should Improve the Accuracy of Risk Score Adjustments for Diagnostic Coding Practices



G A O

Accountability * Integrity * Reliability

GAO-12-51

GAO
Accountability * Integrity * Reliability

Highlights

Highlights of GAO-12-51, a report to congressional requesters

January 2012

MEDICARE ADVANTAGE

CMS Should Improve the Accuracy of Risk Score Adjustments for Diagnostic Coding Practices

Why GAO Did This Study

The Centers for Medicare & Medicaid Services (CMS) pays plans in Medicare Advantage (MA)—the private plan alternative to Medicare fee-for-service (FFS)—a predetermined amount per beneficiary adjusted for health status. To make this adjustment, CMS calculates a risk score, a relative measure of expected health care costs, for each beneficiary. Risk scores should be the same among all beneficiaries with the same health conditions and demographic characteristics. Policymakers raised concerns that differences in diagnostic coding between MA plans and Medicare FFS could lead to inappropriately high MA risk scores and payments to MA plans. CMS began adjusting for coding differences in 2010. GAO (1) estimated the impact of any coding differences on MA risk scores and payments to plans in 2010 and (2) evaluated CMS's methodology for estimating the impact of these differences in 2010, 2011, and 2012. To do this, GAO compared risk score growth for MA beneficiaries with an estimate of what risk score growth would have been for those beneficiaries if they were in Medicare FFS, and evaluated CMS's methodology by assessing the data, study populations, study design, and beneficiary characteristics analyzed.

What GAO Recommends

GAO recommends that CMS should improve the accuracy of its MA risk score adjustments by taking steps such as incorporating adjustments for additional beneficiary characteristics, using the most current data available, accounting for all relevant years of coding differences, and incorporating the effect of coding difference trends.

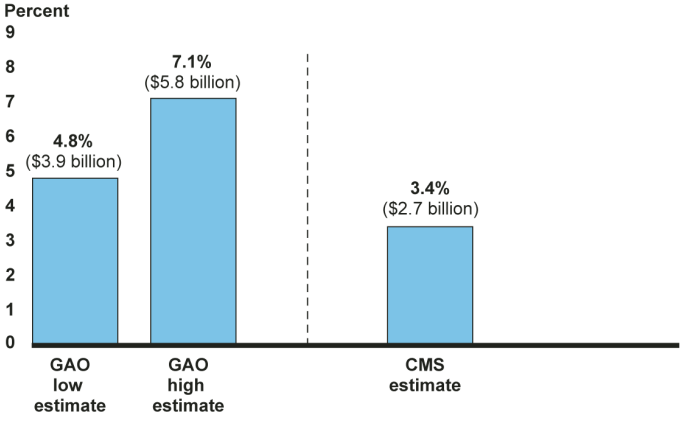
View GAO-12-51. For more information, contact James C. Cosgrove at (202) 512-7114 or cosgrove@gao.gov.

What GAO Found

GAO found that diagnostic coding differences exist between MA plans and Medicare FFS. Using data on beneficiary characteristics and regression analysis, GAO estimated that before CMS's adjustment, 2010 MA beneficiary risk scores were at least 4.8 percent, and perhaps as much as 7.1 percent, higher than they likely would have been if the same beneficiaries had been continuously enrolled in FFS. The higher risk scores were equivalent to $3.9 billion to $5.8 billion in payments to MA plans. Both GAO and CMS found that the impact of coding differences increased over time. This trend suggests that the cumulative impact of coding differences in 2011 and 2012 could be larger than in 2010.

In contrast to GAO, CMS estimated that 3.4 percent of 2010 MA beneficiary risk scores were attributable to coding differences between MA plans and Medicare FFS. CMS's adjustment for this difference avoided $2.7 billion in excess payments to MA plans. CMS's 2010 estimate differs from GAO's in that CMS's methodology did not include more current data, did not incorporate the trend of the impact of coding differences over time, and did not account for beneficiary characteristics other than age and mortality, such as sex, health status, Medicaid enrollment status, beneficiary residential location, and whether the original reason for Medicare entitlement was disability.

Percentage of 2010 MA Risk Scores Attributable to Coding Differences and Effect on Payments to MA Plans



Source: GAO analysis of Medicare data.

CMS did not update its coding adjustment estimate in 2011 and 2012 to include more current data, to account for additional years of coding differences, or to incorporate the trend of the impact of coding differences. By continuing to implement the same 3.4 percent adjustment for coding differences in 2011 and 2012, CMS likely underestimated the impact of coding differences in 2011 and 2012, resulting in excess payments to MA plans.

GAO's findings underscore the importance of both CMS continuing to adjust risk scores to account for coding differences and ensuring that those adjustments are as complete and accurate as possible.

In its comments, CMS stated that it found our findings informative. CMS did not comment on our recommendation.

Contents

Abbreviations

CMS	Centers for Medicare & Medicaid Services
CY	calendar year
ESRD	end-stage renal disease
FFS	fee-for-service
HCC	hierarchical condition category
HCERA	Health Care and Education Reconciliation Act of 2010
HMO	health maintenance organization
MA	Medicare Advantage
PFFS	private fee-for-service
PPO	preferred-provider organization
PSO	provider-sponsored organization

GAO
Accountability * Integrity * Reliability

United States Government Accountability Office
Washington, DC 20548

January 12, 2012

Congressional Requesters

In 2010, the federal government spent about $114 billion on the Medicare Advantage (MA) program, a private plan alternative to the original Medicare fee-for-service (FFS) program that covers about a quarter of all Medicare beneficiaries.[1] The Centers for Medicare & Medicaid Services (CMS), the agency that administers Medicare, pays MA plans a monthly amount to provide health care services for each beneficiary enrolled in these plans. CMS adjusts the payment to account for a beneficiary's health status, a process known as risk adjustment.[2] For example, beneficiaries in poorer health are generally expected to use more health care services relative to beneficiaries in better health. Therefore, CMS's risk adjustment tends to increase payments to those plans serving beneficiaries in poorer health to compensate for the expected higher health care spending by those plans. Risk adjustment helps ensure that a plan's financial incentive to enroll and care for beneficiaries is similar for all beneficiaries regardless of their health status or the resources they are likely to consume.

To risk adjust payments, CMS calculates a risk score for every Medicare beneficiary, including those in MA plans and the FFS program. A beneficiary's risk score is the ratio of expected health care expenditures for that beneficiary under Medicare FFS relative to the average health care expenditures for all Medicare FFS beneficiaries.[3] Information on a beneficiary's age, sex, Medicaid enrollment status, original reason for

[1]Medicare FFS consists of Medicare Parts A and B. Medicare Part A covers hospital and other inpatient stays. Medicare Part B is optional insurance and covers hospital outpatient, physician, and other services. Medicare beneficiaries have the option of obtaining coverage for Medicare Part A and B services from private health plans that participate in the MA program—also known as Medicare Part C. Medicare beneficiaries may purchase optional coverage for outpatient prescription drugs under Medicare Part D.

[2]The payment to an MA plan is based on a plan's bid—the projected revenue required by the plan to provide Medicare coverage—and a benchmark—the maximum amount Medicare will pay the plan to provide Medicare coverage in each county within the plan's service area.

[3]For example, a beneficiary with a risk score of 1.05 would have expected expenditures that were 5 percent greater than the average Medicare FFS beneficiary, who is assigned a risk score of 1.00.

Medicare entitlement (i.e., age or disability), and major medical conditions all factor into the calculation of the risk score.[4] To gather information on medical diagnoses for beneficiaries in Medicare FFS, CMS analyzes the claims that FFS providers submit for payment. For beneficiaries enrolled in MA plans, instead of submitting claims, CMS requires plans to submit certain diagnosis codes for each beneficiary.

Risk scores for beneficiaries with the same health conditions, age, and other characteristics should be identical, regardless of whether the beneficiaries are in an MA plan or Medicare FFS. This will be true if MA plans and FFS providers code medical diagnoses with the same level of reliability and completeness. However, MA plans and FFS providers may code medical diagnoses differently. Since 2004, when CMS transitioned from using only a beneficiary's principal inpatient diagnosis to using a larger set of major medical conditions to risk adjust MA payments, MA plans have had a financial incentive to ensure that all relevant diagnoses are coded, as this can increase beneficiaries' risk scores and ultimately the payment plans receive. In contrast, CMS pays many Medicare FFS providers for services provided rather than beneficiaries' diagnoses.[5] FFS providers that are paid based on services provided have less of a financial incentive to code all relevant diagnoses. If patterns of diagnostic coding differ systematically between MA plans and Medicare FFS, it is possible for beneficiaries in MA plans to be assigned higher risk scores, and appear to be sicker, than identical beneficiaries in Medicare FFS. Because payment adjustments are estimated using FFS data, higher MA risk scores due to diagnostic coding that is more comprehensive than FFS would result in MA plan payments that are too high.

Policymakers have expressed concern that risk scores for MA beneficiaries have grown at a faster rate than those for Medicare FFS beneficiaries and that systematic coding differences have contributed to

[4]See Pope et al., "Risk Adjustment of Medicare Capitation Payments Using the CMS-HCC Model," *Health Care Financing Review*, vol. 25, no. 4, 2004, pp. 119-141.

[5]One important exception is hospital acute inpatient services, for which Medicare payment is based on Medicare severity diagnosis related groups rather than services.

such growth.[6] Under the Deficit Reduction Act of 2005, CMS was required to adjust risk scores for MA beneficiaries in 2008, 2009, and 2010 to take into account differences in treatment and diagnostic coding between MA plans and Medicare FFS providers to the extent that the impact of such differences on risk scores could be identified.[7] CMS did not adjust MA risk scores in 2008 or 2009. However, for 2010, CMS estimated that 3.41 percent of MA beneficiary risk scores were attributable to differences in diagnostic coding over the previous 3 years and reduced MA beneficiaries' 2010 risk scores by 3.41 percent. This adjustment, intended to ensure that individuals with identical health conditions and other characteristics have the same risk score regardless of whether they were in an MA plan or FFS, resulted in an estimated $2.7 billion in savings to Medicare.[8]

The Health Care and Education Reconciliation Act of 2010 (HCERA) required CMS to continue adjusting risk scores for coding differences until CMS implements risk adjustment using MA diagnostic, cost, and use data.[9] CMS reduced 2011 MA beneficiary risk scores by 3.41 percent, the same amount that the agency estimated and used for 2010, and will use for 2012.[10] In addition, HCERA required CMS to reduce MA risk scores by at least 1.3 percent more than the 2010 adjustment (a total of 4.71 percent) in 2014 and that the annual minimum percentage reduction

[6]CMS estimated that from 2004 through 2006, the risk scores of beneficiaries in MA plans rose more than twice as fast as risk scores of beneficiaries in Medicare FFS, increasing an average of 4.5 percent compared to 2 percent per year, respectively. See CMS, "Announcement of Calendar Year (CY) 2008 Medicare Advantage Capitation Rates and Payment Policies," p. 16 (Apr. 2, 2007).

[7]Pub. L. No. 109-171, §5301(b), 120 Stat. 4, 51.

[8]The Medicare savings estimate is based on our analysis of Medicare data. To estimate the savings to Medicare we calculated the difference between total projected payments to MA plans with and without an adjustment for coding differences applied.

[9]CMS will begin collecting the additional data necessary for risk adjustment based on diagnostic, cost, and use data from MA plans in 2012. Pub. L. No. 111-152, §1102(e), 124 Stat. 1029, 1046 (codified at 42 U.S.C. §1395w-23(a)(1)(C)(ii)).

[10]CMS had proposed that it would reduce 2011 MA risk scores by 3.41 percent before HCERA was enacted. See CMS, "Advance Notice of Methodological Changes for Calendar Year (CY) 2011 for Medicare Advantage (MA) Capitation Rates, Part C and Part D Payment Policies and 2011 Call Letter" (Feb. 19, 2010).

gradually increase to not less than 5.70 percent in 2019 and subsequent years.[11]

The accuracy of the adjustments to risk scores can have important consequences for both Medicare spending and MA plans. If CMS does not accurately estimate the effect on MA beneficiary risk scores of coding differences between MA plans and Medicare FFS, then payments to MA plans will not accurately reflect the health status of MA beneficiaries. For example, if the adjustment to account for differences in coding is too small, then MA payments would be set too high and plans would be overpaid due to differences in coding patterns. In contrast, if the adjustment is larger than the actual impact of coding differences on risk scores, then payments to MA plans would be set too low and MA plans would be underpaid for the beneficiaries they served.

You asked us to analyze differences in diagnostic coding practices between MA and Medicare FFS and review CMS's methodology for quantifying differences in coding practices and associated payment adjustments. This report (1) determines the extent to which differences, if any, in diagnostic coding between MA plans and Medicare FFS affected risk scores and payments to MA plans in 2010; and (2) evaluates CMS's methodology for estimating the percentage of MA beneficiary risk scores in 2010, 2011, and 2012 that was attributable to differences in diagnostic coding between MA plans and Medicare FFS.

To determine the extent to which differences in diagnostic coding between MA plans and Medicare FFS affected 2010 risk scores and payments to MA plans, we compared actual risk score growth for beneficiaries in MA plans with the estimated risk score growth MA beneficiaries would have had if they were enrolled in Medicare FFS, and then estimated the impact on payments to MA plans. To do this we calculated changes in disease scores—the portion of the risk score that is based on a beneficiary's coded diagnoses—for MA beneficiaries and used regression analysis to estimate what changes in disease scores would have been if those beneficiaries were enrolled in Medicare FFS. In our regression analysis, we accounted for beneficiary characteristics that could affect disease score growth, including characteristics that may affect the frequency with which beneficiaries interact with health care

[11]42 U.S.C. §1395w-23(a)(1)(C)(ii)(III).

providers and therefore the completeness with which providers code diagnoses. We attributed differences between actual and estimated disease score growth to differences in coding practices between MA plans and Medicare FFS.[12]

We estimated the extent to which differences in diagnostic coding between MA plans and Medicare FFS affected 2010 risk scores by estimating the cumulative impact of coding differences over the 3 year period from 2007 to 2010. Our use of 2007 risk scores, based on prior year diagnoses, as the first risk scores to contribute to our cumulative coding estimate assumes that MA plans and Medicare FFS had similar coding patterns at that time.[13]

Because 2008 data were the most recent available at the time of our analysis, we projected the estimated impact of coding differences to 2010. We analyzed a retrospective cohort by using risk score data to identify MA beneficiary risk scores in 2008 and following them back to 2005.[14,15] To estimate the impact of coding differences on risk scores for 2005 to 2008, we estimated the risk score growth due to coding differences for those beneficiaries over three 2-year periods (2005 to 2006, 2006 to 2007, and 2007 to 2008). We then projected risk score growth due to coding differences for 2008 through 2010 and calculated

[12]We accounted for the following beneficiary characteristics: age, sex, diagnoses as a proxy for health status, mortality, Medicaid enrollment status, beneficiary residential location, and whether the original reason for Medicare entitlement was disability.

[13]CMS estimated the cumulative impact of coding differences on risk scores over the same period.

[14]Risk scores are based on data collected for services provided during the prior calendar year. By analyzing 2005 to 2008 risk scores, we addressed diagnoses coded during 2004 to 2007.

[15]We analyzed beneficiaries enrolled in health maintenance organization (HMO), preferred-provider organization (PPO), and private fee-for-service (PFFS) plans, as well as plans offered by provider-sponsored organizations (PSO). Coverage for beneficiaries in HMOs is generally restricted to services from providers within a network, while beneficiaries in PPOs are covered for services from both in-network and out-of-network providers but must pay higher cost-sharing amounts if they use out-of-network services. Prior to 2011, PFFS plans generally did not have provider networks, and beneficiaries were able to see any provider that accepted the plan's payment terms. However, beginning in 2011, the Medicare Improvement for Patients and Providers Act of 2008 requires most PFFS plans to have provider networks in certain areas. Pub. L. No. 110-275, § 162, 122 Stat. 2494, 2569 (codified at 42 U.S.C. § 1395w-22(d)(5)-(6)). PSOs offer MA plans with provider networks that are operated by a provider or providers.

the weighted sum of the estimated impact for 2007 to 2008 and the projections of the estimated impact for 2008 to 2010, which were based on trends from 2005 to 2008. We made two different projections for 2008 to 2010 using different assumptions of trends: the lower projection assumed that the impact of coding differences on risk scores for 2008 to 2010 was the same as it was for 2007 to 2008, while the higher projection assumed that the trend of impact on our study population from 2005 through 2008 continued through 2010. Finally, we estimated the impact of coding differences on MA risk scores when we restricted our sample of MA beneficiaries to those who were enrolled in MA plans with provider networks since these plans may be better able to influence provider coding patterns.[16]

We also performed an additional analysis to determine how sensitive our results were to our assumption that coding patterns for MA and FFS were similar in 2007. CMS believes that MA coding patterns may have been less comprehensive than FFS when the CMS-Hierarchical Condition Categories (CMS-HCC) model was first implemented, and that coding pattern differences caused MA risk scores to grow faster than FFS; therefore, there may have been a period of "catch-up" before MA coding patterns became more comprehensive than FFS coding patterns. While the length of the "catch-up" period is not known, we evaluated the impact of assuming the actual "catch-up" period was shorter, and that MA and FFS coding patterns were similar in 2005.[17]

To evaluate CMS's methodology for estimating the percentage of MA beneficiary risk scores in 2010, 2011, and 2012 that was attributable to differences in diagnostic coding between MA plans and Medicare FFS,[18] we reviewed documentation on CMS's methodology and interviewed CMS officials. We assessed the data, study population, and study design that CMS used in its calculation and examined the extent to which CMS accounted for relevant beneficiary characteristics that could affect the estimate.

[16]Plans with provider networks include HMOs, PPOs, and plans offered by PSOs.

[17]Specifically, we evaluated the impact of analyzing two additional years of coding differences by estimating the impact of coding differences from 2005 to 2010.

[18]CMS calls this percentage the Coding Pattern Difference Adjustment factor.

To quantify the impact of both our and CMS's estimates of coding differences on payments to MA plans, we estimated the risk score growth attributable to coding differences, as described above, and using data MA plans submitted to CMS that were used to determine payments to MA plans, calculated total risk-adjusted payments for each MA plan before and after applying a coding adjustment. We then calculated the difference between the two payment levels.

The CMS data we analyzed on Medicare beneficiaries are collected from Medicare providers and MA plans. We assessed the reliability of the CMS data we used by interviewing officials responsible for using these data to determine MA payments, reviewing relevant documentation, and examining the data for obvious errors. We determined that the data were sufficiently reliable for the purposes of our study. (See app. I for more details on our scope and methodology.)

We conducted this performance audit from October 2009 through December 2011 in accordance with generally accepted government auditing standards. Those standards require that we plan and perform the audit to obtain sufficient, appropriate evidence to provide a reasonable basis for our findings and conclusions based on our audit objectives. We believe that the evidence obtained provides a reasonable basis for our findings and conclusions based on our audit objectives.

Background

CMS's method of adjusting payments to MA plans to reflect beneficiary health status has changed over time. Prior to 2000, CMS adjusted MA payments based only on beneficiary demographic data. From 2000 to 2003, CMS adjusted MA payments using a model that was based on a beneficiary's demographic characteristics and principal inpatient diagnosis.[19] In 2004, CMS began adjusting payments to MA plans based on the CMS-HCC model.[20] HCCs, which represent major medical conditions, are groups of medical diagnoses where related groups of diagnoses are ranked based on disease severity and cost. The CMS-HCC model adjusts MA payments more accurately than previous models

[19]This model was called the Principal Inpatient Diagnostic Cost Group model.

[20]CMS published the details of the CMS-HCC risk adjustment model on March 28, 2003, and May 12, 2003. CMS-HCC model adjustments to MA payments were phased in from 2004 to 2010. Payments to MA plans in 2011 are adjusted solely by the CMS-HCC model.

because it includes more comprehensive information on beneficiaries' health status.

The CMS-HCC risk adjustment model uses enrollment and claims data from Medicare FFS. The model uses beneficiary characteristic and diagnostic data from a base year to calculate each beneficiary's risk scores for the following year.[21] For example, CMS used MA beneficiary demographic and diagnostic data for 2007 to determine the risk scores used to adjust payments to MA plans in 2008.

CMS estimated that 3.41 percent of 2010 MA beneficiary risk scores was attributable to differences in diagnostic coding between MA and Medicare FFS since 2007. To calculate this percentage, CMS estimated the annual difference in disease score growth between MA and Medicare FFS beneficiaries for three different groups of beneficiaries who were either enrolled in the same MA plan or in Medicare FFS from 2004 to 2005, 2005 to 2006, and 2006 to 2007. CMS accounted for differences in age and mortality when estimating the difference in disease score growth between MA and Medicare FFS beneficiaries for each period. Then, CMS calculated the average of the three estimates.[22] To apply this average estimate to 2010 MA beneficiaries,

- CMS multiplied the average annual difference in risk score growth by its estimate of the average length of time that 2010 MA beneficiaries had been continuously enrolled in MA plans over the previous 3 years,[23] and

[21]The CMS-HCC model uses one calendar year of data to estimate each beneficiary's expected Medicare expenditures for the following year. Expected Medicare expenditures for each beneficiary are divided by the average Medicare expenditures for all Medicare FFS beneficiaries to generate a risk score.

[22]The average was weighted by the number of beneficiaries enrolled in the same MA plan during each time period.

[23]CMS used MA enrollment data for MA beneficiaries in 2009 and the previous 3 years to estimate the average length of time that 2010 MA beneficiaries had been continuously in their MA plan during the previous 3 years.

- CMS multiplied this result by 81.8 percent, its estimate of the percentage of 2010 MA beneficiaries who were enrolled in an MA plan in 2009 and therefore were exposed to MA coding practices.[24]

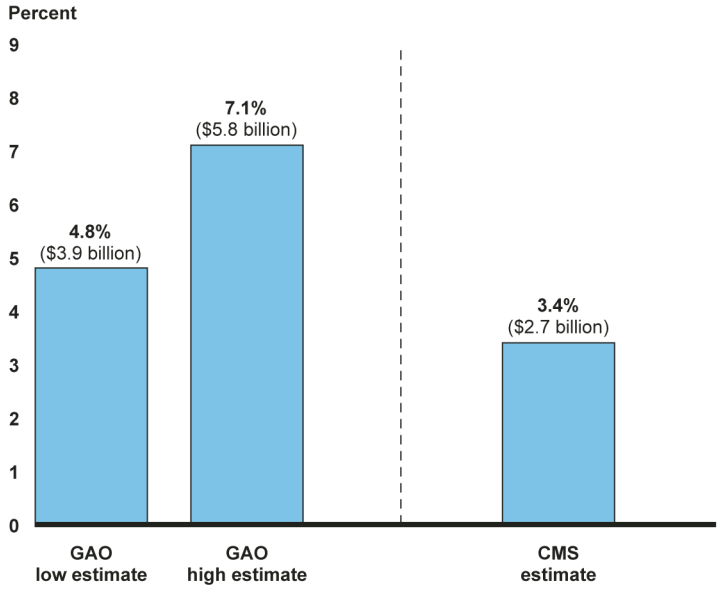
CMS implemented this same adjustment of 3.41 percent in 2011 and has announced it will implement this same adjustment in 2012.

Diagnostic Coding Differences Accounted for Estimated MA Risk Score Growth of at Least $3.9 Billion in 2010, with Likely Larger Impacts in 2011 and 2012

We found that diagnostic coding differences exist between MA plans and Medicare FFS and that these differences had a substantial effect on payment to MA plans. We estimated that risk score growth due to coding differences over the previous 3 years was equivalent to $3.9 billion to $5.8 billion in payments to MA plans in 2010 before CMS's adjustment for coding differences. Before CMS reduced 2010 MA beneficiary risk scores, we found that these scores were at least 4.8 percent, and perhaps as much as 7.1 percent, higher than the risk scores likely would have been as a result of diagnostic coding differences, that is, if the same beneficiaries had been continuously enrolled in FFS (see fig. 1). Our estimates suggest that, after accounting for CMS's 3.4 percent reduction to MA risk scores in 2010, MA risk scores were too high by at least 1.4 percent, and perhaps as much as 3.7 percent, equivalent to $1.2 billion and $3.1 billion in payments to MA plans.

[24]CMS's estimate of the percentage of 2010 MA beneficiaries whose risk scores reflected MA diagnostic coding was based on the percentage of 2009 MA beneficiaries who were also in MA plans in 2008.

Figure 1: Percentage of 2010 MA Risk Scores Attributable to Coding Differences and Effect on Payments to MA Plans



Source: GAO analysis of Medicare data.

Notes: To estimate the percentage of 2010 MA risk scores attributable to coding differences between MA and Medicare FFS over the previous 3 years, we analyzed a retrospective cohort of beneficiaries from 2005 to 2008. We used two different assumptions of the effect of coding differences on risk scores from 2008 to 2010. GAO's low estimate assumes that the percentage of risk score growth attributable to coding differences from 2008 to 2010 was the same as it was from 2007 to 2008. GAO's high estimate assumes that the percentage of risk score growth attributable to coding differences from 2008 to 2010 continues the trend for our study population from 2005 to 2008.

Our two estimates were based on different assumptions of the impact of coding differences over time. We found that the annual impact of coding differences for our study population increased from 2005 to 2008. Based on this trend, we projected risk score growth for the period 2008 to 2010 and obtained the higher estimate, 7.1 percent, of the cumulative impact of differences in diagnostic coding between MA and FFS. However, coding differences may reach an upper bound when MA plans code diagnoses as comprehensively as possible, so we produced the lower estimate of 4.8 percent by assuming that the impact of coding differences on risk scores remained constant and was the same from 2008 to 2010 as it was from 2007 to 2008.[25]

[25]See app. I for more detail on our methodology.

Plans with networks may have greater potential to influence the diagnostic coding of their providers, relative to plans without networks. Specifically, when we restricted our analysis to MA beneficiaries in plans with provider networks (HMOs, PPOs, and plans offered by PSOs), our estimates of the cumulative effect of differences in diagnostic coding between MA and FFS increased to an average of 5.5 or 7.8 percent of MA beneficiary risk scores in 2010, depending on the projection assumption for 2008 to 2010.[26]

Altering the year by which MA coding patterns had "caught up" to FFS coding patterns, from our original assumption of 2007 to 2005, had little effect on our results. Specifically, we estimated the cumulative impact of coding differences from 2005 to 2010 and found that our estimates for all MA plans increased slightly to 5.3 or 7.6 percent, depending on the projection assumption from 2008 to 2010.[27]

Our analysis estimating the cumulative impact of coding differences on 2010 MA risk scores suggests that this cumulative impact is increasing. Specifically, we found that from 2005 to 2008, the impact of coding differences on MA risk scores increased over time (see app. 1, table 1). Furthermore, CMS also found that the impact of coding differences increased from 2004 to 2008.[28] While we did not have more recent data, the trend of coding differences through 2008 suggests that the impact of coding differences in 2011 and 2012 could be larger than in 2010.

[26]Prior to 2011, PFFS plans were not required to have a network; however, beginning in 2011, PFFS plans in certain areas were required to have a provider network. In 2011, 72 percent PFFS enrollees were in counties where PFFS plans were required to have a network.

[27]We found the cumulative impact of coding differences from 2005 to 2010 for plans with provider networks (HMOs, PPOs, and PSOs) to be 6.1 or 8.4 percent of MA beneficiary risk scores in 2010, depending on the projection assumption from 2008 to 2010.

[28]CMS analysis provided to us showed annual risk score growth due to coding differences to be 0.015 from 2004 to 2005, 0.015 from 2005 to 2006, 0.026 from 2006 to 2007, and 0.038 from 2007 to 2008.

CMS's Adjustment for Coding Differences Likely Resulted in Excess Payments to MA Plans

CMS's estimate of the impact of coding differences on 2010 MA risk scores was smaller than our estimate due to the collective impact of three methodological differences described below. For its 2011 and 2012 adjustments, the agency continued to use the same estimate of the impact of coding differences it used in 2010, which likely resulted in excess payments to MA plans.

Three major differences between our and CMS's methodology account for the differences in our 2010 estimates. First, CMS did not include data from 2008. CMS initially announced the adjustment for coding differences in its advance notice for 2010 payment before 2008 data were available. While 2008 data became available prior to the final announcement of the coding adjustment, CMS decided not to incorporate 2008 data into its final adjustment. In its announcement for 2010 payment, CMS explains that it took a conservative approach for the first year that it implemented the MA coding adjustment. Incorporating 2008 data would have increased the size of CMS's final adjustment. Second, CMS did not take into account the increasing impact of coding differences over time. However, without 2008 data, the increasing trend of the annual impact of coding differences is less apparent, and supports the agency's decision to use the average annual impact from 2004 to 2007 as a proxy for the annual impact from 2007 to 2010. Third, CMS only accounted for differences in age and mortality between the MA and FFS study populations. We found that accounting for additional beneficiary characteristics explained more variation in disease score growth, and consequently improved the accuracy of our risk score growth estimate.[29,30]

CMS did not update its estimate in 2011 and 2012 with more current data, even though data were available. CMS did not include 2008 data in its 2010 estimate due to its desire to take a conservative approach for the first year it implemented a coding adjustment, and the agency did not

[29]Specifically, our model explained less than 1 percent of the variation in disease score growth when we accounted only for differences in age and mortality (the only two factors that CMS included); however, our model explained about 20 percent of the variation when we also accounted for additional characteristics, including: sex, diagnoses as a proxy for health status, Medicaid enrollment status, beneficiary residential location, and whether the original reason for Medicare entitlement was disability.

[30]We also assessed the impact of including only MA beneficiaries who remained in the same plan for each time period, as CMS did in its analysis, as opposed to including all MA beneficiaries and found that this methodological difference had little impact on our estimates.

update its estimate for 2011 or 2012 due to concerns about the many MA payment changes taking place. While maintaining the same level of adjustment for 2011 and 2012 maintains stability and predictability in MA payment rates, it also allows the accuracy of the adjustment to diminish in each year. Including more recent data would have improved the accuracy of CMS's 2011 and 2012 estimates because more recent data are likely to be more representative of the year in which an adjustment was made.

By not updating its estimate with more current data, CMS also did not account for the additional years of cumulative coding differences in its estimate: 4 years for 2011 (2007 to 2011) and 5 years for 2012 (2007 to 2012). While CMS stated in its announcement for 2011 payment that it would consider accounting for additional years of coding differences, CMS officials told us they were concerned about incorporating additional years using a linear methodology because it would ignore the possibility that MA plans may reach a limit at which they could no longer code diagnoses more comprehensively. We think it is unlikely that this limit has been reached. Given the financial incentives that MA plans have to ensure that all relevant diagnoses are coded, the fact that CMS's 3.41 percent estimate is below our low estimate of 4.8 percent, and considering the increasing use of electronic health records to capture and maintain diagnostic information, the upper limit is likely to be greater than the 3 years CMS accounted for in its 2011 and 2012 estimates.

In addition to not including more recent data, CMS did not incorporate the impact of the upward trend in coding differences on risk scores into its estimates for 2011 and 2012. Based on the trend of increasing impact of coding differences through 2008, shown in both CMS's and our analysis, we believe that the impact of coding differences on 2011 and 2012 MA risk scores is likely to be larger than it was on 2010 MA risk scores. In addition, less than 1.4 percent of MA enrollees in 2011 were enrolled in a plan without a network, suggesting that our slightly larger results based on only MA plans with a network are more accurate estimates of the impact of coding differences in 2011 and 2012. By continuing to implement the same 3.41 percent adjustment for coding differences in 2011 and 2012, we believe CMS likely substantially underestimated the impact of coding differences in 2011 and 2012, resulting in excess payments to MA plans.

Conclusions

Risk adjustment is important to ensure that payments to MA plans adequately account for differences in beneficiaries' health status and to maintain plans' financial incentive to enroll and care for beneficiaries regardless of their health status or the resources they are likely to consume. For CMS's risk adjustment model to adjust payments to MA plans appropriately, diagnostic coding patterns must be similar among both MA plans and Medicare FFS. We confirmed CMS's finding that differences in diagnostic coding caused risk scores for MA beneficiaries to be higher than those for comparable Medicare FFS beneficiaries in 2010. This finding underscores the importance of continuing to adjust MA risk scores to account for coding differences and ensuring that these adjustments are as accurate as possible. If an adjustment for coding differences is too low, CMS would pay MA plans more than it would pay providers in Medicare FFS to provide health care for the same beneficiaries. We found that CMS's 3.41 percent adjustment for coding differences in 2010 was too low, resulting in $1.2 billion to $3.1 billion in payments to MA plans for coding differences. By not updating its methodology in 2011 or in 2012, CMS likely underestimated the impact of coding differences on MA risk scores to a greater extent in these years, resulting in excess payments to MA plans. If CMS does not update its methodology, excess payments due to differences in coding practices are likely to increase.

Recommendations for Executive Action

To help ensure appropriate payments to MA plans, the Administrator of CMS should take steps to improve the accuracy of the adjustment made for differences in diagnostic coding practices between MA and Medicare FFS. Such steps could include, for example, accounting for additional beneficiary characteristics, including the most current data available, identifying and accounting for all years of coding differences that could affect the payment year for which an adjustment is made, and incorporating the trend of the impact of coding differences on risk scores.

Agency Comments and Our Evaluation

CMS provided written comments on a draft of this report, which are reprinted in appendix II.

In its comments, CMS stated that it found our methodological approach and findings informative and suggested that we provide some additional information about how the coding differences between MA and FFS were calculated. In response, we added additional details to appendix I about the regression models used, the calculations used to generate our

cumulative impact estimates, and the trend line used to generate our high estimate.

CMS did not comment on our recommendation for executive action.

As agreed with your offices, unless you publicly announce the contents of this report earlier, we plan no further distribution until 30 days from the report date. At that time, we will send copies to the Secretary of HHS, interested congressional committees, and others. In addition, the report is available at no charge on the GAO website at http://www.gao.gov.

If you or your staff has any questions about this report, please contact me at (202) 512-7114 or cosgrovej@gao.gov. Contact points for our Offices of Congressional Relations and Public Affairs may be found on the last page of this report. GAO staff who made major contributions to this report are listed in appendix III.

James C. Cosgrove
Director, Health Care

List of Requesters

The Honorable Henry A. Waxman
Ranking Member
Committee on Energy and Commerce
House of Representatives

The Honorable Frank Pallone, Jr.
Ranking Member
Subcommittee on Health
Committee on Energy and Commerce
House of Representatives

The Honorable Pete Stark
Ranking Member
Subcommittee on Health
Committee on Ways and Means
House of Representatives

The Honorable John D. Dingell
The Honorable Charles B. Rangel
House of Representatives

Appendix I: Scope and Methodology

This appendix explains the scope and methodology that we used to address our objective that determines the extent to which differences, if any, in diagnostic coding between Medicare Advantage (MA) plans and Medicare fee-for-service (FFS) affect risk scores and payments to MA plans in 2010.

Estimating the Impact on MA Risk Scores

To determine the extent to which differences, if any, in diagnostic coding between MA plans and Medicare FFS affected MA risk scores in 2010, we used Centers for Medicare & Medicaid Services (CMS) enrollment and risk score data from 2004 to 2008, the most current data available at the time of our analysis, and projected the estimated impact to 2010. For three periods (2005 to 2006, 2006 to 2007, and 2007 to 2008), we compared actual risk score growth for beneficiaries in our MA study population with the estimated risk score growth the beneficiaries would have had if they were enrolled in Medicare FFS. Risk scores for a given calendar year are based on beneficiaries' diagnoses in the previous year, so we identified our study population based on enrollment data for 2004 through 2007 and analyzed risk scores for that population for 2005 through 2008.

Our MA study population consisted of a retrospective cohort of MA beneficiaries. We included MA beneficiaries who were enrolled in health maintenance organization (HMO), preferred provider organization (PPO), and private fee-for-service (PFFS) plans as well as plans offered by provider-sponsored organizations (PSO). Specifically, we identified the cohort of MA beneficiaries who were enrolled in MA for all of 2007 and followed them back for the length of their continuous enrollment to 2004. In addition, for beneficiaries who were enrolled in Medicare FFS and switched to MA in 2005, 2006, or 2007, we included data for 1 year of Medicare FFS enrollment immediately preceding their MA enrollment.[1] Our MA study population included three types of beneficiaries, each of which we analyzed separately for each period:

[1]We included 1 year of FFS data for beneficiaries who were enrolled in FFS in 2004 and MA in 2005 to 2007; in FFS in 2005 and MA in 2006 to 2007; and FFS in 2006 and MA in 2007. By including 1 year of baseline FFS data in our study period for MA beneficiaries who had been enrolled in FFS prior to joining an MA plan, we were able to analyze the impact of coding differences for MA beneficiaries during their first year in an MA plan.

- MA joiners—beneficiaries enrolled in Medicare FFS for the entire first year of each period and then enrolled in MA for all of the following year,
- MA plan stayers—beneficiaries enrolled in the same MA plan for the first and second year of the period, and
- MA plan switchers—beneficiaries enrolled in one MA plan for the first year of the period and a second MA plan in the following year.

Our control population consisted of a retrospective cohort of FFS beneficiaries who were enrolled in FFS for all of 2007 and 2006. We followed these beneficiaries back to 2004 and included data for all years of continuous FFS enrollment. For both the study and control populations, we excluded data for years during which a beneficiary (1) was diagnosed with end-stage renal disease (ESRD) during the study year; (2) resided in a long-term care facility for more than 90 consecutive days; (3) died prior to July 1, 2008; (4) resided outside the 50 United States; Washington, D.C.; and Puerto Rico; or (5) moved to a new state or changed urban/rural status.

We calculated the actual change in disease score—the portion of the risk score that is based on a beneficiary's coded diagnoses—for the MA study population for the following three time periods (in payment years): 2005 to 2006, 2006 to 2007, and 2007 to 2008.[2] To estimate the change in disease scores that would have occurred if those MA beneficiaries were enrolled continuously in FFS, we used our control population to estimate a regression model that described how beneficiary characteristics

[2]We calculated disease scores using the 2007 version of the CMS-Hierarchical Condition Category (CMS-HCC) risk adjustment community model (used for payment in 2007 and 2008), and summing the appropriate coefficients for each of the HCC variables. We normalized disease scores for each year to 2005 by using the FFS normalization factor that CMS used to normalize risk scores in 2008. Normalization keeps the average Medicare FFS risk score constant at 1.0 over time and is necessary to compare disease scores across years.

influenced change in disease score.[3] In the regression model we used change in disease score (year 2 - year 1) as our dependent variable and included age, sex, hierarchical condition categories (HCC), HCC interaction variables, Medicaid status, and original reason for Medicare entitlement was disability as independent variables as they are specified in the CMS-HCC model. We also included one urban and one rural variable for each of the 50 United States; Washington, D.C.; and Puerto Rico as independent variables to identify beneficiary residential location.[4,5] Then we used these regression models and data on beneficiary characteristics for our MA study population to estimate the change in disease scores that would have occurred if those MA beneficiaries had been continuously enrolled in FFS.[6]

We identified the difference between the actual and estimated change in disease scores as attributable to coding differences between MA and FFS because the regression model accounted for other relevant factors affecting disease score growth (see table 1). To convert these estimates of disease score growth due to coding differences into estimates of the impact of coding differences on 2010 MA risk scores, we divided the disease score growth estimates by the average MA risk score in 2010. Because 2010 risk scores were not available at the time we conducted our analysis, we calculated the average MA community risk score for the most recent data available (risk score years 2005 through 2008) and projected the trend to 2010 to estimate the average 2010 MA risk score.

[3]The regression model explained 22.05 percent of the variation (adjusted R-squared) in disease scores when it was run on 2005-2006 data. It explained 22.79 percent of the variation when run on 2006-2007 data, and 18.67 percent when run on 2007-2008 data. In all three models, nearly all of the independent variables in the regression were statistically significant at the 5 percent level. We also performed an additional analysis to determine how sensitive our results were to the variables we accounted for. Specifically, we evaluated the impact on our results of only accounting for age and mortality.

[4]Beneficiary residential location is a proxy for other factors that vary with geography and that may affect the frequency with which beneficiaries interact with health care providers and therefore the completeness with which providers code diagnoses, such as physician practice patterns.

[5]Except for rural variables for Washington, D.C.; New Jersey; and Rhode Island because these locations are entirely urban.

[6]Our analysis also accounted for mortality by requiring all beneficiaries in our study populations to be alive through July 1, 2008.

Table 1: Annual Risk Score Growth Due to Coding Differences for GAO Study Population

Period	MA joiners	MA plan stayers	MA plan switchers	All MA beneficiaries
2005-2006	0.0079	-0.0086	-0.0080	-0.0082
2006-2007	-0.0027	0.0211	0.0288	0.0200
2007-2008	0.0122	0.0253	0.0330	0.0249

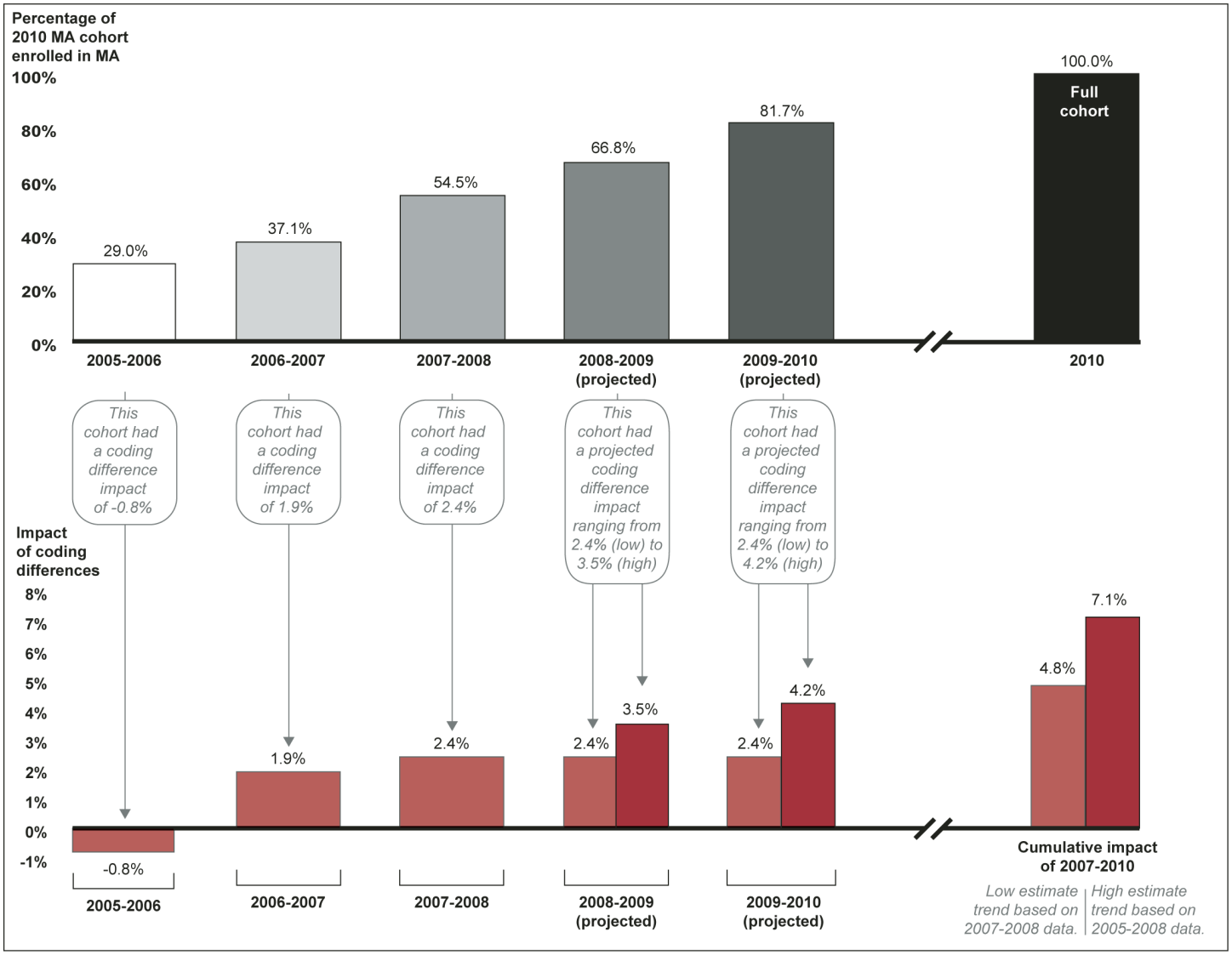
Source: GAO.

Notes: We analyzed a retrospective cohort of beneficiaries from 2005 to 2008 to estimate the impact of coding differences between MA and Medicare FFS on MA risk scores. MA joiners are beneficiaries enrolled in Medicare FFS for the entire first year of each period and then enrolled in MA for all of the following year, MA plan stayers are beneficiaries enrolled in the same MA plan for the first and second year of a given period, and MA plan switchers are beneficiaries enrolled in one MA plan for the first year of a time period and a second MA plan in the following year.

We projected these estimates of the annual impact of coding difference on 2010 risk scores through 2010 using two different assumptions. One projection assumed that the annual impact of coding differences on risk scores was the same from 2008 to 2010 as it was from 2007 to 2008. The other projection assumed that the trend of increasing coding difference impact over 2005 to 2008 continued through 2010 (see fig. 2).[7]

[7]For the latter projection, we fit a log-linear trend line to 2005-2006, 2006-2007, and 2007-2008 impact estimates and used the resulting expression to extrapolate impact estimates to 2008-2009 and 2009-2010. We used the following coordinates (annual impact, period) from table 1 for all MA beneficiaries to estimate the model: (-0.0082, 1), (0.0200, 2), and (0.0249, 3).

Figure 2: Annual Impact of Coding Differences on 2010 MA Risk Scores for GAO's Study Population, 2005 to 2010



Source: GAO.

Notes: We analyzed a cohort of beneficiaries from 2005 to 2008 to estimate the impact of coding differences between MA and Medicare FFS on MA risk scores. We used two different assumptions of the effect of coding differences on risk scores from 2008 to 2010. GAO's low estimate assumes that the percentage of risk score growth attributable to coding differences from 2008 to 2010 was the same as it was from 2007 to 2008. GAO's high estimate assumes that the percentage of risk score growth attributable to coding differences from 2008 to 2010 continues the trend from 2005 to 2008. To calculate the cumulative impact of coding differences on MA risk scores for 2007 through 2010, we summed the annual impact estimates for that period and adjusted each impact estimate to account for beneficiaries who disenrolled from the MA program before 2010.

To calculate the cumulative impact of coding differences on MA risk scores for 2007 through 2010, we summed the annual impact estimates for that period and adjusted each impact estimate to account for beneficiaries who disenrolled from the MA program before 2010.[8] The result is the cumulative impact of coding differences from 2007 to 2010 on MA risk scores in 2010.[9] We separately estimated the cumulative impact of coding differences from 2007 to 2010 on MA risk scores in 2010 for beneficiaries in MA plans with provider networks (HMOs, PPOs, and PSOs) because such plans may have a greater ability to affect provider coding patterns.

We also performed an additional analysis to determine how sensitive our results were to our assumption that coding patterns for MA and FFS were similar in 2007. CMS believes that MA coding patterns may have been less comprehensive than FFS when the CMS-HCC model was implemented, and that coding pattern differences caused MA risk scores to grow faster than FFS; therefore, there may have been a period of "catch-up" before MA coding patterns became more comprehensive than FFS coding patterns. While the length of the "catch-up" period is not known, we evaluated the impact of assuming the actual "catch-up" period was shorter, and that MA and FFS coding patterns were similar in 2005. Specifically, we evaluated the impact of analyzing two additional years of coding differences by estimating the impact of coding differences from 2005 to 2010.

[8]For 2006 and 2007, we used the actual disenrollment rates from our retrospective cohort of MA beneficiaries, while for 2008, 2009, and 2010 we used an annual disenrollment rate of 18.3 percent. To calculate our low and high estimates, we summed the annual impact estimates for 2007 to 2008, 2008 to 2009, and 2009 to 2010, each weighted by the percent of the 2010 MA cohort enrolled in that time period (see fig. 2):

GAO's Low Estimate: 4.8 % = (54.5 % x 2.4 %) + (66.8 % x 2.4 %) + (81.7 % x 2.4 %)

GAO's High Estimate: 7.1 % = (54.5 % x 2.4 %) + (66.8 % x 3.5 %) + (81.7 % x 4.2 %)

Weighted annual estimates may not sum to cumulative estimates due to rounding.

[9]Our use of 2007 risk scores, based on prior year diagnoses, as the first risk scores to contribute to our cumulative coding estimate assumes that MA plans and Medicare FFS had similar coding patterns at this time. CMS estimated the cumulative impact of coding differences on risk scores over the same period.

Estimating the Impact on Payments to MA Plans in 2010

To quantify the impact of both our and CMS's estimates of coding differences on payments to MA plans in 2010, we used data on MA plan bids—plans' proposed reimbursement rates for the average beneficiary—which are used to determine payments to MA plans. We used these data to calculate total risk-adjusted payments for each MA plan before and after applying a coding adjustment, and then used the differences between these payment levels to estimate the percentage reduction in total projected payments to MA plans in 2010 resulting from adjustments for coding differences.[10] Then we applied the percentage reduction in payments associated with each adjustment to the estimated total payments to MA plans in 2010 of $112.8 billion and accounted for reduced Medicare Part B premium payments received by CMS, which offset the reduction in MA payments (see table 2).[11]

Table 2: Impact of Adjustments for Coding Differences on Total Payments to MA Plans in 2010

	Reduction in MA payments in 2010	
Adjustment applied to reduce MA risk scores in 2010 (source)	**Percentage**	**Dollars**
3.4 percent (CMS)	2.4	2.7 billion
4.8 percent (GAO)[a]	3.4	3.9 billion
7.1 percent (GAO)[b]	5.2	5.8 billion

Source: GAO analysis of Medicare data.

Notes: We analyzed a retrospective cohort of beneficiaries from 2005 to 2008 to estimate the impact of coding differences on MA risk scores and used two different assumptions of the effect of coding differences on risk scores from 2008 to 2010. The percentage reduction in 2010 MA payments is less than the adjustment applied to 2010 MA risk scores because the impact of the adjustment to risk scores is reduced by additional payments some MA plans are eligible to receive.

[a]GAO low estimate assumes the annual impacts from 2008 to 2010 are the same as the impact from 2007 to 2008.

[b]GAO high estimate assumes the annual impacts from 2008 to 2010 continue the trend of increasing annual impacts from 2005 to 2008.

[10]We assumed that MA plans did not adjust their bids in 2010 as a result of the adjustment for coding differences.

[11]We estimated $112.8 billion to be the total payments to MA plans without adjustments CMS made in 2010 for budget neutrality and for coding differences. Each estimate in table 2 does not incorporate the impact of CMS's 2010 adjustment. All estimates of the dollar impact of the adjustment for coding differences account for an 11.73 percent offset due to reduced Medicare Part B premiums received by Medicare, and do not include Medicare savings for a small number of beneficiaries with ESRD whose risk scores were adjusted for coding differences.

The CMS data we analyzed on Medicare beneficiaries are collected from Medicare providers and MA plans. We assessed the reliability of the CMS data we used by interviewing officials responsible for using these data to determine MA payments, reviewing relevant documentation, and examining the data for obvious errors. We determined that the data were sufficiently reliable for the purposes of our study.

Appendix II: Comments from the Centers for Medicare & Medicaid Services

DEPARTMENT OF HEALTH & HUMAN SERVICES | OFFICE OF THE SECRETARY

Assistant Secretary for Legislation
Washington, DC 20201

DEC 29 2011

James Cosgrove
Director, Health Care
U.S. Government Accountability Office
441 G Street NW
Washington, DC 20548

Dear Mr. Cosgrove:

Attached are comments on the U.S. Government Accountability Office's (GAO) draft report entitled, "MEDICARE ADVANTAGE: CMS Should Improve the Accuracy of Risk Score Adjustments for Diagnostic Coding Practices" (GAO 12-51).

The Department appreciates the opportunity to review this report before its publication.

Sincerely,

Jim R. Esquea
Assistant Secretary for Legislation

Attachment

GENERAL COMMENTS OF THE DEPARTMENT OF HEALTH AND HUMAN SERVICES (HHS) ON THE GOVERNMENT ACCOUNTABILITY OFFICE'S (GAO) DRAFT REPORT ENTITLED, "MEDICARE ADVANTAGE: THE CENTERS FOR MEDICARE & MEDICAID SERVICES (CMS) SHOULD IMPROVE THE ACCURACY OF RISK SCORE ADJUSTMENTS FOR DIAGNOSTIC CODING PRACTICES" (GAO-12-51)

The Department appreciates the opportunity to review and comment on this draft report, which examines the extent of diagnostic coding differences between Medicare Advantage (MA) and fee-for-service (FFS), and its impact on MA payment.

While GAO used a different methodology than the Centers for Medicare and Medicaid Services (CMS) in calculating the impact of coding differences and arrived at similar results, we found its methodological approach and findings informative.

In describing its methodology, CMS recommends that GAO provide some additional information to better convey how the coding differences between MA and FFS were calculated. For example, where the draft discusses the factors that influence coding differences, it would be helpful to document the results of the regression model used to estimate the effect of different beneficiary characteristics on disease score change. It would also be helpful if GAO shared the calculations underlying the 4.8 percent and 7.1 percent estimates of 2010 coding differences (for the low estimate of 4.8 percent the draft starts with an estimated 2.4 percent difference for 2007-2008 and then applies this difference to 2008-2009 and 2009-2010), and show the adjusted numbers that were used to calculate the 4.8 percent. Similarly, it would be beneficial to understand the trend that was calculated and how it was used in calculating the draft's high estimate of 7.1 percent.

Appendix III: GAO Contact and Staff Acknowledgments

GAO Contact

James C. Cosgrove, (202) 512-7114 or cosgrovej@gao.gov

Staff Acknowledgments

In addition to the contact named above, Christine Brudevold, Assistant Director; Alison Binkowski; William Black; Andrew Johnson; Richard Lipinski; Elizabeth Morrison; and Merrile Sing made key contributions to this report.