How Do the Experiences of Medicare Beneficiary Subgroups Differ between Managed Care and Original Medicare?

Marc N. Elliott, Amelia M. Haviland, Nate Orr, Katrin Hambarsoomian, and Paul D. Cleary

Objective. To examine whether disparities in health care experiences of Medicare beneficiaries differ between managed care (Medicare Advantage [MA]) and traditional fee-for-service (FFS) Medicare.

Data Sources. 132,937 MA and 201,444 FFS respondents to the 2007 Medicare Consumer Assessment of Health Care Providers and Systems (CAHPS) survey.

Study Design. We defined seven subgroup characteristics: low-income subsidy eligible, no high school degree, poor or fair self-rated health, age 85 and older, female, Hispanic, and black. We estimated disparities in CAHPS experience of care scores between each of these groups and beneficiaries without those characteristics within MA and FFS for 11 CAHPS measures and assessed differences between MA and FFS disparities in linear models.

Principal Findings. The seven subgroup characteristics had significant ($p<.05$) negative interactions with MA (larger disparities in MA) in 27 of 77 instances, with only four significant positive interactions.

Conclusion. Managed care may provide less uniform care than FFS for patients; specifically there may be larger disparities in MA than FFS between beneficiaries who have low incomes, are less healthy, older, female, and who did not complete high school, compared with their counterparts. There may be potential for MA quality improvement targeted at the care provided to particular subgroups.

Key Words. Managed care, vulnerable populations, patient experience, Medicare, CAHPS

The Institute of Medicine and the National Academy of Sciences have identified the reduction of disparities in health care quality by race, ethnicity, and socioeconomic status as an important priority of U.S. Health Policy (Institute of Medicine 2002; National Research Council 2004). Seniors and the disabled have high health care needs and may be particularly susceptible to low-quality health care, so examination of disparities among Medicare beneficiaries is of particular interest.

Medicare beneficiaries can choose the fee-for-service (FFS) program, in which a majority of beneficiaries are enrolled, either with or without a free-standing prescription drug plan (Part D). Alternatively, beneficiaries can enroll in a Medicare Advantage (MA) plan, the managed care form of Medicare whereby several hundred managed care plans are under contract with Centers for Medicare and Medicaid Services, with or without a prescription drug option. These various Medicare options have basic structural differences that could result in quite different experiences for beneficiaries obtaining care within each option. For example, in FFS Medicare, beneficiaries are free to choose their own physicians and other providers, but they also are solely responsible for navigating the health care system. Enrollees in the MA plans may have broader benefits at lower costs, but they may be constrained in their provider choices and health care options. Further, given the diversity of the MA plan offerings, beneficiaries may have quite different experiences depending on the plans they choose. In what follows, we use "FFS" to refer to those beneficiaries not enrolled in an MA plan, irrespective of whether they belong to a freestanding prescription drug plan.

Characteristics such as lower socioeconomic status (including income and educational attainment), poor health status, older age, black race, and Hispanic ethnicity have been associated with problems related to health care insurance and access (Shi 2000), and as such they can be considered "vulnerable" (Blumenthal et al. 1995). In this study we analyze these characteristics, as well as gender, and assess the extent to which they are associated with relatively more negative experiences in MA relative to FFS. Evidence of disparities in both clinical process and patient experiences of care are well documented in the Medicare population (Morales et al. 2001; Institute of Medicine 2002; Lurie et al. 2003; Weech-Maldonado et al. 2003; Trivedi et al. 2005; Trivedi et al. 2006; Fongwa et al. 2008; Weech-Maldonado et al. 2008; Goldstein et al. 2010).

There is evidence that disparities in health care for Medicare beneficiaries may vary by plan and coverage type. For example, 2002–2004 black versus white disparities in HEDIS measures have been shown to vary by MA plan (Trivedi et al. 2006). Analyses of data from the Medicare Consumer Assessment of Health Care Providers and Systems (CAHPS) survey have

Address correspondence to Marc N. Elliott, Ph.D., RAND Corporation, 1776 Main Street, PO Box 2138, Santa Monica, CA 90407; e-mail: elliott@rand.org. Amelia M. Haviland, Ph.D., is with the RAND Corporation, Pittsburgh, PA. Nate Orr, M.A., and Katrin Hambarsoomian, M.S., are with the RAND Corporation, Santa Monica, CA. Paul D. Cleary, Ph.D., is with the Yale School of Public Health, New Haven, CT.

found more negative patient experience reports in MA plans than in FFS Medicare (Landon et al. 2004), especially for unhealthy beneficiaries (Keenan et al. 2009). These findings raise the question of whether there may systematically be greater disparities in care in MA than FFS across a range of other beneficiary characteristics indicating potential vulnerability. Findings regarding differential disparities in access to care are mixed; Balsa et al. (2007) found few differences in black–white disparities in access for Medicare beneficiaries in 1996–2001 data between managed care and FFS, but they found evidence of smaller Hispanic-white disparities in access in managed care.

   In the present manuscript we explore the extent to which disparities differ between MA and FFS coverage across a broader set of beneficiary characteristics. We also propose two hypotheses for why there might be greater health care disparities in MA than in FFS, and we use varying predictions across beneficiary characteristics to evaluate these hypotheses.

   The *financial incentive* hypothesis holds that managed care plans, financially incentivized to recruit less costly beneficiaries, create environments that provide relatively positive experiences for lower-costs groups, and relatively less positive experiences for high-costs groups when compared with FFS. Frank et al. (2000) and Keenan et al. (2010) describe possible managed care responses to these financial incentives in terms of underprovision of care to avoid "bad risk," for example, sicker patients, and overprovision of services used to treat the less-seriously ill, thereby attracting "good risk." Accordingly, MA plans may emphasize preventative or active wellness programs that may be more appealing to healthy beneficiaries. Managed care plans might also emphasize lower-cost services, such as physical therapy, and promote access to nutritionists and pharmacists rather than physicians. While these might be effective approaches for those in good health, they may not satisfy the needs of sicker patients. This hypothesis would predict larger disparities in MA than FFS for subgroups with higher expected average costs, such as older beneficiaries, less healthy beneficiaries, and women (Woolhandler and Himmelstein 2007; Owens 2008).

   The *health literacy* hypothesis holds that care within MA varies more across the identified subgroups than it does in FFS because on average, more negotiation is required in MA than FFS for patients to obtain services. That is, even when there are standard procedures for obtaining services for all patients, following those procedures and/or obtaining exemptions to restrictions may require more skill in MA than FFS. We posit that such skills are related to literacy and patient activation, which vary with characteristics such as age, education, race/ethnicity, and English fluency, and which have been strongly

linked to patient experience (Heller et al. 2009; Martin et al. 2009). Thus, there should be more variability across groups in MA than in FFS. Under this hypothesis we would expect to see greater health disparities for groups with lower average health literacy, groups with lower educational attainment, lower income, blacks, and Hispanics. Because women typically have higher health literacy and patient activation than men (Heller et al. 2009; Martin et al. 2009), this hypothesis predicts relatively *better* experiences for women than men within MA than within FFS, unlike the financial incentive hypothesis.

In this study, we investigate disparities in beneficiary assessments of their health care using data from a recent (2007) CAHPS survey of Medicare beneficiaries. The Medicare CAHPS survey is the primary means of assessing care experiences of the 44 million Medicare beneficiaries (Goldstein et al. 2001). The 2007 data include measures of beneficiary experiences with their general health care and health insurance (Medicare Part C) as well as measures of beneficiary experiences with prescription drug coverage (Medicare Part D). We assess the extent to which the earlier finding of greater disparities by health status in MA than in FFS persists and whether there are similar patterns for other groups of potentially vulnerable beneficiaries defined by characteristics such as age, gender, race, and socioeconomic status.

METHODS

The sample frame was all noninstitutionalized Medicare beneficiaries as of October 2006, with minor exclusions of small or ineligible plans.

Our analyses used data from the 132,960 MA beneficiaries and 202,289 FFS beneficiaries who responded to the 2007 Medicare CAHPS survey. The Medicare CAHPS survey was administered in English and Spanish by mail, with bilingual telephone follow-up for nonrespondents.

Dependent Variables: 11 CAHPS Measures of Patient Experience

We analyzed 11 CAHPS scores: all five 0–10 global ratings of care, five composite scores of patient experience derived from multiple report items, and one separate report item. Three of these measures (one rating and two composites) assessed experiences with prescription drug coverage for MA and FFS beneficiaries with Part D coverage. Questions were answered by the subset of beneficiaries to whom they were applicable, with screener items assessing eligibility. Composites were scored as the average of nonmissing items for each individual. The reliability and validity of Part D composites are described

in Martino et al. (2009); corresponding information on the remaining composites appears in Hays et al. (1999) and Hargraves, Hays, and Cleary (2003). These 11 measures are described in the Appendix SA2.

Subgroup Characteristics

We analyzed subgroups defined by seven characteristics: (1) *low-income* beneficiaries eligible for a low-income subsidy (LIS), that is, those whose income is < 150 percent of the Federal Poverty Level; (2) *low-education* beneficiaries with no high school degree or equivalent; (3) *less healthy* beneficiaries who rated their health as poor or fair (as opposed to good, very good, or excellent); (4) *older* beneficiaries who are age 85 and older (compared with beneficiaries aged 65–84, but not to the primarily disabled beneficiaries under age 65); (5) *black* beneficiaries (compared with non-Hispanic white beneficiaries); (6) *Hispanic* beneficiaries (compared with non-Hispanic white beneficiaries); and (7) *female* beneficiaries. Subsequently, we refer to these seven characteristics as defining vulnerable subgroups, although women are not typically considered a subgroup vulnerable to poor health care.

Covariates for Case-Mix Adjustment

Because other beneficiary characteristics are also known to be associated with response tendencies and may differ between the MA and FFS populations, we also include self-rated mental health, dual eligibility for Medicaid, and whether a proxy helped complete the survey as covariates. To control for regional effects, including varying MA penetration, 305 geographic indicators of Hospital Referral Region (Dartmouth Medical School and Center for the Evaluative Clinical Sciences 1998) were included in the models.

Statistical Analysis

Poststratification weights that accounted for the complex survey design and non-response at the level of the plan/contract for MA and at the level of the state for FFS were used for all analyses. Mean beneficiary characteristics and unadjusted CAHPS scores by MA versus FFS status were calculated and differences between them were tested using weighted linear and logistic regressions as appropriate.

Primary multivariate analyses are based on weighted regression models of the Medicare CAHPS measures and include the standard CAHPS case-mix adjustment variables (Elliott et al. 2009a). Each model (one for each CAHPS measure) used as predictors the following: (1) seven subgroup identifiers; (2) an MA indicator; (3) seven MA by subgroup interaction terms; and (4) the

covariates. No dependent variables were imputed; other missing variables were rare and were imputed as the mean of the contract (FFS with Part D or MA) or state (FFS without Part D).

In these models, a statistically significant negative interaction term indicates the hypothesized negative difference-of-differences for the subgroup in MA relative to FFS ($R_{FFS}^{LV} - R_{FFS}^{V} < R_{MA}^{LV} - R_{MA}^{V}$). If the patients in the vulnerable subgroup (V) provided fewer positive ratings or reports than other (LV) beneficiaries (controlling for the other "vulnerability" characteristics simultaneously) in both FFS and MA, but with a larger difference in MA, the negative interaction term indicates a larger disparity in MA than in FFS. There also will be a negative interaction term if the "vulnerable" group tends to rate care more highly and the difference is smaller in MA than in FFS, or they have worse ratings in MA and better ratings in FFS.

Some of these vulnerable subgroups have historically provided more positive evaluations of care than other patients (e.g., persons with less education providing more positive evaluations; Elliott et al. 2001; Zaslavsky et al. 2001; O'Malley et al. 2005). Systematic differences between subgroups of patients in care evaluations can be due to differences in the quality of care received or response tendencies. For example, generally higher ratings from older patients could be due to the fact that older patients generally receive better care, or that older patients generally are more positive than younger patients when they evaluate comparable care experiences. However, because response tendencies should not vary by health plan or insurance type, variability in such differences is usually interpreted as reflecting differences in quality, not confounded by differences in response tendencies (Zaslavsky, Zaborski, and Cleary 2000; Elliott et al. 2001; Elliott et al. 2009a). Because it is unlikely that response tendency would vary by insurance type, a negative difference-of-differences can be interpreted as a *larger disparity* in care quality in MA relative to FFS, even if the ratings and reports of vulnerable beneficiaries are not lower than those of less vulnerable beneficiaries within both FFS and MA.

Linear contrasts were used to estimate and test differences between groups *within* MA and FFS. Because all indicators of vulnerability and all of their interactions with MA enter the models simultaneously, we estimate the interactions of each independent aspect of vulnerability with MA, controlling for the others.

Evidence suggests differences in the characteristics of MA and FFS enrollees with respect to health, SES, race/ethnicity, gender, and age (Keenan et al. 2009; Shimada et al. 2009). As a robustness check regarding the extent to which case-mix adjustors successfully address possible bias due to selection into MA, we repeated the primary analysis using population average

treatment effect propensity score weights. The propensity score model included the variables that define the vulnerable populations as well as the case mix adjustment variables. Each insurance type was weighted to match the distribution of covariates in the population as a whole (using weights equal to the product of the poststratification weights and the propensity score weights). Balance on the covariates included in the propensity score model was assessed using the standardized difference in the propensity score weighted means of each covariate. This approach allows us to consider what would happen if all beneficiaries chose MA instead of only beneficiaries like those who currently choose MA.

RESULTS

The overall MCAHPS response rate for 2007 was 49 percent, with 30 percent of all responses by mail and 3 percent in Spanish.[1] Among respondents with nonmissing gender, 58 percent were female (Table 1). Race/ethnicity was reported as non-Hispanic white, black, and Hispanic by 74.5, 7.0, and 6.3 percent of respondents, respectively, with MA respondents (12 percent) more often Hispanic than FFS respondents (6 percent), $p<.0001$. About 9 percent of respondents were 85 and older, 20 percent did not graduate high school, 17 percent were eligible for LIS (14 percent in MA versus 18 percent in FFS, $p<.0001$), and 32 percent were in poor or fair self-rated health (27 percent in MA versus 33 percent in FFS, $p<.0001$). More than three in four beneficiaries (78 percent) belonged to at least one designated subgroup, with 37 percent belonging to only one subgroup, 22 percent belonging to exactly two, and 7 percent belonging to four or more (data not shown). About half of beneficiaries (54 percent) had one or more characteristics traditionally associated with vulnerability (the six characteristics other than female), with 30 percent belonging to only one, 15 percent to exactly two, and 9 percent to three or more (data not shown). No correlations among subgroup characteristics exceeded 0.26 (data not shown).

Overall unadjusted mean scores of the 11 CAHPS measures we analyzed (Table 2) fell between 80 and 90 on a 0–100 scale, with the exception of getting information about prescription drug coverage (78.2) and paperwork (71.1). Means near the upper end of the response scale are typical of Medicare CAHPS and other surveys of patient experience (see, e.g., Landon et al. 2004). Unadjusted scores were significantly higher for MA than FFS for seven measures and significantly lower for one measure (getting needed care). MA scores were highest compared with FFS on the paperwork item (difference of 9.4

Table 1:   Characteristics of Medicare Advantage (MA) and FFS Beneficiaries

	Weighted%		
	Overall	*MA (n = 132,960), 17.1% Weighted)*	*FFS (n = 202,289, 82.9% Weighted)*
Independent variables designating subgroups of interest[†]			
Gender			
Female	**54.2**	**54.8**	**54.1****
Male	*39.7*	*38.6*	*39.9****
Unknown gender	6.1	6.6	6.0***
Race/ethnicity			
Black	**7.0**	**7.7**	**6.9*****
Hispanic	**6.3**	**12.0**	**5.2*****
White	*74.5*	*68.7*	*75.7****
Native American	1.9	1.5	2.0***
Asian/Pacific Islander	1.8	2.8	1.6***
Unknown race	8.4	7.4	8.6***
Age			
85 or older	**8.9**	**8.6**	**9.0***
65–84	*73.2*	*76.1*	*72.7****
18–64	11.5	8.4	12.2***
Did not graduate high school (versus *did*)	**19.7**	**22.8**	**19.1*****
Deemed eligible for LIS (versus *not eligible*)	**17.1**	**13.8**	**17.8*****
Fair or poor self-rated general health (versus *good, very good*, or *excellent*)	**32.0**	**27.4**	**32.9*****
Covariates			
Mental health			
Excellent	28.8	30.4	28.4***
Very good	31.2	31.7	31.1**
Good	26.9	26.6	26.9
Fair	10.7	9.4	10.9***
Poor	2.5	1.9	2.7***
Proxy status			
Proxy answered	3.1	3.0	3.1
Proxy helped	7.1	6.9	7.2*
No proxy help	89.8	90.2	89.8**
Dually eligible for Medicaid	9.3	5.4	10.1***
Hospital referral region (as 305 fixed effects)	N/A	N/A	N/A

Note. The *p*-values correspond to differences between MA and FFS in a logistic regression for gender, race/ethnicity, age, no diploma, LIS, and dual eligible, and linear regression for the remainder.

[†]Subgroup of interest is in boldface; reference group is in italics; other groups not in the comparison are in standard typeface.

*p < 0.05; **p < 0.01; ***p < 0.001.

FFS, fee-for-service; LIS, low-income subsidy.

Table 2: Unadjusted Mean CAHPS Scores and Case-Mix Adjusted Differences, Medicare Advantage (MA), and Fee-For-Service (FFS) Beneficiaries

	Overall, Unadjusted		MA, Unadjusted		FFS, Unadjusted		Unadjusted Differences (MA − FFS)	Case-Mix Adjusted Coefficients (MA versus FFS)†
	Unweighted N	Weighted Mean	Unweighted N	Weighted Mean	Unweighted N	Weighted Mean	Weighted Mean	Coefficient
0–10 Global ratings								
Rating of personal doctor	252,432	89.9	99,857	90.1	152,575	89.9	0.2*	− 0.2*
Rating of specialist	160,908	88.7	63,380	88.6	97,528	88.7	− 0.1	− 0.5***
Rating of MA plan/Medicare (FFS)	302,352	83.8	122,372	85.6	179,980	83.5	2.1***	1.5***
Rating of health care received	262,075	85.9	102,346	86.4	159,729	85.8	0.6***	0
Rating of PDP	252,547	80.8	107,996	82.8	144,551	80	2.8***	2.2***
Composites								
Doctor communication	258,453	89.2	103,064	89.3	155,389	89.2	0.1	− 0.2
Getting needed care	228,138	85.2	89,853	84.6	138,285	85.3	− 0.7***	− 1.3***
Getting care quickly	262,153	82.5	100,829	82.6	161,324	82.5	0.1	0.2
Paperwork	43,678	71.1	25,431	77.3	18,247	67.9	9.4***	7.9***
PDP: getting Rx drugs	247,192	88.8	104,354	89.7	142,838	88.4	1.3***	0.5***
PDP: getting information	95,866	78.2	36,418	80.2	59,448	77.6	2.6***	1.9***

Note. The *p*-values correspond to differences between MA and FFS in bivariate linear regressions.

†Case mix adjusted for CMS geographic region, educational attainment, age, self-rated general health, self-rated mental health, proxy assistance with survey completion, and eligibility for Low Income Supplement.

*$p<0.05$; **$p<0.01$; ***$p<0.001$.

CAHPS, Consumer Assessment of Health Care Providers and Systems; CMS, Centers for Medicare and Medicaid Services.

points on a 0–100 scale), followed by rate prescription drug plan (2.8-point advantage) and getting information from one's prescription drug plan (2.6-point advantage). Because ratings of MA plan and of Medicare are not exactly parallel, one must use caution in interpreting overall differences between these variables. When we look at case-mix adjusted[2] CAHPS score coefficients for MA versus FFS (last column of Table 2), we find a more mixed pattern where MA scores are significantly higher than FFS on five measures (all three PDP measures, paperwork, and the problematic MA plan versus Medicare measure) and significantly lower than FFS on three measures, two physician measures, and getting needed care.

For each of the 11 CAHPS measures, Table 3 displays disparities within MA associated with each characteristic (adjusted mean differences between the subgroup and its counterpart, with a negative value indicating fewer positive experiences for the vulnerable subgroup), disparities within FFS, and the (MA–FFS) differences in disparities (an adjusted difference of differences corresponding to the interaction term of MA and the characteristic in question). Disparities within MA are calculated as the sum of the coefficient for the vulnerable characteristic and the coefficient for its interaction with MA. Disparities within FFS are simply the model coefficient for the vulnerable characteristic.

We briefly report the absolute differences in measures between beneficiaries with or without each of the seven characteristics within insurance type (MA or FFS), with the caveat that they may in part reflect differences in response tendency as noted above. Whereas Hispanic and black race/ethnicity were associated with less positive absolute evaluations than white race/ethnicity, and fair/poor health with less positive absolute evaluations than good to excellent health in more than half of the comparisons, other subgroups of interest tended to report more positive absolute experiences than their counterparts. Less educated, lower income, older, and female beneficiaries tended to report more positive absolute experiences than their counterparts.

The interactions of these subgroup identifiers with MA when significant were predominantly negative. There were significant ($p < .05$) negative interactions with MA in 27 of 77 instances. Significant negative interactions with MA were found for 8 of 11 measures for poor/fair health, with LIS eligibility, age 85 or older, and black race for 4 of 11 measures, and for female and no high school diploma for 3 of 11 measures. Among these six characteristics, there were only two instances of significant positive interactions. The exceptions to this pattern were Hispanic beneficiaries, where two positive and one negative interaction were observed; the two positive interactions involved Part D measures.

Table 3: Adjusted Disparities in Medicare Advantage (MA) and Fee-For-Service (FFS), by Vulnerability Characteristic[†]

	Female (versus Male)	Hispanic (versus Non-Hispanic White)	Black (versus Non-Hispanic White)	Age 85 or Older	Less Than High School Education	LIS Deemed Eligible	Fair/Poor General Health
Rate doctor							
Disparity within MA[‡]	1.16***	1.05***	0.24	0.11	1.35***	0.08	− 1.29***
Disparity within FFS[‡]	1.26***	1.01***	− 0.03	− 0.11	1.13***	0.65***	− 0.84***
Difference in disparities (MA–FFS)[§]	− 0.10	0.04	0.27	0.21	0.21	− 0.57*	− 0.45*
Rate specialist							
Disparity within MA	0.79***	0.29	− 1.79***	− 0.15	1.32***	0.48	− 1.18**
Disparity within FFS	1.12***	0.20	− 0.55	− 0.65	1.01***	0.11	− 1.02**
Difference in disparities (MA–FFS)	− 0.33	0.09	− 1.24*	0.50	0.31	0.37	− 0.16
Doctor communication							
Disparity within MA	0.49*	0.47	1.04***	− 0.33	0.74***	0.02	− 1.87
Disparity within FFS	0.53***	1.00**	1.07***	− 0.87**	0.27	0.11	− 1.40
Difference in disparities (MA–FFS)	− 0.04	− 0.53	− 0.03	0.54	0.47	− 0.09	− 0.48
Rate MA plan/Medicare (FFS)							
Disparity within MA	1.43***	0.14	− 1.05***	1.68***	1.88***	1.86***	− 2.52***
Disparity within FFS	2.16***	2.22***	0.11	2.68***	2.94***	4.09***	0.33
Difference in disparities (MA–FFS)	− 0.73***	− 2.08***	− 1.16**	− 1.00**	− 1.06***	− 2.22***	− 2.84***
Rate care							
Disparity within MA	1.13***	− 0.74*	− 2.30***	0.19	0.81***	0.15	− 3.00***
Disparity within FFS	1.63***	− 1.13**	− 2.85***	0.49	− 0.11	− 0.24	− 2.33***
Difference in disparities (MA–FFS)	− 0.50*	0.39	0.55	− 0.29	0.92***	0.39	− 0.66**
Getting needed care							
Disparity within MA	− 0.47*	− 1.83***	− 1.90***	0.37	1.47***	− 0.52	− 2.10***
Disparity within FFS	0.07	− 1.11*	− 1.40***	0.41	0.92***	− 0.18	− 0.56
Difference in disparities (MA–FFS)	− 0.54*	− 0.72	− 0.50	− 0.04	0.55	− 0.34	− 1.54***
Getting care quickly							
Disparity within MA	1.04***	− 5.18***	− 4.01***	− 0.06	− 1.34***	0.06	0.15
Disparity within FFS	0.96***	− 4.18***	− 2.27***	0.33	− 1.99***	− 0.42	1.40***
Difference in disparities (MA–FFS)	0.08	− 1.00	− 1.73*	− 0.39	0.65	0.47	− 1.25***

Table 3. Continued

	Female (versus Male)	Hispanic (versus Non-Hispanic White)	Black (versus Non-Hispanic White)	Age 85 or Older	Less Than High School Education	LIS Deemed Eligible	Fair/Poor General Health
Paperwork							
Disparity within MA	0.41	− 4.52***	− 0.75	− 2.81*	− 3.79***	− 1.90	− 2.96*
Disparity within FFS	− 0.08	− 2.96*	0.76	− 0.01	− 4.46***	− 0.26	− 0.55
Difference in disparities (MA–FFS)	0.50	− 1.56	− 1.51	− 2.80*	0.67	− 1.64	− 2.40**
Rate PDP							
Disparity within MA	1.61***	2.35***	1.26***	3.85***	2.34***	6.37***	2.21***
Disparity within FFS	2.09***	− 1.32**	− 0.08	3.41***	3.01***	9.47***	1.18***
Difference in disparities (MA–FFS)	− 0.48	3.66***	1.33*	0.43	− 0.67*	− 3.10***	− 1.02***
PDP: getting Rx drugs							
Disparity within MA	0.89***	− 2.05***	− 2.70***	− 0.49	− 0.78**	− 0.15	− 0.97**
Disparity within FFS	1.33***	− 4.77***	− 2.91***	0.73*	0.40*	0.51**	− 0.77**
Difference in disparities (MA–FFS)	− 0.44	2.72***	0.21	− 1.22**	− 1.17***	− 0.66	− 0.20
PDP: getting information							
Disparity within MA	0.45	− 3.52***	− 4.57***	− 0.03	− 1.61**	1.18	− 2.86***
Disparity within FFS	0.63	− 3.82***	− 1.43*	2.49***	− 0.25	2.99***	− 1.46*
Difference in disparities (MA–FFS)	− 0.18	0.30	− 3.15**	− 2.52*	− 1.37	− 1.81*	− 1.40*
Count of significant negative differences in disparities (MA–FFS)	3	1	4	4	3	4	8
Count of significant positive differences in disparities (MA–FFS)	0	2	1	0	1	0	0

[†]Each row represents a single model predicting a given outcome from all column subgroup indicators, and MA interactions, and additional covariates described above.

[‡]"Vulnerable"–"Less Vulnerable", so a negative value reflects a less positive experience for the "vulnerable" beneficiaries than the less "vulnerable" beneficiaries.

[§]A difference-of-differences such that a negative value reflects a negative disparity that is larger or a positive disparity that is smaller for "vulnerable" beneficiaries in MA when compared with FFS.

*$p < 0.05$; **$p < 0.01$; ***$p < 0.001$.

LIS, low-income subsidy.

With respect to the eight Part C measures, significant ($p<.05$) negative interactions with MA were found in 18 of 56 instances (with one positive). Interactions with MA were negative for six of eight Part C measures for poor/fair health and for three of eight Part C measures for black and female. Physician-related items (doctor rating, specialist rating, and doctor communication) showed the fewest significant interactions (differences in disparities). Specifically, in only 2 of 21 instances did physician-related measures show statistically significant negative interactions, as opposed to 16 of 35 instances for other Part C measures measuring characteristics of care, plans, and paperwork.

With respect to the three Part D measures, there were nine negative and three positive interactions with MA (two of the latter involving Hispanic ethnicity). For the subgroup characteristics of older age, poorer health, lower education, and lower income, 8 of 12 interactions with MA were negative for Part D measures, with no positive interactions.

These interactions were estimated simultaneously, so that for Getting Care Quickly, for example, the disparity between a black beneficiary in fair or poor health and a non-Hispanic white beneficiary in good, very good, or excellent health is estimated to be $1.73+1.25 = 2.98$ points less favorable to the former beneficiary in MA than in FFS, where the two beneficiaries would have more similar experiences with Getting Care Quickly.

The same models shown in Table 3 also allowed us to estimate the absolute difference between MA and FFS within each subgroup of interest (results not shown). We compared MA and FFS within these seven subgroups for each of 10 measures, omitting Rate MA Plan/Medicare for these comparisons. The absolute scores for beneficiaries with each of the seven subgroup characteristics of interest were generally higher in MA than in FFS for Part D measures ($p<.05$ for 7 of 21 instances, no significant examples of FFS$>$MA) and paperwork ($p<.05$ for seven of seven instances), and lower in MA than FFS for most other measures ($p<.05$ for 17 instances where FFS$>$MA and five instances where MA$>$FFS of 42 total instances).

The propensity score weights used for the sensitivity analysis were successful in balancing all covariates at well below the commonly used 0.20 standard deviations threshold (Cochran 1968). The sensitivity analysis replicated the analysis shown in Table 3 using population average treatment effect propensity score weights, as shown in Appendix SA3. The results are largely similar to those shown in Table 3, if somewhat smaller in magnitude on average. In particular, the tendency for larger disparities in MA for African Americans is no longer significant, while the tendency for larger disparities in MA for LIS deemed strengthens.

DISCUSSION

Unlike Landon et al. (2004)[3], who found that FFS scored higher in the aggregate than MA on all measures except immunization, our case-mix adjusted findings present a much more mixed picture, with MA outperforming FFS on five measures and FFS outperforming MA on three measures. We note that Landon and colleagues analysis did not include prescription drug measures (which did not exist in 2004)—which account for three of the five instances of MA outperforming FFS, consistent with findings from Neuman et al. (2007) regarding higher beneficiary ratings for MA plans in the area of prescription drugs. In the time since Landon and colleagues, Medicare payments to MA HMOs have increased as compared with payments for FFS (MedPAC 2009), and therefore the capability of MA plans to provide enhanced benefits to beneficiaries has also increased compared with FFS plans. These increased payments may help explain the somewhat different pattern of MA versus FFS performance that we now observe.

In some instances, beneficiary characteristics often associated with "vulnerable" status in a health care setting (Hispanic and black, relative to non-Hispanic white; fair or poor health, relative to better health) tended to be associated with less positive CAHPS ratings and reports. Other characteristics often associated with "vulnerable" status showed the opposite pattern (no high school degree, income below 150 percent of the federal poverty line, age 85 or older). Because differences in response tendency, and perhaps differences in expectations of care, may underlie some of these patterns in absolute differences (Zaslavsky et al. 2000; Elliott et al. 2001; Elliott et al. 2009a), we focus on the difference-of-differences results to assess relative disparities within MA and FFS Medicare, relying on an assumption that response tendencies should not substantially interact with managed care status.

Differences-of-differences calculations show that the disparities for lower income, less healthy, female, less educated, and black Medicare beneficiaries, relative to their counterparts, tend to be larger (more negative) or less positive in MA than in FFS. While female gender is not traditionally considered a characteristic of vulnerability by itself in a health care setting, women had greater disparities (relative to men) in MA than in FFS, similar to what was seen for the other characteristics tested. No such pattern was evident for Hispanic beneficiaries, consistent with findings of better access for Hispanics in MA by Balsa et al. (2007). Propensity score analysis, which attempted to model selection into MA, found generally similar results, although the tendency for greater disparities for blacks in MA compared with FFS was no longer significant under this sensitivity test.

The financial incentive hypothesis predicted larger disparities in MA than FFS for older, sicker, and female beneficiaries. These predictions were well supported in each case. The health literacy hypothesis correctly predicted larger disparities in MA than FFS for beneficiaries with lower income and less educational attainment. On the other hand, there was not clear evidence for the predicted greater disparities for blacks and Hispanics in MA than in FFS, and the prediction of smaller disparities for females in MA than in FFS was contradicted. It is possible that financial incentives may play a larger role, and health literacy/negotiation a smaller but important role, in greater disparities in MA than FFS for groups defined according to health, socioeconomic status, age, and gender.

Among the measures studied, this pattern tended to be strongest for Part D measures. This suggests that obtaining information and prescriptions under Part D may be especially subject to individual variation, perhaps because it is a new program with less standardized protocols, in line with the health literacy hypothesis. The tendency for similar disparities in MA and FFS for experiences with physicians may reflect a situation in which MA and FFS beneficiaries see a similar set of physicians and physician-beneficiary interactions are relatively unaffected by plans once patients access physicians. It is also likely that significant differences in experiences, and perhaps disparities, occur at the medical group level. For example, there is evidence that for many patient experience measures there is more variation at the group than health plan level (Solomon et al. 2002; Safran et al. 2006), and that medical group characteristics may be strongly predictive of patient assessments of individual physicians (Rodriguez et al. 2009).

Larger disparities in MA do not necessarily mean worse absolute performance for the vulnerable beneficiaries in MA as compared with FFS. The absolute scores for the seven subgroup characteristics of interest were generally higher in MA than in FFS for Part D measures and paperwork and lower in MA than FFS for most other measures. This is consistent with recent work (Neuman et al. 2007) that finds generally more positive experiences with Part D coverage overall in MA than in freestanding PDPs, so that negative interactions of MA with characteristics associated with vulnerability in these instances often mean smaller MA advantages for them than for less vulnerable beneficiaries. For most Part C CAHPS measures of care, where FFS performance tends to be more positive overall than MA (Landon et al. 2004), these larger disparities more often translated into larger absolute disadvantages for vulnerable beneficiaries within MA when compared with FFS.

Our study has several limitations. First, those who select MA plans may differ from FFS beneficiaries in unobserved ways that are related to how they

assess health care. Second, our results could be influenced by nonresponse bias. The response rate, however, is similar to other large surveys of this type (e.g., 56 percent for Neuman et al. 2007) and there were only modest differences between responders and nonresponders (Zaslavsky and Zaborski 2002; Elliott et al. 2005). Third, a higher proportion of blacks and Hispanics than whites responded by phone, rather than mail. Given evidence of more positive responses by phone to CAHPS items in a randomized mode experiment (Elliott et al. 2009b), this might mean that the experiences of blacks and Hispanics are overestimated relative to non-Hispanic whites. Nonetheless, given that these phone proportions did not differ by MA status, such effects should not bias estimates of the relative differences of MA and FFS for these subgroups. Fourth, the differences we observed may appear to be small on 0–100 scales. We note, however, that many of the differences approach or exceed one standard deviation of the MA plan means. Differences of this magnitude have been associated with important measures such as disenrollment in prior analyses of CAHPS data (Lied and Sheingold 2001; Lied et al. 2003). Finally, additional heterogeneity among Hispanics by language preference may exist but is not explored here.

Our findings suggest that there may be opportunities for improving the experiences of vulnerable beneficiaries within MA plans. This effort might begin with efforts to monitor and improve the uniformity of insurer and provider response to vulnerable and less vulnerable beneficiaries. For example, stratified analyses and presentation of results for different subgroups might help providers identify subsets of patients for whom special efforts should be made. Such efforts might be complemented by efforts to teach vulnerable beneficiaries how to most effectively interact with Medicare providers and insurers and to overcome the differential treatment they might otherwise face.

ACKNOWLEDGMENTS

NOTES

1. More blacks (41 percent) and Hispanics (37 percent) than non-Hispanic whites (28 percent) responded by phone ($p < .001$ for each). Phone response was similar for MA (29 percent) and FFS (30 percent).
2. Case-mix adjusted for CMS geographic region, educational attainment, age, self-rated general health, self-rated mental health, proxy assistance with survey completion, and eligibility for Low Income Supplement.
3. Our analysis differs from Landon's in a number of important ways. First, Landon and colleagues excluded the Plan/Medicare rating, and prescription drug plans did not exist at the time. Landon considers immunization measures—we do not.

REFERENCES

Balsa, A. I., Z. Cao, and T. G. McGuire. 2007. "Does Managed Health Care Reduce Health Care Disparities between Minorities and Whites?" *Journal of Health Economics* 26 (1): 101–21.

Blumenthal, D., E. Mort, and J. Edwards. 1995. "The Efficacy of Primary Care for Vulnerable Population Groups." *Health Services Research* 30 (1, part 2): 253–73.

Cochran, W. G. 1968. "The Effectiveness of Adjustment by Subclassification in Removing Bias in Observational Studies." *Biometrics* 24 (2): 295–313.

Dartmouth Medical School and Center for the Evaluative Clinical Sciences. 1998. *The Dartmouth Atlas of Health Care.* Hanover, NH: J. E. Wennberg, American Hospital Publishing Inc.

Elliott, M. N., C. Edwards, J. Angeles, and R. D. Hays. 2005. "Patterns of Unit and Item Non-Response in the CAHPS Hospital Survey." *Health Services Research* 40 (6, part 2): 2096–119.

Elliott, M. N., A. M. Haviland, D. E. Kanouse, K. Hambarsoomian, and R. D. Hays. 2009a. "Adjusting for Subgroup Differences in Extreme Response Tendency in Ratings of Health Care: Impact on Disparity Estimates." *Health Services Research* 44 (2, part 1): 542–61.

Elliott, M. N., R. Swartz, J. Adams, K. L. Spritzer, and R. D. Hays. 2001. "Consumer Evaluations of Health Plans and Health Care Providers: Case-Mix Adjustment of the National CAHPS Benchmarking Data 1.0: A Violation of Model Assumptions?" *Health Services Research* 36 (3): 555–73.

Elliott, M. N., A. M. Zaslavsky, E. Goldstein, W. Lehrman, K. Hambarsoomian, M. Beckett, and L. Giordano. 2009b. "Effects of Survey Mode, Patient Mix, and Nonresponse on CAHPS Hospital Survey Scores." *Health Services Research* 44 (2, part 1): 501–08.

Fongwa, M. N., W. Cunningham, R. Weech-Maldonado, P. R. Gutierrez, and R. D. Hays. 2008. "Reports and Ratings of Care: Black and White Medicare Enrollees." *Journal of Health Care Poor Underserved* 19 (4): 1136–47.

Frank, R. G., J. Glazer, and T. G. McGuire. 2000. "Measuring Adverse Selection in Managed Health Care." *Journal of Health Economics* 19 (6): 829–54.

Goldstein, E., P. D. Cleary, K. M. Langwell, A. M. Zaslavsky, and A. Heller. 2001. "Medicare Managed Care CAHPS: A Tool for Performance Improvement." *Health Care Financing Review* 22 (3): 101–7.

Goldstein, E., M. N. Elliott, W. G. Lehrman, K. Hambarsoomian, and L. A. Giordano. 2010. "Racial/Ethnic Differences in Patients' Perceptions of Inpatient Care Using the HCAHPS Survey." *Medical Care Research Review* 67 (1): 74–92.

Hargraves, J. L., R. D. Hays, and P. D. Cleary. 2003. "Psychometric Properties of the Consumer Assessment of Health Plans Study (CAHPS) 2.0 Adult Core Survey." *Health Services Research* 38 (6): 1509–23.

Hays, R. D., J. A. Shaul, V. S. Williams, J. S. Lubalin, L. D. Harris-Kojetin, S. F. Sweeny, and P. D. Cleary. 1999. "Psychometric Properties of the CAHPS 1.0 Survey Measures." *Medical Care* 37 (3, suppl): MS22–31, li.

Heller, A., M. N. Elliott, A. M. Haviland, D. J. Klein, and D. E. Kanouse. 2009. "Patient Activation Status as a Predictor of Patient Experience among Medicare Beneficiaries." *Medical Care* 47 (8): 850–7.

Institute of Medicine. 2002. *Unequal Treatment: Confronting Racial and Ethnic Disparities in Health Care.* Washington, DC: The National Academies.

Keenan, P. S., P. D. Cleary, A. J. O'Malley, B. E. Landon, L. Zaborski, and A. M. Zaslavsky. 2010. "Geographic Area Variations in the Medicare Health Plan Era." *Medical Care* 48 (3): 260–6.

Keenan, P. S., M. Elliott, P. Cleary, A. Zaslavsky, and B. Landon. 2009. "Quality Assessments by Sick and Healthy Beneficiaries in Traditional Medicare and Medicare Managed Care." *Medical Care* 47 (8): 882–8.

Landon, B., A. Zaslavsky, S. Bernard, M. Cioffi, and P. Cleary. 2004. "Comparison of Performance of Traditional Medicare vs Medicare Managed Care." *Journal of American Medical Association* 291 (14): 1744–52.

Lied, T. R., and S. H. Sheingold. 2001. "Relationships among Performance Measures for Medicare Managed Care Plans." *Health Care Financing Review* 22 (3): 23–33.

Lied, T. R., S. H. Sheingold, B. Landon, J. A. Shaul, and P. D. Cleary. 2003. "Beneficiary Reported Experience and Voluntary Disenrollment in Medicare Managed Care." *Health Care Financing Review* 25 (1): 55–66.

Lurie, N., C. Zhan, J. Sangl, A. S. Bierman, and E. S. Sekscenski. 2003. "Variation in Racial and Ethnic Differences in Consumer Assessments of Health Care." *American Journal of Management Care* 9 (7): 502–9.

Martin, L. T., T. Ruder, J. J. Escarce, B. Ghosh-Dastidar, D. Sherman, M. Elliott, C. E. Bird, A. Fremont, C. Gasper, A. Culbert, and N. Lurie. 2009. "Developing Predictive Models of Health Literacy." *Journal of General Internal Medicine* 24 (11): 1211–6.

Martino, S. C., M. N. Elliott, P. D. Cleary, D. E. Kanouse, J. A. Brown, K. L. Spritzer, A. Heller, and R. D. Hays. 2009. "Psychometric Properties of an Instrument to Assess Medicare Beneficiaries' Prescription Drug Plan Experiences." *Health Care Financing Review* 30 (3): 41–53.

MedPAC. 2009. *A Data Book: Healthcare Spending and the Medicare Program 2009.* Washington, DC: MedPAC 149–59.

Morales, L. S., M. N. Elliott, R. Weech-Maldonado, K. L. Spritzer, and R. D. Hays. 2001. "Differences in CAHPS Adult Survey Reports and Ratings by Race and

Ethnicity: An Analysis of the National CAHPS Benchmarking Data 1.0." *Health Services Research* 36 (3): 595–617.

National Research Council. 2004. *Eliminating Health Disparities: Measurement and Data Needs.* Washington, DC: National Academies Press.

Neuman, P., M. K. Strollo, S. Guterman, W. H. Rogers, A. Li, A. M. C. Rodday, and D. G. Safran. 2007. "Medicare Prescription Drug Benefit Progress Report: Findings from a 2006 National Survey of Seniors." *Health Affairs* 26 (5): w630–43.

O'Malley, A. J., A. M. Zaslavsky, M. N. Elliott, L. Zaborski, and P. D. Cleary. 2005. "Case-Mix Adjustment of the CAHPS Hospital Survey." *Health Services Research* 40 (6, part 2): 2162–81.

Owens, G. M. 2008. "Gender Differences in Health Care Expenditures, Resource Utilization, and Quality of Care." *Journal of Managed Care Pharmacy* 14 (3, suppl): 2–6.

Rodriguez, H. P., T. v. Glahn, W. H. Rogers, and D. G. Safran. 2009. "Organizational and Market Influences on Physician Performance on Patient Experience Measures." *Health Services Research* 21 (1): 13–21.

Safran, D. G., M. Karp, K. Coltin, H. Chang, A. Li, J. Ogren, and W. H. Rogers. 2006. "Measuring Patients' Experiences with Individual Primary Care Physicians. Results of a Statewide Demonstration Project." *Journal of General Internal Medicine* 21 (1): 13–21.

Shi, L. 2000. "Vulnerable Populations and Health Insurance." *Medical Care Research and Review* 57 (1): 110–34.

Shimada, S. L., A. M. Zaslavsky, L. B. Zaborski, A. J. O'Malley, A. Heller, and P. D. Cleary. 2009. "Market and Beneficiary Characteristics Associated with Enrollment in Medicare Managed Care Plans and Fee-For-Service." *Medical Care* 47 (5): 517–23.

Solomon, L. S., A. M. Zaslavsky, B. E. Landon, and P. D. Cleary. 2002. "Variation in Patient-Reported Quality among Health Care Organizations." *Health Care Financing Review* 23 (4): 85–100.

Trivedi, A. N., A. M. Zaslavsky, E. C. Schneider, and J. Z. Ayanian. 2005. "Trends in the Quality of Care and Racial Disparities in Medicare Managed Care." *New England Journal of Medicine* 353 (7): 692–700.

———. 2006. "The Relationship between Quality of Care and Racial Disparities within Medicare Health Plans." *Journal of American Medical Association* 296 (16): 1998–2004.

Weech-Maldonado, R., M. N. Fongwa, P. Gutierrez, and R. D. Hays. 2008. "Language and Regional Differences in Evaluations of Medicare Managed Care by Hispanics." *Health Services Research* 43 (2): 552–68.

Weech-Maldonado, R., L. S. Morales, M. Elliott, K. Spritzer, G. Marshall, and R. D. Hays. 2003. "Race/Ethnicity, Language, and Patients' Assessments of Care in Medicaid Managed Care." *Health Services Research* 38 (3): 789–808.

Woolhandler, S., and D. U. Himmelstein. 2007. "Consumer Directed Healthcare: Except for the Healthy and Wealthy It's Unwise." *Journal of General Internal Medicine* 22 (6): 879–81.

Zaslavsky, A. M., L. Zaborski, and P. D. Cleary. 2000. "Does the Effect of Respondent Characteristics on Consumer Assessments Vary across Health Plans?" *Medical Care Research and Review* 57 (3): 379–94.

Zaslavsky, A. M., L. B. Zaborski, L. Ding, J. A. Shaul, M. J. Cioffi, and P. D. Cleary. 2001. "Adjusting Performance Measures to Ensure Equitable Plan Comparisons." *Health Care Financing Review* 22 (3): 109–26.

Zaslavsky, A. M., L. B. Zaborski, and P. D. Cleary. 2002. "Factors Affecting Response Rates to the Consumer Assessment of Health Plans Study Survey." *Medical Care* 40 (6): 485–99.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

    Appendix SA1: Author Matrix.
    Appendix SA2: 2007 Medicare CAHPS Global Ratings and Composites.
    Appendix SA3: Adjusted Disparities in MA and FFS (using Propensty Weights), by Vulnerability Characteristic.[#]

    Please note: Wiley-Blackwell is not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.