Do health plans influence quality of care?

LAURENCE C. BAKER[1,2,6], DAVID HOPKINS[1,3], RICHARD DIXON[4], JEFFREY RIDEOUT[5]
AND JEFFREY GEPPERT[2,6]

[1]Department of Health Research and Policy, Stanford University, Stanford, CA, [2]National Bureau of Economic Research, Cambridge, MA, [3]Pacific Business Group on Health, San Francisco, CA, [4]RLA Technologies, Jackson, CA, [5]Blue Shield of California Foundation, San Francisco, CA, [6]Center for Primary Care and Outcomes Research, Stanford University, Stanford, CA, USA

Abstract

Objective. To investigate the relative impact of physician groups and health plans on quality of care measures.

Design. Secondary data analysis of receipt of preventive care services included in the Health Plan Employer Data and Information Set (HEDIS) among 10 758 patients representing 21 health maintenance organizations and 22 large provider groups in the San Francisco and Los Angeles, California, areas in 1997. Each patient was eligible for (at least) one of six HEDIS-measured services. Data identify whether or not the service was provided, the patient's health plan, and the provider group responsible for the care. We used logistic regression to examine variations across plans in HEDIS rates, and whether variations persist after controls for provider groups are included.

Setting. Patients from 21 health maintenance organizations serving San Francisco and Los Angeles, California, in 1997.

Main outcome measures. Breast cancer screening, childhood immunizations, cervical cancer screening, diabetic retinal exam, prenatal care in the first trimester, and check-ups after delivery among patients for whom these services are appropriate.

Results. There are statistically significant differences across health plans in utilization rates for the six services examined. These differences are not substantially affected when we control for the provider group that cared for the patient. That is, controlling for provider group does not explain variations across plans, consistent with the view that health plans have an impact on HEDIS quality measures independent of the providers that they contract with.

Conclusions. There are activities that plans can undertake which influence their HEDIS scores. On the face of it, these results suggest that plans can independently improve quality, in contrast to hypotheses that plans would be 'too far' from patients to have an influence. Continued attention to collecting plan-level data is warranted. Further work should address other possible sources of variations in HEDIS scores, such as variability in the quality of plan administrative databases.

Keywords: health plans, HEDIS scores, preventive care, provider groups, quality

Improving the quality of health care is at the center of current health policy discussions in the USA [1–7]. Providing measures of quality that will enable purchasers of health care to select health plans and providers based on evidence about their performance is a key part of strategies to make the health care marketplace work better. The most widespread existing data collection efforts aimed at enhancing quality are those that, like the Health Plan Employer Data and Information Set (HEDIS) from the National Committee on Quality Assurance (NCQA) [8], focus on collecting data on health plans. Despite the prominence of health plan measures, some argue that collecting data on the performance of health care providers could be more valuable [7,9]. In many cases these arguments concern the entities best held accountable for health care delivery, but another set of issues arises around the ability of plans and providers to influence quality of care. It is sometimes suggested that health-plan measures are targeted at the wrong level, since it is really physicians and physician organizations, which interact more directly with patients, that are in the best position to influence processes of care. A strong version of this suggestion might go so far as to say that health plans can have little or no impact on the quality of care, and that differences in quality across health plans sometimes observed in studies of HEDIS rates and other quality measures simply reflect differences in the provider groups with which plans contract. If this is the case, existing data on health plans is of limited use and further data collection efforts should be focused on providers. Efforts to get consumers to better select health plans should also be targeted away from health plans *per se*, and instead aimed at getting consumers to pay attention to the providers with which plans have contracted.

While the hypothesis that providers are more important than plans has some prima facie appeal, it is not immediately

Address reprint requests to L. C. Baker, Department of Health Research and Policy, Stanford University, Stanford, CA 94305-5405. USA. E-mail: laurence.baker@stanford.edu

clear that it is correct. There are plausible avenues by which health plans can influence the quality of care. For example, health plans can use reminder systems to encourage the use of preventive services, or utilization monitoring with feedback to physicians to encourage better processes of care. Thus, physicians and plans may both be in a position to influence care delivery and quality.

Despite the importance of understanding the locus of control over quality, we are aware of no evidence on the separate importance of health plans and providers. We study a unique database that, for the first time, provides an opportunity to learn about the effects of plans and providers on quality of care. Our study focuses on six common quality measures that are part of HEDIS: breast cancer screening, cervical cancer screening, childhood immunizations, diabetic retinal exams, prenatal care use, and check-ups after delivery. Although they are not perfect measures of the true but ethereal 'quality of care', they are widely used indicators that are commonly accepted as valid and have been used in previous analyses of quality of care [10]. Our data on these measures come from a collaboration of health plans in California, called the California Cooperative Healthcare Reporting Initiative (CCHRI). The plans contract with numerous provider groups, including many organized medical groups and IPAs, to care for their enrolled patients, so the database represents a wide range of providers. Moreover, the California managed care marketplace is dominated by large network-model health maintenance organizations (HMOs) that contract with many of the same provider groups. This database thus contains data on a large number of patients in a variety of health plans and provider groups that have numerous interrelationships, an ideal situation for examining the relative impact of plans and providers on quality of care.

Data and methods

Data

We used data on 1997 HEDIS rates, collected in 1998, that were originally developed for CCHRI, a statewide collaborative of purchasers, plans, and providers in California. For our use, the data were blinded to ensure that no patients were individually identifiable. The data report measures of health care quality for major commercial and Medicare health plans in the state. Nineteen health plans serving commercial markets and 14 plans serving the Medicare market participated in CCHRI in 1998. In some cases, the commercial and Medicare plans are operated by the same company, but they are treated as separate plans in this analysis. All of the plans are HMOs. All but two are network-model HMOs with significantly overlapping physician networks. Two are group model HMOs. The plans participating in CCHRI represent approximately 95% of commercial HMO enrollees in the state and the majority of Medicare managed care enrollees.

For each participating plan, CCHRI compiled information about eight HEDIS measures of effectiveness of care: adolescent immunization, beta-blocker use after heart attack, breast cancer screening, childhood immunizations, cervical cancer screening, diabetic retinal exam, prenatal care in the first trimester, and check-ups after delivery. For each measure, CCHRI collected data for each participating plan's eligible membership using a cluster sampling methodology. For each plan and each measure, they first sampled physician practices (including groups and solo practices) with which the plan contracted, with probability proportional to the number of patients in the eligible population assigned to the practice. Within the sampled practices, they then randomly sampled individual patients meeting the criteria for the measure (e.g. women who were continuously enrolled in the health plan for 2 years and were aged 52–69 years for the mammography measure) with probability inversely proportional to the number of eligibles assigned to the practice. The sampling process was designed to return data on approximately 400 patients per plan per measure.

The HEDIS 'hybrid' methodology was used to determine whether the measured event (e.g. mammogram in the past 2 years) occurred. In the first step, computerized administrative databases maintained by plans, which contain data received from their contracted providers, were queried to search for evidence of service provision. Next, information was sought through chart review for individuals for whom administrative records did not verify provision of the HEDIS service. In our data, 24% of the patients sampled were positively identified in administrative records as having received the service, 32% were identified as 'positive events' using medical charts, and 28% were confirmed negative. No administrative or chart data for the remaining 16% of cases could be found, and thus these cases could not be confirmed as positive or negative. HEDIS specifications require that these missing data cases be counted as negative. The sampling and data collection process was administered by an independent organization identically for all of the plans in CCHRI.

In addition to an indicator of whether or not the service in question was provided and basic demographic information about the patient, the database also identifies the plan to which the individual belonged, and the provider group to which the patient was assigned. In some cases, the database identified multiple entities that might be considered the provider group for a given patient. Where this occurred, we chose the entity that we determined to be most proximate to the physician delivering the care. For example, in cases where the database identified as providers both an IPA and a smaller group that contracted with the IPA, we chose the smaller group.

Altogether, the database contains information for approximately 70 000 patients. 71% were enrolled in commercial plans and 29% in Medicare plans. We developed our analysis sample from this database to meet two criteria. Firstly, since there are geographical variations in practice patterns around the state of California, it is important to work with patients from well-defined areas to minimize the potential for differences across plans or groups to arise only because of variations in the locations of their patient population. We chose the Los Angeles metropolitan area and the San Francisco Bay areas, both of which are urbanized areas with large numbers of competing health plans and provider groups. We thus identified patients in the database from the Los Angeles and San Francisco areas for further consideration.

Secondly, to maximize our ability to separately identify the importance of plans and provider groups as predictors of HEDIS rates, we identified individuals who met the following criteria, implemented sequentially: (i) they were enrolled in a health plan that contracted with at least two provider groups in the same geographical area; (ii) they were enrolled in a provider group that contracted with at least five plans in the same geographical area (the results are robust if we use a two plan limit instead); and (iii) they were in a plan–provider cell that contained at least five patients total. This ensured that we had data with enough overlap between provider groups and health plans to separately identify the influence of plans and providers. We implemented this procedure separately by measure and by geographical area. Because of this, the specific set of health plans and provider groups varies somewhat from measure to measure and across geographical areas.

After implementing this procedure, we determined that sample sizes for the beta-blocker and adolescent immunization measures were too small for analysis. For the remaining six measures, we had samples of 7892 patients in the Los Angeles area and 2866 patients in the San Francisco area. Taking the two areas together, these patients were enrolled in 15 commercial and 12 Medicare plans. Seventy-nine per cent of patients were in commercial plans and 21% in Medicare plans. All of the plans represented in the analysis sample are network-model plans (all of the patients from the two group model plans in the base dataset were excluded because, by definition, their plan did not contract with two or more provider groups). Because of the criteria for sample inclusion, all of the providers represented in our sample are large physician groups. Table 1 defines the measures we use. Table 2 summarizes the numbers of patients and the number of plans and providers represented for each measure in each area.

Our analysis samples are subsets of the original CCHRI database. They are not random samples of all patients, health plans, or providers in California. Rather, by identifying this particular group of patients, plans, and providers for study, we hope to have defined a setting that best facilitates a statistically valid comparison of the impact of plans and providers on quality of care.

Methods

We began by examining the influence of plans alone as predictors of HEDIS rates, and then considered the influence of plans and providers together. We did this using two patient-level logistic regressions per measure per area. In both, the

Table 1 HEDIS measures collected by the CCHRI

Measure	Definition
Breast cancer screening	The percentage of women aged 52–69 years old who had a mammogram screening during the last 2 years
Check-ups after delivery	The percentage of women who delivered a baby who saw their health care provider within 8 weeks after giving birth
Childhood immunizations	The percentage of 2-year-olds who are up to date on all standard Immunizations
Cervical cancer screening	The percentage of women aged 21–64 years who had a Pap smear test in the last 3 years
Diabetic retinal exam	The percentage of diabetics who received eye screenings to prevent diabetes-related eye diseases in the last year
Prenatal care in the first trimester	The percentage of pregnant women who began their prenatal care within the first 3 months of pregnancy

Table 2 Analysis sample sizes, and number of plans and groups by measure

Measure	Los Angeles			San Francisco		
	Sample size	Plans *n*	Providers *n*	Sample size	Plans *n*	Providers *n*
Check-ups after delivery	1051	11	13	499	8	5
Cervical cancer screening	1063	12	13	449	9	6
Childhood immunizations	1204	12	15	412	9	4
Diabetic retinal exams	2430	14	16	678	8	5
Breast cancer screening	1675	13	17	599	9	5
Prenatal care	469	10	7	229	8	3

dependent variable is a binary variable indicating documented receipt of the measured service. In the first regression, the main set of independent variables was a series of dummy variables for health plans. In the second regression, the independent variable set included both dummies for health plans and providers.

Although standard HEDIS measurement procedures do not specify controlling for the characteristics of patients, our models included controls for patient age and, where appropriate, sex. This controlled for some degree for variability in patient characteristics across plans.

If plans had no impact independent of the providers groups they contract with, the first regression would show significant differences across plans in the measured HEDIS rates, but the second would show no differences across plans. That is, statistically controlling for providers would eliminate differences across plans in the HEDIS rates. If the provider groups explained part of the effect of plans on HEDIS rates, then the odds ratios for plans would move closer to 1 when the provider controls were added, although some differences across plans may still be apparent. If the impact of plans was independent of the impact of the providers they contract with, there would be no substantial change in the odds ratios associated with plans when the provider controls were added. As in any analysis that uses group dummies, the specific odds ratios we observed would vary with the reference plan chosen to be omitted from the analysis. The overall assessment of whether the effects of plans were smaller after inclusion of provider controls, however, should be invariant to the choice of the reference group.

We used several tests to investigate the impact of the addition of the provider controls on the plan odds ratios. First, we used χ^2 statistics to test the hypothesis that each individual plan odds ratio was different from 1 in the model that used only plan dummies, and in the model that used both plan and provider dummies. Second, we tested the equality of the odds ratios estimated in the plan only models, and in the plan and provider models. Third, we conducted likelihood ratio tests of the hypothesis that the plan odds ratios were jointly equal to 1 in the plan only model, and in the plan and provider model, and compared the results. These tests provided information about the extent to which significant differences are observed across plans before and after the addition of the provider controls, and whether the addition of the provider controls significantly changed the measured effects of plans. As another measure of the extent to which the addition of the provider coefficients significantly altered the measured plan odds ratios, we computed the correlation between the plan odds ratios before and after the addition of the provider controls. Finally, since we were interested in whether the addition of the provider measures to the model moves the measured plan odds ratios closer to 1, we computed the average movement in the coefficients (not the odds ratios) towards 0. For each plan, we measured the amount by which the coefficient moves closer to 0 with the addition of the provider controls. Positive values reflect movements towards 0; negative values reflect cases where the coefficient moved away from 0. We then took the average across all plans. Higher values reflect stronger coefficient movements towards 0, which correspond to stronger odds ratio movements toward 1.

We conducted large numbers of statistical tests in this analysis. We investigated the use of Bonferroni adjustments, but they did not alter our conclusions, and we therefore report traditional hypothesis test results.

We investigated the extent to which provider effects are independent of plans using techniques similar to those just described. We began with a model that included only provider controls, and then added plan controls to examine whether the odds ratios for providers changed with the addition of the plan controls. If providers had no influence on HEDIS rates independent of the health plans with which they contract, we would expect to observe an influence of provider group in the first equation, but not the second.

Results

Overall, we conducted 12 logistic regression analyses, one for each of six measures in each of two geographical areas. Tables 3 and 4 report results for two example measures: diabetic retinal exams in Los Angeles and mammography screening in San Francisco. The tables report the average HEDIS rate by plan and odds ratios from models in which plan variables only, provider variables only, and then both plan and provider variables are simultaneously included. For diabetic retinal examinations, for example, plan LA2 had a raw rate of 26%, indicating that 26% of the sampled patients eligible for the service were identified as having received it. In a logistic regression controlling only for plan, its odds ratio was 0.56 ($P < 0.05$), suggesting that it was statistically significantly less likely to provide documented diabetic retinal exams than the reference plan, plan LA1. In a regression controlling for plan and provider, its odds ratio dropped to 0.49 ($P < 0.05$).

In Tables 3 and 4, the regressions controlling only for plan indicate that some plans are significantly better than other plans. If the reason that plans vary in their HEDIS rates is that some plans were better than others at identifying and contracting with higher-scoring provider groups, then we would expect adding provider controls to explain the differences across plans, moving the plan odds ratios closer to 1. This did not occur. Plans that had odds ratios significantly different from 1 in the models without provider controls frequently continued to have odds ratios significantly different from 1 even after provider controls were included in the models. For no plan were the odds ratios obtained without and with the provider controls statistically significantly different from one another. Likelihood ratio tests of the hypothesis that the plan odds ratios were jointly equal to 1 were significant at the $P < 0.05$ level in both the plan only, and the plan and provider models.

Similar patterns were observed for all of the conditions in both areas. Figures 1–4 illustrate the complete set of results for all six conditions in both areas, plotting the odds ratios for plans before and after the inclusion of provider controls. In each case, the plans are ordered from lowest odds ratio to highest odds ratio from the model with no provider controls.

Table 3 Odds ratios from logistic regression analysis of receipt of diabetic retinal exams, Los Angeles sample

	Odds ratio from model including:			
	Unadjusted rate	Plan variables only	Provider variables only	Plan and provider variables
Plans				
LA1	0.37	–[1]	–	–[1]
LA2	0.26	0.56*	–	0.49*
LA3	0.45	1.32	–	1.30
LA4	0.30	0.58*	–	0.67
LA5	0.60	2.04*	–	1.40
LA6	0.32	0.63*	–	0.60*
LA7	0.36	0.83	–	0.76
LA8	0.50	1.71	–	2.41*
LA9	0.55	1.52	–	1.26
LA10	0.24	0.55*	–	0.50*
LA11	0.42	1.07	–	1.36
LA12	0.43	1.18	–	0.97
LA13	0.43	1.24	–	0.88
LA14	0.33	0.62	–	0.71
Groups				
LA G1	0.41	–	–[1]	–[1]
LA G2	0.33	–	0.77	0.62
LA G11	0.42	–	1.18	1.13
LA G12	0.49	–	1.58	1.68
LA G29	0.61	–	2.30*	1.41
LA G35	0.39	–	1.00	0.94
LA G36	0.47	–	1.43	1.20
LA G39	0.46	–	1.34	1.19
LA G45	0.26	–	0.52*	0.50*
LA G48	0.28	–	0.59*	0.55*
LA G56	0.44	–	1.22	1.12
LA G58	0.45	–	1.21	1.14
LA G67	0.39	–	1.02	0.91
LA G71	0.43	–	1.25	1.24
LA G95	0.33	–	0.80	0.83
LA G96	0.37	–	0.96	0.85
Joint likelihood ratio test for plan variables				
$\chi^2(10)$		18.74	–	37.42
P-value		0.044		<0.001
n		2430	2430	2430

[1]Index category.
*Significantly different from 1 at $P < 0.05$.

If inclusion of the provider controls substantially altered the inferences about plan effects, we would expect to see noticeable changes in the odds ratio profiles across plans with the addition of the provider controls. If inclusion of the provider controls explained all of the variation across plans, we would expect to see the profile lines become flat at odds ratio 1.0 with the addition of the provider controls. This did not occur. Across the graphs, there is not a noticeable trend toward flattening of the odds ratio profile with the addition of the provider controls, and in some cases the profile becomes steeper.

As further evidence that the addition of the provider controls did not substantially change the observed plan odds ratios, the correlations between the odds ratios obtained before and after the use of the provider controls are generally quite high: always about 0.8 and frequently above 0.9 (Table 5). We also computed the average amount by which the regression

Table 4 Odds ratios from logistic regression analysis of receipt of mammography screening, San Francisco sample

	Odds ratio from model including:			
	Unadjusted rate	Plan variables only	Provider variables only	Plan and provider variables
Plans				
SF1	0.69	–[1]	–	–[1]
SF2	0.50	0.50	–	0.47
SF3	0.66	0.98	–	1.09
SF4	0.89	3.62*	–	3.70*
SF5	0.80	1.39	–	1.22
SF8	0.83	2.78	–	2.83
SF9	0.83	1.73	–	1.91
SF10	0.67	1.17	–	1.34
SF11	0.86	3.34*	–	3.80*
Groups				
SF G5	0.69	–	–[1]	–[1]
SF G13	0.81	–	1.72	2.44
SF G43	0.69	–	0.93	1.56
SF G47	0.73	–	1.20	1.40
SF G80	0.70	–	0.97	1.55
Joint likelihood ratio test for plan variables				
$\chi^2(8)$		20.77	–	19.51
P-value		0.008		0.012
n		599	599	599

[1]Index category.
*Significantly different from 1 at $P < 0.05$.

coefficients moved towards 0 with the addition of the provider controls (Table 5). For illustration, consider the case of diabetic retinal exams in Los Angeles. Here, the average movement in the coefficients towards 0 was 0.04. For reference, we also computed the value that would have been observed if all of the coefficients moved all the way to 0, which is 0.40. That is, the average movement towards 0 is substantially less than would have been observed if all of the coefficients had become 0, or equivalently if all the odds ratios had become 1, with the addition of the provider controls. Results for other measures are similar, with all showing small movements and some showing negative movements, which indicate that, on average, the coefficients moved further from 0 with the addition of the provider controls.

Finally, we find that the χ^2 values from likelihood ratio tests of the hypothesis that the plan odds ratios are jointly equal to 1 are generally similar between the two models, rejecting the null hypothesis in both models in many cases (Table 5). There is one case, check-ups after delivery in San Francisco, where the addition of the provider controls moves the joint test for the plan coefficients from being statistically significant to being insignificant. There are also cases where addition of the provider controls appears to increase the statistical strength of the plan effect. Across all of the measures and areas, the changes in the χ^2 values between the plan only and the plan and provider models are not large, nor do they suggest a pattern of reduction in the importance of the plan variables with the addition of the provider controls.

We take these results to suggest that plans do have an influence on measured HEDIS rates which is statistically independent of the influence of providers. Providers also have an important and measurable influence on HEDIS rates, independent of the plans they contract with. In Tables 3 and 4, as in the other cases not shown, statistically significant differences between providers are observed, with and without plan controls. Addition of plan controls never moves provider odds ratios substantially closer to 1.

Discussion

We find significant differences between plans in HEDIS rates for preventive care services, even after controlling for the provider groups that cared for their patients. The fact that plans have an effect on HEDIS scores that is statistically independent of the effect of providers strongly suggests that there are activities that some plans engage in that result in higher HEDIS scores, independent of the provider groups with which they contract. The persistent influence of plans

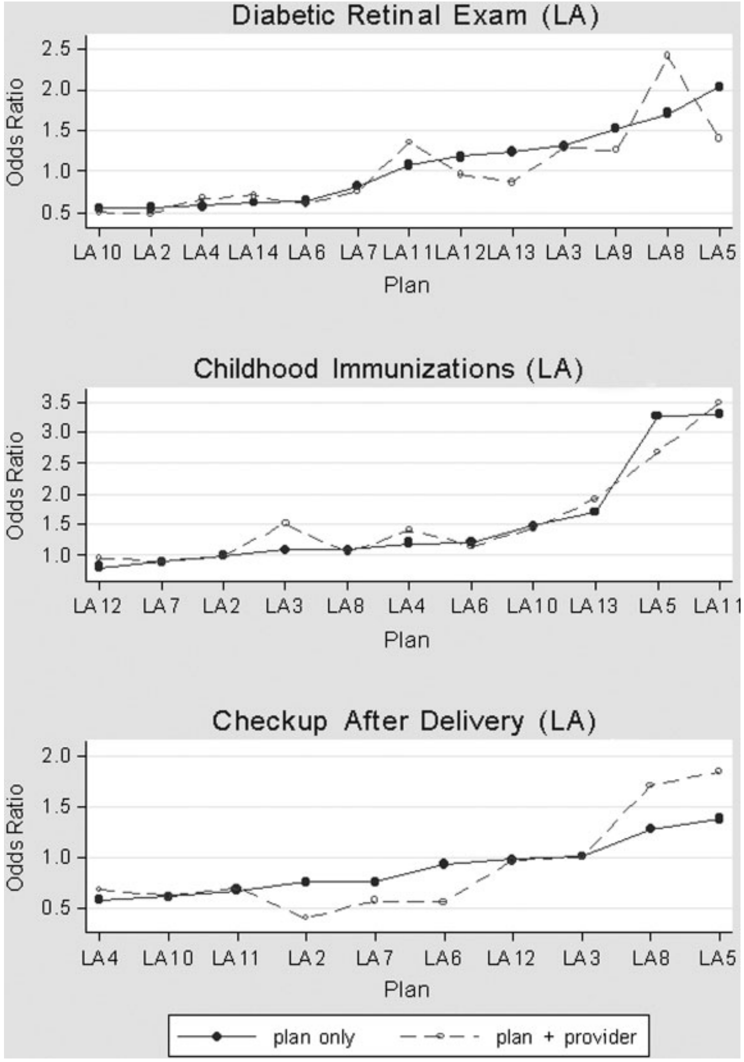

Figure 1 Plan odds ratios, before and after inclusion of provider controls, for diabetic retinal examination, childhood immunizations, and check-up after delivery, in the Los Angeles area.

contrasts with the hypothesis that health plans are 'too far' from patients to impact their care and thus would have minimal influence on HEDIS rates once provider choice was taken into account.

On the face of it, these findings suggest that plans can contribute to health care quality. It is, in fact, quite plausible that plans could have an impact on HEDIS scores by influencing processes of care. Particularly for the preventive services on which we focus, there are a number of activities plans can undertake that could improve scores, including efforts to educate patients, the development of reminder systems, and the use of financial incentives. Plans may also work to educate physicians, although for education efforts to play a role in our findings their impact would have to be specific to an individual plan's patients, whereas it may be more likely that physicians would respond to education campaigns by changing their practices for all of their patients, not just those in the plan conducting the campaign. Regardless of the specific source, if some plans are better than others at improving the delivery of services measured by HEDIS rates for their patients, this would produce variation in plan scores that are independent of the providers with which plans contract.

There are also other explanations for the persistent variations across plans. One derives from the nature of the widely used 'hybrid' HEDIS data collection process. There may be differences in the ability of plans to collect and manage data, either because their administrative databases are more or less complete, or because they are more or less proficient at tracking down and reviewing medical charts. Plans that are not adept at collecting or maintaining data could end up with more cases with missing data, and thus, because HEDIS specifications count cases with missing data as negative, with worse rates. This, in turn, could lead us to observe variations in scores due to variations in data collection ability, rather than the ability of plans to actually influence processes of

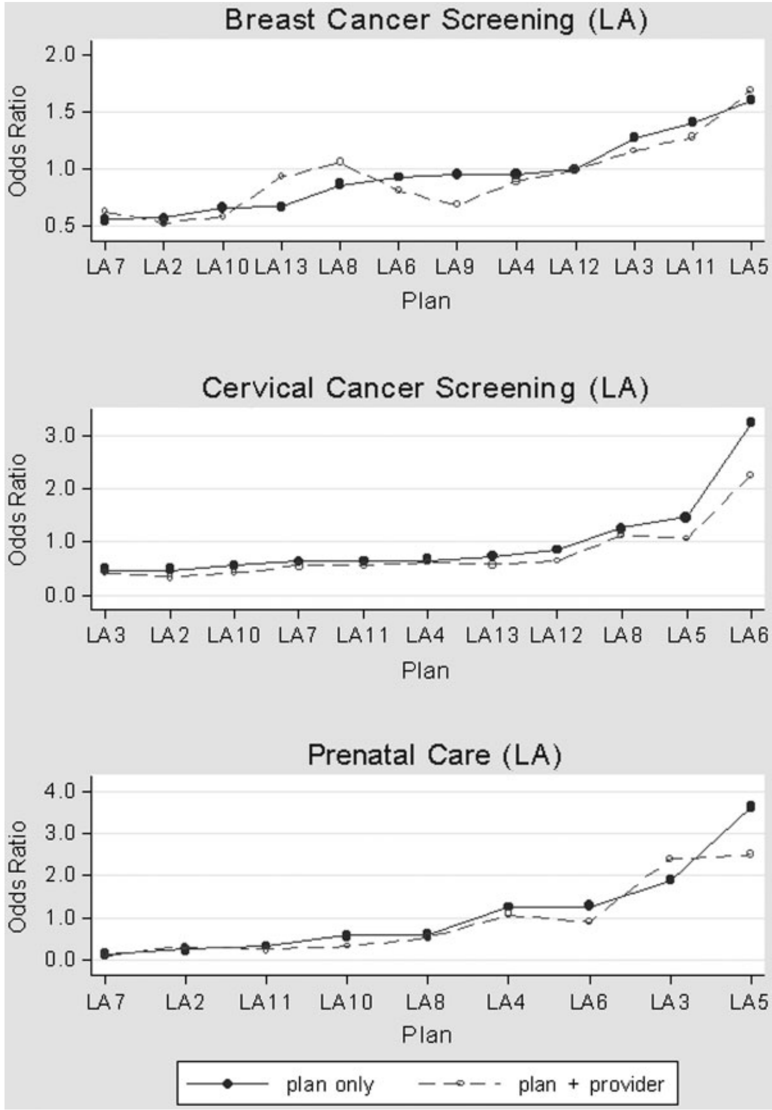

Figure 2 Plan odds ratios, before and after inclusion of provider controls, for breast cancer screening, cervical cancer screening, and prenatal care, in the Los Angeles area.

care. In fact, a simple analysis of the relationship between plans' published HEDIS scores and their published percentages of sampled cases for which definitive data were obtained [11] shows large positive correlations for five of the six measures on which we have focused here. Put another way, plans that produce definitive data for more of the sampled cases also have higher HEDIS rates, suggesting that data collection abilities can influence scores.

If it is the case that our results are due in part to variations in data collection ability, this need not imply that there is no associated variation in quality of care across plans. It may be that data capture proficiency is a critical prerequisite for a health plan's ability to effectively manage the health status of its membership. If an organization responsible for the health of a given population is unable to effectively identify and document care for a targeted population subset, it is very unlikely that such an organization will be able to effectively manage their care, whether the organization is a health plan or provider group. This would appear to be supported by NCQA's recent attempts to highlight not only HEDIS rates for individual plans, but also the rate of administrative positive data capture for those plans.

There are two other factors that could play roles in our finding of persistent variations in HEDIS scores across plans after controlling for providers. We controlled for provider groups, not individual physicians. Moreover, because of our sampling algorithm, all of the provider organizations represented in our sample are relatively large. If it is really individual physicians or small groups of physicians that influence care, we might not observe this because our data on providers are not sufficiently detailed.

Finally, there may be differences in patient characteristics across plans. The official HEDIS specifications do not allow for risk adjustment. The HEDIS measures we studied, however,

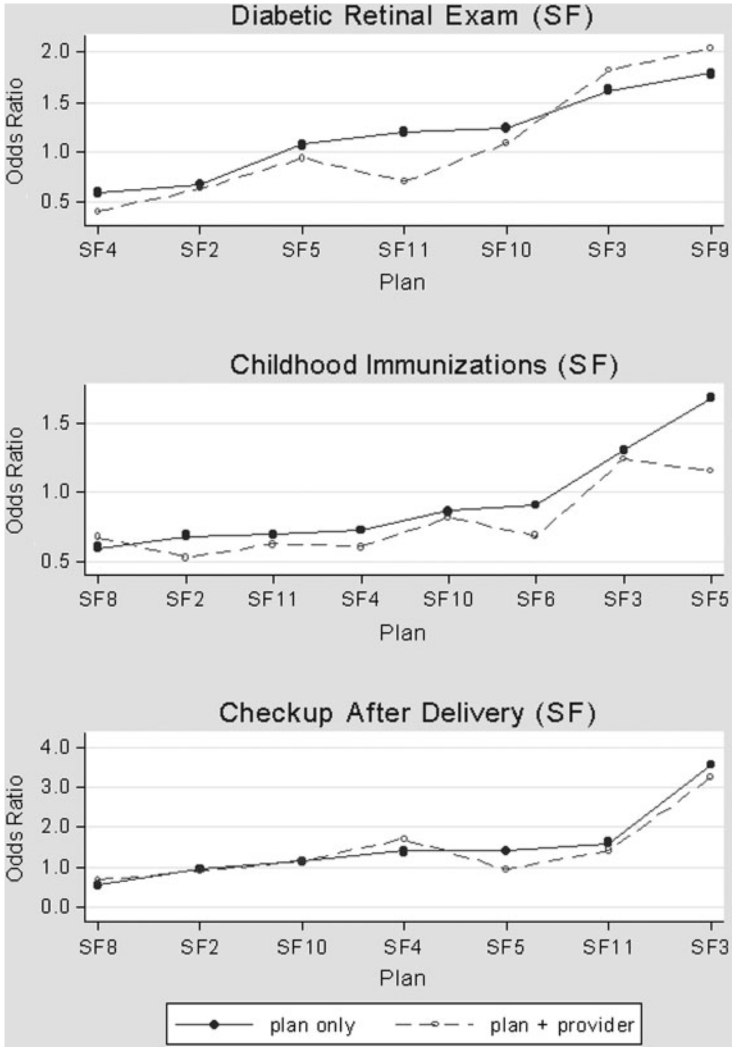

Figure 3 Plan odds ratios, before and after inclusion of provider controls, for diabetic retinal examination, childhood immunizations, and check-up after delivery, in the San Francisco area.

were selected to minimize bias in measures across plans from variation in patient characteristics, so it is not clear that this would lead to substantial bias in our findings. Furthermore, we included controls for patient age and sex, which will reduce difficulties associated with any patient selection. It nonetheless remains possible that there are unmeasured differences across plans in the characteristics of their enrollee populations, and this could be associated with differences in plan scores.

Even in the presence of alternative explanations, we find the fact that there are persistent variations in plan scores after controlling for providers intriguing. In particular, this finding leaves distinctly open the possibility that plans could have an impact on quality of care independent of providers, in sharp contrast with hypotheses that plans can have little to do directly with quality of care. We stress, nonetheless, that this is the first analysis of which we are aware that presents data on the roles of plans and providers in determining quality, and further efforts are clearly needed to investigate the importance of alternate explanations.

Our findings suggest that continued attention to plans in quality data collection efforts is warranted. Although there are sometimes calls to concentrate data collection efforts on providers rather than plans, the existence of persistent variation across plans, even with caveats, suggests that moving away from plan-level data collection is premature. We would also note that an independent contribution of health plans to quality is not the only thing that makes collecting data at the plan level useful. It could make sense to collect data at the plan level in order to hold plans accountable to employers or others who purchase their services, even if some or all of the impact of plans were due to the set of providers that they contracted with rather than plan actions *per se*.

Our findings should do nothing to deter continued efforts to collect data on provider groups. In fact, our results also indicate that provider groups have impacts on measured quality, independent of plans. Efforts to identify the best provider groups could continue to help patients select providers and help plans identify the best groups with which to contract.

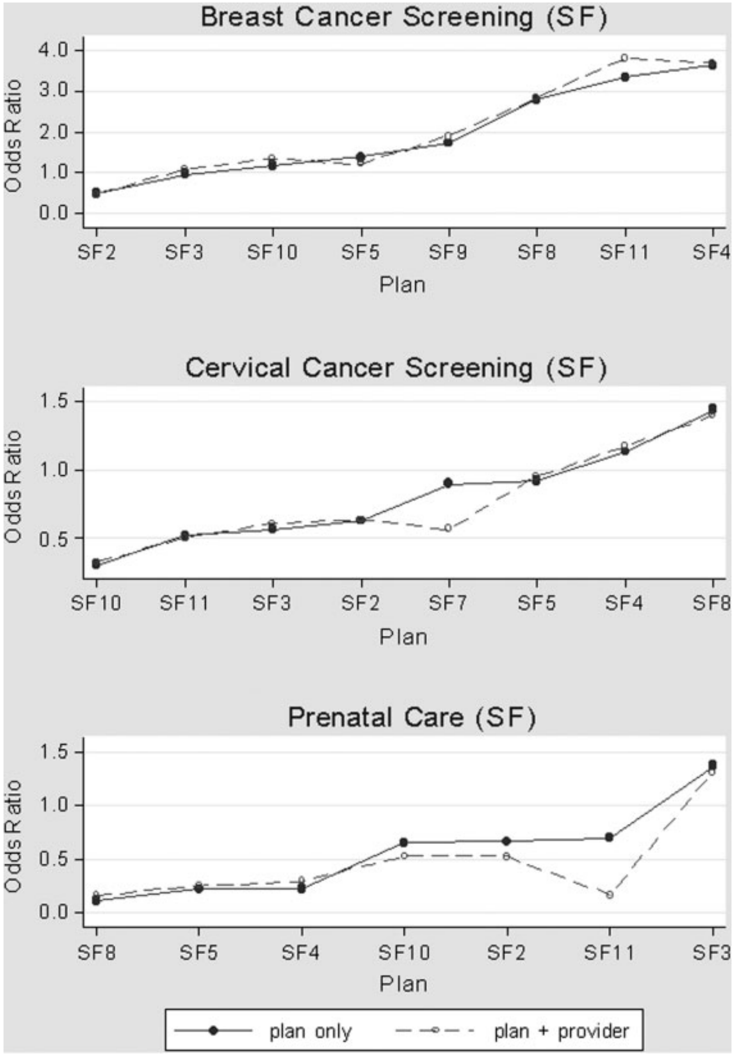

Figure 4 Plan odds ratios, before and after inclusion of provider controls, for breast cancer screening, cervical cancer screening, and prenatal care, in the San Francisco area.

Another issue raised by the findings that plans can independently influence HEDIS scores is whether it would still benefit plans to focus on careful selection of the providers with which they contract. Our analysis in no way suggests that plans should rely on their own efforts to improve quality to the exclusion of efforts to identify and contract with the best provider groups. In fact, given the independent impacts of plans and providers, our results would suggest that improvements in HEDIS scores could be obtained if the best plans and providers were to systematically work together.

There is one limitation of this work that we wish we could better address. A question that naturally follows our analysis is whether plans are more or less important than providers. We are not able to offer information about the relative importance of plans and providers as determinants of HEDIS scores. After experimenting with a variety of methods, we have not been able to obtain clear and rigorous answers to this question.

Finally, we stress that the measures we studied are for a specific set of six preventive care services. It may be easier for plans to have an impact on some services than on others. For example, there may be more mechanisms by which plans can affect the provision of preventive care services than services associated with acute events or with other measures of quality. Plans may be able to influence primary care providers in different ways than specialists. More generally, it is not clear that the presence of persistent plan effects here necessarily implies that they would also be observed if other services were studied.

Overall, we believe that our finding of significant associations between plans and HEDIS rates independent of providers should lead to continued attention to collecting data on health

Table 5 Summary measures of odds ratio changes

	Los Angeles				San Francisco			
	Correlation between odds ratios from plan only and plan+ provider models	Average coefficient move toward 0 (scale)[1]	LR χ^2 for joint significance of plan odds ratios in…[2]		Correlation between odds ratios from plan only and plan+provider models	Average coefficient move toward 0 (scale)[1]	LR χ^2 for joint significance of plan odds ratios in…[2]	
			plan only model (*P*)	plan+prov model (*P*)			plan only model (*P*)	plan+prov model (*P*)
Check-ups after delivery	0.89	–0.18 (0.26)	18.74 (0.044)	37.42 (<0.001)	0.96	0.08 (0.46)	14.32 (0.046)	10.34 (0.170)
Cervical cancer screening	0.99	–0.08 (0.52)	30.85 (<0.001)	31.93 (<0.001)	0.94	–0.04 (0.44)	15.04 (0.058)	12.93 (0.114)
Childhood immunizations	0.96	–0.02 (0.38)	42.50 (<0.001)	32.04 (<0.001)	0.90	–0.04 (0.33)	6.05 (0.642)	6.75 (0.563)
Diabetic retinal exams	0.81	0.04 (0.40)	138.53 (<0.001)	54.79 (<0.001)	0.94	0.01 (0.35)	17.48 (0.015)	28.22 (<0.001)
Mammography screening	0.90	0.00 (0.28)	28.87 (0.004)	22.29 (0.034)	0.99	–0.03 (0.66)	20.77 (0.008)	19.51 (0.012)
Prenatal care	0.93	–0.02 (0.87)	49.13 (<0.001)	45.56 (<0.001)	0.88	–0.16 (0.94)	13.04 (0.071)	9.66 (0.209)

[1]The scale value is the average change that would be observed if all of the coefficients moved to 0.
[2]LR = likelihood ratio.

plan quality, and further discussion and analysis of the role of plans and providers in influencing care.

Acknowledgements

We wish to express our appreciation the California Cooperative Healthcare Reporting Initiative for providing the data used in this study. This project was supported by the California Health Care Foundation and the Pacific Business Group on Health.

References

1. Bodenheimer T. The American health care system: the movement for improved quality in health care. *N Engl J Med* 1999; **340**: 488–492.
2. Brook RH, McGlynn EA, Cleary PD. Measuring quality of care. *N Engl J Med* 1996; **335**: 966–970.
3. Blumenthal D. Quality of care: what is it? *N Engl J Med* 1996; **335**: 891–893.
4. Chassin MR. Improving the quality of care. *N Engl J Med* 1996; **335**: 1060–1062.
5. Chassin MR, Galvin RW. The urgent need to improve health care quality. *J Am Med Assoc* 1998; **280**: 1000–1005.
6. Eddy DM. Performance measurement: problems and solutions. *Health Aff* 1998; **17**: 7–25.
7. McGlynn EA. Six challenges in measuring the quality of health care. *Health Aff* 1997; **16**: 7–21.
8. National Committee on Quality Assurance. *HEDIS 2000*. Washington, DC: NCQA, 2000.
9. Galvin RS. Are performance measures relevant? *Health Aff* 1998; **17**: 29–31.
10. Himmelstein DU, Woolhandler S, Hellander I, Wolfe SM. Quality of care in investor-owned vs. not-for-profit HMOs. *J Am Med Assoc* 1999; **282**: 159–163.
11. California Cooperative Healthcare Reporting Initiative. *1998 Report on Quality*. San Francisco, CA: California Cooperative Healthcare Reporting Initiative, 1998.

Accepted for publication 24 September 2003